

**UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO**  
**CENTRO TECNOLÓGICO**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA**

**MARCOS ALÉCIO SPALENZA**

**APLICAÇÃO DE SELEÇÃO DE  
CARACTERÍSTICAS PARA MELHORIA DE  
UM SISTEMA SEMIAUTOMÁTICO DE  
AVALIAÇÃO DE QUESTÕES DISCURSIVAS**

**VITÓRIA-ES**  
**2017**

**MARCOS ALÉCIO SPALENZA**

**Dissertação de MESTRADO**

**- 2017**

**MARCOS ALÉCIO SPALENZA**

**APLICAÇÃO DE SELEÇÃO DE  
CARACTERÍSTICAS PARA MELHORIA DE  
UM SISTEMA SEMIAUTOMÁTICO DE  
AVALIAÇÃO DE QUESTÕES DISCURSIVAS**

Dissertação apresentada ao Programa de Pós-Graduação em  
Informática do Centro Tecnológico da Universidade Federal  
do Espírito Santo, como requisito parcial para a obtenção de  
Título de Mestre em Informática.

Orientador: Prof. Dr. Elias de Oliveira

Co-orientador: Prof<sup>ª</sup>. Dr. Márcia Gonçalves de Oliveira

**VITÓRIA-ES**

**2017**

Para fazer a ficha catalográfica, que ficará no lugar desta página, na versão final, ligue ou escreva para Arlete Franco

Bibliotecária/BC/UFES

franco@bc.ufes.br

27 4009-2405

Há um formulário específico para solicitar tal ficha, que é enviado à bibliotecária via e-mail

(ver [www.ele.ufes.br/paginappgee.html](http://www.ele.ufes.br/paginappgee.html))

**MARCOS ALÉCIO SPALENZA**

**APLICAÇÃO DE SELEÇÃO DE CARACTERÍSTICAS EM UM  
SISTEMA SEMIAUTOMÁTICO DE AVALIAÇÃO DE QUESTÕES  
DISCURSIVAS**

Dissertação submetida ao programa de Pós-Graduação em Informática do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para a obtenção do Grau de Mestre em Informática.

Aprovada em 20 de abril de 2017.

**COMISSÃO EXAMINADORA**

---

Prof. Dr. Elias de Oliveira

Universidade Federal do Espírito Santo

Orientador

---

Prof<sup>a</sup>. Dr. Márcia Gonçalves de Oliveira

Centro de Referência em Formação e em Educação a Distância

Co-orientadora

---

Prof. Dr. Davidson Cury

Universidade Federal do Espírito Santo

---

Prof. Dr. Pável Calado

Instituto de Engenharia de Sistemas e Computadores - Investigação  
e Desenvolvimento em Lisboa

*Dedico esse trabalho aos meus pais, Marcos e Sirlene, e ao meu irmão, Murilo, pelo apoio em todos os momentos, inclusive durante esse tempo de imersão nos estudos.*

*Sem a ajuda deles, nada seria possível.*

## **Agradecimentos**

Agradeço a Deus.

Agradeço aos meus pais, Marcos e Sirlene, pela oportunidade e suporte incondicional.

Agradeço ao meu irmão, Murilo, pela grande amizade e por me apoiar sendo sempre presente.

Agradeço meus tios, Vanderlei e Solange, que muito me ajudaram, possibilitando meus estudos em Vitória.

Minha gratidão ao Professor Elias de Oliveira, pela disponibilidade e por cada orientação acreditando no andamento desse trabalho.

Agradeço a Professora Márcia Gonçalves de Oliveira, por todas orientações, contribuições e dicas.

Agradeço a todos os amigos, principalmente os que fiz durante o mestrado, por cada conversa, incentivo e desafio passados.

Agradeço a todos os membros do PPGI e do LCAD, pelo suporte, pelos conhecimentos adquiridos e pelo profissionalismo ímpar.

Também sou grato à todos que fizeram parte, direta ou indiretamente, da minha formação.

*“I don’t believe in pessimism. If something doesn’t come up the way you want, forge ahead.*

*If you think it’s going to rain, it will.”*

*(Clint Eastwood)*



# Publicações

Como parte desse trabalho, foram desenvolvidos e publicados os seguintes trabalhos, no período do mestrado, que apresentam, em maior ou menor grau, relação com o tema proposto.

Publicação de trabalhos em anais de congressos:

- Spalenza, M. A.; Nogueira, M. A.; Oliveira, M. G.; Oliveira, E. **Uso de Mapa de Características na Avaliação de Textos Curtos nos Ambientes Virtuais de Aprendizagem**. XXVII Simpósio Brasileiro de Informática na Educação (SBIE). Uberlândia, MG, Brazil. 2016.
- Spalenza, M. A.; Nogueira, M. A.; Oliveira, M. G.; Oliveira, E. **Construção de Mapas de Características em Classes de Respostas Discursivas**. XXI Congreso Internacional de Informática Educativa (TISE). Santiago, Chile. 2016.

## *Resumo*

O processo de avaliação é muito importante para a verificação de aprendizagem dos estudantes e manutenção do andamento do ensino conforme o currículo previsto. A avaliação formativa, garante ao professor a visão geral do ensino de forma a medir a efetividade do aprendizado. Porém, conforme é ampliado o acesso à educação, a aplicação de métodos avaliativos se torna ainda mais necessária e a carência de apoio em sua aplicação é vista como um fator limitante.

O suporte computacional dos métodos pedagógicos visa melhorar a qualidade dos materiais para impactar diretamente na aprendizagem. Neste trabalho, apresentamos uma ferramenta de apoio à correção de respostas discursivas curtas segundo os termos e sua respectiva representatividade por classe de nota. A adoção da seleção de características proporciona a criação de modelos de avaliação mais próximos ao critério do professor. Com os trechos de respostas identificados esperamos padrões que tornem o sistema um avaliador em potencial. Assim, o conjunto da modelagem do método avaliativo e seu *feedback* para suporte ao ensino é aqui chamada de mapas de características.

**Palavras-chave:** Mapa de Características, Avaliação Assistida por Computador, Mineração de Dados Educacionais, Aprendizado Semiautomático, Extração de Informação.

## *Abstract*

The evaluation process is an important step for student's learning verification to keep the teaching program as expected. The formative evaluation ensures the teacher a general view of the lesson to determine their effectiveness. Therefore, according to the improvements in education access, the evaluation methods have become increasingly needed and its support deficiency should be seen as a limiting factor.

The computational support for pedagogical methods want to improve the quality of materials causing a direct and positive impact on learning. In this work, we present a tool to supports short answer evaluation based on terms and its representativeness inside grade classes. Feature selection adoption provides models closer than teacher's criteria. With identified terms as patterns, turns system into a potential answer evaluator. Thus, the criteria modelling and its feedback to teaching support is called features map.

**Keywords:** Features Map, Computer Assisted Assessment, Educational Data Mining, Semiautomatic Learning, Information Extraction.

# Lista de Figuras

2.1	Tradução da categorização feita por Burrows, Gurevych e Stein (2015) de cada tipo de atividade segundo seu modelo de extração de conhecimento. . .	27
2.2	Tradução da descrição feita por Bailey e Meurers (2008) das possibilidades computacionais de processamento das questões discursivas e a limitação dado ao escopo da atividade. . . . .	39
3.1	Arquitetura do sistema de extração de dados dos AVA. . . . .	45
4.1	Linha de execução do G.A. na otimização de cada classe de nota quanto ao número de características e a densidade dos agrupamentos. . . . .	59
4.2	Estrutura do G.A. para seleção de características nas classes de nota. . . . .	59
4.3	Visão geral do mapa de características como resultado da atividade Tec-II-5-186. . . . .	63
4.4	Visualização do mapa de características como resultado individual no campo de <i>feedback</i> do professor. . . . .	64
5.1	Erro Médio Absoluto - MAE para os sete experimentos encontrados no trabalho tomado como referência, o sistema aqui apresentado e cada avaliador, utilizando como <i>baseline</i> o erro entre os próprios especialistas. . . . .	70

5.2	Desvio Padrão do Erro - SEM para os sete experimentos encontrados no trabalho tomado como referência, o sistema aqui apresentado e cada avaliador, utilizando como <i>baseline</i> o erro entre os próprios especialistas. . . . .	71
5.3	Similaridade interna por classe de nota em atividades de disciplinas da UFES antes e depois da seleção de características. . . . .	73
5.4	Mapa de características de todos os alunos para a atividade <i>DS CC UNT 1-5</i> . . . . .	78
5.5	Mapa de características de todos os alunos para a atividade <i>DS CC UNT 9-7</i> . . . . .	80
5.6	Mapa de características de todos os alunos para a atividade <i>DS CC UNT 2-6</i> . . . . .	81
5.7	Mapa de características de todos os alunos para a atividade <i>DS CC UNT 7-5</i> . . . . .	82
5.8	Mapa de características de todos os alunos para a atividade <i>Tec-II-5-170</i> . . . . .	84
C.1	Gráficos para aumento de similaridade por nota. . . . .	130
C.2	Gráficos para aumento de similaridade por nota das atividades da base de dados de Disciplinas da UFES. . . . .	131

# Lista de Tabelas

2.1	Propriedades para distinção dos formatos de atividades textuais. . . . .	29
5.1	Informações de treinamento da base de dados após particionamento durante a análise de distribuição na etapa de <i>clustering</i> . . . . .	76
5.2	Avaliação da classificação para cada atividade da base de dados conforme 6 métricas e dois classificadores antes e depois da seleção de características. .	77
A.2	Calibração do Algoritmo Genético conforme os resultados da seleção de características para quatro classes [100, 93, 90 e 0]. . . . .	111
B.1	Lista de <i>stopwords</i> utilizado no pré-processamento da língua inglesa. . . . .	113
B.2	Lista de <i>stopwords</i> utilizado no pré-processamento da língua portuguesa. . .	114
C.2	<i>Datasets</i> coletados em disciplinas da Universidade Federal do Espírito Santo com as respectivas informações de esforço de correção e redução de dimensionalidade. . . . .	122
C.4	Avaliação da classificação para cada atividade da base de dados conforme 6 métricas e dois classificadores. . . . .	125
C.6	Similaridade interna anterior e posterior à seleção de características em atividades da base de dados com mais de uma classe de nota. . . . .	129

---

C.8	<i>Datasets</i> coletados por (MOHLER; BUNESCU; MIHALCEA, 2011) em turmas de Ciência da Computação durante as aulas de Estrutura de Dados na Universidade do Norte do Texas. . . . .	134
C.10	Avaliação da classificação para cada atividade da base de dados conforme 6 métricas e dois classificadores. . . . .	145
C.12	Similaridade interna anterior e posterior à seleção de características em atividades da base de dados com mais de uma classe de nota. . . . .	163
C.14	Tamanho dos documentos para cada atividade de todas as bases de dados após a etapa de pré-processamento. . . . .	170

# Sumário

<b>1</b>	<b>Introdução</b>	<b>18</b>
1.1	Contexto . . . . .	19
1.2	Motivação . . . . .	20
1.3	Objetivos . . . . .	21
1.4	Metodologia do Trabalho . . . . .	21
1.5	Estrutura do Trabalho . . . . .	22
<b>2</b>	<b>Trabalhos Relacionados</b>	<b>24</b>
2.1	Breve Histórico da Predição de Nota . . . . .	24
2.2	Modelos Preditivos . . . . .	26
2.2.1	Técnicas para Apoio à Avaliação . . . . .	29
2.2.2	Aplicação no Brasil . . . . .	37
2.2.3	Desafios . . . . .	38
2.3	Visualização dos Resultados . . . . .	41



Sumário	15
<b>3 Avaliação Semi-Automática de Questões Discursivas</b>	<b>43</b>
3.1 Coleta de Dados . . . . .	43
3.1.1 Ambientes Virtuais de Aprendizagem - AVA . . . . .	45
3.1.2 Sistema de Transferência de Dados - STD . . . . .	45
3.1.3 Sistema de Apoio - SA . . . . .	46
3.2 Modelo Vetorial . . . . .	47
3.2.1 Similaridade de Cosseno . . . . .	48
3.2.2 Pré-Processamento . . . . .	49
3.3 Algoritmo Genético . . . . .	51
3.4 Aprendizado Semi-Supervisionado . . . . .	52
3.4.1 Combinando Clusterização e Classificação . . . . .	53
3.4.2 Identificação de Treino e Teste . . . . .	54
3.4.3 <i>K-Nearest Neighbors - K-NN</i> . . . . .	55
3.4.4 <i>Centroid Based Classifier - CBC</i> . . . . .	55
3.4.5 Métricas de Avaliação . . . . .	56
<b>4 Mapa de Características</b>	<b>58</b>
4.1 Identificação das Características Relevantes . . . . .	58
4.2 <i>Feedbacks</i> e a Visualização das Características . . . . .	62

Sumário	16
<b>5 Experimentos e Resultados</b>	<b>66</b>
5.1 Bases de Dados . . . . .	66
5.1.1 Base de Dados do Vestibular UFES . . . . .	67
5.1.2 Base de Dados das Disciplinas UFES . . . . .	67
5.1.3 Base de dados da Universidade do Norte do Texas (Inglês) . . . . .	67
5.1.4 Base de Dados do Concurso ASAG-SAS no <i>Kaggle</i> (Inglês) . . . . .	68
5.2 Experimentos Realizados . . . . .	69
5.2.1 Experimentos com o Vestibular da UFES . . . . .	69
5.2.2 Experimentos com as Disciplinas da UFES . . . . .	72
5.2.3 Experimentos com o <i>Texas Dataset</i> . . . . .	74
5.2.4 Experimentos com Dados do <i>Kaggle</i> . . . . .	75
5.2.5 Experimentos com Mapa de Características . . . . .	76
5.3 Discussão . . . . .	83
<b>6 Conclusões e Trabalhos Futuros</b>	<b>87</b>
<b>A Algoritmo Genético</b>	<b>95</b>
A.1 Calibração com a base de dados TEC-II_5-8 . . . . .	95
<b>B Processamento de Linguagem Natural</b>	<b>112</b>
B.1 Lista de <i>Stopwords</i> . . . . .	112

Sumário	17
<b>C Experimentos</b>	<b>115</b>
C.1 Base de Dados VestUfes . . . . .	115
C.1.1 Questão 1 . . . . .	115
C.1.2 Questão 2 . . . . .	116
C.1.3 Questão 3 . . . . .	118
C.1.4 Questão 4 . . . . .	118
C.1.5 Questão 5 . . . . .	120
C.2 Base de Dados Disciplinas UFES . . . . .	121
C.3 Base de Dados Ciência da Computação UNT . . . . .	132
C.4 Caracterização das Questões Curtas . . . . .	164

# Capítulo 1

## Introdução

As avaliações de aprendizado são fundamentais para o ensino ao apontar o desempenho da turma diante dos conteúdos. Com aplicações frequentes, as atividades permitem ao professor interagir com os alunos e com os materiais pedagógicos para reformulação e aperfeiçoamento da sua metodologia. Desse modo, é com o acompanhamento da disciplina e o apoio ao educando que as atividades formativas permitem a reformulação do processo de ensino-aprendizagem (BARREIRA; BOAVIDA; ARAÚJO, 2006). Por meio das atividades podemos identificar o domínio dos estudantes sobre o contexto e sua capacidade de realizar inferências sobre o assunto. O papel da avaliação, portanto, é diagnosticar, apreciar e verificar a proficiência dos alunos para que o professor atue no processo de formação de modo a consolidar o aprendizado (OLIVEIRA; SANTOS, 2005).

Ao observar problemas no modelo de ensino-aprendizagem, o professor pode agir de forma a contorná-los, personalizando a estrutura curricular. É através das atividades, portanto, que observamos paralelos do conhecimento individual dos alunos. Dessa forma, a mediação tecnológica para a criação, avaliação, recomendação e visualização em dados educacionais visa apoiar o professor na melhoria e no acompanhamento do currículo do aluno (PAIVA et al., 2012). Com ferramentas de apoio, o tutor pode verificar a aptidão dos estudantes, de forma individual ou coletiva, para adequá-los à disciplina.

## 1.1 Contexto

A aplicação das atividades visa orientar o professor a agir de forma corretiva na disciplina para administrar o conteúdo assimilado pelos alunos. Segundo Bezerra (2008), as verificações de aprendizado podem ser divididas em três modelos para diferentes fundamentos do conhecimento: atividades de identificação, de inferência e de opinião. As questões de identificação requisitam do aluno a localização de trechos do conteúdo apresentado sem remontar a pesquisa realizada na criação do conceito. As questões de inferência realizam exatamente esse laço entre a tarefa a ser realizada e a disciplina, onde o leitor deve associar o conhecimento adquirido em aula na compreensão do contexto apresentado. Por último, as atividades de opinião dependem do contato do avaliado para caracterizar e julgar partes chave do conteúdo, definindo pontualmente sua interpretação das informações.

Cada um dos métodos avaliativos apresentados nas atividades auxilia o professor na verificação da cobertura do conhecimento e do exercício cognitivo. Diante das tarefas exige-se do aluno uma perícia numérica, motora, criativa, linguística ou didática para representar sua aptidão com o conhecimento. Dessa forma, as atividades buscam agir para desenvolvimento e avaliação das habilidades do discente com o conteúdo aplicado.

Dentre as atividades para avaliação do aprendizado podemos destacar as questões discursivas pois se enquadram nos três modelos de verificação citados. Esse tipo de questão têm papel especial por desenvolver habilidades essenciais nos alunos enquanto formulam ideias, interpretam o problema, constroem o texto e põe em prática a leitura. É por esse formato de atividade que diversos esforços são criados para incentivo de práticas para leitura e escrita, como os descritos periodicamente no PNLL<sup>1</sup>. Porém, esse tipo de atividade demanda muito tempo de análise do professor.

---

<sup>1</sup> Plano Nacional do Livro e Leitura / Ministério da Cultura - Ministério da Educação - PNLL/MinC-MEC  
Disponível em: <<http://www.cultura.gov.br/pnll>>

## 1.2 Motivação

A análise textual consome muito tempo do professor sobrecarregando-o. A falta de tempo para correção de atividades reduz a qualidade da docência pois representa um dos principais indicadores de estresse (ASSUNÇÃO; OLIVEIRA, 2009). Enquanto isso, aos alunos, a falta de prática com as atividades causa defasagem nas habilidades básicas, tal como a leitura e escrita. A esse déficit, é atribuído graves problemas de aprendizagem e baixo rendimento acadêmico (MARTENS, 2016). Por isso, é essencial evitar que o tempo do docente seja limitado pelo controle dos métodos de avaliação de aprendizagem, ampliando os planejamentos de conteúdo, aperfeiçoamentos profissionais e a gerência das necessidades dos alunos. Há um empenho portanto para retificar o tempo docente, ou seja, permitir que o tempo reservado para a atividade profissional seja distribuído para as práticas que tenham como foco o aprendizado dos estudantes.

Um método de retificar o tempo docente e contribuir para análise de aprendizado é o emprego de novas tecnologias de apoio ao tutor e a modelagem dos dados educacionais. A tecnologia permite apoiar o desenvolvimento e a modelagem das salas de aula por meio de ambientes virtuais que controlam os materiais e permitem elucidar o professor sobre o desenvolvimento educacional. Conforme Lima e Giraffa (2006), a introdução de sistemas computacionais como recurso pedagógico torna o processo de ensino e de aprendizagem mais efetivo, incentivando a interação dos alunos com a disciplina.

Além da modelagem, os métodos de apoio ao tutor podem suportar o processo de avaliação. Esses coletam informações das atividades para compreender as formas de avaliação do especialista, interpretar os padrões da questão e gerar sua própria avaliação, tornando o professor seu supervisor. Para isso, assim como o proposto, vários sistemas buscam redução do esforço do tutor. O algoritmo desse trabalho, então, coleta padrões de atividade para compreender o modelo avaliativo do professor, aplicando-o no processo avaliativo.

### 1.3 Objetivos

O objetivo desse trabalho é apresentar um modelo de seleção de características que melhore os resultados de sistemas de avaliação de questões discursivas. O modelo utilizado auxilia o especialista com a redução de esforço de correção e identificação dos termos relevantes para compreensão dos documentos de resposta. Por isso, o sistema foi comparado antes e depois do método de seleção com técnicas conhecidas de classificação na tentativa de formar padrões para a atividade.

Os estudos realizados nessa dissertação, em certos aspectos, são uma continuação do trabalho desenvolvido por (PISSINATI, 2014) com o acréscimo de algumas melhorias. Esse trabalho visa identificar informações necessárias do especialista de forma a melhorar a densidade dos grupos de documentos e sua classificação. Tal objetivo geral pode ser elencado como os seguintes objetivos específicos:

- Reproduzir os experimentos realizados por (PISSINATI, 2014);
- Selecionar os itens necessários que aproximem o modelo à correção do especialista;
- Desenvolver sistemas classificadores para comparação da efetividade da seleção de características;
- Verificar a validade da proposta conforme os resultados da avaliação.
- Discutir os resultados obtidos apresentando a eficácia dos modelos;
- Gerar um modelo de *feedback* que apresente aos usuários os modelos criados.

### 1.4 Metodologia do Trabalho

O primeiro passo da metodologia foi realizar a revisão da literatura com o objetivo de identificar o histórico da predição de nota, estudos recentes da classificação desses dados

educacionais e os formatos de visualização de resultados apresentados pela comunidade de pesquisa.

Com essa revisão foram identificadas as técnicas comuns para extração da informações e a predição de notas. Definindo as ferramentas e os métodos podemos verificar as representações de informação associadas a esses modelos preditivos. Verificado o estado da arte da pesquisa tanto em avaliação assistida pela computação e visualização da informação desenvolvemos os algoritmos de predição de nota para questões discursivas curtas.

O sistema foi testado em mais de 50 bases de dados da Universidade Federal do Espírito Santo - UFES, além de bases de dados utilizadas por algumas referências. A maioria dessas, enviadas através de um Ambiente Virtual de Aprendizagem - AVA Moodle<sup>2</sup> e coletadas por um sistema de transferência de informações.

Após os testes nessas bases de dados, por demonstrar-se efetivo na classificação, foi desenvolvido o formato de visualização de informações pelo mapa de características. Os mapas apresentam a forma de identificação do conjunto de respostas utilizada pelo sistema e caracterizam o padrão de avaliação por nota atribuída pelo professor.

## 1.5 Estrutura do Trabalho

Nesse Capítulo são apresentados os conteúdos trabalhados nessa dissertação, descrevendo suas motivações, suas metodologias e seus problemas. Além dessa Introdução, também compõe esse trabalho os seguintes capítulos:

- **Capítulo 2 (Trabalhos Relacionados):** Apresenta trabalhos semelhantes no âmbito da correção automática e do apoio à avaliação do professor.
- **Capítulo 3 (Avaliação Semiautomática de Questões Discursivas):** Define a estrutura do sistema desenvolvido para apoio à verificação de aprendizagem.

---

<sup>2</sup>Modular Object-Oriented Dynamic Learning - Moodle. Disponível em: <https://moodle.net/>



- **Capítulo 4 (Mapa de Características):** Apresenta o modelo de identificação de informações relevantes para seleção de características e suporte ao método de classificação e visualização de resultados.
- **Capítulo 5 (Experimentos e Resultados):** Detalha os experimentos realizados, mostrando os resultados obtidos e os desafios encontrados.
- **Capítulo 6 (Conclusões e Trabalhos Futuros):** Discute os resultados obtidos e as ações posteriores disponíveis para elaborar ainda mais as contribuições desse trabalho.

## Capítulo 2

# Trabalhos Relacionados

Esse capítulo tem como objetivo apresentar os trabalhos publicados que têm abordagens em volta da correção automática e semiautomática de atividades. Dentre os modelos de atividades trabalhados ainda há distinção entre o modelo de escrita livre, como nas redações, e a objetividade das questões curtas alvo desse trabalho. Assim, em torno desse cenário, podemos analisar a influência da forma textual e das técnicas aplicadas em diferentes situações na reprodução do critério de correção.

Várias técnicas foram empregadas para a modelagem de respostas e interpretação dos documentos dos estudantes. O desenvolvimento de formas de descoberta de padrões e/ou busca estatística para compreensão das respostas foram foco do estudo na literatura. Através dos trabalhos elencados em cada metodologia, discutimos os problemas encontrados nas pesquisas inclusive com estudos brasileiros desse tipo de ferramenta.

### 2.1 Breve Histórico da Predição de Nota

A avaliação automática é um processo de classificação de documentos muito comum à partir da inclusão da informática no meio educacional. Um trabalho muito relevante considerado o início desse processo foi o estudo realizado por Page (1966). Nesse, Page descreve

o *Project Essay Grade - PEG*, que consiste no primeiro corretor completo de redações, apresentado com detalhes posteriormente (PAGE; PAULUS, 1968). Além desse tipo de questão, o trabalho iniciou o processo de estudo textual para a avaliação automática, observando apenas os padrões de escrita do aluno. Nesse conjunto observado de padrões estavam o tamanho e quantidade de palavras, frases e parágrafos, número de preposições, conectivos e erros gramaticais, frequência dos sinais gráficos, dentre outros, totalizando 30 variáveis. Basicamente, identificando essas características da escrita em geral e a função das palavras utilizadas pelo autor, o PEG apresentou uma correlação múltipla de 71% das variáveis coletadas para a nota atribuída.

Dessa forma, iniciaram diversas pesquisas para desenvolvimentos dos sistemas inteligentes para apoio ao processo de avaliação. Conforme Wresch (1993), após os trabalhos pioneiros, o avanço da interpretação do texto, com as análises linguísticas, e das técnicas de aprendizado de máquina possibilitaram a criação de melhores ferramentas. Posteriormente, com aplicação de tais técnicas, tornou-se ainda mais evidente a predisposição dos algoritmos no apoio ao processo de correção de questões e avaliação de aprendizagem.

Em 1993, Wresch já destacava o grande aumento do desempenho dos computadores para aplicação desse tipo de reconhecimento. Onde, anteriormente, Page gastava horas de processamento com cartões perfurados foi possível a aplicação em computadores pessoais. Atualmente, com a evolução das técnicas e a especialização dos modelos preditivos, podemos processar grupos de textos em larga escala, sejam redações ou questões discursivas e classificá-las em um curto período de tempo. Além disso, a eficiência do aprendizado nos cursos *online* em massa (GAMAGE; PERERA; FERNANDO, 2016), os MOOC's - *Massive Open Online Courses*, aumentaram o potencial da coleta de dados educacionais e favoreceram a produção de sistemas inteligentes que suportem ao aprendiz e ao tutor.

## 2.2 Modelos Preditivos

À partir do momento que foram desenvolvidos os primeiros métodos, a investigação do conteúdo, os problemas da necessidade de interpretação textual e a integração da computação cognitiva inovaram a pesquisa. A Avaliação Assistida pela Computação (ou *Computer Assisted Assessment*) - CAA (CONOLE; Warburton, 2005) ganhou outros aspectos e foi integrada com duas grandes áreas de pesquisa em Educação: *Educational Data Mining* - EDM (Romero et al., 2010) e *Learning Analytics* - LA (Siemens; Baker, 2012).

Além da preocupação com o processo de correção, a proximidade entre essas pesquisas definiu um eixo de novas abordagens para os dados educacionais. Conforme Spalenza et al. (2016a), CAA é definido como qualquer processo que objetive apoiar a correção e a verificação de aprendizagem como tarefa dos professores. Enquanto isso, EDM geralmente é associado aos métodos de identificação de padrões nos dados educacionais de forma a utilizá-los em benefício do aprendiz. Por fim, LA é formado por um conjunto de técnicas para descoberta de informações e recomendação de aprendizado, lidando diretamente com o aluno e técnicas de ajuste curricular. Então, apesar de distintas em seu objetivo, vários sistemas permeiam entre as áreas e usualmente trabalham com os mesmos conjuntos de informações, quando não funcionam em conjunto.

Segundo Conole e Warburton (2005), a aplicação das tecnologias educacionais para apoio à avaliação de aprendizagem, como as de CAA, adicionam eficiência e eficácia para o processo de ensino. Porém, para cada modelagem é atribuído um tipo de atividade e uma forma de interpretação diferente sobre os dados. Esses diferentes tipos de atividade são apresentados na Figura 2.1.

Dentre as atividades apresentadas na Figura 2.1, podemos identificar seis categorias de avaliação e dois modelos de aprendizado. Dentre as categorias, as três que se enquadram na recuperação de informação (discurso, estrutura e fala) exigem maior esforço computacional na busca de informações que permeiam as tarefas dos alunos. Existem subdivisões como as

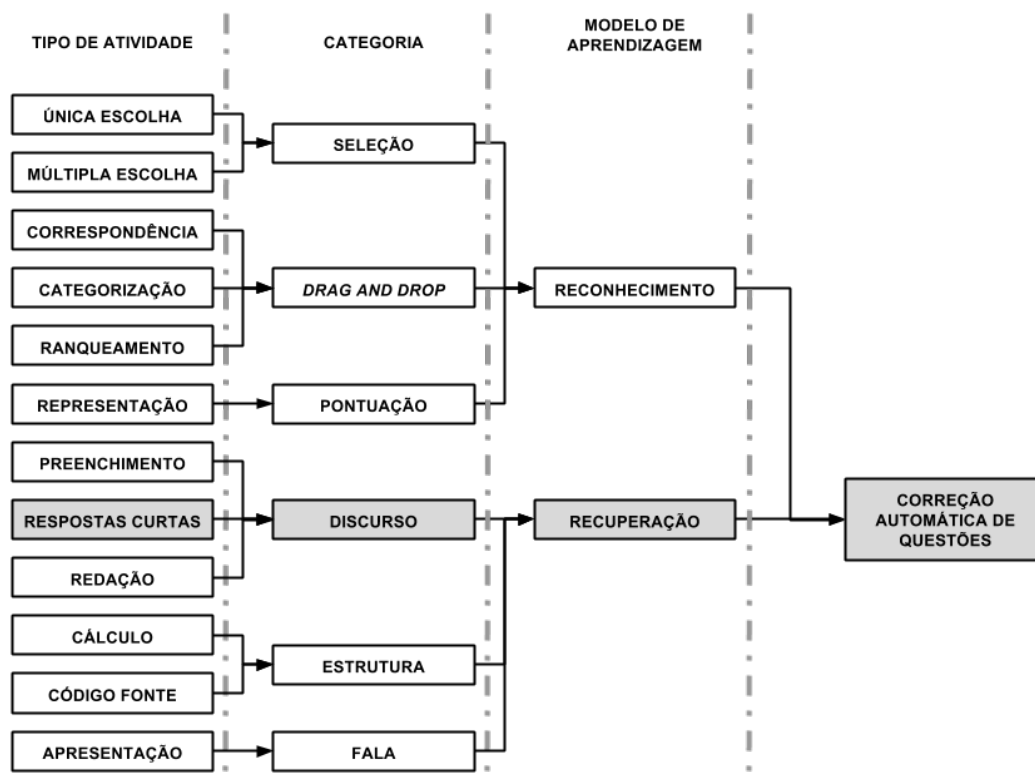


Figura 2.1: Tradução da categorização feita por Burrows, Gurevych e Stein (2015) de cada tipo de atividade segundo seu modelo de extração de conhecimento.

questões de preenchimento, redações e respostas curtas na categoria discurso, pelas diferentes estruturas textuais produzidas pelos alunos. Assim, no nível de automação gerado pelos sistemas, individualmente os tipos de atividade devem passar por análises adequadas para produção de bons resultados.

Estudado por Burrows, Gurevych e Stein (2015), a profundidade do aprendizado, em tradução literal do “*depth of learning*”, divide conforme a literatura, cada tipo de atividade em dois grupos: de reconhecimento ou de recuperação. No Brasil, conhecemos por questões abertas e fechadas, como também foi citado pelo autor. Essa divisão então estabelece a diferença entre as atividades que exploram apenas a necessidade de identificação e organização de conteúdo e as que dependem de construção de ideias visando respostas próprias e originais. Definimos, então, a liberdade do aluno na criação do seu conjunto de resposta como a chave para separar atividades que necessitem de maior ou menor conhecimento factual, ou respectivamente, questões abertas ou fechadas.

Ainda na Figura 2.1, a categoria recuperação apresenta três tipos de atividades que dependem do discurso do aluno para aquisição das informações. Nessas atividades, especificamente, trabalhamos para encontrar soluções que apoiem o professor na interpretação da escrita dos alunos. Porém, se observarmos as subdivisões respostas curtas, redações e de preenchimento, existem diferenças entre as abordagens e os conhecimentos requisitados. Entretanto, muitos sistemas na literatura não descrevem sua atuação entre os Sistemas Automáticos de Correção de Respostas Curtas - *Automatic Short Answer Grader* - ASAG e os Sistemas Automáticos de Correção de Redações - *Automatic Essay Scorer* - AES. Então, a modelagem computacional é um problema amplo apesar da interpretação de cada modelo na visão do professor.

Dentro da especialidade dos modelos preditivos, as questões de linguagem natural podem ser distintas pelo tamanho, foco e liberdade do aluno em sua escrita. O tamanho é o principal fator que distingue redações e textos curtos. Esse fator é relevante pois não é claro o limiar entre eles, onde consideramos redações os textos com muitos parágrafos, Siddiqi, Harrison e Siddiqi (2010) define respostas curtas como “frases de três ou quatro sentenças” enquanto Sukkarieh e Stoyanchev (2009) como algo “próximo à 100 palavras”. Para as atividades de preenchimento essa divisão é visível pois delimitam por espaços as poucas palavras que correspondem ao conteúdo.

Em outra propriedade, o foco, podemos verificar que ao lidar com a produção de texto dos estudantes, as redações são tratadas pelos sistemas nas características de estilo (GÜTL, 2007). Diferentemente das redações, os sistemas ASAG focam no reconhecimento de conteúdo (SHERMIS; BURSTEIN, 2013). Enquanto isso, no modelo de preenchimento, as atividades não baseiam-se no estilo nem no conteúdo em geral, mas sim em palavras específicas. Esse modelo, o de liberdade da questão, é livre de contexto enquanto apenas o essencial precisa ser expresso. Além disso, especificamente no caso das redações, o escritor é livre para a discussão da proposta. Essa liberdade contrasta com os textos curtos, que são inclinados à delimitar o conteúdo abordado no enunciado levando o estudante à uma escrita

mais direta e menos opinativa. Essa modelagem textual é resumida por Burrows, Gurevych e Stein (2015) para cada tipo de atividade, com a versão traduzida apresentada na Tabela 2.1.

Tipo de Atividade	Tamanho	Foco	Liberdade Textual
Preenchimento	De uma palavra até poucas palavras	Palavras	Fixo
Respostas Curtas	De uma frase até um parágrafo	Conteúdo	Fechado
Redações	De dois parágrafos até algumas páginas	Estilo	Aberto

Tabela 2.1: Propriedades para distinção dos formatos de atividades textuais.

Assim, dependendo do modelo de atuação são consideradas determinadas características da produção discente para avaliação do conhecimento. Então, para trabalhar especificamente sobre as atividades textuais dos alunos observamos os modelos de avaliação de respostas curtas e redações, apesar de notável a diferença entre estilos.

### 2.2.1 Técnicas para Apoio à Avaliação

Devido a objetividade implícita nas respostas das atividades discursivas curtas, certos métodos foram desenvolvidos para seu processamento. Apesar do carácter avaliativo do texto ser alvo de pesquisas desde a década de 60, o desenvolvimento de ferramentas específicas para cada uma de suas subdivisões ocorreu pouco depois dos anos 2000.

Os bons resultados dos trabalhos focados em modelos particulares de escrita, mudaram a visão sobre os critérios de correção. Os modelos com análise de padrões de escrita, como as respostas únicas e rígidas aguardadas pelos sistemas, foram melhorados para melhor interpretação do texto e análise do conteúdo. Os novos padrões adotados, alteraram a visão mecânica para modos refinados de compreensão computacional da linguagem, muito mais próximos da tarefa humana. Mitchell et al. (2002) é um dos autores que detalharam a avaliação das respostas curtas. Nesse artigo, o autor apresenta a ferramenta *Automark* na correção de questões de ciência do *Science National Test* através da extração de modelos com análise

sintática e semântica. Porém, muitas técnicas foram apresentadas na literatura para identificação dos conjuntos de resposta de cada atividade.

As técnicas aplicadas para localizar relações entre os documentos criados pelos alunos geralmente baseiam-se em duas grandes pesquisas. Uma delas é o Processamento de Linguagem Natural - NLP, que visa definir regras baseadas nas relações estabelecidas na escrita para validá-las junto ao professor ou o material de apoio. Outra grande área trabalhada é o Aprendizado de Máquina - ML. Com essa segunda, a identificação de relações é feita indiretamente pelos algoritmos através de modelos matemáticos e estatísticos. Estabelecendo como são os padrões de avaliação de um conjunto de dados repassados para calibração das informações espera-se que a máquina seja capaz de comparar cada resultado e realizar a avaliação. Geralmente, as técnicas encontradas na literatura misturam a aplicação dessas para reconhecer a relação entre escrita e os padrões. À seguir, descrevemos em alguns grupos de técnicas a influência do ML e do NLP na avaliação de textos curtos até hoje.

### **Estatística e Base de Dados**

É muito comum para a avaliação dos dados a requisição de um padrão de resposta esperado para reconhecer regras de escrita e relação entre os objetos citados. Para isso são utilizadas bases de dados como uma forma de extração de conhecimentos sobre o tema proposto. Essas técnicas foram integradas após Page e Paulus (1968) empregar estatística pura na pesquisa de avaliação de textos, para realizar uma análise minuciosa das referências encontradas na escrita dos alunos. O objetivo das bases de dados então, é permitir que o *software* tenha conhecimentos em volta do tema trabalhado para ponderar se as afirmações são corretas.

Para estabelecer informações como regras, os sistemas usam relações entre a questão, as respostas e a base de dados. Porém, a base de dados pode ser obtida de diferentes formas: resposta(s) aguarda(s) pelo professor, atividades avaliadas anteriormente por um especialista, bases relacionais (do conteúdo e da linguagem) ou fontes de pesquisa dos alunos. Assim, o uso das bases de dados visam dar maior compreensão em torno do assunto para delinear



o tema, comparando modelos segundo a avaliação do professor. Dentre as pesquisas que partem desse pressuposto, Datar et al. (2013), faz uso da base relacional Wordnet <sup>1</sup> para identificar a relação entre os substantivos encontrados nas respostas. O resultado de sua pesquisa foi o *Essay Grading and Analysis Logic - EGAL*, que faz uso dessa metodologia justamente para encontrar estruturas não articuladas na escrita através do *Gibberish Detector*. Já no processo delineado por Gabrilovich e Markovitch (2007), foram localizados 389,202 termos distintos distribuídos em 241,393 artigos do Wikipedia <sup>2</sup> para relacionar os conceitos encontrados na escrita dos estudantes. Outra abordagem com bancos de dados para aquisição de conhecimento foi utilizada por Kakkonen et al. (2005), com um banco de redações já avaliadas. Interpretando por pares relacionais previamente avaliados o autor gera análises análogas nos textos.

A proposta de Gütl (2007) apresenta muitos dos fatores associados às bases de conhecimento para o sistema, chamado de *e-Examiner*. Nesse trabalho ele criou um repositório com questões e respostas com todas as possíveis fontes para elaboração de atividades, correção automática e a criação de comentários avaliativos. De forma parecida, a base de atividades criada e disponibilizada por Mohler, Bunescu e Mihalcea (2011), foi coletada em turmas de introdução à ciência da computação e estrutura de dados. Para cada atividade discursiva, foi elaborada uma chave de resposta e cada submissão dos alunos foram avaliadas por dois especialistas.

### **Processamento de Linguagem Natural - NLP**

A avaliação das questões discursivas por meio de características da linguagem adentra-se na formação de modelos para aplicação de técnicas de reconhecimento de padrões para avaliar de modo qualitativo e quantitativo a escrita. Ao deparar-se com todas as estruturas que uma linguagem contém, como o português, reconhecidamente complexo, o estudo tende à ser dividido para ser melhor trabalhado. Aos sistemas uma outra linguagem geralmente

---

<sup>1</sup><https://wordnet.princeton.edu/>

<sup>2</sup><https://www.wikipedia.org/>

amplia bastante sua biblioteca ou seu conjunto de dados para compreensão do texto à esse nível. Dessa forma, outras técnicas geralmente não esbarram com esse problema mas não reconhecem diretamente o conteúdo. Conforme Manning e Schütze (1999), para além da análise da sintaxe linguística temos a morfologia que estuda a formação das palavras e a semântica que estuda seu significado. O autor ainda usa a frase de Chomsky “*colorless green ideas sleep furiously*” para representar que as figuras de linguagem podem acrescentar complexidade para a compreensão linguística por meio computacional.

Porém, a percepção de que há essa diversidade de informações possíveis através da escrita dos alunos tornou foco o estudo do conteúdo. Burstein et al. (1998) fez uso de NLP no sistema *e-rater* para compreender a adequação de textos escritos em testes de proficiência em língua estrangeira. Com a aplicação em provas como o GMAT (*Graduate Management Admissions Test*) e o TWE (*Test of Written English*) utilizados pelo autor, geraram bases de dados de referências situacionais e carentes de avaliação. Esse trabalho ainda serviu como referência para o desenvolvimento das pesquisas de Wang, Chang e Li (2008) para o chinês, Ishioka e Kameda (2004) para o japonês e o proposto por Alfonseca e Pérez (2004) para o espanhol.

O reconhecimento de padrões também esteve muito atrelado a pesquisa de NLP para as redações. A extração de regras e a identificação de conteúdos necessários para o conjunto de respostas permitiu melhorias nos critérios de avaliações conforme existência de trechos chave. Semelhantes as máquinas especialistas, em NLP são criadas estruturas e representações que colaborem com todo o processo de identificação textual. Em geral, trabalhos anteriores com base em NLP extraíam através de regras previamente descritas, como as técnicas estatísticas, o conhecimento inato ao conjunto de dados processado. Hoje, durante a observação textual os sistemas de NLP reconhecem por regras, identificam relações e funções das palavras (PULMAN; SUKKARIEH, 2005) ou o próprio conteúdo apresentado pelo aluno em seu conjunto de resposta (CUTRONE; CHANG; KINSHUK, 2011), gerando análises para suporte ao método avaliativo.

O processo delineado por Pulman e Sukkarieh (2005), utiliza um marcador léxico baseado em Processamento de Linguagem Natural - NLP, conhecido por *tagger*, marcando o texto conforme o *PennTreebank dataset* para a formação de padrões buscados com técnicas de aprendizado de máquina. Cutrone, Chang e Kinshuk (2011) busca modelos na aplicação das palavras, observando as referências em conjunto. Estudo semelhante também foi realizado por Levy et al. (2013), onde para que trechos vinculados sejam encontrados aplica-se inferência léxica e sintática no reconhecimento das construções e palavras que mantêm o entendimento das orações.

Um exemplo que mistura Processamento de Linguagem Natural com reconhecimento de padrões em diversos níveis foi produzido por Bailey e Meurers (2008). Nesse trabalho, cada modelo compara as respostas do especialista com a do aprendiz através de um módulo chamado de *CAM - Content Assessment Module*. Através da árvore de dependências produzida por um *chunker* vinculando as diferentes formas de escrita segundo a compatibilidade com a estrutura aguardada. Assim, cada frase escrita da resposta curta é comparada com a do especialista, nos níveis de *tokens*, estrutural e por dependência.

Dessa forma, reconhecimento de padrões quando associado à regras de NLP atua nos sistemas de avaliação de questões discursivas encontrando padrões nas atividades dos estudantes. Porém, existem trabalhos que contornam a aplicação de NLP através do aproveitamento do esforço humano. A proposta de Kulkarni et al. (2014), por exemplo, utiliza do conhecimento dos alunos com a avaliação por pares para a criação de *rubrics* (ARTER; CHAPPUIS, 2006), ou seja, para elaborar descrições sobre o modelo de correção da questão. Após identificar características relevantes para avaliação, a própria ferramenta torna-se autossuficiente.

### **Análise Semântica Latente - LSA**

O *Latent Semantic Analysis - LSA*, (DEERWESTER et al., 1990) é uma técnica que visa reduzir o conjunto textual através da decomposição vetorial para obter as relações semân-

ticas mais significativas. De forma numérica, portanto, a análise de co-ocorrências textuais retorna referências internas aos grupos de documentos, apresentando de modo sumarizado o contexto presente nas respostas (LANDAUER; FOLTZ; LAHAM, 1998). A análise semântica entregue, apesar dos fundamentos matemáticos apresenta características próximas ao cognitivismo humano com termos bem relacionados aos documentos processados. Esse modelo foi utilizado por Foltz, Laham e Landauer (1999) na construção do *Automated Essay Assessor*. Para isso, o LSA foi treinado com o domínio das atividades tornando a semântica foco da avaliação. Com base nesse treinamento os termos selecionados são componentes inter-relacionados com o contexto dos documentos.

Algumas variantes do LSA foram criadas para melhor atender o problema de avaliação. Em sua proposta, Kakkonen et al. (2005) utilizou o PLSA na melhoria dos problemas encontrados em redações para o método básico. No caso, o autor sugere mudanças no algoritmo tradicional visando contornar a perda de grande parte das informações na seleção de variáveis latentes para muitos documentos. Outros fatores elencados como problemas no método foram limitação do uso de métricas de similaridade e na geração de modelos linguísticos. O *Probabilistic Latent Semantic Analysis* em conjunto com o *Expectation Maximization*, apresentados pelo autor, elaboram modelos linguísticos para uma análise semântica mais profunda do contexto das redações. De forma auxiliar, essa função apresenta resultados conforme dois passos: (I) o de expectativas probabilísticas conforme as variáveis latentes e (II) o de maximização com a atualização das variáveis segundo as probabilidades do primeiro passo. Como meta, esse algoritmo visa construir padrões que melhor caracterizem as redações para uma posterior classificação em técnicas de *Machine Learning*.

O LSA é utilizado na avaliação de textos curtos principalmente pela sua redução de dimensionalidade controlada e a seletividade das características, com *feedback* baseado no contexto (PÉREZ et al., 2005). Isso ocorre porque a verificação de conexão semântica entre os poucos pares de palavras definidos com a transformação vetorial apresenta resultados muito bons em respostas diretas. Principalmente quando a relação semântica é obser-

vada com apoio de bases de conhecimento como a Wordnet para similaridade entre palavras (MOHLER; MIHALCEA, 2009). Mais uma modificação ao modelo básico foi estudada por Gabrilovich e Markovitch (2007) através do *Explicit Semantic Analysis - ESA*. Esse método apoia-se nas técnicas de NLP e uma grande base de conhecimento, o Wikipedia como citado na Seção 2.2.1. Assim, cada conteúdo extraído é associado a conceitos da base de dados. O conjunto de textos é então classificado conforme um mapa de conceitos ponderados, com um interpretador semântico e especialmente mais abrangentes para o domínio da base adotada.

### **Aprendizado de Máquina - ML**

*Machine Learning* é um conjunto de técnicas que se apoiam na identificação de padrões e extração de informação. Baseiam-se em algoritmos computacionais que produzem modelos de acordo com a distribuição dos dados. Na avaliação de textos curtos, essas técnicas visam compreender a forma de organização textual no espaço  $n$ -dimensional para reproduzir o comportamento. Então, quanto melhores as informações repassadas pelo especialista humano, há uma tendência de criação de modelos de avaliação mais representativos.

O uso do teorema probabilístico bayesiano foi estudado por Rudner e Liang (2002) para avaliação de redações. De acordo com as características extraídas com os resultados do LSA (LANDAUER; FOLTZ; LAHAM, 1998) e do PEG (PAGE; PAULUS, 1968), a análise probabilística da rede bayesiana efetua a avaliação no formato da classificação de documentos clássica. Posteriormente, o trabalho realizado por Pulman e Sukkarieh (2005) associou as técnicas de processamento linguístico com o aprendizado de máquina. Foi utilizado *HMM - Hidden Markov Models* treinado como *parser* para reconhecimento das dependências na estrutura da frase. Com a árvore de dependências extraída, os algoritmos *Inductive Logic Programming - ILP*, *Decision Tree Learning - DTL* e *Naive Bayes Classifier - NBC* realizam a avaliação. Porém, para as duas últimas técnicas probabilísticas foram adotadas duas maneiras de análise dos dados. Inicialmente os testes foram realizados sem os dados anotados pelo especialista. Posteriormente, os autores aumentaram o desempenho com indicações da parte

essencial da atividade segundo o domínio do especialista tornando a busca bem definida. Porém, esse esforço atribuído ao especialista é maior do que o de correção.

Outras abordagens foram dadas durante a competição da *Hewlett Foundation* no *Kaggle*<sup>3</sup>, descrita posteriormente na Seção 5.1. Na época, o competidor Conort (2012), utilizou o *ensemble Gxav*, ou seja, os resultados oficiais originam de uma combinação das saídas dos algoritmos de *Machine Learning* na tentativa de obter modelos de classificação mais consistentes do que os individuais. Sua solução foi a quarta colocada da competição. Para alcançar essa posição foram combinados 81 métodos distintos de pré-processamento. Em seu relatório técnico, o autor apresenta a mistura dos métodos *Support Vector Machines - SVM*, *Random Forest - RF* e *Gradient Boosting Machines - GBM* na classificação do *Gxav*. As técnicas, apoiadas em diferentes formas de aprendizado e modelagem, definem por parâmetros de escolha dos resultados a ação tomada para um conjunto geral de atividades. Outro participante da competição, Zbontar (2012), também utiliza de um modo de parrear os resultados dos algoritmos para obter melhores modelos. Utilizou da técnica de *ensemble* chamada *stacking*, para grupos de termos de 4-grams e 6-grams com a combinação dos resultados dos modelos de aprendizado estudados. Como modelos de aprendizado, foram utilizados o *Ridge Learner - RL*, *Support Vector Regression - SVR*, *Gradient Boosting Machines - GBM*, *Random Forests - RF* e o *K-Nearest Neighbors*. Com a combinação dos classificadores em subconjuntos de termos, esse autor conseguiu o terceiro lugar da competição.

A metodologia utilizada por Roy, Bhatt e Narahari (2016) também é um *ensemble* simples. Apenas dois classificadores: um guiado quantitativamente, relacionando a classificação diretamente à frequência das palavras, e um qualitativamente, com ponderações sobre as melhores e piores características encontradas. Dessa forma, as informações vistas com e sem ponderações enfatizam a importância de características em ambos contextos. A resultante de cada avaliação, portanto, é estimada com a atribuição de pesos entre os dois valores obtidos. Outra maneira de observar a influência das características foi proposta por Zehner, Sälzer e Goldhammer (2016). O algoritmo proposto estabelece um método de *clustering* para iden-

---

<sup>3</sup>Kaggle Inc. <https://www.kaggle.com/>

tificação de respostas relacionadas. Uma análise de centroides em  $n$  *clusters* pré-definidos pelo professor é utilizada para identificação de grupos de conceitos. Os conceitos extraídos são avaliados pelo professor para validar com “correto” ou “incorreto” cada centroide de respostas. A classificação final é dada então através da compatibilidade das respostas com os conceitos de cada centroide.

### 2.2.2 Aplicação no Brasil

No Brasil, poucos estudos tratam especificamente da avaliação de questões discursivas. A metodologia mais abordada, é o LSA, presente em muitos estudos da literatura. Nessa abordagem, a proposta de Santos et al. (2012) compara a avaliação humana das características com variações da ponderação de termos e da métrica de similaridade. Posteriormente, de forma mais direta, Santos e Favero (2015) compara as divergências entre a avaliação de dois especialistas paralelamente aos resultados obtidos pelo sistema inteligente. Outro autor Passero, Filho e Dazzi (2016), compara os efeitos do LSA com alguns modelos de análise semântica. Para esse estudo, o modelo com essa técnica teve resultados superiores se comparados com as demais implementações da literatura.

Além do LSA, os estudos nacionais fizeram testes com metodologias alternativas para o problema. Oliveira et al. (2010) desenvolveu um modelo de *clustering* com grande redução de esforço na avaliação da atividades. Nesse artigo, os padrões encontrados em seis itens avaliados pelo professor foram capazes de reproduzir o modelo de correção para as outras 6000 respostas.

A lógica *fuzzy* foi testada por Vilela et al. (2012) na produção do *SCATeDi*. O ambiente desenvolvido é um sistema especialista que associa um conjunto esperado de termos corretos segundo uma construção de regras do professor. A avaliação ocorre conforme a compatibilidade das respostas com os padrões de conteúdo aguardados. O problema associado aos sistemas especialistas é o esforço maior para organizar regras de correção, enquanto

a verificação de completude das respostas, alvo dos demais métodos, geralmente atinge esse objetivo.

O estudo de NLP para análise de respostas discursivas foi alvo da proposta de Júnior, Paiva e Bittencourt (2015). Porém, como é usual na aplicação dessa técnica, concentraram-se no problema da verificação da sintaxe e da morfologia da construção dos alunos. Esse detalhe é incomum nas atividades curtas pela objetividade da escrita e o caráter restrito das informações à serem abordadas. Entretanto, o estudo é muito influente para a verificação minuciosa da língua como apresentado na Seção 2.2.1. Esse ponto de vista da avaliação contribui muito com a análise do discurso e tem influência significativa na correção de redações.

Uma ferramenta que trabalha no âmbito da correção de questões discursivas para ensino de SQL foi descrita por Figueira et al. (2013). Nesse estudo, o autor realiza a avaliação através de regressores lineares para questões discursivas e códigos SQL. O diferencial dessa análise é a busca pelo comportamento linear das notas atribuídas pelo especialista aos documentos. Nesse método de *Machine Learning*, o coeficiente linear que define cada característica encontrada nos textos dos alunos relaciona-se diretamente à nota recebida.

### 2.2.3 Desafios

Apesar de ser um estudo encontrado na literatura há décadas, dentro da avaliação automática de questões discursivas existem desafios que permeiam as técnicas utilizadas e a interpretação do conteúdo. Nos primeiros sistemas, a modelagem de questões discursivas era um trabalho realizado com o texto bruto. A partir disso, a busca por equivalência entre resposta e texto dos estudantes falhou por inúmeras vezes na padronização dos documentos e na identificação de sinônimos (LEFFA, 2003). O estudo dessa problemática derivou discussões em torno da identificação do conhecimento na escrita dos estudantes, realizada atualmente em boa parte dos algoritmos.

A liberdade dada ao aluno, típica das atividades discursivas, permite que todo tipo de



texto seja apresentado como resposta, incluindo informações inesperadas. Toda informação que fuja do padrão das respostas é considerado um *outlier*. Esses dados confundem os algoritmos de correção automática e comumente são a grande causa do nível de erro dos sistemas. Existem várias fontes desse tipo de conteúdo como a falta de objetividade das questões, avaliações não criteriosas dos professores, uso de figuras de linguagem e a inclusão de referências externas, encontrados após a categorização por formas distintas de processamento textual. Bailey e Meurers (2008), por exemplo, ao aplicar modelos de linguagem natural a textos curtos, definiu o comportamento do estilo textual segundo a capacidade de compreensão computacional. O estudo gerou a representação da Figura 2.2, onde especifica a relação da restrição do discurso das questões de preenchimento e a liberdade das redações.

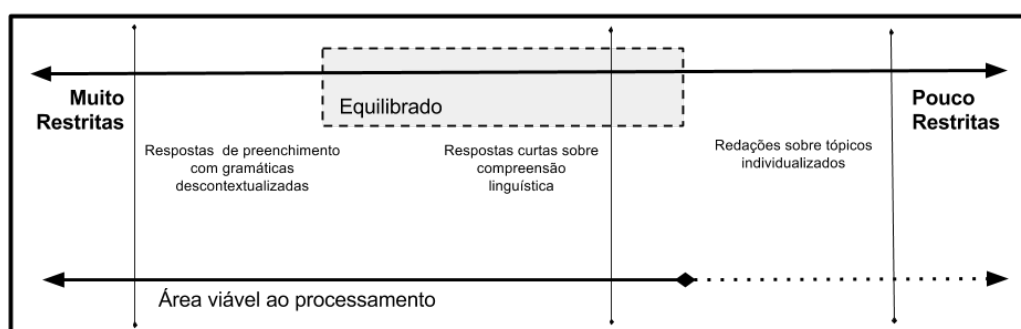


Figura 2.2: Tradução da descrição feita por Bailey e Meurers (2008) das possibilidades computacionais de processamento das questões discursivas e a limitação dado ao escopo da atividade.

Observando o cenário delineado pelo autor, a Figura 2.2 mostra os problemas nas premissas gerados pela falta de contexto nas respostas e inviabilidade da análise de conteúdo pela liberdade completa de alguns estilos de redação. Assim, o ideal para o processamento de textos é a compreensão do discurso contextualizado com controle do conhecimento ali inserido. Porém, além das nuances da escrita livre como as figuras de linguagem (metáforas, pleonasmos, hipérboles, etc.), o estilo, a padronização, os regionalismos e o sentimento da escrita (negativismo, partidarismo, criticismo, etc.), devemos deixar explícito que definir enunciados que modelem a resposta aguardada faz parte do processo de adequação para a correção automática. Desse modo, definindo questões únicas, focadas e direcionadas ao material estudado é o ideal para esse tipo de análise de dados. Enquanto isso, respostas indi-

vidualizadas, opinativas e irrestritas ao aluno geralmente não seguem o conteúdo e estrutura e dificultam os processos computacionais.

Modelos baseados em *Machine Learning*, como o apresentado no Capítulo 3 desse trabalho, podem ter problemas como as demais aplicações. A criação de modelos irregulares por *overfitting* ou *underfitting* remetem à padrões que são verossímeis porém representam apenas parcialmente a resposta esperada. Isso implica diretamente na avaliação pelo corte ou adição de informações. É comum aos modelos de definição de padrões a tendência ao erro no processamento de *outliers*, pois não deveriam corresponder à nota ou conteúdo ao qual foi classificada pela avaliação professor. Esse caso ocorre, quando o professor aceita respostas não condizentes com o critério ou quando erroneamente o aluno cria alusões ao conhecimento sem realmente citá-lo.

Apesar de ser uma área de pesquisa já consolidada, a falta de *datasets* abertos ainda é uma limitação para a análise dos algoritmos (BURROWS; GUREVYCH; STEIN, 2015). Os dados pessoais presentes no conteúdo e a falta de abordagens tecnológicas para coleta são fatores limitantes. Desse modo, com bases de dados, padrões e métricas distintas cria-se uma barreira para comparação entre sistemas semelhantes.

Na abordagem apresentada neste trabalho, tentamos obter os critérios da classificação do professor. Desse modo, através do processo de seleção de características, contornamos os problemas de representação dos dados e explicamos a distribuição das notas. Gerando uma visualização conseguimos associar termos à notas com modelos avaliativos próximos do especialista e quando a seleção de características não foi efetiva. Assim, podemos avaliar quantitativamente, segundo as métricas de classificação, ao mesmo tempo que é possível interpretar os termos relevantes por nota no papel de usuários do sistema.

O número de *datasets* adquiridos para testes do sistema foram um diferencial. A capacidade do algoritmo de adaptação do pré-processamento para múltiplas linguagens permitiu o uso de bases nacionais e estrangeiras. A falta de bases de dados em português foi contornada com a coleta de dados e suporte aos professores da própria universidade. Enquanto

isso, o uso de bases de dados estrangeiras permitiu a comparação do método semiautomático proposto com os resultados de outros sistemas já relatados na literatura.

## 2.3 Visualização dos Resultados

As técnicas de reconhecimento de padrões, ao serem associadas ao processamento de linguagem natural, frequentemente têm sucesso na representação visual do conhecimento. Alguns trabalhos aplicam dessas técnicas na busca de informações para organizá-las de forma a apoiar o processo de avaliação. No artigo apresentado por Gütl (2007), o sistema avaliou cada aluno segundo as respostas sumarizadas, comparando-as com uma base produzida pelo professor do conteúdo através do módulo *ROGUE*. Ao realizar comparações com essa estrutura, é produzido um *feedback* com visualização dos resultados por meio das características relevantes identificadas durante a correção.

Outro sistema de marcações foi proposto juntamente com o *Willow* (PEREZ-MARÍN et al., 2006). Os autores desse estudo definiram uma forma de relacionamento direto entre grupos de resposta do especialista e a produção escrita dos alunos. Com uso das técnicas de NLP, são feitas marcações no texto após a avaliação das respostas curtas pelo LSA, representando a análise semântica em um relatório com as submissões dos alunos.

O trabalho realizado por Spalenza et al. (2016b) detalha o *mapa de características* como um método de seleção das principais características e encaminhamento de *feedbacks* com as respostas marcadas. Nesse, temos como diferencial a seleção de informações com base apenas na avaliação parcial dada para treinamento dos algoritmos. Dessa forma a representatividade dos termos por nota não é atribuída diretamente pelo professor, mas através do reconhecimento de padrões para trechos de grande relevância. Além de indicar a possível resposta para a questão, as marcações identificam tendências para todas as classes de nota. A representatividade de cada termo nas marcações se dá pela frequência que foi citada pelos alunos por classe.

A visualização dos resultados pode ser caracterizada ainda através de observações não textuais. É o caso do *TreeMap*, utilizado por Pissinati, Oliveira e Oliveira (2014) para caracterizar o desempenho da avaliação nos grupos de sala de aula. Essa ferramenta gera uma visualização da situação das turmas como apoio para a observação de desempenho dos estudantes. Nesse artigo, o autor ainda elenca e compara seu método com as técnicas clássicas da visualização de dados educacionais. O *TreeMap*, assim como demais modelos, colabora para a tomada de decisões e a autoavaliação do professor.

O *mapa de características*, como sistema aplicado para melhoria e suporte da avaliação de questões discursivas curtas, é apresentado posteriormente no Capítulo 4.

## **Capítulo 3**

# **Avaliação Semi-Automática de Questões Discursivas**

Nesse capítulo, apresentamos a abordagem do sistema na correção de atividades discursivas. Definimos então a extração de informações dos AVA, a padronização da escrita, o reconhecimento de padrões de resposta, a avaliação semiautomática de questões discursivas e o retorno de resultados aos participantes da disciplina para após compararmos os efeitos da seleção de características isoladamente. Dessa forma, com a inclusão do mapa de características no sistema, podemos analisar a influência da redução de dimensionalidade na melhoria das formas de classificação e visualização da informação para cada atividade.

### **3.1 Coleta de Dados**

Certas ferramentas de mineração de dados auxiliam a aquisição de conhecimento em inúmeros procedimentos do cotidiano. Entre eles, a educação não é uma exceção, apesar de ser pequena a aplicação nas salas de aula tradicionais. Com o emprego dos Ambientes Virtuais de Aprendizagem e a popularização dos Cursos Online Abertos e em Massa - MOOC's a quantidade de dados produzidos e a necessidade de apoio aos profissionais dessa área se

tornam cada vez maiores. Porém, para processamento desses dados é necessária a extração das plataformas de ensino e o transporte para servidores pré-configurados com os Sistemas de Apoio - SA.

Segundo o site oficial com estatísticas sobre o Moodle<sup>1</sup>, atualmente são 72 mil sistemas cadastrados em 232 países do mundo. Com 95 milhões de usuários, esse AVA é um grande exemplo de produtividade acadêmica que certamente têm alto custo aos profissionais de ensino que, individualmente, avaliam todos os alunos. Com uma estrutura de coleta de informações, tornamos possível o processamento das bases de dados para apoio ao processo de ensino-aprendizagem. Com a extração desses dados, beneficiamos os participantes dos cursos *online* incentivando a discussão de resultados, elaborando relatórios para manter o professor atento ao andamento do curso.

Para habilitar o apoio ao tutor, um Sistema de Transferência de Dados - STD interage diretamente com o banco de dados da plataforma de ensino para coletar os documentos produzidos pelos estudantes. Por meio do *web service*, o STD, interliga ambientes virtuais e o servidor de processamento. Cada SA desenvolvido torna-se um serviço acionado pelo professor na configuração da tarefa, recebendo resultados diretamente na plataforma. Essa arquitetura é apresentada na Figura 3.1 de forma a manter os serviços em alta disponibilidade para todos os AVA.

Como pode ser visto na Figura 3.1, a interação dos dois tipos de servidores ocorre através do *web service* do Sistema de Transferência de Dados. Nessa arquitetura, os dados são extraídos dos Ambientes Virtuais de Aprendizagem para processamento nos Sistemas de Apoio. Cada *software* é descrito em detalhes à seguir.

---

<sup>1</sup> Estatísticas do uso do Moodle. Disponível em: <https://moodle.net/stats/>

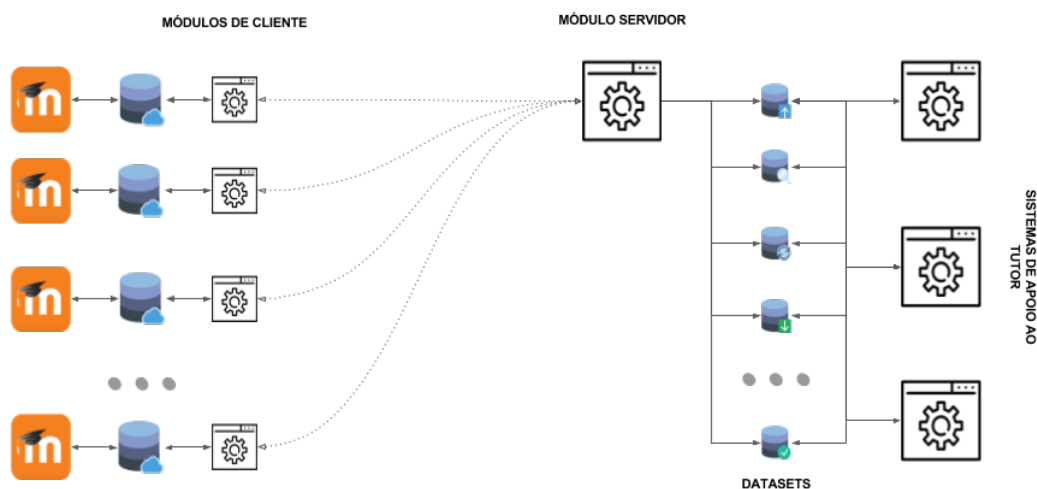


Figura 3.1: Arquitetura do sistema de extração de dados dos AVA.

### 3.1.1 Ambientes Virtuais de Aprendizagem - AVA

Os AVA's são ambientes de suporte ao ensino que modelam as informações para auxiliar na evolução da disciplina. O seu uso permite a organização de atividades e materiais didáticos para os participantes terem amplo acesso ao conteúdo. A alta disponibilidade atrelada a esses sistemas amplia o tempo de disciplina para além das aulas presenciais viabilizando o EaD e a extensão dos planos de estudo.

Além de impactar diretamente na metodologia de ensino e aprendizagem, o uso dos AVA é muito relevante para coleta de dados educacionais. A sua organização deixa à disposição todos os dados, permitindo a criação de ferramentas que observem cada tarefa, nota, aluno ou professor. Com esse tipo de conhecimento, é possível compreender as avaliações, dificuldades e recomendações dos usuários para melhoria de desempenho.

### 3.1.2 Sistema de Transferência de Dados - STD

Os sistemas de transferência de dados são dois módulos integrados aos servidores que conectam os SA aos AVA. As ferramentas interagem com o servidor para transformar informações em arquivos estruturados a serem transportados através de *web services*. Basicamente,

a colaboração entre tais sistemas é o que permite a fluidez das informações em trânsito nos servidores.

Nessa arquitetura, o primeiro módulo é integrado ao AVA e aguarda solicitações de coleta e inclusão de dados. As informações da tarefa, notas, submissões e *feedbacks* são utilizados para obter o conhecimento necessário para as técnicas de mineração de dados acopladas no segundo módulo. Porém, no segundo módulo, os dados coletados são interpretados por um sistema específico já conhecido e designado pelo professor. O SA, com a requisição, observa o comportamento dos usuários e corresponde conforme suas ações.

Dessa forma, os módulos em determinados espaços de tempo interagem para passagem de informações e processamento sob demanda. Os dados são atualizados a cada *download* e os resultados são diretamente enviados às respectivas plataformas de ensino. No AVA, todos os participantes dos cursos recebem simultaneamente os resultados de modo transparente enquanto o tutor pode inserir suas próprias notas e/ou *feedbacks* a qualquer momento.

### 3.1.3 Sistema de Apoio - SA

Os Sistemas de Apoio - SA realizam toda a parte de processamento e atendimento de necessidades específicas aos usuários dos AVA. Com a coleta de dados sendo realizada para todos os ambientes virtuais cadastrados, as bases de dados são processadas pelas ferramentas indicadas pelo professor durante a elaboração da tarefa. Nessa configuração é designado o sistema para apresentar detalhes relevantes sobre a base de dados de forma a suportar os métodos de ensino e aprendizado.

Esse tipo de ferramenta têm como objetivo realizar a mineração de dados para informar aos participantes sobre o andamento do curso. Com relatórios e conhecimentos em torno da atividade ficam evidentes os critérios de avaliação para recuperação de conteúdo, verificação colaborativa e recomendações de aprendizado.

O SA proposto nesse trabalho é o mapa de características, descrito detalhadamente no



Capítulo 4, para redução de dimensionalidade e representação dos dados. Essa ferramenta, além de ser um método de visualização dos resultados ainda facilita os processamentos eliminando características irrelevantes. Tais características, diretamente extraídas da escrita dos alunos, são as informações que o sistema mantém pela ocorrência na avaliação do professor. Sendo automático, a melhoria com visualização dos resultados proporciona aos usuários a discussão colaborativa e classificações mais adequadas.

## 3.2 Modelo Vetorial

Após a coleta e o armazenamento de resultados o sistema controla a aquisição de informações sobre a base de dados. Para isso é utilizado o modelo vetorial descrito por (SALTON; WONG; YANG, 1975). Como forma de interpretação computacional de documentos, o tipo de vetorização adotado ocorre através da frequência dos termos indexados à partir dos textos, ou *Term Frequency - TF* em inglês. Esse modelo se baseia na técnica de processamento de textos “*bag of words*” onde cada índice é associado a um termo distinto encontrado no conjunto de documentos.

No modelo vetorial, cada documento  $d$  de um conjunto de documentos  $D = \{d_1, d_2, d_3, \dots, d_{|D|}\}$  é representado como um vetor de termos  $t$ . Durante a vetorização, é verificada a frequência de ocorrência (TF) de cada um dos  $t$  termos em cada documento  $d$  em  $|D|$ , sendo  $|D|$  o total de documentos desse conjunto. Para cada  $d$ , portanto, é computada a frequência individual dos  $t = \{t_1, t_2, t_3, \dots, t_k\}$  termos que existem na base de dados, sendo  $k$  o número de termos distintos encontrados na coleção. Assim, cada documento  $d$  é representado como um vetor com as frequências individuais  $n$  de cada um dos  $k$  termos distintos, como o exemplo visto na Equação 3.1.

$$d_0 = \{n_{t_1}, n_{t_2}, n_{t_3}, \dots, n_{t_k}\}$$

$$d_1 = \{n_{t_1}, n_{t_2}, n_{t_3}, \dots, n_{t_k}\}$$

$$\vdots$$

$$d_{|D|} = \{n_{t_1}, n_{t_2}, n_{t_3}, \dots, n_{t_k}\} \quad (3.1)$$

Na forma numérica, o documento passa a ser interpretado no espaço vetorial. Com os valores  $n$  de cada característica  $t$  no determinado temos  $n_{t_d}$ , ou seja, o valor da frequência desse termo em um determinado documento  $d$ . Após a transformação, os vetores são comparados segundo a ocorrência dos seus termos, a similaridade dos vetores e a correlação entre as frequências.

Outros modelos com a inclusão de ponderações sobre as características são utilizados para estudo da distribuição dos termos entre documentos ou classes. O método mais tradicional *Inverse Document Frequency - IDF* (BAEZA-YATES; RIBEIRO-NETO, 2008) é um exemplo disso, onde termos que ocorrem em vários documentos recebem peso menor do que termos raros, o que prioriza os que possivelmente são diferenciais entre classes. Neste trabalho porém, foi usada a distribuição original (TF) sem pesos, avaliando a redução de dimensionalidade diretamente sobre a classificação do especialista. Estudos com as várias ponderações devem ser feitas para uma análise da distribuição de cada característica, vendo a adequação dos termos com o interesse desse trabalho, como as esperadas pelo usuário através de sua avaliação.

### 3.2.1 Similaridade de Cosseno

Para as análises computacionais, a similaridade de cosseno permite parear vetores de documentos para análise estatística, reconhecimento de padrões e definição de agrupamentos. O cálculo é dado em pares de documentos para compará-los através do conteúdo normalizado. Então, a similaridade adotada é um modelo matemático que quantifica a distância entre documentos no espaço vetorial.

A distância de cosseno, segundo (BAEZA-YATES; RIBEIRO-NETO, 2008), é a métrica

mais usual para a comparação entre documentos textuais. Nessa análise de similaridade, um valor entre 0 e 1 representa o quanto pares de documentos são idênticos, onde os exatamente iguais têm similaridade 0 e, por consequência, os completamente dissimilares 1. A similaridade de cosseno é representada pela Equação 3.2.

$$Sim_{cos}(d_i, d_j) = \frac{d_i \cdot d_j}{||d_i \cdot d_j||} \quad (3.2)$$

Na Equação de distância de cosseno 3.3, é uma transformação do resultado da razão entre a multiplicação e a intercessão do par  $(d_i, d_j)$ , dado pela similaridade, (Equação 3.2) em uma métrica de distância. A diferença está na inversão dos valores com menores para vetores equivalentes e maiores para distintos. Na métrica de distância valores também estão entre 0 e 1. Porém, quanto mais próximos de 0 maior a equivalência dos documentos, sendo zero retornado para vetores de mesma norma. Em contrapartida, valores próximos à 1 representam vetores distintos.

$$D_{cosseno(d_i, d_j)} = 1 - Sim_{cos}(d_i, d_j) \quad (3.3)$$

### 3.2.2 Pré-Processamento

A padronização dos documentos é um procedimento comum para a extração de informações e redução de dimensionalidade. Para isso, algumas análises por NLP do conteúdo visam manter termos equivalentes e correlacionados, removendo palavras que não adicionam informações relevantes ao texto. A padronização de escrita utilizada pelo sistema é dada pela remoção de *stopwords*, *stemming* e tokenização, descritas individualmente abaixo.

#### Remoção de *Stopwords*

A remoção de *stopwords* (MANNING; SCHÜTZE, 1999) foi utilizada para extrair da avaliação e recuperação de informação as palavras que possivelmente nada adicionam ao

contexto. Nesse processo todas as palavras de uma lista pré-definida são excluídas do texto antes da vetorização. As palavras excluídas, geralmente, são conectivos textuais já conhecidos da linguagem. Esses termos, durante a interpretação computacional, não influenciam nas informações passadas pelo interlocutor e são removidos para estudo do conteúdo.

As listas de *stopwords* para inglês e português são apresentadas nas Tabelas B.1 e B.2.

### *Stemming*

Stemmização é outra importante padronização do conjunto de documentos, que consiste na radicalização de todas as ocorrências das palavras através da remoção de afixos (WEISS; INDURKHYA; ZHANG, 2010). Marcantes na língua portuguesa, com várias sintaxes e modificadores, esse processo remove prefixos e sufixos para destacar apenas o morfema básico nos grupos textuais e nas referências de avaliação.

Com a aplicação desse método, palavras em pessoa, número ou tempo diferentes são normalizadas para um mesmo radical. Dessa forma, espera-se que uma frase escrita em todas as possíveis conjugações sejam interpretadas de forma equivalente pelo sistema, apesar da perda de informação.

### **Tokenização**

A tokenização é a separação das palavras dentro de uma frase para posterior vetorização (MANNING; SCHÜTZE, 1999), sendo a etapa responsável pela inclusão de métodos de tratamento do conteúdo. Nessa etapa os dados indesejáveis como as *stopwords*, acentuação e a pontuação são eliminados.

Porém durante a tokenização, para identificar conexões textuais as palavras foram extraídas em faixas de *n-grams*. Os *n-grams* são a quantidade de palavras analisadas em uma

distância  $n$ . Assim, dentro de cada documento  $d$  é verificada a frequência de grupos de palavras de  $n$  distância. Caso seja 1 (*unigram*), a resultante da vetorização será igual a gerada por TF.

Nesse caso, especificamente, foram coletados de 1 a 3-grams (*unigrams*, *bigrams* e *trigrams*). Por exemplo, se temos uma frase formada por seis termos com índices  $\{t_{45}, t_{32}, t_{123}, t_{1024}, t_{256}, t_{13}\}$ , essa coleta é dada da seguinte forma:

1-gram:  $\{t_{45}\} \{t_{32}\} \{t_{123}\} \{t_{1024}\} \{t_{256}\} \{t_{13}\}$

2-grams:  $\{t_{45}; t_{32}\} \{t_{32}; t_{123}\} \{t_{123}; t_{1024}\} \{t_{1024}; t_{256}\} \{t_{256}; t_{13}\}$

3-grams:  $\{t_{45}; t_{32}; t_{123}\} \{t_{32}; t_{123}; t_{1024}\} \{t_{123}; t_{1024}; t_{256}\} \{t_{1024}; t_{256}; t_{13}\}$

Esse modelo de extração apresenta os grupos de termos como os destacados acima, onde são organizados conjuntos conforme a ocorrência por vizinhança. Simultaneamente, com a definição das características textuais no conjunto  $D$ , é realizada a contagem da frequência para cada documento  $d$ . Desse modo, torna-se possível a realização das comparações entre os documentos através das Equações 3.2 e 3.3 para análise por pares.

### 3.3 Algoritmo Genético

Aplicado na seleção de características para redução de dimensionalidade, o Algoritmo Genético - GA (HAUPT; HAUPT, 2004) é inspirado na teoria da evolução darwiniana. Sua otimização funciona através de combinações vetoriais para a melhoria das soluções candidatas com avaliação qualitativa através da função de *fitness*. Cada solução candidata remete à um cromossomo, sendo conjuntos de valores denominados genes. Cada um passa pelo processo de seleção, *crossover* e mutação. Na seleção é definido o par da população que realizará a combinação. O *crossover* é a função responsável pela combinação do par de cromossomos candidatos selecionados para criação de novos indivíduos. A etapa de mutação altera os dados (ou genes) para adicionar variações genéticas na população.

Para o GA, é necessário criar uma população que será a distribuição de valores disponíveis para combinação. A otimização é cíclica, ou seja, ocorre com um dado número de iterações. Para cada ciclo completo de etapas, chamado de gerações, o melhor item resultante é comparado com os vetores populacionais. Se esse indivíduo de melhor *fitness* gerado tiver valor superior ao pior elemento existente da população ele o substitui.

Aplicado à seleção de características (RAYMER et al., 2000), os vetores das  $k$  características são combinadas para eliminação de informações conforme a otimização da métrica. No caso do mapa de características, foi aplicada uma razão da quantidade de termos selecionados pela similaridade interna da classe. Espera-se, portanto, reduzir a quantidade de termos do conjunto enquanto se amplia a similaridade média do grupo, tornando-o mais denso. Com maior densidade, os documentos são sumarizados e espera-se que representem a essência do conjunto de informações.

### 3.4 Aprendizado Semi-Supervisionado

Na inteligência artificial, sistemas são classificados quanto ao seu modelo de aprendizado de forma dependente ou independente do especialista. Nos sistemas supervisionados o especialista treina o algoritmo com uma quantidade de dados para atuação desse na tarefa. Esse modelo criado direciona a qualidade dos resultados através da interpretação e completude dos padrões nos exemplos apresentados. Por outro lado, os não-supervisionados são independentes da classificação de um especialista para descoberta de padrões e a aquisição de conhecimento. O paradigma desse modelo é exploratório, ou seja, o próprio sistema busca nos dados as informações para a especialização do trabalho.

Para a avaliação de questões discursivas, a definição do critério de correção é fundamentada na busca conhecimento de um especialista. Nesse caso, o conhecimento pode ser a avaliação do professor, porém atentando-se à redução do esforço do humano. Pensando na sobrecarga desse profissional utilizamos da supervisão humana apenas durante a avaliação

dos padrões identificados automaticamente. Dessa forma, diferentemente dos dois modelos clássicos de aprendizado, os exemplos requisitados para treinamento são selecionados pelo próprio sistema. Nesse sistema, utilizamos do aprendizado semi-supervisionado, onde através de uma combinação de clusterização (não-supervisionado) e classificação (supervisionado), buscamos identificar previamente os requisitos da avaliação. Nesse formato, o fator de avaliação humana é um passo intermediário entre a descoberta de padrões e a obtenção dos resultados pós-classificação.

### 3.4.1 Combinando Clusterização e Classificação

Para definição dos padrões para os métodos de avaliação analisamos a distribuição espacial dos agrupamentos. Os grupos de documentos foram formados conforme os resultados da clusterização com distância de cosseno (Equação 3.3). Esse processo utiliza o CLUTO<sup>2</sup> (KARYPIS, 2002) para testes do número de grupos visando reduzir a taxa de sobreposição, ou *overlapping* em inglês. Essa taxa de *overlapping* é obtida através da Equação 3.4 do *silhouette score* (ROUSSEEUW, 1987). Os testes de número ideal de grupos se dão pela execução do sistema para todas possibilidades de 3 até  $c$  de grupos, sendo  $c$  o máximo definido por  $\sqrt{\frac{|D|}{2}} \times 2$ . O agrupamento de melhor *silhouette score* é utilizado como ideal.

$$s_{cluster_c} = \frac{b - a}{\max(a, b)} \quad (3.4)$$

No coeficiente *silhouette* para um determinado número de agrupamentos  $s_{cluster_c}$ , a variável  $a$  refere-se à distância intra-*cluster*, enquanto  $b$  à média de distância de cada amostra até o *cluster* mais próximo. Essa métrica retorna valores entre -1 e 1, sendo que os valores negativos representam arranjos ruins de amostras e os positivos *clusters* bem definidos. Valores próximos à 0 representam *clusters* em *overlapping*.

Essa métrica foi utilizada justamente na tentativa de contornar altas taxas de *overlapping*

<sup>2</sup><http://glaros.dtc.umn.edu/gkhome/views/cluto>

esperando o aumento da densidade dos agrupamentos. Com grupos bem selecionados nessa etapa, o treinamento do algoritmo deve ser mais representativo para a diversidade textual das submissões, permitindo a análise de *clusters* mais objetiva para a classificação.

### 3.4.2 Identificação de Treino e Teste

Na clusterização resultante, com base na análise intra-*cluster* estudada por Oliveira et al. (2014), escolhemos itens que definam as melhores características para avaliação. Seleccionamos para cada agrupamento os itens mais similares, mais dissimilares e de maior e menor norma para a correção do professor. Os de maior similaridade tendem à referenciar conteúdos equivalentes para um possível grupo de nota. Os mais dissimilares, ao contrário, representam as ideias divergentes do mesmo grupo. Os vetores de maior e menor norma estabelecem, respectivamente, os conteúdos extensos e resumidos associados para a avaliação. A extensão dos documentos é aplicada essencialmente no cálculo de representatividade das características, influenciando diretamente no método avaliativo.

Após a avaliação dos itens de treinamento pelo professor, são seleccionadas as características, definindo conjuntos mínimos para cada nota. Então, cada nota atribuída pelo professor torna-se um rótulo (classe) para as amostras, desconsiderando os grupos obtidos na etapa de clusterização. Esse processo de seleção com o mapa de características é descrito detalhadamente no Capítulo 4.

Com os conceitos mínimos por classe, o número de características total diminui consideravelmente reduzindo o esforço de interpretação e correção feito pelo algoritmo. Dois classificadores são utilizados para o processo de avaliação o *k-vizinhos mais próximos*- *K*-NN e o classificador baseado em centroide - CBC.



### 3.4.3 *K-Nearest Neighbors - K-NN*

O algoritmo *K-NN* é um classificador escolhido pela classificação por associação aos padrões de treinamento. Nesse algoritmo, o modelo é criado através dos dados de treinamento, onde as características dos  $k$  vizinhos mais próximos definem a classe de cada item de teste (BAEZA-YATES; RIBEIRO-NETO, 2008). A comparação é realizada pela distância de cosseno entre pares. Esse classificador tende a criar zonas de influência sinuosas no espaço vetorial através dos  $K$  termos de menor distância para o elemento sem classificação. Para  $K = 1$ , como nesse caso, o documento recebe o rótulo do item de treino mais similar.

Nesse problema, o *K-NN* com  $K = 1$  é utilizado por conta das atividades com poucas respostas e as notas raras no conjunto de avaliações. O algoritmo, portanto, mantém a distribuição da base de treino classificando os demais conforme documentos análogos (padrões conhecidos).

### 3.4.4 *Centroid Based Classifier - CBC*

A classificação baseada em centroide, também conhecida por *Rocchio*, é uma forma de extração de modelo de classes baseada no ponto médio (centroide) (BAEZA-YATES; RIBEIRO-NETO, 2008). Para cada classe é calculada a média entre os itens e, para cada elemento sem classificação, é atribuída a classe do centroide mais próximo. Assim, cada classe é representada pelo item médio na etapa de classificação.

Nesse caso, a escolha do *Centroid Based Classifier - CBC* da-se pela recomposição de *agrupamentos* por classes tal como os formados na etapa de *clustering*. Para esse classificador são criadas zonas de influência por centroide no espaço vetorial o que o torna menos suscetível à *outliers* do que o *K-NN*.

### 3.4.5 Métricas de Avaliação

Para avaliar a qualidade das classificações do sistema (Capítulo 5), utilizamos algumas métricas para validação dos resultados (COHEN, 1995). Nos métodos onde o sistema identifica que o treinamento é dado em notas contínuas, o sistema compara as notas com o especialista através do Erro Médio Absoluto - MAE, do Erro Quadrático Médio - MSE e/ou do Desvio Padrão do Erro - SEM.

O Erro Médio Absoluto - MAE (*Mean Absolute Error*), apresentado na Equação 3.5, calcula a diferença média relativa às predições do classificador ou regressor quando comparados com os valores reais atribuídos pelo especialista.

$$MAE(y_{real}, y_{pred}) = \frac{1}{n_{amostras}} \sum_{i=0}^{n_{amostras}-1} |y_{i_{real}} - y_{i_{pred}}| \quad (3.5)$$

O Erro Quadrático Médio - MSE (*Mean Squared Error*), apresentado na Equação 3.6, calcula o quadrado da diferença média para as predições do sistema se comparados com os valores atribuídos pelo avaliador humano. A diferença entre o MSE e o MAE está na penalização de erros maiores causados principalmente pelos *outliers*.

$$MSE(y_{real}, y_{pred}) = \frac{1}{n_{amostras}} \sum_{i=0}^{n_{amostras}-1} (y_{i_{real}} - y_{i_{pred}})^2 \quad (3.6)$$

O Desvio Padrão do Erro - SEM (*Standard Error of the Mean*), apresentado na Equação 3.7, calcula o desvio das amostras da média (erro) para os valores preditos pelo sistema comparado-os com os valores atribuídos pelo avaliador humano. O valor retornado pode ser associado às métricas de erro para análise da distribuição do erro conforme os valores médios obtidos através do MAE.

$$SEM(y_{real}, y_{pred}) = \frac{\sigma}{\sqrt{n_{amostras}}} \quad (3.7)$$

Os problemas de classificação são analisados qualitativamente com as métricas de *precision*, *recall*, *accuracy* e *f1*, detalhadas abaixo.

A acurácia, ou *accuracy*, dada pela Equação 3.8, avalia a proximidade das notas atribuídas para os valores conhecidos, resultando no erro. O resultado desse cálculo é o percentual dos itens avaliados de forma equivalente.

$$Accuracy(y_{real}, y_{pred}) = \frac{1}{n_{amostras}} \sum_{i=0}^{n_{amostras}-1} 1(y_{i_{real}} = y_{i_{pred}}) \quad (3.8)$$

Precisão, ou *precision*, dada pela Equação 3.9, avalia a quantidade de notas atribuídas pelo classificador que realmente equivalem aos valores conhecidos. Nesse cálculo, é analisado o percentual de valores corretamente avaliados (verdadeiros positivos  $T_P$ ) dentre os avaliados de forma equivalente (verdadeiros positivos e falsos positivos  $T_P + F_P$ ).

$$Precision(y_{real}, y_{pred}) = \frac{T_P}{T_P + F_P} \quad (3.9)$$

Revocação, ou *recall*, dada pela Equação 3.10, avalia a quantidade de notas atribuídas pelo classificador das que têm mesmo valor conhecido. Nesse cálculo, é analisado o percentual de notas que deveriam ter mesma avaliação (verdadeiros positivos e falsos negativos  $T_P + F_N$ ) que foram corretamente avaliados (verdadeiros positivos  $T_P$ ) pelo sistema.

$$Recall(y_{real}, y_{pred}) = \frac{T_P}{T_P + F_N} \quad (3.10)$$

O *f1* ou *f-score*, dada pela Equação 3.11, é uma ponderação entre *precision* e *recall*. No formato utilizado, ponderado (*weighted*), as métricas são recalculadas para cada classe e é extraída a média ponderada pelo suporte (número de instâncias conhecidas de cada classe).

$$F1 = 2 \times \frac{precision \times recall}{precision + recall} \quad (3.11)$$

## Capítulo 4

### Mapa de Características

O mapa de características é uma ferramenta de extração e visualização de informações que visa identificar os principais elementos no texto do aluno com a finalidade de reduzir os esforços de verificação de aprendizagem e revisão do conteúdo. Através desse sistema, de forma automática e independente do professor, são definidos o termo ou conjunto de termos mais relevantes para os grupos de nota da atividade. Com esse material, geramos um relatório do conteúdo elaborado pelos participantes com marcações nos documentos, para autoavaliação, discussão do tema e comparação dos resultados.

#### 4.1 Identificação das Características Relevantes

As atividades com foco no conteúdo ministrado em sala de aula, em geral, têm como objetivo identificar o saber do aluno. Por conta disso, o sistema observa, sob o cunho avaliativo, trechos similares nos grupos de respostas com notas idênticas atribuídas pelo professor para uma atividade. Nos textos selecionados para treino, o sistema busca identificar de forma automática o *modus operandi* do professor especialista na correção. Através do Algoritmo Genético Binário (HAUPT; HAUPT, 2004), são selecionadas partes do texto que possivelmente caracterizam tal avaliação. Para isso, partimos do pressuposto de que a avaliação do

professor tem por consequência a equivalência textual. O sistema, então, seleciona termos correlacionados com o grupo de análise, no caso, por classe de nota. A Figura 4.1 apresenta a linha de execução do sistema incluindo a seleção de termos pelo GA.

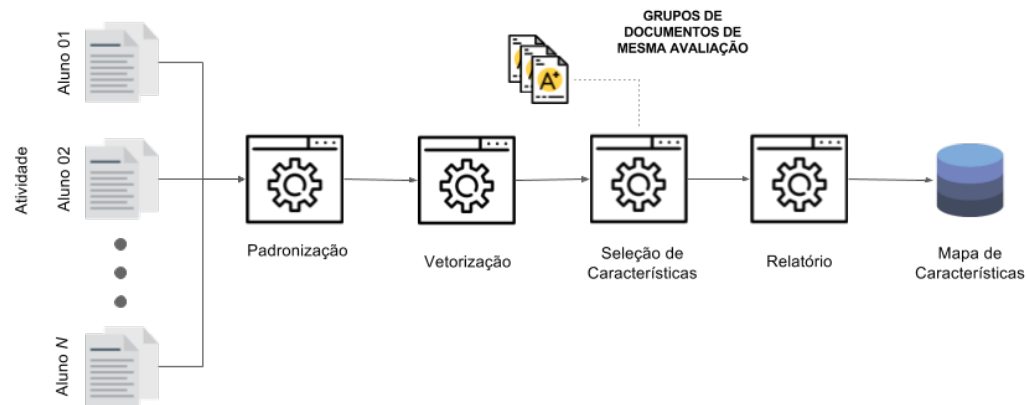


Figura 4.1: Linha de execução do G.A. na otimização de cada classe de nota quanto ao número de características e a densidade dos agrupamentos.

A representação apresentada na Figura 4.1 mostra o processo do texto com pré processamento, vetorização e a divisão em classes, individualmente entregues ao Algoritmo Genético objetivando a redução de dimensionalidade e seleção de características. Essas duas metas são necessárias para alcançar melhores resultados de classificação, pressupondo a existência de características chave para a resposta. Consideramos assim que existe uma resposta ótima possível que seja referência para a nota dada pelo professor. A estrutura do GA utilizado é apresentada na Figura 4.2.

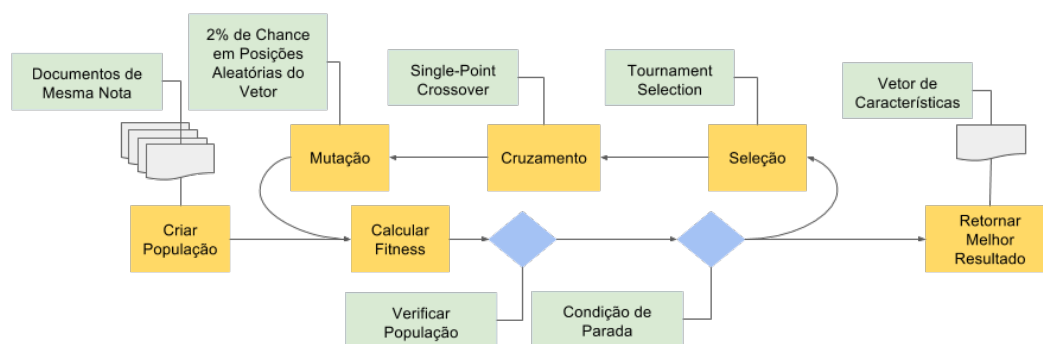


Figura 4.2: Estrutura do G.A. para seleção de características nas classes de nota.

Conforme representa a Figura 4.2, para início do GA o vetor dos documentos dos alunos

é passado junto com o vetor de menor norma de cada grupo. Uma população aleatória é definida com vetores de tamanho  $k$  características. A otimização inicia-se por esse menor que, em tese, deveria conter o requisito da avaliação tal como os outros itens do grupo de nota. Em cada geração do GA é executada a seleção *por torneio*, aonde um par de cromossomos é selecionados aleatoriamente dentre a população e o de menor *fitness* é retornado como um dos candidatos para *cruzamento*. Após a definição de um par de vencedores é realizado o *crossover*. Para isso, um índice randômico define o corte dos vetores, gerando dois novos vetores que correspondam à soma das partições desses primeiros. Esses vetores resultantes são submetidos à mutação. A chance de ocorrência é de 2% e pode ocorrer em um número aleatório de genes do indivíduo, alternando o *bit* dessas posições. Sem haver melhoria durante 700 gerações, o cromossomo de melhor *fitness* é retornado como resultado.

No GA, para atribuir a cada indivíduo seu *fitness*  $f_{ind}$ , primeiro calculamos a média de similaridade do agrupamento levando em conta as características selecionadas. Sendo o grupo  $S$  formado por todos os documentos  $d$  com mesma classificação na base de dados  $D$ , é calculada a similaridade interna do agrupamento para os termos do indivíduo  $t_{ind}$ . Então, o *fitness*, dado pela Equação 4.1, é o resultado da divisão do número de características selecionadas  $t_{ind}$  pela média de similaridade interna quando consideradas apenas as  $t_{ind}$  características.

$$Fitness_{f_{ind}} = \frac{t_{ind}}{\bar{x}_{sim_s}} \quad (4.1)$$

A métrica acima tende a selecionar características pela média de similaridade dos documentos, ou seja, através da densidade do grupo classificado pelo especialista. Assim, esperamos tornar os termos referenciados por essa classe mais próximos da avaliação. Assim, localizar quais características foram observadas pelos usuários é um desafio que deve ser aprofundado pois contém aspectos de classificação subjetiva ou interpretativa para atribuição de relevância por grupo de nota.

Segundo How e Narayanan (2004), algumas das formas de qualificar a seleção de carac-

terísticas são: a análise de *clusters*, métricas de distância, medidas estatísticas ou baseadas em teorias da informação. Um trabalho que realizou a comparação entre alguns métodos tradicionais de análise de relevância dos termos foi feito por Algarni e Tairan (2014). Nesse artigo os autores estudam o impacto na seleção de características de algumas técnicas de ponderação de relevância dos termos, semelhante à técnica de *fitness* adotada, como *Mutual Information*, *Chi-Square*, *Information Gain* e *Odds Ratio*.

Na Equação 4.1 foi utilizada portanto a medida de distância de similaridade interna para otimização da quantidade do conjunto de termos. A complexidade dessa seleção é dada pela atenção para um conjunto de termos objetivo (resposta) do usuário para além do modelo gerado pelo classificador. Os termos selecionados então, além do modelo avaliativo pretende representar uma descrição da atividade segundo o mapa de características. Experimentos que demonstram a aplicação dessa visão qualitativa das marcações segundo a resposta-base do professor são apresentados na Seção 5.2.5.

Portanto, a calibração do GA foi observada com cautela. Existem bases de dados aos quais a apresentação de resultados de grande minimização têm muitos trechos visualmente equivalentes removidos. O contrário também ocorre, com pequenas minimizações e a marcação de grande porção de texto para a classe. Ambos os casos são importantes dependendo da atividade e do seu critério de avaliação.

Para contornar problemas de seleção textual, maximizado ou minimizado, foi implementada a inserção do documento de menor norma da classe como indivíduo inicial. Se o processo de avaliação for adequado ao critério, espera-se que esse elemento tenha o melhor *fitness* dos documentos. Caso contrário, espera-se que os ciclos sejam capazes de contornar *outliers* realizando testes com as demais características. Outra melhoria para acréscimo de liberdade de seleção ao GA foi o limite do número de gerações relativo a última melhoria. Essa alteração visava manter a influência de vetores selecionados para os próximos ciclos, mantendo a possibilidade de melhorias consecutivas.

Para definição dos parâmetros de população e número de gerações do GA, realizamos

ciclos de testes com cinco processos paralelos. A população variou de 10 a 100 indivíduos em intervalos de 10 e o número de gerações de 100 a 1000 em intervalos de 100 com o objetivo de estudar a incidência do algoritmo genético nos extremos de densidade por classes. Para isso, foi escolhida a atividade *TecII-5-8* com cinco classes sendo dois grupos desses de grande densidade e dois de pequena densidade. As notas maiores que representavam as classes '0' e '100' eram de otimização mais complexa pela menor densidade, ou seja, era maior a distância intra-cluster Equação 3.4. Enquanto as classes '90' e '93' eram respostas mais curtas e com maior densidade. Os resultados da otimização são listados na Tabela A.2.

Observamos pela Tabela A.2 os 15 melhores *fitness* obtidos para cada classe. As de maior densidade apresentaram média de 700 gerações para retorno dos resultados. As de menor densidade por outro lado apresentaram resultados em média de 300 gerações. Para a população, foi comum a ocorrência das maiores independente do problema observado, tendo médias 90 e 100 em ambos. Assim, um número máximo foi definido como padrão, 100 de população e 700 gerações, para que o algoritmo tente atender, pelo menos, esses dois casos.

## 4.2 *Feedbacks* e a Visualização das Características

A etapa de identificação dos fatores relevantes elabora chaves de resposta que direcionam cada classe de nota ao respectivo conjunto de termos. Com a seleção do mapa de características através do GA, medindo cada ciclo de *fitness*, o aumento da densidade das respostas torna o grupo mais conciso e melhora a identificação e a validação dos resultados. Porém, é importante salientar que os termos analisados de forma individualizada podem não caracterizar a resposta completa aguardada pelos professores por co-ocorrência de termos. Para isso, durante a produção desse sistema o pré processamento realizou seleções com faixas de 1 a 3-grams para identificação de trechos compatíveis por classe.

Com os resultados em *n-grams*, conseguimos identificar grandes trechos onde a grafia é semelhante. Com a seleção, os termos eliminados tornam os classificadores mais objetivos



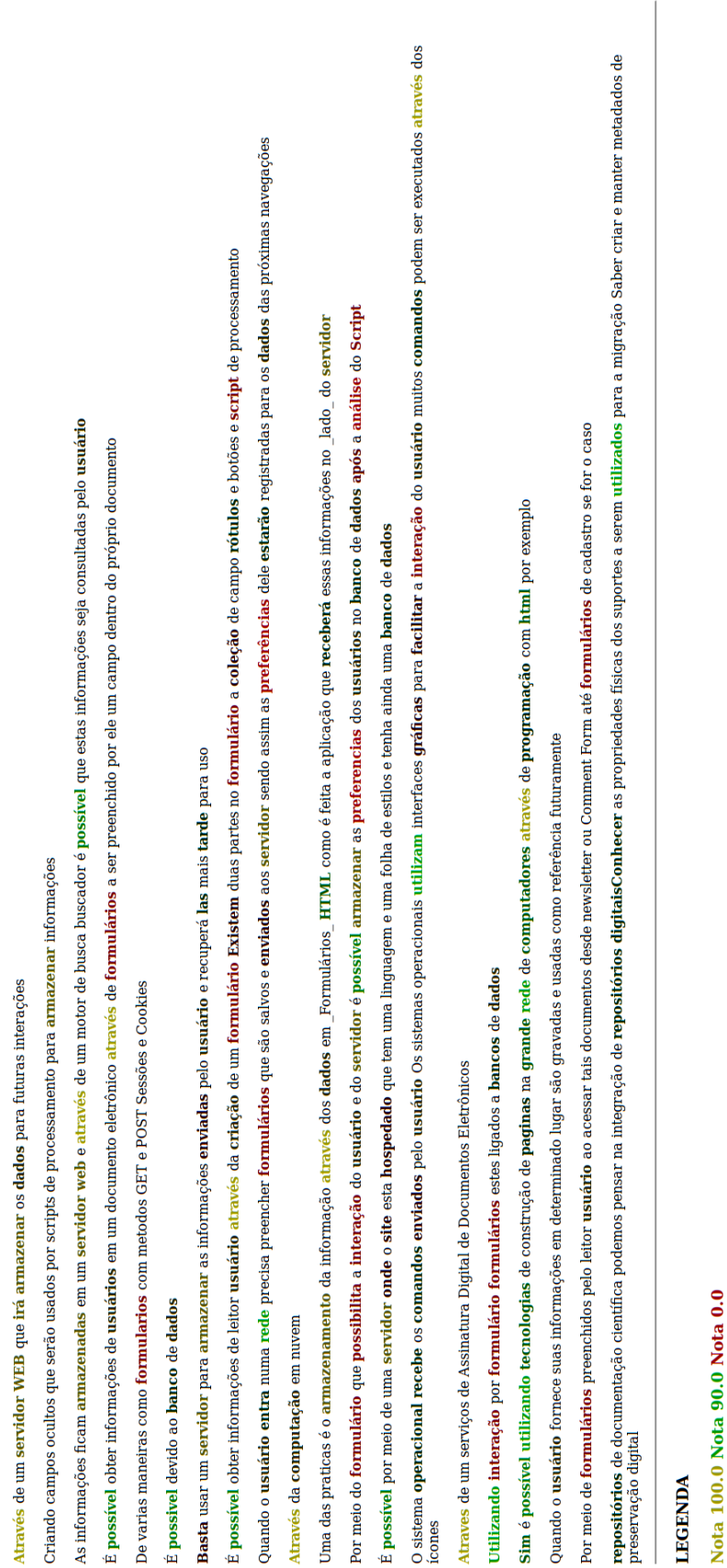


Figura 4.3: Visão geral do mapa de características como resultado da atividade Tec-II-5-186.

na avaliação por classe. Com isso, as características restantes resumem o grupo ao qual é associado de tal forma que os descreve. Assim, a seleção de características além de melhorar o desempenho dos sistemas de classificação é uma ferramenta de marcação e visualização dos resultados para visão sistêmica da questão e do conhecimento obtido.

Cada termo é associado ao grupo que ele descreve, gerando relatórios em HTML e em  $\text{\LaTeX}$  (GOOSSENS; MITTELBAACH; SAMARIN, 1994). Esses termos são marcados para todas as atividades, sendo treino ou teste dos algoritmos de avaliação. Em  $\text{\LaTeX}$ , o relatório é individual e apresenta ao professor criteriosamente sua avaliação. Em HTML é apresentando em uma escala de dez níveis para cada cor segundo uma nota, até no máximo seis notas. O nível é dado pela representatividade do termo na classe com as escalas RGB de 00 à 99. Um exemplo de marcação é dado na Figura 4.3 pela atividade 186 da disciplina de Tecnologia da Informação II da Universidade Federal do Espírito Santo.

Para a Figura 4.3, a atividade *Tec-II-5-186* contém 4 classes distintas. As tonalidades das cores por nota, 100.0 em amarelo, 90.0 em verde e 0.0 em vermelho foram definidas conforme o peso dos fatores. Para cada palavra, o sistema apresentou o nível de representatividade do termo para a classe associada.

O *feedback* gerado pelo mapa de características é disponibilizado através do AVA no campo disponível para comentários do avaliador e o professor e os alunos podem utilizar as marcações para correções colaborativas. Com o reconhecimento das regiões, ambos têm

### Feedback



Nota	100,00 / 100,00
Avaliado em	domingo, 12 Feb 2017, 11:08
Avaliado por	 Moodle WS
Comentários de feedback	 Exercício 5 Os <b>elementos base</b> no <b>modelo proposto</b> pelo <b>DCC Digital Curation Center</b> são <b>Dados objetos digitais base</b> de <b>dados</b> e também...

Figura 4.4: Visualização do mapa de características como resultado individual no campo de *feedback* do professor.

os resultados como uma ferramenta para argumentar com o tema da questão e revisar o conteúdo pelo conhecimento dos participantes. Desse modo, além de melhoria dos processos internos de avaliação, o mapa formado durante o processo explica as soluções propostas para os participantes. Além disso, a ferramenta dá poder a todos os avaliados e avaliadores da discussão dos resultados. A resposta no Moodle, como nosso AVA padrão, pode ser vista na Figura 4.4.

Os testes da classificação através da seleção de características e as comparações de resultados são apresentadas no Capítulo 5.

# Capítulo 5

## Experimentos e Resultados

Neste capítulo detalhamos os experimentos realizados em quatro bases de dados. A primeira base de dados, coletada do vestibular da UFES, é utilizada para comparação de notas contínuas com o sistema proposto anteriormente por Pissinati (2014). A base de dados de disciplinas da UFES e o *Texas Dataset* foram análises do mapa de características como parte do *software* na avaliação por classes antes e depois da redução de dimensionalidade.

A última base de dados foi utilizada para testes de robustez do sistema. A base de dados do Prêmio de Avaliação Automática de Estudantes - ASAP (*Automated Student Assessment Prize*) do *Kaggle*, foge muito do padrão das demais correções do sistema pela quantidade de documentos e características. Assim, o *dataset* foi aplicado na verificação da busca de padrões através da clusterização e do efeito da redução de dimensionalidade nos métodos de classificação em bases de dados maiores.

### 5.1 Bases de Dados

Utilizamos nos testes quatro tipos de bases de dados de questões discursivas: do vestibular da UFES, de disciplinas ministradas por professores da UFES, da disciplina de Estrutura

de Dados da Universidade do Norte do Texas (*Texas Dataset*) e da competição da *Hewlett Foundation* no *Kaggle*. Cada base de dados é descrita à seguir.

### 5.1.1 Base de Dados do Vestibular UFES

Conjunto de atividades discursivas transcritas do vestibular de 2012 da Universidade Federal do Espírito Santo - UFES da prova de língua portuguesa. No total são 460 respostas para as cinco questões avaliadas por dois avaliadores, contendo 92 cada. Caso houvesse divergências de mais de 1 ponto entre as duas correções, um terceiro avaliador era acionado.

O enunciado das cinco atividades é apresentado na Seção C.1 nos Apêndices.

### 5.1.2 Base de Dados das Disciplinas UFES

Essa base de dados foi coletada de algumas disciplinas ministradas na Universidade Federal do Espírito Santo - UFES entre 2015 e 2016 através do Moodle do Laboratório Computação de Alto Desempenho - LCAD. Entre as disciplinas estão Metodologia e Técnicas de Pesquisa Científica, Filosofia e Tecnologia da Informação II. Totalizando 45 atividades, foram recebidas 1162 submissões com média 25,82 e desvio padrão de 13,54 respostas por atividade.

O diferencial dessa base de dados é a mudança de características nas respostas encontradas nas múltiplas disciplinas, professores e alunos. Observam-se alterações consideráveis quanto ao tamanho, número de grupos, critério avaliativo ou objetividade.

### 5.1.3 Base de dados da Universidade do Norte do Texas (Inglês)

*Dataset* coletado com alunos de Estrutura de Dados do curso de Ciência da Computação na Universidade do Norte do Texas. Resultado do trabalho realizado por (MOHLER;

BUNESCU; MIHALCEA, 2011), a base de dados, em sua segunda versão, constitui-se de dez atividades com sete questões e dois exames com dez questões cada. As provas foram avaliadas em intervalos de 0 a 10 e as atividades de 0 a 5 pontos. A avaliação foi feita por dois professores para as respostas de, em média, 30 alunos participantes. Após a avaliação foi extraída a média entre os dois corretores. No total, essa base de dados contém 630 respostas curtas e para cada atividade existe uma chave de resposta como o critério de correção do professor.

#### 5.1.4 Base de Dados do Concurso ASAG-SAS no *Kaggle* (Inglês)

A competição de avaliação automática de questões discursivas curtas - *ASAP - Automated Student Assessment Prize*, foi uma competição organizada no *Kaggle*. O *Kaggle* é uma plataforma de competições em mineração de dados que reúne tarefas complexas formadas por problemas reais enfrentados por grandes empresas. A sua comunidade era de 536 mil usuários registrados em 2016, que disputam prêmios e compartilham aprendizados em análise estatística e ciência de dados.

No *Kaggle*, a organização *Hewlett Foundation* patrocinou três competições. As três foram denominadas como as seguintes fases da *ASAP*:

- Fase 1: Demonstração em respostas longas (redações);
- Fase 2: Demonstração em respostas curtas (discursivas);
- Fase 3: Demonstração simbólica matemática/lógica (gráficos e diagramas).

A Avaliação Automática de Respostas Curtas (*Automatic Short Answer Grader*) *ASAG - SAS* premiou com um total de 100 mil dólares os cinco primeiros colocados. A base de dados de respostas curtas, contém dez questões de artes a ciências. Ao final do concurso, os dados disponíveis avaliados por dois professores totalizam 17043 itens com aproximadamente 1700 itens por atividade.

## 5.2 Experimentos Realizados

Para cada base de dados alguns experimentos foram realizados para efeito de comparação com os resultados da literatura. A base de dados do vestibular da UFES (Seção 5.1.1), é utilizada para comparação desse trabalho com os experimentos realizados por (PISSINATI, 2014) para avaliação em notas contínuas.

O mapa de características e os sistemas de classificação foram testados em classes de nota de disciplinas universitárias. A base de dados coletada na UFES (Seção 5.1.2) foi construída durante os desenvolvimento do algoritmo de avaliação e do mapa de características. Para maior diversidade nos testes, esses também são realizados na base de dados de Estrutura de Dados da UNT (Seção 5.1.3), em linguagem estrangeira. Essa última base de dados ainda conta com chaves de correção para parearmos as marcações formadas no mapa de características com as respostas aguardadas pelos professores.

Também em língua estrangeira, a base de dados do concurso ASAG-SAS no *Kaggle*, é a com maior número de submissões e taxa de *overllaping* testada. Essas duas informações são fundamentais para verificação da eficiência dos classificadores em situações de maior complexidade. Os resultados e as discussões envolvendo os métodos de classificação e a seleção de características são apresentados abaixo.

### 5.2.1 Experimentos com o Vestibular da UFES

O primeiro experimento realizado tem como objetivo analisar a equivalência do sistema desenvolvido com os resultados obtidos por Pissinati (2014). Para isso, foi utilizada a mesma base de dados do vestibular da UFES com 460 respostas de alunos para cinco questões de Literatura/Português, avaliada por dois especialistas. No trabalho, o autor realizou sete experimentos para cada um dos avaliadores humanos do vestibular, também realizando uma comparação entre eles. O resultado dos trabalhos com base no Erro Médio Absoluto - MAE,

tendo a divergência entre avaliadores como *baseline*, é dado pelas Figuras 5.1a e 5.1b para o primeiro e segundo avaliador respectivamente.

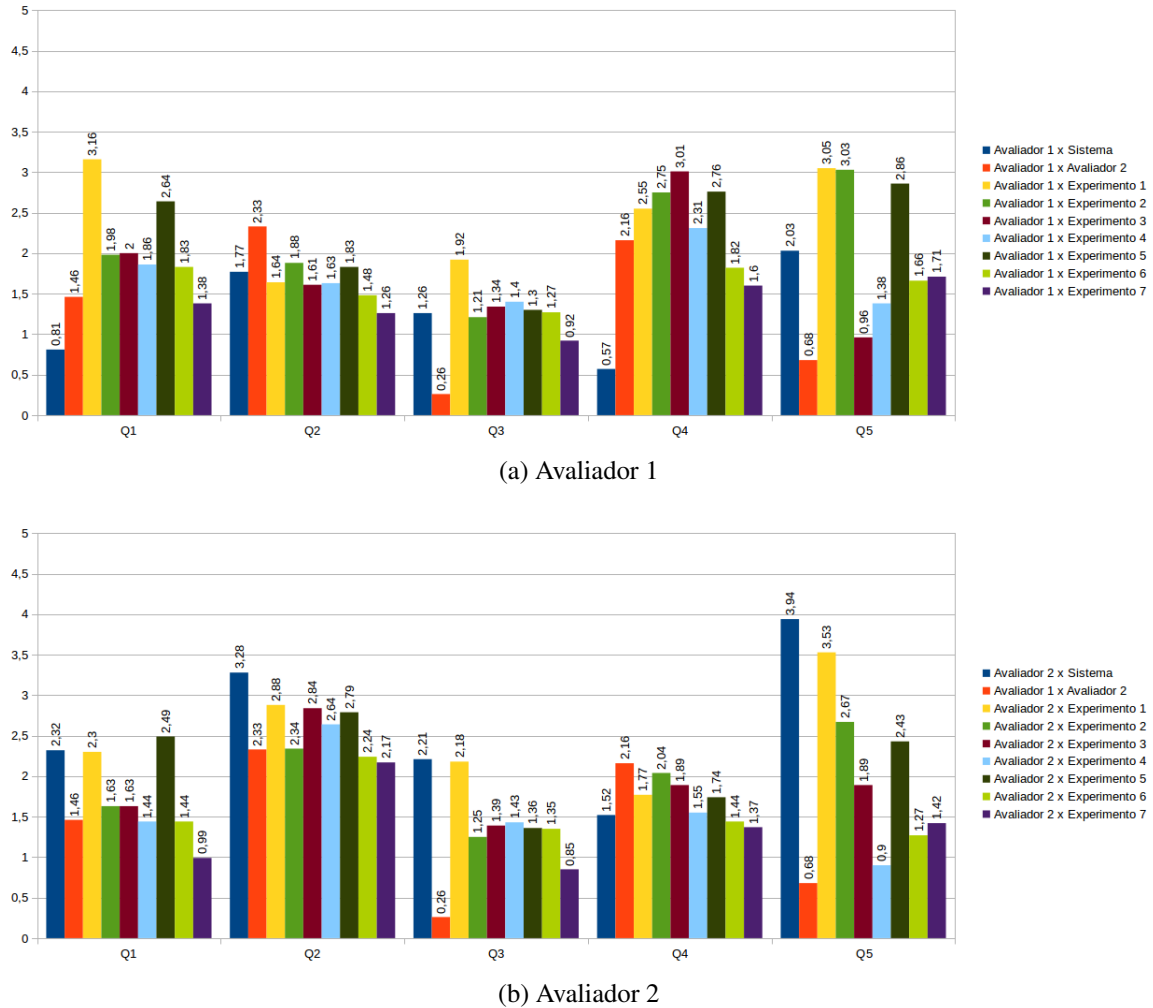


Figura 5.1: Erro Médio Absoluto - MAE para os sete experimentos encontrados no trabalho tomado como referência, o sistema aqui apresentado e cada avaliador, utilizando como *baseline* o erro entre os próprios especialistas.

Nas Figuras 5.1a e 5.1b, é possível visualizar a adequação do sistema com os avaliadores e com os melhores resultados apresentados nos sete experimentos. O único caso crítico observado, principalmente na adequação aos resultados do Avaliador 2, foi encontrado para questão 5. Nas demais, o sistema obtém resultados semelhantes aos melhores experimentos de (PISSINATI, 2014), sendo equiparável ao erro resultante dos dois avaliadores humanos. As Figuras 5.2a e 5.2b mostram o Desvio Padrão do Erro - SEM para os corretores humanos.

Nas Imagens 5.2a e 5.2b, o desvio padrão médio de 1,202 pontos para o Avaliador 1



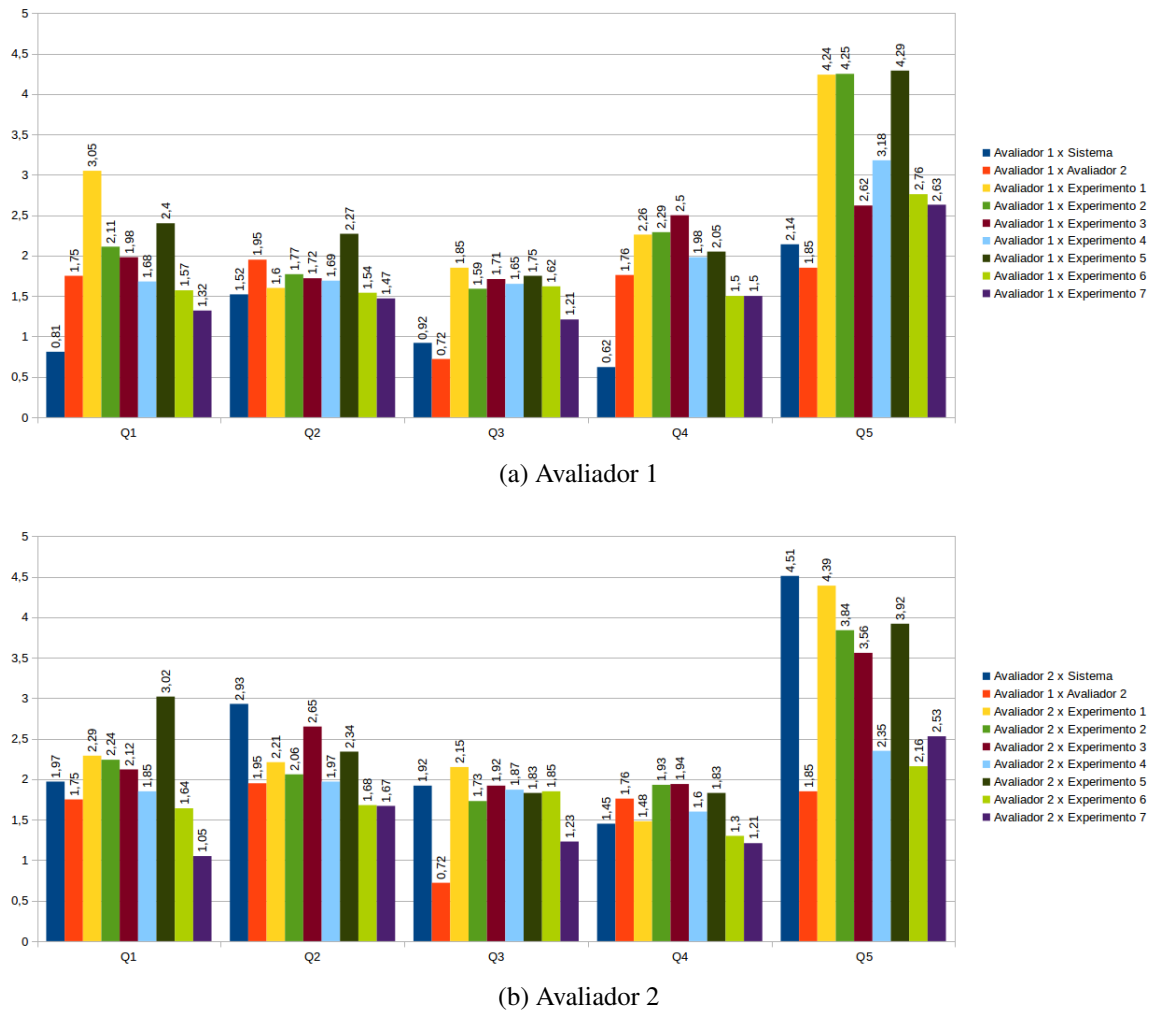


Figura 5.2: Desvio Padrão do Erro - SEM para os sete experimentos encontrados no trabalho tomado como referência, o sistema aqui apresentado e cada avaliador, utilizando como *baseline* o erro entre os próprios especialistas.

mostra que o sistema tem critérios de regulares de avaliação semelhantes a esse corretor, enquanto para o Avaliador 2 o desvio médio foi de 2,556 pontos. Quanto ao erro médio de 1,25 pontos para o primeiro avaliador e 2,32 pontos para o segundo avaliador mostramos que além de regulares temos bons modelos de avaliação. Os resultados são muito bons tendo em vista que o erro entre os dois humanos é, em média, 1,46 e os desvios 1,57 pontos.

Abaixo, vemos a descrição da atividade 5, para identificar o que possivelmente gerou os piores resultados:

“*QUESTÃO 5: Com base nos elementos constitutivos do ato de comunicação, Roman Ja-*

*kobson estabeleceu seis funções da linguagem (e a ênfase de cada uma delas): referencial (ênfase no assunto; no conteúdo), emotiva (ênfase no emissor; no sujeito), conativa (ênfase no receptor; no interlocutor), poética (ênfase na forma; na construção), metalinguística (ênfase no código; na autorreferência) e fática (ênfase no canal; no contato).*

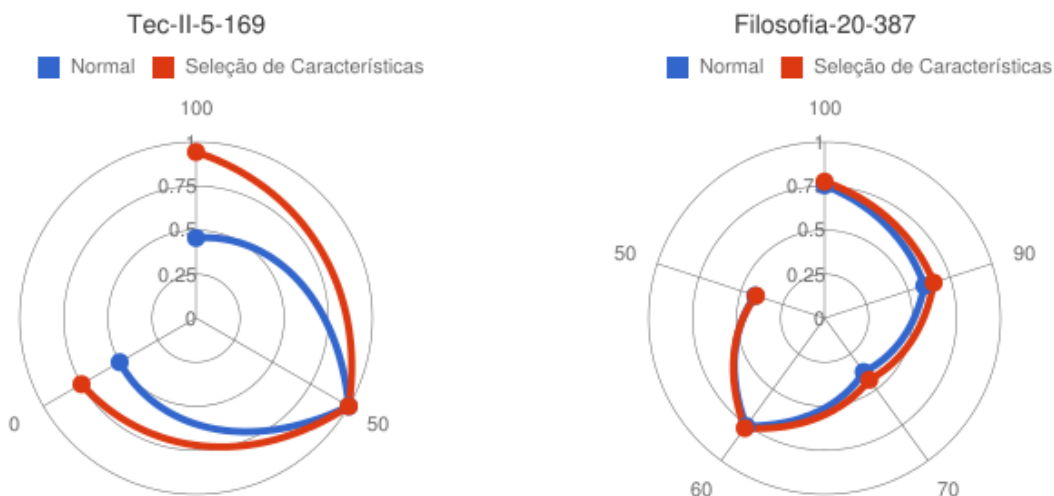
- “O navio negreiro” - Castro Alves
- “7” - Mário de Sá-Carneiro
- “Os arredores florem” - Paulo Roberto Sodré

*Escolha um dos textos da 4ª Questão indique e explique a ocorrência de uma dessas funções.”*

Nessa atividade, como vemos ao analisar o enunciado, esperam-se textos personalizados em cada resposta. Isso se deve a liberdade dada para cada estudante para escolher entre três textos e sete funções linguísticas, tornando a correção muito específica para cada resposta. A combinação de funções linguísticas com os textos disponíveis tornam a base de 92 textos transcritos pequena para a quantidade de padrões de resposta disponíveis. Possivelmente, o nível de erro apresentado é uma consequência dessa diversidade de padrões de resposta.

### **5.2.2 Experimentos com as Disciplinas da UFES**

Após a validação do modelo básico por similaridade direta, essa segunda base de dados também em língua portuguesa, usa atividades de professores da UFES para a análise de aumento de similaridade e visualização da informação por classe de nota. O aumento de similaridade interna pode ser exemplificado através de duas tarefas (Figuras 5.3a e 5.3b). Nas figuras, estão as atividades *TECH-5-169* e *Filosofia-20-387* mostrando a similaridade interna de 0 à 1 para as classes de nota. A medida foi obtida pela média da similaridade interna por pares de documentos em cada nota antes e depois da seleção de características.



(a) Média de similaridade interna para as classes da atividade TEC-II-5-169.

(b) Média de similaridade interna para as classes da atividade Filosofia-20-387.

Figura 5.3: Similaridade interna por classe de nota em atividades de disciplinas da UFES antes e depois da seleção de características.

Como podemos observar nas Figuras 5.3a e 5.3b, existem ganhos superiores à 30% de similaridade. Especificamente, se analisarmos os resultados apresentados nas Tabelas C.2 e C.6 do Apêndice, podemos ver que para as atividades apresentadas temos reduções de, respectivamente, 79,23%, 54,05% do total de características.

Para a atividade 1, *Tec-II-5-169*, temos 42 características selecionadas para a classe 100, e apenas 1 para a classe 0. Para a classe 50, não tivemos redução por conta do grupo ser formado por apenas uma instância. Nesses casos, a análise de similaridade interna por pares não é aplicável e não ocorre redução. Assim, das 207 características distintas apenas 43 foram levadas em consideração. A similaridade interna nesse caso foi alta para a classe 100. Nessa classe a similaridade inicial de 0,4554 foi ampliada para 0,9431. Assim, os 42 termos retornados devem representar bem a resposta esperada pelo professor pois faz parte, praticamente, de todas as respostas avaliadas dessa forma.

Para a atividade 2, *Filosofia-20-387*, apesar da redução para apenas 635 das 1382 iniciais, não observamos aumento da similaridade interna. Assim, as características foram identificadas para cada classe normalmente, mas o modelo de avaliação não torna a equivalência dos

documentos exata para todas as respostas. Apesar de não ter sido ampliada pela seleção de características, a similaridade interna das amostras por classe se mantiveram em níveis altos. Para a classe 100, por exemplo, a similaridade interna foi de 0,7537 para 0,775. Normalmente, em casos onde ocorrem problemas na identificação das características relevantes os resultados são visivelmente ruins. Ao analisar a Tabela C.4 do Apêndice, encontramos para os dois classificadores utilizados resultados melhores para a atividade *Tec-II-5-169*, com *precision* e *recall* de 100%. Para a atividade *Filosofia-20-387* com a questão “Para o filósofo Kant, o que é ser livre?”, até por conta da abertura dada ao aluno, as métricas de classificação têm queda de *accuracy* em 6% para o K-NN e 8% para o CBC.

Para as atividades dessa base de dados (Tabela C.4), o classificador K-NN obteve resultados médios de *precision* de 94,29% e *recall* de 88,64%. Enquanto isso, para o CBC os resultados médios de *precision* são de 92,86% e de *recall* 80,00%.

### 5.2.3 Experimentos com o *Texas Dataset*

Outra base de dados testada, formada por respostas de questões da aula de Estrutura de Dados da Universidade do Norte do Texas. Para as atividades dessa base de dados, em geral as classificações foram melhores antes da seleção de características, que apresentou para o classificador KNN médias de *precision* de 87,85%, *recall* de 81,80% e *accuracy* de 81,80%. Com o CBC os resultados foram melhores, com *precision* de 92,45%, *recall* de 87,50% e *accuracy* de 87,50%. Esses resultados foram superiores aos valores apresentados por (MOHLER; BUNESCU; MIHALCEA, 2011), com *precision* de 85,00% e *recall* de 62,00%.

Na classificação prevaleceram as classificações iniciais. Uma redução em *precision* e *recall* ocorreu para algumas atividades e são associadas à perda de informação da redução de dimensionalidade como apresenta a Tabela C.12 do Apêndice. Isso ocorre quando o mapa de características não encontra boa intercessão por classe de nota. Das 86 atividades, para

o classificador K-NN 27 tiveram uma mudança positiva em *precision* enquanto em 34 a mudança foi negativa. Nas demais 25 atividades não tiveram alterações maiores de 0,1%. Porém, o classificador CBC apresentou maiores problemas quanto a isso. Após a redução, em 56 atividades tiveram queda de *precision* enquanto 18 apresentaram melhoria. Nas outras 12 atividades não ocorreram alterações significativas. Apesar disso, o sistema não influencia negativamente na construção dos mapas de características.

Em geral, são bons os resultados médios de *precision* e *recall* mesmo após a redução de dimensionalidade com, respectivamente, 86,3% e 74,4% para o KNN 82,3% e 73,3% para o CBC. As métricas de classificação para os dois algoritmos antes e depois da seleção de características em cada atividade também estão disponíveis na Tabela C.10 do Apêndice.

#### 5.2.4 Experimentos com Dados do Kaggle

A base de dados do Kaggle foi utilizada de forma distinta da perspectiva da competição. Apenas são utilizados os dados de treinamento para aquisição de informações por *clustering*. Os demais dados não selecionados pela aquisição de padrões de correção formaram o conjunto de teste da classificação. A única alteração no processo dos demais experimentos foi o aumento do número inicial dos testes de melhor *clustering*. O início da busca foi de 3 para em 30 *clusters*, visando contornar as altas taxas de *overlapping* dado o *silhouette score* negativo em boa parte dos agrupamentos. Além disso, essa mudança causa aumento no percentual de treino para alcançar números significativos. Os dados de treinamento e o número de *clusters* para cada atividade é apresentada na Tabela 5.1.

Podemos ver através da Tabela 5.1 que temos um esforço muito baixo de correção apesar do aumento do número mínimo de *clusters*. Devemos observar a representatividade desse treino conforme os resultados de classificação. Os detalhes da avaliação segundo *precision* e *recall* estão disponíveis na Tabela 5.2, observando o impacto da seleção de características e a compatibilidade com a avaliação de cada avaliador.

<i>Dataset</i> KAGGLE	1	2	3	4	5	6	7	8	9	10
Nº Grupos	38	46	30	30	41	59	39	56	44	45
Treino (em %)	12,56	19,25	8,99	9,90	12,48	16,92	12,12	16,51	13,40	15,24
Treino (em respostas)	210	246	170	172	224	304	218	297	241	250
Total (em respostas)	1672	1278	1891	1738	1795	1797	1799	1799	1798	1640

Tabela 5.1: Informações de treinamento da base de dados após particionamento durante a análise de distribuição na etapa de *clustering*.

A Tabela 5.2 apresenta a compatibilidade da correção do sistema para os dois avaliadores. Podemos, com base nas métricas aferir que algumas atividades têm problemas de representatividade pois *precision* é muito próximo aos 50%. Nos casos de *precision* baixa, a extração de informações *intra-cluster* provavelmente não é suficientemente abrangente, o que é esperado com um treinamentos em torno de 10% dos documentos. O que não ocorre quando mais exemplos são utilizados para treinamento. Uma análise refinada desses grandes *clusters* deverá ser realizada para uma subdivisão ou a aquisição de uma quantidade maior de treino buscando dados mais representativos.

Para a classificação dos demais itens é relevante salientar o seu bom desempenho em sete dos dez *datasets*. Os classificadores alcançam *precision* médio de 76,51% para o primeiro avaliador e 76,00% para o segundo, com destaque para as atividades 4, 5 e 6 que apresentaram valores acima dos 90%. Sabendo do grande número de submissões se comparada com as demais bases de dados, a adequação ao método avaliativo de ambos corretores mostra a eficácia do *software* e seu baixo de esforço de correção necessário. Sendo que, a base de dados particionada foi disponibilizada inteiramente para treinamento dos algoritmos participantes da competição, ao qual usamos menos de 20% dos itens para correção.

### 5.2.5 Experimentos com Mapa de Características

O mapa de características, como apresentado no Capítulo 4, é uma forma de visualização dos resultados segundo a representatividade dos termos. As marcações propostas são au-

Dataset	Avaliador	Antes da Seleção de Características				Depois da Seleção de Características			
		KNN		CBC		KNN		CBC	
		Precision	Recall	Precision	Recall	Precision	Recall	Precision	Recall
KAGGLE 1	1	0,5260	0,3768	0,5129	0,4839	0,5592	0,3451	0,5613	0,4605
	2	0,5319	0,3780	0,5176	0,4934	0,5587	0,3397	0,5787	0,5042
KAGGLE 2	1	0,4361	0,3529	0,4918	0,4898	0,3944	0,3850	0,5063	0,4867
	2	0,4306	0,3513	0,4847	0,4812	0,4746	0,4249	0,5184	0,4969
KAGGLE 3	1	0,5529	0,4823	0,4444	0,4469	0,5759	0,4014	0,4901	0,4617
	2	0,6063	0,5098	0,4980	0,4944	0,5865	0,4971	0,5334	0,5056
KAGGLE 4	1	0,8777	0,4931	0,6721	0,6358	0,9007	0,4810	0,7284	0,6133
	2	0,8890	0,4781	0,6253	0,6197	0,9043	0,4764	0,7024	0,5875
KAGGLE 5	1	0,9760	0,7989	0,8213	0,8123	0,9779	0,7972	0,8962	0,8290
	2	0,9762	0,7967	0,8315	0,8078	0,9749	0,7961	0,8858	0,8340
KAGGLE 6	1	0,9579	0,8731	0,7172	0,7323	0,9608	0,8809	0,8490	0,8548
	2	0,9593	0,8709	0,7293	0,7490	0,9720	0,8709	0,7888	0,7997
KAGGLE 7	1	0,7644	0,6459	0,6521	0,6165	0,5774	0,5264	0,7418	0,6348
	2	0,7630	0,6465	0,6515	0,6115	0,5936	0,5192	0,7190	0,6092
KAGGLE 8	1	0,6798	0,5431	0,5957	0,6009	0,7831	0,4247	0,6633	0,5264
	2	0,6182	0,5425	0,6101	0,6087	0,7569	0,4330	0,6070	0,5520
KAGGLE 9	1	0,7659	0,4149	0,6379	0,6307	0,7268	0,3849	0,6420	0,5512
	2	0,8016	0,4705	0,6410	0,6346	0,8497	0,3826	0,6624	0,5423
KAGGLE 10	1	0,6550	0,4537	0,6547	0,6415	0,7081	0,3909	0,6427	0,6311
	2	0,6591	0,4579	0,6432	0,6299	0,7133	0,4128	0,6074	0,5854

Tabela 5.2: Avaliação da classificação para cada atividade da base de dados conforme 6 métricas e dois classificadores antes e depois da seleção de características.

automaticamente relacionadas conforme a avaliação do professor e o modelo de avaliação do *software*. O resultado das marcações são termos destacados pela correlação com uma nota (com a cor específica) ou neutra (com a cor preta), esse último caso ocorre em baixa ou nula correlação com a avaliação. Para o mapa de características os demais termos não marcados não fazem parte do conteúdo essencial de nenhuma classe de nota.

Para as atividades do *Texas Dataset* temos, além da avaliação, respostas aguardadas apresentados pelos corretores. Por exemplo, na atividade 1.5 apresentada na Figura 5.4, os alunos responderam a seguinte pergunta: “*what is a variable?*”. Nas respostas vemos marcações que

relacionam quatro notas distintas com os termos. Essas marcações variam em 10 tons de coloração de 0 equivalente à neutro até 9 com alta correlação à nota. Os termos não marcados são desconsiderados pelo sistema. As alterações de tonalidade são dispostas em até 6 cores vinculadas a uma nota específica.

**Variable** can be a integer or a **string** in a **program**

In **programming** a **structure** that **holds data** and is **uniquely named** by the **programmer** It **holds** the **data assigned** to it until a **new value** is **assigned** or the **program** is **finished** br

It is a **location** in the **computer** s **memory** where it can be **stored** for **use** by a **program**

A **location** in **memory** where **value** can be **stored**

a **value word** that can assume any of a set of **values**

A pointer to a **location** in **memory**

A **variable** is the **memory** address for a specific type of **stored data** or from a mathematical perspective a **symbol** representing a fixed definition with **changing values**

A **variable** is a **value** that is subject to **change** in a **computer** s **memory** that can be **used** by **programs** **Programs** can **change** the **value** of the **variable** and recall it later or act on it directly

a **symbol** that **stands** in for a **value** that **may** or **may not change** depending on the **program**

a placeholder to **hold** information **used** in the **program** br for **example** br **int** can **hold** 1 2 3 4 68 72 256 etc br **float** can **hold** 1 54 55 55 1 24 5 657 8 8123 et br **char** can **hold** A B C D E F 4 5 6 P etc br

a **stored value** used by the **program**

A way to **store** different **values** into the **program** such as **numbers words** letters etc

A **variable** is a **location** in **memory** where a **value** can be **stored**

An object with a **location** in **memory** where **value** can be **stored** br

**location** in **memory** where a **value** can be **stored**

a **variable** is an object where **data** is **stored**

**Location** in **memory** where a **value** can be **stored**

it is a **location** in **memory** where **value** can be **stored** br

A **variable** is the **location** in a **computer** s **memory** where a **value** can be **stored** for **use** by a **program**

A **variable** is a **location** in a **computers memory** where a **value** can be **stored** br br for **use** by a **program**

a **location** in **memory** where **data** can be **stored** and retrieved

Is a method or identifier I would say we **use** to bind a **data** object to **memory location** which is then **stored** in a **location** that can be accessed when and manipulated later when the **variable name** is called

**Variable** is a **location** in the **computer** s **memory** in which a **value** can be **stored** and later can retrieve that **value**

A **variable** is a **location** in the **computer** s **memory** where a **value** can be **stored** for **use** by a **program** Each **variable** has a **name** a **value** a type and a **size**

A **named** object that can **hold** a numerical or letter **value**

It s a sybol or **name** for a **value number** Example a\_used\_number can **stand** for any given **number** and the **programmer** can refer to that **number** by **using** the **variable name**

A **variable** is a **location** in the **computers memory** where a **value** can be **stored** for **use** by a **program**

An identifier that **holds** a **location** in **memory**

a block of **memory** that **holds** a specific type of **data**

#### LEGENDA

Nota 5.0 Nota 4.0 Nota 3.0 Nota 2.0

Figura 5.4: Mapa de características de todos os alunos para a atividade *DS CC UNT 1-5*.

Na Figura 5.4 podemos ver que as palavras *location*, *memory* e *value* são fortemente ligadas com a nota máxima 5,0 em azul. Enquanto isso *program* e *string* pertencem ao grupo



de nota 4,0 em amarelo, *variable name* e *stored* ao de nota 3,0 em verde e *data* e *assigned* remetem à nota 2,0 em vermelho. Palavras pouco correlacionadas como *symbol*, *example*, *computers* e *programmer* são consideradas neutras. Porém, se compararmos as marcações com a resposta esperada apresentada pelo professor “*a location in memory that can store a value*”, temos as principais palavras diretamente relacionadas ao critério de correção.

Apesar disso, para apresentar a influência da avaliação na ponderação dos termos podemos analisar uma questão de nota única como a 9.7, marcada em vermelho. A pergunta apresentada aos alunos foi “*what data structure is more appropriate for scheduling printing jobs at a printer, a stack or a queue?*”. A resposta aguardada para essa pergunta é “*queue*” e pode ser vista em todos os documentos da Figura 5.5.

Através da Figura 5.5 podemos descrever a função de seleção de forma simplificada. O fato da atividade toda ser avaliada com a nota 5 e todas as respostas conterem *queue* a correlacionam fortemente com a nota. As demais palavras marcadas, são realçadas pelo sistema apenas por aparecerem com certa frequência, mesmo não sendo o núcleo da resposta. Isso ocorre por exemplo com *appropriate*, *first*, *job*, *printer*. Os termos neutros marcados em negrito não são tão representativos à ponto de serem relacionados com uma nota. Na atividade 2-6, por exemplo, ocorrem avaliações não polarizadas (neutras) para a pergunta “*what is the difference between a function prototype and a function definition?*”. Os documentos dessa questão são apresentados na Figura 5.6.

Podemos analisar a Figura 5.6 paralelamente com a chave de correção do professor para verificar os possíveis motivos da quantidade de marcações neutras. A resposta aguardada era “*a function prototype includes the function signature, i.e., the name of the function, the return type, and the parameters’ type. The function definition includes the actual body of the function*”. Pelos documentos da figura e a chave de correção vemos que a enumeração de partes de uma função era a essência da resposta. Porém, ao contrário do aguardado, não temos intercessão observável exceto em *function prototype* e *function definition*. Tais trechos são encontrados nas respostas das três notas disponíveis (5,0 3,0 e 1,0). Assim, o processo

A stack some **printing jobs might** have a **higher priority** than **others** and they can be easily **inserted** at the front of the **list** or **anywhere** between

a **queue** otherwise some things **would** never be **printed**

**queue**

a **queue**

**Queue** s are more **appropriate** as **printers** are **usually first come first serve**

**queue**

**queue** because of the **first in first out** principles and if it was a stack it **would** be **last in last out** because it could possibly not **let** the **first** item that was added get removed in a timely **manner**

**Queue** br

a **queue**

A **Queue data structure** is more **appropriate** for **scheduling printing jobs** at a **printer** The **first job sent** to the **printer** is the **first job printed**

a **queue** It wouldn't be **unfair** for the **first job** to **finish last**

a **queue** because it **takes** the **job** that has **waited** the **longest**

A **queue** is more **appropriate** for **scheduling printing jobs** because **queues** have **first in first out** behavior

A **queue** is more **appropriate** because the clients which have **waited** the **longest** should be the **first served**

**Queue**

a **queue**

**Queue**

i **would** say that a **queue** is better because the **first** thing you **try to print** should be the **first one to come** out of the **printer FIFO**

a **queue** because you **want** the **jobs** to **execute** in the **order** that they are **received**

**queue**

**Queue** its better to be **FIFO** than **LIFO** stack

**Queue** of course

**Queue** the **person** who **wanted to print first** should be **first served**

**queue** because it **guarantees** that **job** are **executed** in the **order** they **come in**

**queue**

**Queues would** be preferable to stacks for **use** as **scheduling print jobs** because it **would print jobs** in the **order** that they were **sent** to the **printer**

a **queue**

---

#### LEGENDA

**Nota 5.0**

Figura 5.5: Mapa de características de todos os alunos para a atividade *DS CC UNT 9-7*.

de avaliação foi criterioso quanto a diferenciação entre as funções, independentemente do aspecto elencado. Isso confunde o algoritmo pela ausência de termos semelhantes por classes de nota ao coexistirem várias respostas possíveis.

Apesar da ocorrência de “questões neutras” quanto ao conteúdo podemos ver em mais um exemplo de questão através da capacidade de interpretação das marcações. Para o enunciado “*what is the difference between a circular linked list and a basic linked list?*”, temos respostas bem consistentes. A Figura 5.7 apresenta as submissões dos alunos avaliadas em duas notas 5,0 e 2,0 pelo professor.

A **function definition** does not **require** any **additional information** that **needs** to be **passed inside** its **parenthesis** **br** to **execute** While a **definition** **prototype** **requires** more than **one** **parameters** to be **passed in order** to **complete** its **br** **task**

The **FUNCTION PROTOTYPE** is where the **programmer declares** that he she is **using** a **function** other than **main** This is **like declaring** a **variable** the **programmer knows** that he she will be **using** in the **future** but has **yet** to **say** where they are **going** to **use** it or how This is **answers** the **question** who it **gives** the **function** a **name** and **character** The **function prototype** by **common practice** is **placed** at the **beginning** of the **program** after the **includes** and before **main** **br** The **FUNCTION DEFINITION** is the **guts** of the **function** This is where the **programmer decides** what the **function** is **going** to do and **tells** it how to do it It **takes whatever information** it is **given** and **performs** the **operations** It works sort of like the **brain** the **brain** takes in **input** and **based upon** that **input** **performs** in some way **producing** an **output** The **function definition** is **placed outside** of **main** and any other **functions** A **function** is its own **entity** and should be **thought** of as such

**Function definitions** are just that the **definition** The **prototype** is what the **compiler** uses to **check** that **calls** to **function** are **correct**

a **prototype** does not **include** any **actual code** where the **function** has all the **code** that is **executed** in the **program**

a **prototype** **declares** what will be **used** in the **program** and the **definition**

A **function prototype** **lays out** the **name** **return type** and the **number** and **types** of **parameters** the **function** **expects** to **receive** in a **certain order** The **details** for **function prototypes** are in the **function definition**

A **function definition** is the **code** that **defines** the **function** **placed** in the **brackets** that **determines** that **function** s **operation** **br** **br** A **function prototype** shows the **function** s **public interface** without **exposing implementation** It shows **name** **return type** and **type** of **parameters**

A **function prototype** is just a **declaration** of the **function** **existing** and **cant** be **used** as an **actual function** A **function** has to be **created** with a **definition** **within** to **tell** the **compiler** what the **function** does

the **Prototype** creates a **framework** to **call** the **function definition** While a **function definition** is where the **function** is **actually** **programmed** out and **created** into a **final product**

A **function prototype** describes the **class** s **public interface** without **providing** how the **function** works **br** A **function definition** contains the **inner workings** of the **function**

**prototype** states all **functions** in that **class** before **compilation** where the **definition** actually **holds** the **source** for the **functions**

**Function prototypes** tell the **compiler** the **function** **names** its **return type** and the **types** of its **parameters** where as **function definitions** actually **implement** the **member functions**

A **function prototype** tells the **compiler** the **function** **name** **return type** and the **number** and **type** of **parameters** without **revealing** the **implementations** contained in the **function definition**

A **function prototype** just **specifies** **parameters** A **function definition** includes **parameters** and a **code body**

**Function prototypes** describe the **class** s **public interface**

A **prototype** only **declares** **name** **return type** and **input type** **br** A **definition** also **defines** the **scope** **variables** **process** and **return function**

a **function prototype** simply **declares** the **functions** **parameters** the **function definition** includes any **necessary** **variables** and the **function** s **actual code**

A **function prototype** is a **declaration** of a **function** while **function definition** specifies what a **function** does

in a **function prototype** you **include** the **return type** the **name** of the **function** and its **parameters** if any are **needed** **br** **br** in a **function definition** you **write** the **code** of what the **function** will do

A **function definition** contains all the **code** for a **function** to **work** A **function prototype** just shows the **output** **input** and **function name**

**function prototype** describe the **class** s **public interface** without **revealing** **br** **br** the **class** s **member function** **implementations** **function definitions** show what **br** **br** **implementations** are being **done**

A **function prototype** only **names** the **function** its **return type** and its **argument list** while a **definition** defines the above as **well** as what the **function** actually does

**function prototype** describes a **classes** **interface** without **reviling** whatever is **inside** as for the **function definition** can t do that **br**

**Function prototype** is a **declaration** **Function definitions** w **multiple** **parameters** often **require** more than **one** **piece** of **information** to **perform** their **tasks**

A **function prototype** is a **declaration** of a **function** that **tells** the **compiler** the **function** s **name** its **return type** and the **types** of its **parameters**

A **prototype** shows only **return types** and **necessary** **parameters** The **definition** includes **names** for those **parameters** and **defines** what the **object** is **actually capable** of doing

**prototype** only **tells** the **user** what **data** **types** **go** into a **function** and what **type** is **returned**

**Function prototype** is **located** in the **h file** and only **contains** the **access** **function** **name** and **paramater type** **br** **br** **Function definition** contains the **code** for the **function** to **perform** its **activity**

a **function prototype** is **used** to **reference** the **compiler** to a **function** that will be **defined** later on a **function definition** is the **actual** **function** itself **complete** with **return type** **parameters** etc

A **function prototype** describes a **class** s **public interface** without **revealing** the **class** s **member function** **implementations**

---

#### LEGENDA

Nota 5.0 Nota 4.0 Nota 1.0

Figura 5.6: Mapa de características de todos os alunos para a atividade DS CC UNT 2-6.

They are passed by reference because you want the function to change the pointer br

In a **circular linked list** every **node** has a **successor** which means that the **last node**'s **successor** is the **head** or **first node**

**Basic linked list** points to a **NULL** at the end and **circular points** to the **head** at the end

in a **circular linked list** the **last** item **points** to the **first** item

The **circular linked list**'s **tail** **points** to the **head** whereas the **basic linked list**'s **tail** **points** to a **NULL**

a **basic linked list** has an end with a **null** value where a **circular linked list** has a pointer from the end to the **beginning**

how to determine the end of the **list** in **basic linked lists** the **last element** **links** to a **null** pointer while **circular linked lists** **link** to the **head element** at the end

No **Node** in a **circular linked list** contains **NULL** the **last Node** **points** back to a **Node** within the **list**

The difference between a **circular linked list** and a **basic linked list** is that a **circular linked list** the **last node** references the **first node** Every **node** in a **circular linked list** has a **successor** and no **node** in a **circular linked list** contains **NULL**

No **node** in a **circular linked list** contains **NULL**

in a **circular linked list** the **last element** **points** to the **head** of the **list**

In a regular **linked list** the **last object** **points** to **NULL** In a **circular linked list** the **last object** **points** back to the **beginning** of the **list**

A **circular linked list** is a **linked list** that **links** items in a continuous **circle** This allows the full **list** to be traversed by starting at any item in the **list** a fixed **head** or **tail** is not required

In **circular linked list** each **node** has **successor** no **node** contains **NULL** while in **basic linked list** **last node** contain **NULL**

in a **circular linked lists** the **last element** **points** to the **first**

A **circular linked list** does not have a **last element** Instead it's **last** item **points** to the **head** of the **list**

in a **circular linked list** the **last node** contains a pointer that goes back to the **first node** in a **basic linked list** the **last node** contains a **null** pointer br

in a **circular linked list** the **last object** is **linked back** to the **first object**

In the case of a non **circular link list** if you are at the **last node** of the **list** and u want to move to **first** u need to go to one step back till u reach to **first** But in **circular link list** from **last** to **first** you need to make only one move br In a single **link list** u can move only in one direction but in double **link list** u can move in any direction back or forward or forward br br

In a **circular linked list** the **last node** **points** to the **first node**

On a **circular linked list** the end **element** **points** to the **head** thus making the **circle** A **basic link list** has nothing pointing to the **head** and the **last element** **points** to **null**

the **circular linked list**'s **last node** **points** to the **begining** of the **list** while a standard **linked list** **points** to **NULL**

The **last** item in the **list** **points** to the **head**

In a **circular linked list** the **last node** **points** back to the **first node** there is no **NULL**

**Circular linked lists** are most useful for describing naturally **circular** structures and have the advantage of regular structure and being able to traverse the **list** starting at any **point**

A **circular linked list** has the **last node** point back to the **head** of the **linked list**

---

#### LEGENDA

Nota 5.0 Nota 2.0

Figura 5.7: Mapa de características de todos os alunos para a atividade DS CC UNT 7-5.

Podemos ver na Figura 5.7 que para uma lista encadeada (*linked list*) em vez de nulo (*null*) a referência do último nó (*last node*) deve ser para o primeiro nó (*first node* ou *head*). E exatamente essa resposta era aguardada pelo professor: “*the last element in a circular linked list points to the head of the list*”. A palavra *back* atribuiu fator negativo de avaliação pois em uma estrutura de dados circular a sequência do último item é o inicial. Esse termo quando utilizado para uma estrutura de dados passa a ideia de retorno por todos os itens, ao contrário do que efetua o código, direcionando-o por referência.

As atividades em português da base de dados de disciplinas da UFES tiveram resultados semelhantes. Podemos, ao observar a Figura 5.8, ver comportamentos da marcação equivalente ao da Figura 5.7. Tanto para a atividade *DS CC UNT 7-5* quanto para a *Tec-II-5-170* um detalhe foi crucial para diferenciar dois níveis de nota. Para essa segunda, existe maior esforço do sistema no reconhecimento das características para a questão “Cite algumas linguagens para o lado do servidor”. Isso ocorre por conta da liberdade dos alunos para escolher qualquer linguagem de programação executadas no servidor.

Nesse caso, a Figura 5.8 deixa bem explícito o decréscimo da nota associado ao termo “mysql”. Obviamente, por ser um sistema gerenciador de banco de dados e não uma linguagem de programação, como aguardado nas respostas, era avaliado negativamente pelo professor. Nas marcações, esse termo é associado à classe 90.0 enquanto as demais “ASP”, “PHP”, “.Net”, “Python” “Pearl” e “Java” remetem à nota 100.0.

### 5.3 Discussão

Nas atividades do Vest-UFES a comparação do sistema com os experimentos de pré-processamento de (PISSINATI, 2014) temos resultados equivalentes. O sistema em todas as atividades adquire métodos avaliativos semelhantes ao Avaliador 1, tendo como *baseline* o erro entre especialistas. O autor, contudo, alcança resultados próximos aos dois avaliadores apenas após sete experimentos cumulativos. Nos modelos de avaliação, ambos os sistemas superam a correção humana nas métricas avaliadas, sendo que a análise das diferenças deve ser um fator descritivo quanto ao erro. Erros de um grau de nota, ou seja, apenas um nível de classe de diferença geralmente são considerados comuns até entre especialistas (BAZELATO; AMORIM, 2013). O baixo desvio padrão indica a ocorrência desse tipo de erro.

A comparação com os experimentos do autor, com o intuito de validar o modelo de avaliação, apresenta que o sistema aqui proposto tem boa atuação até mesmo em notas contínuas. Além disso, atividades de língua portuguesa, quando voltadas para a análise textual, induzem

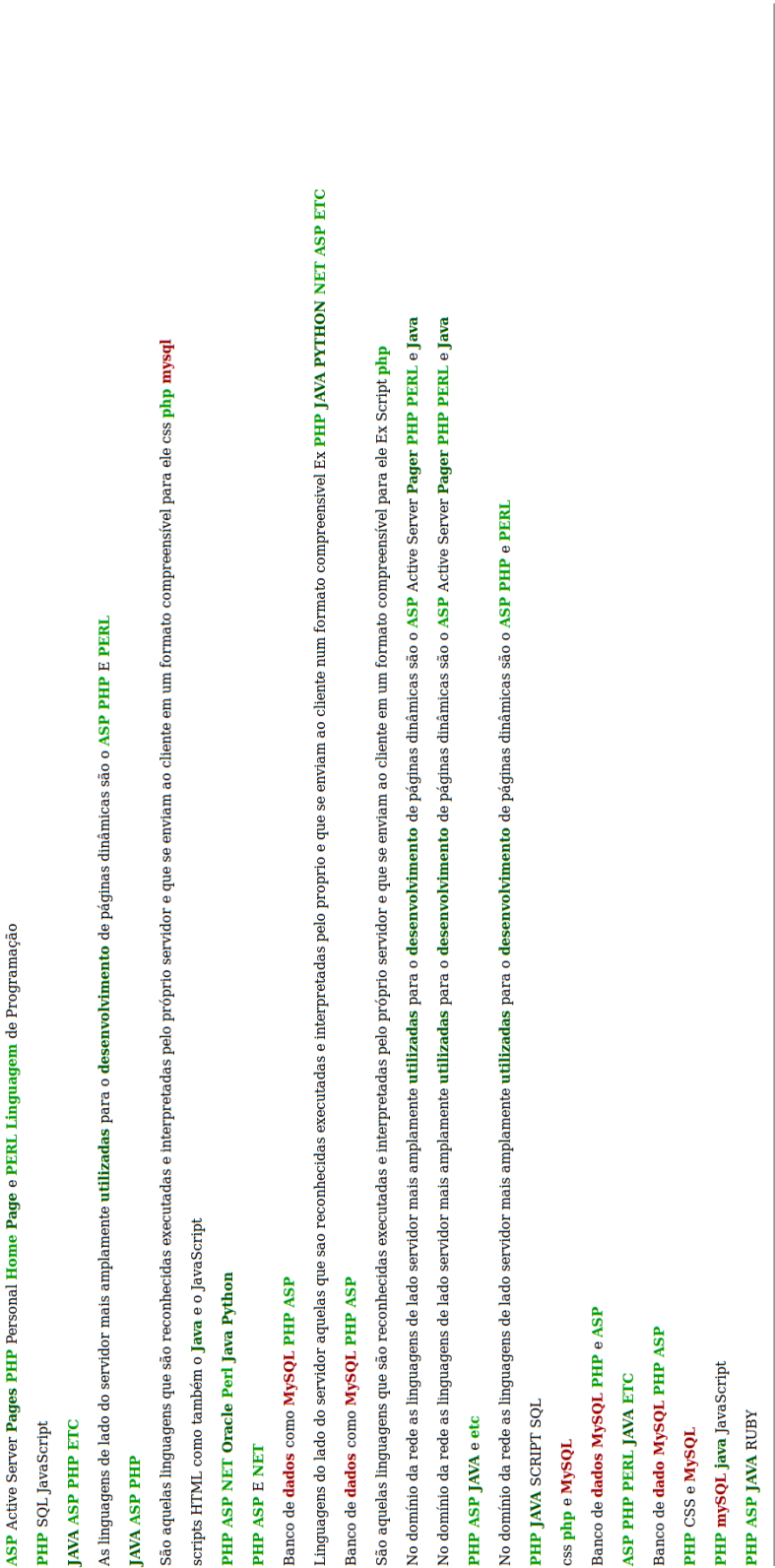


Figura 5.8: Mapa de características de todos os alunos para a atividade Tec-II-5-170.

os estudantes a interpretações distintas. Assim, com exceção da questão 5 para o segundo avaliador, o sistema tem grande qualidade resultando um baixo nível de erro.

Para as atividades das disciplinas da UFES temos como foco a avaliação por notas discretas (classes). Essa base de dados colaborou com as formas de análise de similaridade e estudo de distribuição das respostas. Com o trabalho em classes de nota, podemos observar a classificação do especialista de forma a buscar critérios específicos. Com a redução de ruídos e a sumarização através do algoritmo genético, grupos muito similares melhoraram a classificação. Enquanto isso, nas situações onde ocorreu certa dificuldade de extração dos termos ideais indicaram que o critério de correção foi além do conteúdo analisado pelo sistema ou os termos não foram correlacionados à nota tal como deveriam.

Nos casos onde algum problema referencial ao texto foi encontrado os resumos dos documentos após seleção de características não foram tão representativos. Por exemplo, se na etapa de correção o professor for indiferente quanto ao conteúdo do texto são gerados *outliers*. Outro caso são as múltiplas respostas por classe e as incompletudes na avaliação. Nesse primeiro, as várias respostas possíveis, como ocorre na atividade 5 do vestibular da UFES, algumas respostas são consideradas como corretas e recebem mesma classe, porém, são distintas para a análise do sistema. No segundo caso, como a seleção não elimina *outliers*, a ocorrência de respostas inconsistentes retorna conjuntos muito piores pela falta de um único referencial. Dessa forma, as incompletudes geram *overfitting* ou *underfitting* nos modelos. À partir do momento que estão dispostos na mesma classe, os termos excedentes são desconsiderados para o aumento de similaridade interna. Concluimos então que, se algumas respostas estão incompletas ou têm certo modelo avaliativo distinto, por consequência os termos referenciais são descartados por conta do treinamento (SPALENZA et al., 2016b).

As atividades do *Texas Dataset*, além de serem interessantes para verificação dos modelos de classificação, foram fundamentais para testes do mapa de características. Apesar dos resultados de otimização reduzirem a qualidade da classificação em determinadas questões, a disponibilidade de respostas permitiu boas comparações com a marcação das atividades.

Podemos ver isso se compararmos os resultados das métricas da Tabela C.10 com os mapas de características das Figuras 5.4, 5.5, 5.6 e 5.7, onde o mapa apoia a compreensão do modo de avaliação.

Dessa forma, essa base de dados além de bons resultados de classificação apresenta boas marcações. Essas marcações devem auxiliar tanto os professores na visão geral do professor quanto aos alunos na discutir dos critérios de avaliação como o realizado para a questão 7.5. Então, A redução das métricas ocorrida para alguns elementos dessa base de dados pode ser justificada pelos termos não correlacionados à nota como ocorreu para a atividade 2.6.

A base de dados do *Kaggle* também foi um experimento muito importante por destoar das demais três bases de dados pela quantidade de respostas e número de características encontradas. Vemos essa diferença através da Tabela C.14. Nela, Apesar dos documentos manterem o padrão de tamanho das demais respostas curtas temos um aumento considerável de características. Essa alteração apresenta o efeito da redução de dimensionalidade na classificação onde segundo a Tabela 5.2 há efeitos de até 10% de melhoria em *precision*. A *precision*, como descrito na Seção 3.4.5, representa a adequação do sistema com o método avaliativo de um especialista. Dessa forma, esse fator é essencial para mostrar que a qualidade de captação de uma boa modelagem dos critérios influencia diretamente na avaliação dos alunos.

Com resultados significativos nessa base, bem como nas demais, temos que a qualidade de correção é suficientemente adequada ao aguardado por um professor especialista. As marcações em texto do critério remetem ao modelo avaliativo do algoritmo e suporta a verificação de seu desempenho durante o processo. Através das representações coloridas, também temos a visualização dos resultados auxiliando na análise de conteúdo dos documentos e no processo de correção colaborativa independente do professor.



## Capítulo 6

### Conclusões e Trabalhos Futuros

A avaliação estudantil vêm sendo estudada de perto pelas pesquisas em informática na educação. Hoje, buscamos compreender as necessidades dos alunos, acompanhar o desempenho nas avaliações, recomendar práticas de estudo e, como nesse trabalho, apoiar o processo de avaliação.

A comparação inicial com o *dataset* do Vestibular da UFES mostra que o sistema tem erro baixo para o humano ao qual se pretende imitar, quando comparado com a distância entre dois humanos que deveriam manter certa coerência avaliativa para uma mesma atividade. Pela avaliação de um dos avaliadores vemos que a ferramenta é tão eficiente quanto os especialistas ao atribuir notas tendo como base o erro resultante nas notas contínuas.

Com precisão média acima de 90% nas bases de dados de disciplinas da UFES e *Texas Dataset*, podemos então dizer que o sistema reconhece adequadamente o modelo avaliativo. Esses resultados portanto, tornam-o um avaliador em potencial. Deste modo, podemos afirmar que, na avaliação por classes com bons resultados obtidos de *precision* e *recall*, o uso dessa ferramenta de avaliação semiautomática é capaz de apoiar de forma equivalente aos especialistas no processo avaliativo.

Além da eficiência na modelagem avaliativa, ao apresentar criteriosamente sua meto-

dologia através do mapa de características, o *software* pode ser controlado de perto pelo especialista. Partindo da avaliação dada pelo próprio humano, a identificação de modelos de correção, o mantém no controle do processo principalmente com o mapa de características para validação após o *feedback*. Dessa forma o mapa, como resultado do processo, torna-se mais um importante instrumento para a discussão de resultados e a correção colaborativa.

Como trabalhos futuros, esperamos aprimorar a seleção de características conforme a distribuição analisada na etapa de *clustering*. Remontando os agrupamentos da etapa inicial podemos identificar pares distintos de resposta dentro de cada classe. Com análises mais complexas desses documentos, estabelecemos critérios mais próximos do professor através dos termos que descrevem melhor cada conteúdo.

Segundo a distribuição, uma melhoria em potencial dados os experimentos é a adição de funções de remoção de *outliers*. Com a remoção desse tipo de documento espera-se tornar mais qualitativa a seleção dos termos ao aumentar a consistência das classes. Se os grupos são mais densos pela similaridade entre os documentos a seleção tende a evitar elementos de baixa correlação representando a nota com os termos mais correlacionados.

São importantes também, para incentivar o uso do sistema com especialistas (independentemente da área), a construção de relatórios que se aproximem com seus critérios de correção. Um passo que deverá ser incorporado para garantir isso seria a criação do quadro de *rubrics* (ARTER; CHAPPUIS, 2006). Portanto, com relatórios objetivos, representamos melhor o conhecimento abordado na tarefa e nos aproximamos da metodologia do professor.

# Referências Bibliográficas

- ALFONSECA, E.; PÉREZ, D. Automatic Assessment of Open Ended Questions with a Bleu-Inspired Algorithm and Shallow NLP. In: \_\_\_\_\_. *Advances in Natural Language Processing: 4th International Conference, EsTAL 2004, Alicante, Spain, October 20-22, 2004. Proceedings*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004. p. 25–35.
- ALGARNI, A.; TAIRAN, N. Feature Selection and Term Weighting. In: *2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*. Warsaw, Poland: IEEE Computer Society Press, 2014. v. 1, p. 336–339.
- ARTER, J. A.; CHAPPUIS, J. *Creating & Recognizing Quality Rubrics*. Nee York, USA: Pearson Education, 2006. (Assessment Training Institute, Inc Series).
- ASSUNÇÃO, A. Á.; OLIVEIRA, D. A. Intensificação do Trabalho e Saúde dos Professores. *Educação & Sociedade, SciELO*, v. 30, p. 349 – 372, 08 2009.
- BAEZA-YATES, R.; RIBEIRO-NETO, B. *Modern Information Retrieval*. 2nd. ed. USA: Addison-Wesley Publishing Company, 2008.
- BAILEY, S.; MEURERS, D. Diagnosing Meaning Errors in Short Answers to Reading Comprehension Questions. In: *Proceedings of the 3rd Workshop on Innovative Use of NLP for Building Educational Applications, held at ACL 2008*. Columbus, Ohio, USA: Association for Computational Linguistics, 2008. p. 107–115.
- BARREIRA, C.; BOAVIDA, J.; ARAÚJO, N. Avaliação Formativa: Novas Formas de Ensinar e Aprender. *Revista Portuguesa de Pedagogia*, Universidade de Coimbra, v. 40, n. 3, p. 95–133, 2006.
- BAZELATO, B. S.; AMORIM, E. C. F. A Bayesian Classifier to Automatic Correction of Portuguese Essays. In: *XVIII Congreso Internacional de Informática Educativa (TISE)*. Porto Alegre, Brazil: Centro de Computación y Comunicación para la Construcción del Conocimiento, 2013. v. 18.
- BEZERRA, M. A. Questões Discursivas para Avaliação Escolar. *Acta Scientiarum Language and Culture*, Universidade Estadual de Maringá, v. 30, p. 149–157, 2008.
- BURROWS, S.; GUREVYCH, I.; STEIN, B. The Eras and Trends of Automatic Short Answer Grading. *International Journal of Artificial Intelligence in Education*, v. 25, n. 1, p. 60–117, 2015.

- BURSTEIN, J. et al. Automated scoring using a hybrid feature identification technique. In: *Proceedings of the 17th International Conference on Computational Linguistics - Volume 1*. Stroudsburg, PA, USA: Association for Computational Linguistics, 1998. (COLING '98), p. 206–210.
- COHEN, P. R. *Empirical Methods for Artificial Intelligence*. Cambridge, MA, USA: MIT Press, 1995.
- CONOLE, G.; Warburton, B. A review of computer-assisted assessment. *Research in Learning Technology*, Co-Action Publishing, v. 13, n. 1, p. 17–31, 2005.
- CONORT, X. *Short Answer Scoring: Explanation OF Gxav Solution*. San Francisco, CA, USA: Kaggle, 2012. <[http://github.com/Gxav73/Gxav\\_Sol\\_ASAP\\_round2](http://github.com/Gxav73/Gxav_Sol_ASAP_round2)>. ASAP '12 SAS Methodology Paper.
- CUTRONE, L.; CHANG, M.; KINSHUK. Auto-Assessor: Computerized Assessment System for Marking Student's Short-Answers Automatically. In: *2011 IEEE International Conference on Technology for Education*. Chennai, India: IEEE Computer Society Press, 2011. p. 81–88.
- DATAR, A. et al. *EGAL - Essay Grading and Analysis Logic*. Duluth, MN, USA: University of Minnesota Duluth, 2013. <<http://egal.sourceforge.net>>.
- DEERWESTER, S. et al. Indexing by Latent Semantic Analysis. *Journal of the American Society for Information Science*, Wiley Subscription Services, Inc., A Wiley Company, v. 41, n. 6, p. 391–407, 1990.
- FIGUEIRA, A. S. et al. Module of Evaluation Automatic Essay Questions on Virtual Learning Environment LabSQL. In: *2013 8th Iberian Conference on Information Systems and Technologies (CISTI)*. Lisboa, Portugal: Institute of Electrical and Electronics Engineers (IEEE), 2013. p. 1–5.
- FOLTZ, P. W.; LAHAM, D.; LANDAUER, T. K. Automated Essay Scoring: Applications to Educational Technology. In: COLLIS, B.; OLIVER, R. (Ed.). *Proceedings of EdMedia: World Conference on Educational Media and Technology 1999*. Seattle, WA USA: Association for the Advancement of Computing in Education (AACE), 1999. p. 939–944.
- GABRILOVICH, E.; MARKOVITCH, S. Computing Semantic Relatedness Using Wikipedia-based Explicit Semantic Analysis. In: *Proceedings of the 20th International Joint Conference on Artificial Intelligence*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2007. (IJCAI'07), p. 1606–1611.
- GAMAGE, D.; PERERA, I.; FERNANDO, S. Evaluating Effectiveness of MOOCS Using Empirical Tools: Learners Perspective. In: *INTED2016 Proceedings*. Valencia, Spain: IATED, 2016. (10th International Technology, Education and Development Conference), p. 8276–8284.
- GOOSSENS, M.; MITTELBAACH, F.; SAMARIN, A. *The Latex Companion*. 1st. ed. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1994.

- GÜTL, C. e-Examiner: Towards a Fully-Automatic Knowledge Assessment Tool Applicable in Adaptive e-Learning Systems. In: *Proceedings of the 2nd International Conference on Interactive Mobile and Computer Aided Learning*. Amman, Jordan: Princess Sumaya University for Technology (PSUT), 2007. p. 1–10.
- HAUPT, R.; HAUPT, S. *Practical Genetic Algorithms*. 2. ed. New York, NY, USA: John Wiley & Sons, Inc., 2004. (Wiley InterScience Electronic Collection).
- HOW, B. C.; NARAYANAN, K. An Empirical Study of Feature Selection for Text Categorization based on Term Weightage. In: *Web Intelligence, 2004. WI 2004. Proceedings. IEEE/WIC/ACM International Conference on*. Beijing, China: IEEE Computer Society Press, 2004. p. 599–602.
- ISHIOKA, T.; KAMEDA, M. Automated Japanese Essay Scoring System: Jess. In: *Proceedings. 15th International Workshop on Database and Expert Systems Applications, 2004*. Zaragoza, Spain: IEEE Computer Society Press, 2004. p. 4–8.
- JÚNIOR, J. J. S.; PAIVA, R.; BITTENCOURT, I. I. Avaliação Automática de Atividades Escritas Baseada em Algoritmo Genético e Processamento de Linguagem Natural: Avaliador Ortográfico-Gramatical. In: *Anais dos Workshops do Congresso Brasileiro de Informática na Educação*. Maceió, AL, Brasil: Sociedade Brasileira de Computação, 2015. v. 4.
- KAKKONEN, T. et al. Automatic essay grading with probabilistic latent semantic analysis. In: *Proceedings of the Second Workshop on Building Educational Applications Using NLP*. Stroudsburg, PA, USA: Association for Computational Linguistics, 2005. (EdAppsNLP 05), p. 29–36.
- KARYPIS, G. *CLUTO-a clustering toolkit*. Minneapolis, MN, USA: University of Minneapolis Dept. of Computer Science, 2002. DTIC Document <<http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA439508>>.
- KULKARNI, C. E. et al. Scaling Short-Answer Grading by Combining Peer Assessment with Algorithmic Scoring. In: *Proceedings of the First ACM Conference on Learning @ Scale Conference*. New York, NY, USA: ACM, 2014. (L@S '14), p. 99–108.
- LANDAUER, T. K.; FOLTZ, P. W.; LAHAM, D. An introduction to latent semantic analysis. *Discourse Processes*, Taylor & Francis, v. 25, n. 2-3, p. 259–284, 1998.
- LEFFA, V. J. Análise Automática da Resposta do Aluno em Ambiente Virtual. *Revista Brasileira de Linguística Aplicada*, SciELO, v. 3, p. 25 – 40, 00 2003.
- LEVY, O. et al. Recognizing Partial Textual Entailment. In: *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*. Sofia, Bulgaria: Association for Computational Linguistics (ACL), 2013. p. 451–455.
- LIMA, J. d. O. d.; GIRAFFA, L. M. M. Um Estudo Sobre Funcionalidades Dos Softwares Educacionais Para Suporte A Aprendizagem De Matemática Do Ensino Médio. Porto Alegre, RS, 4 2006.
- MANNING, C. D.; SCHÜTZE, H. *Foundations of Statistical Natural Language Processing*. Cambridge, MA, USA: MIT Press, 1999.

- MARTENS, V. M. Nivelamento de Habilidades de Leitura e Escrita: Um Fazer Pedagógico nas Escolas Públicas de Ensino Integral Frente ao Fracasso Escolar. *Construção psicopedagógica*, SciELO Periódicos Eletrônicos de Psicologia (PePSIC), v. 24, p. 119–132, 2016.
- MITCHELL, T. et al. Towards Robust Computerised Marking of Free-Text Responses. In: *Proceedings of the 6th CAA Conference*. Loughborough, England: Loughborough University, 2002. p. 233–249.
- MOHLER, M.; BUNESCU, R.; MIHALCEA, R. Learning to Grade Short Answer Questions Using Semantic Similarity Measures and Dependency Graph Alignments. In: *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1*. Stroudsburg, PA, USA: Association for Computational Linguistics, 2011. (HLT '11), p. 752–762.
- MOHLER, M.; MIHALCEA, R. Text-to-text semantic similarity for automatic short answer grading. In: *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics*. Stroudsburg, PA, USA: Association for Computational Linguistics, 2009. (EACL '09), p. 567–575.
- OLIVEIRA, E. et al. Uma Tecnologia de Agrupamento de Respostas para Redução de Esforço de Correção de Atividades em Sistema Online de Apoio à Avaliação Formativa em Indexação. In: *XI Encontro Nacional de Pesquisa em Ciência da Informação*. Rio de Janeiro, RJ, Brasil: Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação (Ancib), 2010.
- OLIVEIRA, E. de et al. Combining Clustering and Classification Approaches for Reducing the Effort of Automatic Tweets Classification. In: *Proceedings of the International Conference on Knowledge Discovery and Information Retrieval - Volume 1: KDIR, (IC3K 2014)*. Rome, Italy: ScitePress, 2014. p. 465–472.
- OLIVEIRA, K. L. d.; SANTOS, A. A. A. Compreensão em Leitura e Avaliação da Aprendizagem em Universitários. *Psicologia: Reflexão e Crítica*, SciELO, v. 18, p. 118 – 124, 04 2005.
- PAGE, E. B. The Imminence of... Grading Essays by Computer. *The Phi Delta Kappan*, Phi Delta Kappa International, v. 47, n. 5, p. 238–243, 1966.
- PAGE, E. B.; PAULUS, D. H. *The Analysis of Essays by Computer. Final Report*. Washington, DC: Office of Education (DHEW), 1968.
- PAIVA, R. et al. Mineração de Dados e a Gestão Inteligente da Aprendizagem: Desafios e Direcionamentos. In: *I Workshop de Desafios da Computação Aplicada à Educação (DesafIE/2012)*. Curitiba, PR, Brasil: Sociedade Brasileira de Computação, 2012. v. 1.
- PASSERO, G.; FILHO, A. H.; DAZZI, R. L. S. Avaliação do Uso de Métodos Baseados em LSA e WordNet para Correção de Questões Discursivas. In: *Simpósio Brasileiro de Informática na Educação (SBIE)*. Uberlândia, MG, Brasil: Sociedade Brasileira de Computação, 2016. v. 27.

- PÉREZ, D. et al. About the Effects of Combining Latent Semantic Analysis with Natural Language Processing Techniques for Free-Text Assessment. *Revista Signos*, SciELO, v. 38, p. 325 – 343, 2005.
- PEREZ-MARÍN, D. et al. Willow: Automatic and Adaptive Assessment of Students Free-Text Answers. *Sociedade Espanhola de Processamento de Linguagem Natural(SEPLN)*, v. 37, p. 367–368, 2006.
- PISSINATI, E.; OLIVEIRA, M. G.; OLIVEIRA, E. Direcionando a Atuação do Professor Através de uma Ferramenta de Visualização. In: *Escola de Computação Bahia-Alagoas-Sergipe (ERBASE)*. Feira de Santana, BA, Brasil: Sociedade Brasileira de Computação, 2014. v. 14.
- PISSINATI, E. M. *Uma Proposta de Correção Semi-Automática de Questões Discursivas e de Visualização de Atividades para Apoio à Atuação do Docente*. Dissertação (Mestrado) — PPGI - Universidade Federal do Espírito Santo, Vitória, ES, Setembro 2014.
- PULMAN, S. G.; SUKKARIEH, J. Z. Automatic Short Answer Marking. In: *Proceedings of the Second Workshop on Building Educational Applications Using NLP*. Stroudsburg, PA, USA: Association for Computational Linguistics, 2005. (EdAppsNLP 05), p. 9–16.
- RAYMER, M. L. et al. Dimensionality reduction using genetic algorithms. *IEEE Transactions on Evolutionary Computation*, v. 4, n. 2, p. 164–171, Jul 2000.
- ROMERO, C. et al. *Handbook of Educational Data Mining*. Boca Raton, FL: CRC Press, 2010.
- ROUSSEEUW, P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, v. 20, p. 53 – 65, 1987.
- ROY, S.; BHATT, H. S.; NARAHARI, Y. An Iterative Transfer Learning Based Ensemble Technique for Automatic Short Answer Grading. *The Computing Research Repository (CoRR)*, dblp computer science bibliography, abs/1609.04909, oct 2016.
- RUDNER, L. M.; LIANG, T. Automated Essay Scoring Using Bayes' Theorem. *The Journal of Technology, Learning and Assessment*, eScholarship at Boston College, v. 1, n. 2, p. 22, 2002.
- SALTON, G.; WONG, A.; YANG, C. S. A vector space model for automatic indexing. *Commun. ACM*, ACM, New York, NY, USA, v. 18, n. 11, p. 613–620, nov 1975.
- SANTOS, J. C. A.; FAVERO, E. L. Practical Use of a Latent Semantic Analysis (LSA) Model for Automatic Evaluation of Written Answers. *Journal of the Brazilian Computer Society*, v. 21, n. 1, p. 21, 2015.
- SANTOS, J. C. A. et al. Avaliação do Uso de Métodos Baseados em LSA e WordNet para Correção de Questões Discursivas. In: *I Workshop de Desafios da Computação Aplicada á Educação (DesafIE!2012)*. Curitiba, PR, Brasil: Sociedade Brasileira de Computação, 2012. v. 1.

- SHERMIS, M.; BURSTEIN, J. *Handbook of Automated Essay Evaluation: Current Applications and New Directions*. 1. ed. New York, NY: Routledge, 2013.
- SIDDIQI, R.; HARRISON, C. J.; SIDDIQI, R. Improving Teaching and Learning through Automated Short-Answer Marking. *IEEE Transactions on Learning Technologies*, v. 3, n. 3, p. 237–249, July 2010.
- SIEMENS, G.; BAKER, R. S. J. d. Learning Analytics and Educational Data Mining: Towards Communication and Collaboration. In: *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*. New York, NY, USA: ACM, 2012. (LAK '12), p. 252–254.
- SPALENZA, M. A. et al. Construção de Mapas de Características em Classes de Respostas Discursivas. In: *XXI Congresso Internacional de Informática Educativa (TISE)*. Santiago, Chile: Centro de Computación y Comunicación para la Construcción del Conocimiento, 2016. v. 21.
- SPALENZA, M. A. et al. Uso de Mapa de Características na Avaliação de Textos Curtos nos Ambientes Virtuais de Aprendizagem Classes de Respostas Discursivas. In: *Simpósio Brasileiro de Informática na Educação (SBIE)*. Uberlândia, MG, Brasil: Sociedade Brasileira de Computação, 2016. v. 27.
- SUKKARIEH, J. Z.; STOYANCHEV, S. Automating Model Building in C-rater. In: *Proceedings of the 2009 Workshop on Applied Textual Inference*. Stroudsburg, PA: Association for Computational Linguistics, 2009. (TextInfer '09), p. 61–69.
- VILELA, R. F. et al. SCATeDi: Sistema Inteligente para Avaliação de Desempenho Escolar em Avaliações Discursivas. In: *Workshop de Informática na Escola (WIE)*. Rio de Janeiro, RJ, Brasil: Sociedade Brasileira de Computação, 2012. v. 18.
- WANG, H.-C.; CHANG, C.-Y.; LI, T.-Y. Assessing Creative Problem-Solving with Automated Text Grading. *Computers & Education*, v. 51, n. 4, p. 1450 – 1466, 2008.
- WEISS, S. M.; INDURKHIA, N.; ZHANG, T. *Fundamentals of Predictive Text Mining*. 1st. ed. London, United Kingdom: Springer-Verlag London, 2010.
- WRESCH, W. The Imminence Of Grading Essays by Computer: 25 Years Later. *Computers and Composition*, v. 10, n. 2, p. 45 – 58, 1993.
- ZBONTAR, J. *Short Answer Scoring by Stacking*. San Francisco, CA, USA: Kaggle, 2012. <<http://bitbucket.org/jzbontar/asap>>. ASAP '12 SAS Methodology Paper.
- ZEHNER, F.; SÄLZER, C.; GOLDHAMMER, F. Automatic coding of short text responses via clustering in educational assessment. *Educational and Psychological Measurement*, v. 76, n. 2, p. 280–303, 2016.



# Apêndice A

## Algoritmo Genético

### A.1 Calibração com a base de dados TEC-II\_5-8

Calibração do Algoritmo Genético - G.A. para Populações [*Pop*] de 10 à 100 (intervalos de 10) e Gerações [*Gen.*] de 100 à 1000 (intervalos de 100) avaliando o *Fitness* [*Fit.*] (Equação 4.1) e o número de características selecionadas [*N*].

<i>Pop.</i>	<i>Gen.</i>	Nota 100		Nota 93		Nota 90		Nota 0	
		<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N
10	100	0.52492	185	0.27817	46	0.34842	41	0.49975	194
20	100	0.61292	188	0.17347	33	0.62343	372	0.2606	65
30	100	0.44091	180	0.19947	34	0.23092	32	0.23418	67
40	100	0.35587	179	0.18093	28	0.19326	28	0.18636	59
50	100	0.43357	179	0.16162	23	0.17518	24	0.2338	73
60	100	0.50034	174	0.09936	20	0.14641	21	0.22027	69
70	100	0.62796	200	0.14811	21	0.16095	21	0.1245	58
80	100	0.36678	173	0.10458	17	0.09324	12	0.16368	64
90	100	0.37332	166	0.14312	21	0.10471	22	0.24675	83
100	100	0.88347	171	0.05328	11	0.09557	16	0.17062	60
10	100	0.48685	187	0.28548	42	0.32102	46	0.34146	99

Continua na próxima página

<i>Pop.</i>	<i>Gen.</i>	Nota 100		Nota 93		Nota 90		Nota 0	
		<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N
20	100	0.53292	182	0.25409	41	0.30847	38	0.25329	83
30	100	0.43659	183	0.17735	27	0.48491	79	0.19374	61
40	100	0.45454	178	0.20817	40	0.14002	24	0.29482	121
50	100	1.10625	173	0.05981	11	0.16242	27	0.19788	77
60	100	0.68752	207	0.16447	28	0.16087	29	0.26207	89
70	100	0.46317	173	0.07417	16	0.10382	17	0.17478	61
80	100	0.31605	169	0.1425	22	0.31815	227	0.67368	311
90	100	0.52471	176	0.04696	9	0.06335	13	0.16982	64
100	100	0.32186	172	0.07464	14	0.09765	12	0.12852	59
10	100	0.53897	189	0.27334	41	0.37267	43	0.2846	73
20	100	1.44204	177	0.23405	39	0.25502	32	0.22859	78
30	100	0.46814	176	0.09864	18	0.19228	27	0.22674	63
40	100	0.48752	178	0.15741	23	0.12146	21	0.23589	71
50	100	0.83561	172	0.13474	18	0.1478	18	0.17592	66
60	100	0.48231	180	0.0919	14	0.13991	26	0.19683	65
70	100	0.57842	171	0.10092	15	0.12214	20	0.19774	63
80	100	0.8987	178	0.0869	18	0.09081	17	0.17713	58
90	100	0.36581	174	0.07476	11	0.15038	25	0.20935	71
100	100	0.41101	169	0.0897	17	0.09944	17	0.1539	59
10	100	0.50998	184	0.28028	43	0.33246	48	0.1925	71
20	100	0.47042	185	0.28723	46	0.2632	38	0.30572	81
30	100	1.13399	181	0.22236	33	0.25415	33	0.22765	75
40	100	0.43028	187	0.15007	23	0.16261	24	0.22767	69
50	100	0.39176	171	0.10262	16	0.16754	23	0.19167	64
60	100	0.49735	213	0.12764	19	0.21089	29	0.16317	65
70	100	0.41832	171	0.09306	19	0.12101	16	0.18917	58
80	100	0.48153	187	0.12676	21	0.13118	19	0.12941	54
90	100	0.64046	173	0.07988	18	0.14226	18	0.10423	58
100	100	0.51424	222	0.13493	25	0.10948	18	0.16872	58
10	100	0.48775	520	0.2585	46	0.34246	46	0.24774	84
20	100	0.47057	182	0.16811	25	0.17788	28	0.25954	77
30	100	0.61844	181	0.16759	28	0.18998	26	0.18739	67

Continua na próxima página

<i>Pop.</i>	<i>Gen.</i>	Nota 100		Nota 93		Nota 90		Nota 0	
		<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N
40	100	0.42377	176	0.16113	26	0.20547	28	0.16262	72
50	100	0.30993	173	0.11369	21	0.17123	23	0.21473	63
60	100	0.74537	169	0.10701	19	0.15068	28	0.17464	58
70	100	0.53292	170	0.06881	14	0.11429	19	0.21257	68
80	100	0.29039	174	0.1085	18	0.12361	16	0.13765	57
90	100	0.34716	218	0.04528	10	0.1117	16	0.09552	70
100	100	0.71132	170	0.12488	20	0.11442	16	0.13568	57
10	200	0.60305	218	0.41943	68	0.74945	220	0.33915	120
20	200	0.58371	198	0.28979	62	0.46733	62	0.51448	160
30	200	0.55305	203	0.336	66	0.28211	43	0.4344	153
40	200	0.56548	209	0.23469	46	0.21577	43	0.26739	99
50	200	0.44387	203	0.22165	39	0.14503	32	0.24873	91
60	200	0.6989	197	0.20033	42	0.26634	34	0.203	77
70	200	0.31079	197	0.07504	20	0.13912	31	0.29788	108
80	200	0.58018	191	0.03804	13	0.04557	16	0.11876	81
90	200	0.34646	200	0.06975	16	0.03023	12	0.1382	83
100	200	0.54113	199	0.05445	11	0.07972	21	0.13962	61
10	200	0.63062	213	0.38251	70	0.49901	79	0.31672	105
20	200	0.61712	226	0.43352	71	0.39829	65	0.2648	114
30	200	0.51228	201	0.14353	35	0.32894	50	0.23532	90
40	200	0.50578	205	0.31764	54	0.22381	46	0.1281	86
50	200	0.51791	205	0.20446	43	0.19588	46	0.19223	89
60	200	0.42296	194	0.11546	27	0.17359	35	0.13258	86
70	200	1.1154	197	0.09362	21	0.03877	25	0.26674	111
80	200	0.47046	202	0.06159	13	0.11971	29	0.15033	70
90	200	0.53725	232	0.07588	19	0.11092	27	0.23208	87
100	200	0.35993	185	0.07404	19	0.11243	25	0.21747	101
10	200	0.76061	216	0.45905	92	0.53691	72	0.32271	101
20	200	1.38226	208	0.35372	62	0.3817	52	0.26199	92
30	200	0.47579	207	0.34918	61	0.27165	40	0.45338	146
40	200	0.51116	210	0.12618	24	0.27506	48	0.27013	97
50	200	0.42804	206	0.12975	30	0.21598	33	0.28602	93

Continua na próxima página

<i>Pop.</i>	<i>Gen.</i>	Nota 100		Nota 93		Nota 90		Nota 0	
		<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N
60	200	0.51907	204	0.1166	28	0.07611	31	0.17575	80
70	200	0.46477	204	0.17089	32	0.0915	18	0.19022	75
80	200	0.4837	192	0.08917	19	0.15234	39	0.1056	76
90	200	0.78124	200	0.11756	22	0.10893	26	0.15167	75
100	200	0.65451	187	0.11351	23	0.03284	21	0.18125	75
10	200	0.85286	237	0.56498	88	0.47331	71	0.35327	108
20	200	0.6299	216	0.28465	63	0.38752	61	0.29476	105
30	200	0.54373	256	0.24498	50	0.2314	41	0.30079	105
40	200	0.40794	199	0.17196	41	0.2542	57	0.20987	82
50	200	0.43603	204	0.26307	51	0.16001	40	0.11615	88
60	200	0.51045	197	0.08877	20	0.15608	29	0.20661	116
70	200	0.44356	195	0.15462	35	0.18076	30	0.15888	74
80	200	0.31199	191	0.13268	21	0.05106	16	0.18426	68
90	200	0.43501	198	0.11964	23	0.03596	14	0.19622	70
100	200	0.38935	199	0.08956	20	0.05138	20	0.16561	84
10	200	0.63865	211	0.35329	57	0.57404	94	0.43214	152
20	200	0.64622	204	0.3753	64	0.25187	53	0.21319	97
30	200	0.68798	207	0.28876	46	0.2838	54	0.32502	135
40	200	0.48246	204	0.2406	55	0.33378	49	0.20627	100
50	200	0.37204	212	0.17522	40	0.26364	47	0.41219	168
60	200	0.51101	202	0.09654	21	0.15865	36	0.22721	74
70	200	0.67684	200	0.07626	18	0.08226	18	0.23839	90
80	200	0.4104	195	0.05604	16	0.15693	30	0.14777	82
90	200	0.37051	195	0.04605	17	0.03672	23	0.25206	86
100	200	1.1079	196	0.07981	16	0.04277	19	0.20597	82
10	300	0.77531	264	0.62792	116	0.58446	108	0.35391	119
20	300	0.75313	245	0.50622	100	0.55642	97	0.39552	150
30	300	0.96381	220	0.39527	84	0.30772	74	0.28997	121
40	300	0.62834	233	0.21997	66	0.20893	63	0.46623	207
50	300	0.45924	220	0.20098	46	0.1399	54	0.27147	119
60	300	0.54015	243	0.22032	56	0.16105	47	0.26373	102
70	300	0.49287	221	0.15105	33	0.09928	43	0.30105	145

Continua na próxima página

<i>Pop.</i>	<i>Gen.</i>	Nota 100		Nota 93		Nota 90		Nota 0	
		<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N
80	300	0.48	213	0.12981	33	0.17023	42	0.29491	104
90	300	0.33966	210	0.0956	29	0.08743	38	0.14953	80
100	300	0.32037	221	0.01968	6	0.25971	205	0.22784	95
10	300	0.79754	257	0.6356	107	0.49651	116	0.38639	139
20	300	0.60495	287	0.38932	80	0.43424	76	0.33512	131
30	300	0.65998	243	0.37762	76	0.37094	84	0.2293	116
40	300	0.67459	238	0.28727	58	0.35899	81	0.35294	134
50	300	0.84233	234	0.27359	65	0.14191	55	0.15043	96
60	300	0.51896	231	0.17157	44	0.23077	52	0.27175	102
70	300	0.36642	228	0.13799	45	0.04394	29	0.12273	87
80	300	0.40887	223	0.1021	24	0.10009	42	0.15749	100
90	300	0.34144	227	0.10632	31	0.0818	32	0.11792	83
100	300	0.61296	225	0.15257	31	0.05853	26	0.10421	81
10	300	1.07729	267	0.44314	101	0.59685	123	0.34299	125
20	300	0.72018	251	0.56093	99	0.23504	84	0.35131	141
30	300	1.07385	232	0.31446	70	0.26395	63	0.80889	407
40	300	0.4935	237	0.31532	69	0.34234	73	0.298	128
50	300	0.64527	228	0.24915	55	0.07133	46	0.17781	123
60	300	0.41692	207	0.25324	48	0.09503	44	0.20264	96
70	300	0.40411	242	0.15056	34	0.05465	38	0.22052	101
80	300	0.87715	239	0.14749	30	0.08231	34	0.18257	101
90	300	0.65513	216	0.08297	21	0.02102	14	0.23634	103
100	300	0.31611	218	0.13309	26	0.08049	27	0.18444	87
10	300	0.8129	256	0.67672	119	0.25242	98	0.46977	158
20	300	0.51968	243	0.4629	108	0.4826	87	0.32251	139
30	300	0.77896	246	0.28375	72	0.46504	74	0.28307	122
40	300	0.64073	234	0.29689	60	0.18624	61	0.29153	112
50	300	0.75355	218	0.24068	51	0.18514	48	0.29778	129
60	300	0.51574	228	0.15362	41	0.18047	55	0.14357	91
70	300	0.47171	221	0.14867	40	0.10544	42	0.2009	116
80	300	0.8781	219	0.08834	23	0.07028	46	0.22537	142
90	300	0.43932	224	0.04868	20	0.0569	30	0.1568	90

Continua na próxima página

<i>Pop.</i>	<i>Gen.</i>	Nota 100		Nota 93		Nota 90		Nota 0	
		<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N
100	300	0.60646	252	0.11333	33	0.08974	38	0.44941	248
10	300	0.72048	264	0.46948	91	0.75172	150	0.45433	150
20	300	0.75678	219	0.37232	74	0.33585	79	0.34106	151
30	300	1.02108	244	0.23013	68	0.28033	73	0.28778	116
40	300	0.61316	219	0.2148	51	0.15909	54	0.19051	100
50	300	0.51177	283	0.20081	55	0.30202	57	0.26918	159
60	300	0.85535	228	0.13726	43	0.12564	58	0.17511	97
70	300	0.44075	238	0.13456	24	0.10171	35	0.14135	103
80	300	1.11557	216	0.09999	25	0.14035	45	0.08722	69
90	300	0.30818	314	0.09332	26	0.07877	31	0.12722	90
100	300	0.26749	227	0.03533	10	0.02586	17	0.13843	77
10	400	0.93995	323	0.66921	147	0.41977	142	0.41735	162
20	400	0.61131	260	0.73332	131	0.43516	112	0.47689	164
30	400	0.41699	261	0.40892	104	0.23612	109	0.19109	134
40	400	0.5953	252	0.27728	76	0.26994	84	0.20344	132
50	400	0.69097	263	0.25786	74	0.07966	54	0.23213	138
60	400	0.62498	245	0.11029	44	0.10442	52	0.22577	127
70	400	0.33393	249	0.19173	57	0.16754	57	0.17409	125
80	400	0.98039	246	0.07965	29	0.02666	17	0.32109	194
90	400	0.74223	245	0.1637	51	0.0477	31	0.20958	114
100	400	0.38693	235	0.08378	25	0.17563	43	0.15585	102
10	400	0.92696	272	0.7739	131	0.59994	131	0.44225	154
20	400	0.88365	266	0.44144	119	0.62774	119	0.32966	152
30	400	0.57579	269	0.53963	124	0.23165	96	0.32191	120
40	400	0.97902	258	0.2927	89	0.32336	94	0.26116	127
50	400	0.56255	278	0.30006	94	0.13737	88	0.33257	145
60	400	0.48231	260	0.16948	54	0.17532	84	0.231	134
70	400	0.34876	244	0.19176	54	0.06146	40	0.29472	120
80	400	0.43665	241	0.15395	45	0.13772	52	0.27214	135
90	400	0.37231	248	0.05084	16	0.07334	48	0.17517	105
100	400	0.90963	275	0.05728	22	0.02468	17	0.10372	77
10	400	1.02258	321	0.65981	132	0.71639	160	0.40624	165

Continua na próxima página

<i>Pop.</i>	<i>Gen.</i>	Nota 100		Nota 93		Nota 90		Nota 0	
		<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N
20	400	0.8193	292	0.38833	110	0.5461	115	0.36456	161
30	400	0.4364	279	0.42338	129	0.34256	111	0.55529	279
40	400	0.72119	242	0.36701	105	0.21618	93	0.69142	339
50	400	1.22464	238	0.21668	70	0.18491	67	0.21491	136
60	400	0.40718	257	0.18153	68	0.19959	88	0.26881	121
70	400	0.43995	239	0.17216	37	0.04325	28	0.20027	117
80	400	0.62653	240	0.01886	8	0.02629	18	0.16131	114
90	400	0.37129	237	0.10114	30	0.12078	57	0.10413	78
100	400	0.49865	249	0.0302	13	0.03884	26	0.14811	104
10	400	0.83177	302	0.71703	150	0.6134	141	0.45322	176
20	400	0.59066	262	0.55759	121	0.3661	121	0.3904	163
30	400	0.88444	258	0.43504	85	0.4482	122	0.37365	153
40	400	0.66253	254	0.4081	104	0.15749	106	0.44145	145
50	400	0.60077	455	0.32437	92	0.15261	84	0.19471	133
60	400	0.57626	280	0.24617	65	0.35853	283	0.22679	137
70	400	0.84625	249	0.21658	54	0.07848	52	0.18599	113
80	400	0.7276	559	0.11612	34	0.06501	45	0.17792	127
90	400	0.88292	236	0.08438	27	0.04631	31	0.2108	95
100	400	0.52119	247	0.13516	48	0.06173	42	0.13804	101
10	400	1.09529	250	0.67294	141	0.83947	168	0.40726	179
20	400	0.72187	265	0.66594	143	0.57387	126	0.41885	157
30	400	0.57498	260	0.40478	103	0.30146	108	0.20894	146
40	400	0.61056	264	0.26937	86	0.20547	74	0.37799	185
50	400	0.6779	266	0.20047	61	0.06593	43	0.31257	147
60	400	0.58429	257	0.21926	66	0.10323	68	0.19212	125
70	400	0.53534	248	0.14433	46	0.08707	58	0.64309	333
80	400	0.43704	226	0.15423	47	0.07433	52	0.19564	105
90	400	0.6986	255	0.1709	46	0.06304	34	0.11513	86
100	400	0.42198	232	0.00084	1	0.09319	64	0.117	85
10	500	1.07526	292	0.8856	177	0.55208	176	0.52035	191
20	500	0.9495	305	0.52764	153	0.54885	132	0.40734	176
30	500	0.97873	306	0.53027	144	0.74185	450	0.38981	172

Continua na próxima página

<i>Pop.</i>	<i>Gen.</i>	Nota 100		Nota 93		Nota 90		Nota 0	
		<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N
40	500	0.57345	279	0.42029	111	0.26905	135	0.28682	154
50	500	0.58985	299	0.33596	94	0.12944	86	0.29638	158
60	500	0.38869	274	0.20441	58	0.25525	100	0.19418	126
70	500	0.59234	268	0.06504	42	0.08242	55	0.29245	139
80	500	0.51691	261	0.02777	29	0.06342	44	0.17192	128
90	500	0.482	247	0.06824	30	0.16114	50	0.13822	105
100	500	0.44218	259	0.04769	39	0.09091	61	0.14085	111
10	500	0.94164	322	0.6867	172	0.71362	186	0.72126	289
20	500	1.14893	313	0.47588	134	0.61124	147	0.4609	177
30	500	0.83353	282	0.45984	127	0.19008	135	0.57389	239
40	500	0.55677	279	0.24358	91	0.32516	130	0.28528	176
50	500	0.49055	271	0.34959	89	0.19768	89	0.38176	185
60	500	0.69913	284	0.18357	84	0.24606	89	0.27038	140
70	500	0.45113	288	0.2082	70	0.11774	79	0.17868	127
80	500	0.32171	272	0.15279	67	0.10152	59	0.22333	134
90	500	0.51997	245	0.05759	28	0.03809	26	0.32066	199
100	500	0.64615	268	0.08348	40	0.0038	3	0.16144	119
10	500	0.84447	295	0.81228	171	0.80316	175	0.44305	186
20	500	0.86915	311	0.46346	139	0.52266	168	0.45884	207
30	500	0.81022	312	0.50066	135	0.2871	128	0.4591	242
40	500	0.8108	297	0.3735	122	0.29099	114	0.32464	191
50	500	0.51037	294	0.29225	84	0.22724	109	0.29318	144
60	500	0.41006	293	0.40804	114	0.09206	61	0.24537	140
70	500	0.92034	298	0.16376	63	0.10944	72	0.20498	135
80	500	0.43745	280	0.14634	66	0.04759	32	0.19091	126
90	500	0.99202	266	0.0464	37	0.11311	78	0.15525	110
100	500	0.52986	373	0.08784	40	0.13064	60	0.1594	114
10	500	0.79555	318	0.80495	176	0.73524	183	0.6313	269
20	500	0.70039	308	0.53866	148	0.5046	155	0.2725	171
30	500	0.70563	310	0.44799	125	0.24116	115	0.3979	182
40	500	0.64467	261	0.30064	96	0.36289	104	0.28942	161
50	500	0.61922	283	0.25471	100	0.12919	87	0.26162	121

Continua na próxima página



<i>Pop.</i>	<i>Gen.</i>	Nota 100		Nota 93		Nota 90		Nota 0	
		<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N
60	500	0.39869	276	0.14324	62	0.21201	91	0.424	186
70	500	0.41492	260	0.26526	78	0.08299	57	0.23803	137
80	500	0.37254	292	0.18071	81	0.07801	54	0.26545	152
90	500	1.07627	367	0.14779	43	0.08046	53	0.19207	109
100	500	0.4625	278	0.00553	6	0.0587	39	0.19241	138
10	500	0.99125	303	0.70994	166	0.41489	156	0.58379	212
20	500	0.70013	301	0.52022	140	0.67609	151	0.56393	204
30	500	0.78183	273	0.4777	123	0.248	139	0.39614	187
40	500	0.59083	282	0.26866	88	0.22004	106	0.32988	175
50	500	0.79448	289	0.40112	101	0.24958	116	0.26423	156
60	500	0.61109	257	0.23572	59	0.10583	72	0.19506	137
70	500	0.90385	283	0.16745	54	0.10975	71	0.20916	147
80	500	0.77314	273	0.15904	59	0.09442	62	0.13842	101
90	500	0.31335	264	0.10354	28	0.08118	52	0.16489	122
100	500	0.5405	261	0.02935	11	0.02154	17	0.21206	129
10	600	0.96977	338	1.0131	213	0.78512	191	0.57764	237
20	600	1.03881	339	0.5483	157	0.39445	175	0.4909	206
30	600	0.81959	302	0.46559	149	0.36741	166	0.49416	196
40	600	0.78856	303	0.3813	134	0.2245	153	0.50104	215
50	600	0.61876	302	0.37177	116	0.17513	96	0.3379	180
60	600	0.52222	289	0.26304	99	0.16212	114	0.3038	164
70	600	0.62721	276	0.16748	68	0.09401	67	0.47237	216
80	600	0.38409	285	0.17057	65	0.11445	76	0.18165	133
90	600	0.44878	278	0.03635	26	0.06607	43	0.14264	108
100	600	0.77982	290	0.02655	18	0.00084	1	0.16305	115
10	600	0.9001	298	0.96195	207	0.90322	217	0.52337	237
20	600	0.98354	314	0.47189	169	0.67443	178	0.3865	199
30	600	0.66705	294	0.37494	160	0.41068	134	0.42904	205
40	600	0.85192	294	0.23146	107	0.23815	143	0.36347	191
50	600	0.69031	289	0.27558	108	0.23903	93	0.31198	177
60	600	0.68571	282	0.13977	50	0.20166	111	0.28585	159
70	600	0.49932	290	0.34877	112	0.11153	75	0.17622	129

Continua na próxima página

<i>Pop.</i>	<i>Gen.</i>	Nota 100		Nota 93		Nota 90		Nota 0	
		<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N
80	600	0.46588	291	0.28515	78	0.18091	83	0.19436	133
90	600	0.39732	293	0.00645	7	0.08703	63	0.16438	123
100	600	0.41659	283	0.00459	4	0.04706	31	0.26576	142
10	600	0.96089	323	0.89225	197	0.95879	207	0.6034	227
20	600	0.66733	321	0.68389	173	0.55613	186	0.46711	219
30	600	1.04203	329	0.31691	128	0.40776	146	0.37342	202
40	600	0.7064	288	0.42229	129	0.53986	394	0.34495	176
50	600	0.84182	305	0.48069	132	0.18756	129	0.72691	308
60	600	0.52265	445	0.24013	82	0.10905	73	0.2026	148
70	600	0.60793	285	0.12929	84	0.14525	101	0.18998	135
80	600	0.40633	285	0.13172	50	0.09566	66	0.23829	146
90	600	0.63253	269	0.06333	53	0.07296	52	0.17936	129
100	600	0.55987	281	0.00092	1	0.03835	24	0.15352	100
10	600	1.4716	329	0.86249	189	0.95449	214	0.6613	262
20	600	0.83834	320	0.80951	185	0.84176	188	0.51436	230
30	600	0.90627	327	0.55811	152	0.36917	184	0.77196	354
40	600	0.88853	300	0.43718	153	0.32498	144	0.3121	198
50	600	0.8165	317	0.25924	116	0.15981	112	0.38488	201
60	600	1.00108	292	0.19381	89	0.22142	129	0.17676	128
70	600	0.48564	281	0.28378	85	0.18349	93	0.34502	160
80	600	0.44805	297	0.1683	64	0.08132	52	0.18408	137
90	600	0.55469	300	0.04154	15	0.06841	54	0.18345	128
100	600	0.37042	279	0.01159	8	0.09897	66	0.18775	136
10	600	1.01697	318	0.70995	182	1.01881	201	0.99749	326
20	600	1.5021	312	0.80018	192	1.02294	197	0.43219	212
30	600	0.4824	334	0.39262	124	0.36424	170	0.45598	224
40	600	0.70621	309	0.46257	138	0.2977	210	0.45118	211
50	600	0.89317	309	0.32966	98	0.3362	113	0.26618	188
60	600	0.41141	328	0.22516	74	0.23691	135	0.21122	140
70	600	0.50494	316	0.40374	120	0.00507	4	0.2874	152
80	600	0.55152	287	0.00092	1	0.11801	79	0.25503	157
90	600	0.47448	291	0.14098	61	0.02864	20	0.13088	109

Continua na próxima página

<i>Pop.</i>	<i>Gen.</i>	Nota 100		Nota 93		Nota 90		Nota 0	
		<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N
100	600	0.47188	282	0.01934	9	0.06837	48	0.17907	131
10	700	1.25109	330	0.81736	213	1.0474	264	0.55757	264
20	700	1.17971	325	0.85233	218	0.32552	196	0.60208	260
30	700	1.08773	341	0.36257	157	0.54335	185	0.69377	290
40	700	0.70944	338	0.40902	137	0.36402	173	0.34138	186
50	700	0.48942	338	0.37315	132	0.25023	142	0.34704	198
60	700	0.56797	289	0.25314	112	0.15468	108	0.28505	177
70	700	0.43425	286	0.07995	68	0.18351	128	0.24863	183
80	700	0.40351	309	0.14183	101	0.14485	101	0.25506	178
90	700	0.70669	285	0.16131	61	0.10895	86	0.26694	162
100	700	0.47458	311	0.02541	20	0.10726	74	0.22881	147
10	700	1.01218	352	0.82252	202	0.97918	244	1.34809	509
20	700	1.06719	341	0.65427	203	0.76334	204	0.48362	218
30	700	0.78888	344	0.52331	184	0.46331	184	0.45478	220
40	700	0.96347	351	0.32586	143	0.3456	176	0.30632	219
50	700	0.62761	320	0.33485	119	0.21782	153	0.24483	183
60	700	0.57079	325	0.27314	89	0.205	147	0.2608	188
70	700	0.61225	316	0.28444	95	0.17243	102	0.19853	143
80	700	0.44144	329	0.15964	73	0.06261	43	0.30592	153
90	700	0.54583	293	0.08881	47	0.00084	1	0.18231	126
100	700	0.37621	272	0.00084	1	0.04767	33	0.15861	119
10	700	1.04786	344	1.00884	262	0.86485	237	0.56051	228
20	700	0.78328	350	0.66005	200	0.50324	227	0.56176	224
30	700	0.59225	322	0.56538	187	0.21381	148	0.46263	232
40	700	0.64596	435	0.49189	163	0.31089	173	0.44963	217
50	700	0.51191	308	0.30285	123	0.21818	156	0.25693	186
60	700	0.57895	310	0.31308	125	0.2429	139	0.30069	191
70	700	1.08735	300	0.12708	106	0.13602	93	0.30677	178
80	700	0.51274	304	0.04477	38	0.26351	208	0.21788	155
90	700	0.40147	306	0.12599	62	0.05954	47	0.22479	156
100	700	0.48493	306	0.03792	18	0.3339	245	0.21623	155
10	700	0.89628	327	0.76329	208	1.05558	230	0.652	250

Continua na próxima página

<i>Pop.</i>	<i>Gen.</i>	Nota 100		Nota 93		Nota 90		Nota 0	
		<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N
20	700	1.00314	343	0.67354	197	0.5809	183	0.39545	215
30	700	0.8507	325	0.55926	171	0.41442	208	0.42631	224
40	700	0.83571	340	0.37726	154	0.20923	145	0.38345	199
50	700	1.0378	328	0.3434	147	0.46788	272	0.26247	188
60	700	0.6439	302	0.08573	83	0.16471	115	0.26984	172
70	700	0.67971	349	0.07517	53	0.26955	126	0.17422	133
80	700	0.79681	299	0.196	96	0.11604	84	0.28617	165
90	700	0.45644	299	0.05935	45	0.07728	61	0.16801	122
100	700	0.36027	316	0.00507	2	0.0187	13	0.17363	123
10	700	0.86294	367	0.90923	233	0.77877	225	0.62374	234
20	700	0.90966	336	0.59945	201	0.52153	205	0.52576	242
30	700	0.83769	372	0.61025	179	0.29976	180	0.47765	217
40	700	0.67288	347	0.31399	154	0.30515	144	0.31802	209
50	700	0.86304	322	0.48313	163	0.28935	176	0.42694	223
60	700	1.04314	327	0.31851	98	0.21261	149	0.37957	181
70	700	0.56258	480	0.14185	102	0.06423	43	0.24692	183
80	700	0.488	330	0.13561	59	0.1211	88	0.20358	148
90	700	0.64014	292	0.13915	98	0.09135	62	0.20428	150
100	700	0.39437	287	0.00276	3	0.06068	42	0.21192	156
10	800	1.19726	381	1.15637	258	0.85359	271	0.63762	301
20	800	0.84508	374	0.98062	243	0.75452	230	0.64119	270
30	800	0.88394	370	0.76527	233	0.4681	197	0.37842	224
40	800	0.66232	361	0.56902	184	0.22989	162	0.48253	233
50	800	0.58347	346	0.50665	174	0.17331	122	0.28419	187
60	800	0.64482	327	0.21912	154	0.22041	152	0.38409	209
70	800	1.03303	329	0.24222	118	0.14626	100	0.32428	193
80	800	0.56262	321	0.09581	65	0.11909	94	0.4264	253
90	800	0.49518	323	0.00921	10	0.09033	65	0.23883	168
100	800	0.40858	366	0.03501	38	0.06081	48	0.15693	113
10	800	1.24216	375	0.99549	244	1.15296	263	0.6005	293
20	800	0.88482	360	0.89713	222	0.61693	220	0.64382	261
30	800	0.63412	365	0.82018	211	0.27019	190	0.41003	220

Continua na próxima página

<i>Pop.</i>	<i>Gen.</i>	Nota 100		Nota 93		Nota 90		Nota 0	
		<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N
40	800	0.7312	354	0.56534	202	0.38978	195	0.36363	189
50	800	0.51504	373	0.37345	157	0.17103	135	0.38613	228
60	800	0.83248	349	0.31504	154	0.22893	157	0.32074	196
70	800	0.55079	322	0.07076	38	0.20912	122	0.26641	166
80	800	0.57493	309	0.08617	90	0.16903	115	0.18501	138
90	800	0.66334	325	0.03923	41	0.1452	100	0.27883	182
100	800	0.34099	316	0.09104	76	0.02524	16	0.23271	157
10	800	0.89812	366	0.90108	231	0.94039	259	0.72962	291
20	800	0.99207	372	0.56764	233	0.86027	248	0.4505	224
30	800	1.09486	365	0.8744	227	0.28237	196	0.61599	291
40	800	0.73273	323	0.42141	172	0.23232	163	0.47364	244
50	800	0.80418	348	0.37608	170	0.20837	144	0.37689	190
60	800	0.85743	337	0.30679	135	0.22651	156	0.28106	179
70	800	0.59026	315	0.14036	95	0.19503	139	0.3435	212
80	800	0.48436	308	0.01032	9	0.12778	88	0.30946	174
90	800	0.74252	338	0.08304	66	0.07221	57	0.23742	173
100	800	0.70359	330	0.06369	63	0.08995	71	0.18779	138
10	800	0.809	378	1.13265	261	0.54713	243	0.6552	295
20	800	0.74664	366	0.70141	253	0.55798	229	0.52242	246
30	800	0.68516	353	0.48192	194	0.58266	226	0.42485	249
40	800	0.62265	586	0.42203	185	0.22537	156	0.42751	238
50	800	0.59964	349	0.2177	156	0.25629	151	0.34182	214
60	800	0.78102	317	0.18804	138	0.15629	111	0.35848	228
70	800	0.82939	323	0.19787	106	0.08602	62	0.26024	160
80	800	0.87548	318	0.19335	103	0.06048	41	0.23837	157
90	800	0.45183	542	0.11781	83	0.09375	74	0.2214	159
100	800	0.66962	303	0.00737	8	0.04654	31	0.20197	148
10	800	1.20994	372	1.26348	268	0.90551	256	0.78942	312
20	800	0.88931	376	0.94674	235	0.69836	453	0.54703	247
30	800	0.81129	350	0.65215	208	0.30512	211	0.44558	243
40	800	0.70065	342	0.43346	202	0.29938	173	0.39754	246
50	800	0.6409	356	0.43129	168	0.204	142	0.35898	214

Continua na próxima página

<i>Pop.</i>	<i>Gen.</i>	Nota 100		Nota 93		Nota 90		Nota 0	
		<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N
60	800	0.62877	337	0.22251	110	0.20792	143	0.30318	196
70	800	0.63933	319	0.38814	130	0.20206	147	0.3705	217
80	800	0.81713	317	0.17157	131	0.11616	80	0.30133	177
90	800	0.50731	317	0.08724	81	0.15461	105	0.25876	135
100	800	0.46541	330	0.04388	44	0.36486	288	0.20629	155
10	900	1.1457	447	1.04968	272	0.54524	259	0.8496	330
20	900	0.85704	406	0.78976	259	0.68085	241	0.5733	291
30	900	0.78236	399	0.43253	214	0.65999	250	0.60307	281
40	900	0.70938	374	0.57231	222	0.40425	208	0.54665	261
50	900	0.96194	360	0.3301	164	0.25223	175	0.2971	225
60	900	0.52048	365	0.40496	189	0.15456	122	0.43456	242
70	900	0.65612	364	0.2448	166	0.22391	154	0.27621	191
80	900	0.77595	367	0.20275	128	0.16161	114	0.70948	419
90	900	0.64853	341	0.09245	92	0.1428	99	0.28265	188
100	900	0.38143	319	0.00084	1	0.00084	1	0.21041	151
10	900	1.63553	380	1.11535	279	0.52564	266	0.64469	286
20	900	0.86194	392	0.74839	247	0.52262	222	0.65924	291
30	900	0.73033	375	0.69184	235	0.43808	221	0.50231	256
40	900	0.62882	355	0.33101	146	0.64111	259	0.36397	259
50	900	0.71554	343	0.33207	174	0.2138	157	0.40695	217
60	900	0.64269	359	0.20722	140	0.17658	128	0.32586	229
70	900	0.51417	347	0	0	0.15583	123	0.41589	242
80	900	0.59061	360	0.11923	104	0.11275	89	0.24087	171
90	900	0.52733	331	0.11258	87	0.02154	17	0.28371	190
100	900	0.44416	303	0.00507	6	0.13085	94	0.22529	168
10	900	1.25573	408	0.95249	270	1.02801	275	0.65986	301
20	900	1.02535	383	0.96056	261	0.73457	234	0.57366	289
30	900	0.97445	364	0.59209	229	0.3025	211	0.43945	247
40	900	0.7214	367	0.54565	197	0.40144	219	0.51408	264
50	900	0.7375	406	0.3829	187	0.31945	155	0.37921	232
60	900	0.55091	370	0.38676	161	0.25586	156	0.26965	192
70	900	0.5231	328	0.08858	90	0.19948	137	0.25784	187

Continua na próxima página

<i>Pop.</i>	<i>Gen.</i>	Nota 100		Nota 93		Nota 90		Nota 0	
		<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N
80	900	0.48662	347	0.1751	94	0.17444	128	0.30826	190
90	900	0.52048	333	0.09245	89	0.00084	1	0.40268	210
100	900	0.48358	323	0.01605	19	0.12289	97	0.29625	179
10	900	1.07754	429	1.30866	283	0.75165	283	0.69114	283
20	900	0.91377	374	1.03407	253	0.34816	241	0.50869	273
30	900	0.70945	388	0.73367	241	0.33565	237	0.48686	262
40	900	0.88197	397	0.37412	190	0.35268	195	0.32132	227
50	900	0.64371	379	0.25255	174	0.35342	176	0.4096	246
60	900	0.58715	343	0.21032	144	0.2433	168	0.37646	222
70	900	0.52566	366	0.20155	106	0.19461	139	0.34047	190
80	900	0.81888	354	0.1427	136	0.13936	110	0.24526	173
90	900	1.08822	333	0.08022	36	0.15311	108	0.29449	207
100	900	0.42053	331	0.00084	1	0	0	0.35879	240
10	900	1.00019	402	1.00954	274	1.08872	284	0.68781	262
20	900	1.03694	362	0.72663	242	0.58354	261	0.54156	288
30	900	0.64919	380	0.60528	225	0.32565	223	0.50557	268
40	900	0.9016	365	0.32307	185	0.59291	468	0.5349	262
50	900	0.97449	374	0.48993	209	0.22091	155	0.3312	241
60	900	0.67228	368	0.26792	137	0.45331	331	0.34067	205
70	900	0.50376	352	0.12536	92	0.14189	112	0.28553	207
80	900	0.61256	569	0.2247	142	0.12796	101	0.22724	162
90	900	0.50898	338	0.06497	44	0.39527	312	0.37207	266
100	900	0.43141	381	0.14763	104	0.09375	74	0.25139	182
10	1000	1.31174	412	1.0277	285	1.06443	294	0.80594	339
20	1000	0.74818	386	0.60328	266	0.47961	283	0.68393	296
30	1000	0.95733	387	0.59182	261	0.41	260	0.53692	267
40	1000	0.85598	382	0.70465	246	0.30776	214	0.45261	252
50	1000	0.56072	383	0.23385	177	0.35142	252	0.47857	282
60	1000	0.76124	340	0.27226	136	0.24954	177	0.3292	211
70	1000	0.52881	355	0.40037	163	0.21733	149	0.35619	203
80	1000	0.73076	421	0.15573	105	0.15709	124	0.31845	224
90	1000	0.45338	362	0.09096	95	0.09541	68	0.30147	204

Continua na próxima página

<i>Pop.</i>	<i>Gen.</i>	Nota 100		Nota 93		Nota 90		Nota 0	
		<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N
100	1000	0.73287	336	0.10949	93	0.07205	48	0.18902	139
10	1000	1.45387	419	1.27587	302	1.02379	284	0.83619	348
20	1000	0.89223	400	0.65652	273	0.53578	273	0.71601	314
30	1000	0.74627	439	0.69978	242	0.33545	233	0.61785	284
40	1000	1.24189	354	0.53758	232	0.32469	222	0.41106	256
50	1000	1.06661	376	0.54236	236	0.35964	213	0.32621	239
60	1000	0.74306	361	0.22701	148	0.24629	181	0.38433	225
70	1000	0.57715	365	0.20792	141	0.19804	134	0.37515	234
80	1000	0.93477	347	0.22009	105	0.1723	136	0.25129	184
90	1000	0.44305	333	0.09058	72	0.11402	90	0.2415	172
100	1000	0.99743	350	0.09845	75	0.00084	1	0.19271	145
10	1000	1.19969	430	1.09253	320	1.35961	304	0.83178	304
20	1000	1.11381	390	0.7972	275	0.777	506	0.64425	298
30	1000	0.70241	407	0.66578	258	0.36061	256	0.7063	302
40	1000	0.94675	383	0.71924	253	0.4455	259	0.53721	271
50	1000	0.77543	367	0.49111	190	0.26172	180	0.42115	247
60	1000	0.82707	386	0.26008	171	0.14569	115	0.3917	233
70	1000	0.47023	345	0.33656	162	0.20883	143	0.31886	225
80	1000	0.77715	339	0.09206	97	0.1815	125	0.30115	216
90	1000	0.54569	367	0.05925	62	0.12669	100	0.31305	186
100	1000	0.40041	349	0.0465	46	0.07221	57	0.21627	156
10	1000	1.00488	424	0.86763	286	0.86682	296	0.76942	339
20	1000	1.07129	401	0.75085	266	0.36086	257	0.6072	299
30	1000	0.94161	409	0.46089	253	0.36487	255	0.50835	279
40	1000	0.79475	378	0.41642	210	0.40992	253	0.43451	260
50	1000	0.70511	352	0.60945	188	0.29067	206	0.41533	238
60	1000	0.57926	333	0.33654	188	0.25223	175	0.40876	226
70	1000	0.98144	411	0.35412	147	0.09815	66	0.24967	183
80	1000	0.77105	344	0.22139	155	0.21119	151	0.24016	181
90	1000	0.81574	335	0.13719	129	0.1685	133	0.24431	178
100	1000	0.62162	336	0.00084	1	0.00084	1	0.27271	187
10	1000	1.11155	406	1.02685	297	1.29033	311	0.88185	350

Continua na próxima página



<i>Pop.</i>	<i>Gen.</i>	Nota 100		Nota 93		Nota 90		Nota 0	
		<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N	<i>Fit.</i>	N
20	1000	1.11345	430	0.7391	271	0.77265	292	0.54368	267
30	1000	0.83384	377	0.61273	240	0.48214	279	0.53687	296
40	1000	0.78039	580	0.36452	215	0.36092	254	0.52047	285
50	1000	0.62016	378	0.42897	172	0.31555	218	0.53515	251
60	1000	0.65673	401	0.23724	162	0.22705	162	0.4198	233
70	1000	0.71973	365	0.19457	152	0.17361	126	0.30997	211
80	1000	0.48315	367	0.2185	142	0.18521	132	0.35565	249
90	1000	0.47243	378	0.01605	19	0.16442	120	0.22895	177
100	1000	0.42622	343	0.05252	52	0.10389	82	0.20086	147

Última página

Tabela A.2: Calibração do Algoritmo Genético conforme os resultados da seleção de características para quatro classes [100, 93, 90 e 0].

## Apêndice B

# Processamento de Linguagem Natural

### B.1 Lista de *Stopwords*

Um dos processos para eliminação das informações que são irrelevantes quanto ao contexto e manutenção do conteúdo essencial do texto para interpretação foi a remoção de *stopwords*. Essas, são palavras que podem estar presentes em qualquer texto e são utilizadas como conectivos textuais, sem grande ganho na interpretação dos fatos apresentados. A Tabela B.1 apresenta a lista de *stopwords* para a língua inglesa enquanto a Tabela B.2 contém a lista de *stopwords* para a língua portuguesa.

Lista <i>stopwords</i> para o Inglês				
i	their	doing	above	each
me	theirs	a	below	few
my	themselves	an	to	more
myself	what	the	from	most
we	which	and	up	other
our	who	but	down	some
ours	whom	if	in	such
ourselves	this	or	out	no
you	that	because	on	nor
your	these	as	off	not
yours	those	until	over	only
yourself	am	while	under	own
yourselves	is	of	again	same
he	are	at	further	so
him	was	by	then	than
his	were	for	once	too
himself	be	with	here	very
she	been	about	there	s
her	being	against	when	t
hers	have	between	where	can
herself	has	into	why	will
it	had	through	how	just
its	having	during	all	don
itself	do	before	any	should
they	does	after	both	now
them	did			

Tabela B.1: Lista de *stopwords* utilizado no pré-processamento da língua inglesa.

Lista de <i>stopwords</i> para o Português				
de	ela	teus	havemos	fossem
a	entre	tuas	hã	for
o	depois	nosso	houve	formos
que	sem	nossa	houvemos	forem
e	mesmo	nossos	houveram	serei
do	aos	nossas	houvera	será
da	seus	dela	houvéramos	seremos
em	quem	delas	haja	serão
um	nas	esta	hajamos	seria
para	me	estes	hajam	seríamos
com	esse	estas	houvesse	seriam
não	eles	aquele	houvéssemos	tenho
uma	você	aquela	houvessem	tem
os	essa	aqueles	houver	temos
no	num	aquelas	houvermos	tém
se	nem	isto	houverem	tinha
na	suas	aquilo	houverei	tínhamos
por	meu	estou	houverá	tinham
mais	às	está	houveremos	tive
as	minha	estamos	houverão	teve
dos	numa	estão	houveria	tivemos
como	pelos	estive	houveríamos	tiveram
mas	elas	esteve	houveriam	tivera
ao	qual	estivemos	sou	tivéramos
ele	nós	estiveram	somos	tenha
das	lhe	estava	são	tenhamos
à	deles	estávamos	era	tenham
seu	essas	estavam	éramos	tivesse
sua	esses	estivera	eram	tivéssemos
ou	pelas	estivéramos	fui	tivessem
quando	este	esteja	foi	tiver
muito	dele	estejamos	fomos	tivermos
nos	tu	estejam	foram	tiverem
já	te	estivesse	fora	terei
eu	vocês	estivéssemos	fôramos	terá
também	vos	estivessem	seja	teremos
só	lhes	estiver	sejamos	terão
pelo	meus	estivermos	sejam	teria
pela	minhas	estiverem	fosse	teríamos
até	teu	hei	fôssemos	teriam
isso	tua	há		

Tabela B.2: Lista de *stopwords* utilizado no pré-processamento da língua portuguesa.

# Apêndice C

## Experimentos

### C.1 Base de Dados VestUfes

Para a base de dados do Vestibular da Universidade Federal do Espírito Santo, são apresentados os enunciados das cinco atividades de português avaliadas em 2012. Essas atividades foram utilizadas por (PISSINATI, 2014) para comparar sete modelos de avaliação semi-automática desenvolvidos com as correções de dois humanos especialistas. Cada atividade foi transcrita do documento original escrito à mão pelo aluno no vestibular. Ao total foram coletadas 92 atividades de cada questão, totalizando 460 respostas.

#### C.1.1 Questão 1

*Ao elaborar um panorama da ficção contemporânea, Flávio Carneiro (No país do presente. Rio de Janeiro: Rocco, 2005. p. 305-311) elenca uma série de traços de nossa literatura contemporânea, a saber: 1) o cruzamento da literatura com a linguagem da mídia e da internet; 2) a problematização das relações amorosas, em especial as femininas e homoeróticas; 3) a volta ao campo ou a cidades do interior; 4) o ressurgimento do narrador clássico, cujo relato beira a oralidade; 5) a força da narrativa fantástica; 6) a multiplicação*

*de textos de autoria e temática femininas; 7) a reescritura contínua das histórias; 8) o clima detetivesco; 9) a presença de fatos históricos; 10) o aumento de narrativas memorialísticas; 11) a ficção de cunho social; 12) a hegemonia de personagens anônimos, comuns, anti-heroicos; 13) a mistura de gêneros, literários ou não; 14) o humor como recurso e técnica; 15) a influência do mercado editorial. Considerando um dos três romances seguintes: Boca do inferno, de Ana Miranda (1989); Ensaio sobre a cegueira, de José Saramago (1995); Kitty aos 22: divertimento, de Reinaldo Santos Neves (2006), indique, na obra escolhida, a ocorrência de pelo menos dois dos quinze traços acima citados e explique de que modo eles ocorrem na narrativa.*

### C.1.2 Questão 2

Escolha uma das três obras seguintes: Romanceiro da Inconfidência, de Cecília Meireles; Vidas secas, de Graciliano Ramos; O noviço, de Martins Pena, e faça o que se pede:

- Justifique o título da obra que você escolheu.
- Explique a relevância do trecho da obra escolhida, abaixo transcrito, para a compreensão dessa obra.

#### TEXTO 1

Romance II ou Do ouro incansável

[...]

De seu calmo esconderijo,

o ouro vem, dócil e ingênuo;

torna-se pó, folha, barra,

prestígio, poder, engenho...

É tão claro! ? e turva tudo:

honra, amor e pensamento.

[...]

Mil galerias desabam;

mil homens ficam sepultos;

mil intrigas, mil enredos

prendem culpados e justos;

já ninguém dorme tranquilo,

que a noite é um mundo de sustos.

(Romanceiro da Inconfidência, de Cecília Meireles)

## TEXTO 2

Que iriam fazer? Retardaram-se, temerosos. Chegariam a uma terra desconhecida e civilizada, ficariam presos nela. E o sertão continuaria a mandar gente para lá. O sertão mandaria para a cidade homens fortes, brutos como Fabiano, sinha Vitória e os dois meninos.  
(Vidas secas, de Graciliano Ramos)

## TEXTO 3

AMBRÓSIO : Senhores, denuncio-vos um criminoso.

MEIRINHO : É verdade que tenho aqui uma ordem contra um noviço...

MESTRE : ... Que já de nada vale. (Prevenção.)

TODOS : O Padre-Mestre!

MESTRE (para Carlos) : Carlos, o Dom Abade julgou mais prudente que lá não voltásseis. Aqui tens a permissão por ele assinada para saíres do convento.

CARLOS (abraçando-o) : Meu bom Padre-Mestre, este ato reconcilia-me com os frades.

MESTRE : E vós, senhoras, esperai da justiça dos homens o castigo deste malvado. (Para Carlos e Emília:) E vós, meus filhos, sede felizes, que eu pedirei para todos (ao público:) indulgência! ? AMBRÓSIO : Oh, mulheres, mulheres! (Execução.)

(O noviço, de Martins Pena).

### C.1.3 Questão 3

Reescreva, com as devidas adaptações, as duas últimas falas do personagem MESTRE, de O noviço (TEXTO 3 da 2ª QUESTÃO), fazendo uso do pronome VOCÊ ou VOCÊS, conforme o caso.

### C.1.4 Questão 4

Leia os textos abaixo e faça o que se pede:

TEXTO 1

"O navio negreiro"

Negras mulheres, suspendendo às tetas

Magras crianças, cujas bocas pretas

Rega o sangue das mães:

Outras, moças... mas nuas, espantadas,

No turbilhão de espectros arrastadas,

Em ânsia e mágoa vãs.



(Castro Alves)

TEXTO 2

"7"

Eu não sou eu nem sou o outro,

Sou qualquer coisa de intermédio:

Pilar da ponte de tédio

Que vai de mim para o Outro.

(Mário de Sá-Carneiro)

TEXTO 3

"Os arredores florem"

Os arredores florem:

figos, nervos, libélulas

a criarem nas águas

os brevíssimos movimentos.

(Paulo Roberto Sodré)

- Escolha um dos textos acima ("O navio negreiro"; "7"; "Os arredores florem"), indique e explique a ocorrência de um dos seguintes aspectos: som (aliteração, assonância, paronomásia, etc.), sentido (metáfora, alegoria, ironia, etc.), ritmo (rima, métrica, tonicidade, etc.) ou representação (imagem, descrição, comparação, etc.).
- Nos três versos iniciais do trecho de "O navio negreiro"(TEXTO 1), o sujeito do enunciado é "o sangue das mães". Reescreva, em prosa, esses versos, iniciando o período

com "O sangue das mães", fazendo as adaptações que o texto requer e mantendo o sentido do texto original.

### **C.1.5 Questão 5**

Com base nos elementos constitutivos do ato de comunicação, Roman Jakobson estabeleceu seis funções da linguagem (e a ênfase de cada uma delas): referencial (ênfase no assunto; no conteúdo), emotiva (ênfase no emissor; no sujeito), conativa (ênfase no receptor; no interlocutor), poética (ênfase na forma; na construção), metalinguística (ênfase no código; na autorreferência) e fática (ênfase no canal; no contato).

Escolha um dos textos da Questão 4, indique e explique a ocorrência de uma dessas funções.

## C.2 Base de Dados Disciplinas UFES

Base de dados extraída de disciplinas da Universidade Federal do Espírito Santo entre 2015 e 2016. Dentre as atividades coletadas estão as disciplinas de Filosofia, Métodos e Técnicas de Pesquisa Científica - MTPC e Tecnologia da Informação II. A Tabela C.2 mostra para cada atividade o identificador, o nome, a quantidade e o percentual referente da base de treinamento e total de submissões, total de características selecionadas pelo mapa de características, percentual referente à seleção e o total de características da base de dados.

#	<i>Dataset</i>	Submissões			Características		
		Treino	(em %)	Total	Selecionadas	(em %)	Total
1	Filosofia-20-301	16	32,00	50	273	16,73	1632
2	Filosofia-20-326	18	35,29	51	159	10,11	1572
3	Filosofia-20-327	15	30,00	50	885	51,91	1705
4	Filosofia-20-328	16	32,65	49	409	22,35	1830
5	Filosofia-20-387	23	46,94	49	635	45,95	1382
6	mtpc-13-258	12	66,67	18	690	32,68	2111
7	mtpc-13-281	17	89,47	19	347	37,43	927
8	mtpc-13-283	11	84,61	13	548	18,45	2970
9	mtpc-13-307	8	88,89	9	1178	47,46	2482
10	mtpc-13-321	8	88,89	9	847	36,70	2308
11	mtpc-13-340	10	90,91	11	91	19,32	471
12	mtpc-13-352	11	84,61	13	108	9,16	1179
13	mtpc-13-355	15	93,75	16	4	0,59	676
14	mtpc-13-362	10	62,50	16	12	4,27	281
15	mtpc-13-365	12	80,00	15	18	4,37	411
16	mtpc-13-366	11	78,57	14	143	33,25	430
17	mtpc-13-380	12	70,59	17	210	35,65	589
18	mtpc-13-381	11	68,75	16	20	3,37	593
19	mtpc-13-382	13	86,67	15	5	0,72	692
20	mtpc-13-383	15	88,23	17	100	17,95	557
21	mtpc-13-392	11	84,61	13	8	5,97	134
22	mtpc-13-397	13	81,25	16	571	34,96	1633
23	mtpc-13-405	12	85,71	14	626	40,46	1547
24	mtpc-13-412	16	94,12	17	444	39,71	1118
25	mtpc-13-418	16	94,12	17	74	10,23	723
26	Tec-II-5-11	17	38,64	44	1422	32,74	4343
27	Tec-II-5-115	16	61,54	26	82	7,47	1097
28	Tec-II-5-117	17	89,47	19	23	4,99	461
29	Tec-II-5-131	16	80,00	20	53	9,81	540
30	Tec-II-5-159	15	51,72	29	336	50,22	669

Continua na próxima página

#	<i>Dataset</i>	Submissões			Características		
		Treino	(em %)	Total	Selecionadas	(em %)	Total
31	Tec-II-5-162	15	55,56	27	3	0,50	604
32	Tec-II-5-169	13	52,00	25	43	20,77	207
33	Tec-II-5-179	11	40,74	27	145	27,36	530
34	Tec-II-5-185	12	63,16	19	34	18,89	180
35	Tec-II-5-186	15	75,00	20	98	11,47	854
36	Tec-II-5-187	11	52,38	21	559	52,59	1063
37	Tec-II-5-189	15	60,00	25	160	16,39	976
38	Tec-II-5-190	16	61,54	26	99	8,75	1131
39	Tec-II-5-191	17	70,83	24	140	18,09	774
40	Tec-II-5-3	32	72,73	44	2234	47,82	4672
41	Tec-II-5-4	32	72,73	44	2352	50,34	4672
42	Tec-II-5-6	17	38,64	44	1411	32,49	4343
43	Tec-II-5-77	17	38,64	44	1667	38,38	4343
44	Tec-II-5-8	18	40,00	45	872	49,21	1772
45	Tec-II-5-82	18	40,00	45	707	39,90	1772

Última página

Tabela C.2: *Datasets* coletados em disciplinas da Universidade Federal do Espírito Santo com as respectivas informações de esforço de correção e redução de dimensionalidade.

A Tabela C.4 apresenta para todas as atividades dessa base de dados o desempenho de avaliação antes e depois do processo de seleção de características com as métricas de Erro Quadrático Médio - MSE e Erro Médio Absoluto - MAE, *Accuracy* - ACC, *Precision*, *Recall* e F1 para os dois classificadores testados KNN e CBC.

<i>Dataset</i>	KNN						CBC					
	MSE	MAE	ACC	Precision	Recall	F1	MSE	MAE	ACC	Precision	Recall	F1
Filosofia-20-327												
Inicial	0.0000	0.0000	1.0000	1.0000	1.0000	1.0000	0.0050	0.0100	0.9800	0.9850	0.9800	0.9813
Apos G.A.	0.0000	0.0000	1.0000	1.0000	1.0000	1.0000	0.0050	0.0100	0.9800	0.9850	0.9800	0.9813
Tec-II-5-191												
Inicial	0.0000	0.0000	1.0000	1.0000	1.0000	1.0000	0.0000	0.0000	1.0000	1.0000	1.0000	1.0000
Apos G.A.	0.0000	0.0000	1.0000	1.0000	1.0000	1.0000	0.0104	0.0208	0.9583	0.9609	0.9583	0.9576
Tec-II-5-169												
Inicial	0.0000	0.0000	1.0000	1.0000	1.0000	1.0000	0.0000	0.0000	1.0000	1.0000	1.0000	1.0000
Apos G.A.	0.0000	0.0000	1.0000	1.0000	1.0000	1.0000	0.0000	0.0000	1.0000	1.0000	1.0000	1.0000
mtpc-13-362												
Inicial	0.0000	0.0000	1.0000	1.0000	1.0000	1.0000	0.0013	0.0125	0.8750	0.8929	0.8750	0.8486
Apos G.A.	0.0000	0.0000	1.0000	1.0000	1.0000	1.0000	0.0625	0.0625	0.9375	0.9420	0.9375	0.9282
Tec-II-5-159												
Inicial	0.0062	0.0207	0.9310	0.9376	0.9310	0.9278	0.0031	0.0103	0.9655	0.9672	0.9655	0.9646
Apos G.A.	0.0062	0.0207	0.9310	0.9376	0.9310	0.9278	0.0031	0.0103	0.9655	0.9672	0.9655	0.9646
Tec-II-5-117												
Inicial	0.0132	0.0263	0.9474	0.9522	0.9474	0.9466	0.0132	0.0263	0.9474	0.9561	0.9474	0.9486
Apos G.A.	0.0132	0.0263	0.9474	0.9561	0.9474	0.9445	0.0132	0.0263	0.9474	0.9561	0.9474	0.9486
mtpc-13-281												
Inicial	0.0089	0.0263	0.8947	0.9013	0.8947	0.8927	0.0168	0.0421	0.8947	0.8947	0.8947	0.8947
Apos G.A.	0.0137	0.0316	0.8947	0.9079	0.8947	0.8851	0.0084	0.0211	0.9474	0.9539	0.9474	0.9476
Filosofia-20-328												
Inicial	0.0092	0.0306	0.8571	0.9412	0.8571	0.8828	0.0092	0.0306	0.8571	0.9412	0.8571	0.8828
Apos G.A.	0.0090	0.0612	0.4694	0.7021	0.4694	0.4154	0.0082	0.0367	0.7755	0.8112	0.7755	0.7863

Continua na próxima página

<i>Dataset</i>	KNN						CBC					
	MSE	MAE	ACC	Precision	Recall	F1	MSE	MAE	ACC	Precision	Recall	F1
Tec-II-5-185												
Inicial	0.0126	0.0632	0.6842	0.8565	0.6842	0.6842	0.0084	0.0421	0.7895	0.8660	0.7895	0.7828
Apos G.A.	0.0021	0.0105	0.9474	0.9579	0.9474	0.9491	0.0021	0.0105	0.9474	0.9579	0.9474	0.9491
Filosofia-20-387												
Inicial	0.0176	0.0694	0.6327	0.7406	0.6327	0.6377	0.0147	0.0490	0.7959	0.8605	0.7959	0.8130
Apos G.A.	0.0292	0.1000	0.5714	0.7366	0.5714	0.5712	0.0182	0.0633	0.7143	0.8776	0.7143	0.7533
Tec-II-5-3												
Inicial	0.1136	0.1136	0.8864	0.8871	0.8864	0.8862	0.1818	0.1818	0.8182	0.8182	0.8182	0.8182
Apos G.A.	0.1136	0.1136	0.8864	0.8878	0.8864	0.8866	0.2045	0.2045	0.7955	0.7958	0.7955	0.7952
Tec-II-5-4												
Inicial	0.1136	0.1136	0.8864	0.8871	0.8864	0.8862	0.1818	0.1818	0.8182	0.8182	0.8182	0.8182
Apos G.A.	0.0909	0.0909	0.9091	0.9091	0.9091	0.9091	0.2045	0.2045	0.7955	0.7958	0.7955	0.7952
Tec-II-5-131												
Inicial	0.1162	0.1285	0.8500	0.9250	0.8500	0.8826	0.1162	0.1285	0.8500	0.9250	0.8500	0.8826
Apos G.A.	0.1092	0.1215	0.8500	0.9500	0.8500	0.8848	0.1092	0.1215	0.8500	0.9500	0.8500	0.8848
Tec-II-5-187												
Inicial	0.1310	0.1667	0.7619	0.9286	0.7619	0.8163	0.1786	0.2143	0.7143	0.9107	0.7143	0.7692
Apos G.A.	0.0833	0.1190	0.8095	0.9429	0.8095	0.8709	0.1310	0.1667	0.7619	0.9286	0.7619	0.8163
Tec-II-5-186												
Inicial	0.1620	0.1800	0.8000	0.9000	0.8000	0.8167	0.1620	0.1800	0.8000	0.9000	0.8000	0.8167
Apos G.A.	0.1620	0.1800	0.8000	0.9000	0.8000	0.8167	0.1620	0.1800	0.8000	0.9000	0.8000	0.8167
Tec-II-5-8												
Inicial	0.1779	0.1811	0.7333	0.7852	0.7333	0.7519	0.1557	0.1589	0.7556	0.8020	0.7556	0.7708
Apos G.A.	0.1781	0.1824	0.7333	0.7827	0.7333	0.7533	0.1778	0.1796	0.7556	0.7975	0.7556	0.7708

Continua na próxima página

<i>Dataset</i>	KNN						CBC					
	MSE	MAE	ACC	Precision	Recall	F1	MSE	MAE	ACC	Precision	Recall	F1
Tec-II-5-82												
Inicial	0.1779	0.1811	0.7333	0.7852	0.7333	0.7519	0.1557	0.1589	0.7556	0.8020	0.7556	0.7708
Apos G.A.	0.2003	0.2047	0.7111	0.7537	0.7111	0.7277	0.2000	0.2018	0.7333	0.7807	0.7333	0.7514
Tec-II-5-189												
Inicial	0.2100	0.2200	0.7600	0.7876	0.7600	0.7520	0.1700	0.1800	0.8000	0.8933	0.8000	0.8155
Apos G.A.	0.1700	0.1800	0.8000	0.8933	0.8000	0.8155	0.1700	0.1800	0.8000	0.8933	0.8000	0.8155
Filosofia-20-301												
Inicial	0.0284	0.7200	0.7400	0.9001	0.7400	0.8002	0.0250	0.0700	0.7200	0.8418	0.7200	0.7664
Apos G.A.	0.0304	0.7200	0.7600	0.9500	0.7600	0.8249	0.0306	0.0740	0.7400	0.9115	0.7400	0.8019
Tec-II-5-6												
Inicial	0.2916	0.3250	0.6136	0.8600	0.6136	0.6407	0.3370	0.3705	0.5682	0.8582	0.5682	0.5971
Apos G.A.	0.3598	0.3932	0.5455	0.8600	0.5455	0.5767	0.4052	0.4386	0.5000	0.8691	0.5000	0.5392
Tec-II-5-77												
Inicial	0.2916	0.3250	0.6136	0.8600	0.6136	0.6407	0.3370	0.3705	0.5682	0.8582	0.5682	0.5971
Apos G.A.	0.3143	0.3477	0.5909	0.8582	0.5909	0.6185	0.4052	0.4386	0.5000	0.8691	0.5000	0.5392
Tec-II-5-11												
Inicial	0.2916	0.3250	0.6136	0.8600	0.6136	0.6407	0.3370	0.3705	0.5682	0.8582	0.5682	0.5971
Apos G.A.	0.2689	0.3023	0.6364	0.8636	0.6364	0.6637	0.3825	0.4159	0.5227	0.8636	0.5227	0.5573

Última página

Tabela C.4: Avaliação da classificação para cada atividade da base de dados conforme 6 métricas e dois classificadores.

A Tabela C.6 apresenta para todas as atividades com mais de uma classe dessa base de dados o processo de seleção de características com número de documentos e características - inicial e final. Para cada classe dessas atividades com mais de um item é analisada a similaridade interna - SI inicial e final, o coeficiente de redução e o número de características de classe.

Dataset	C. Final		Documentos	C. Inicial
Filosofia-20-327	885		50	1705
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
100	0,2057	0,2162	242	0,1619
60	1 Item de Classe 60,0			
50	0,4889	1,0343	748	0,3636
Tec-II-5-191	140		24	774
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
100	0,2895	0,0044	3	0,2761
50	0,2565	0,2968	137	0,1692
0	1 Item de Classe 0,0			
Tec-II-5-169	43		25	207
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
100	0,5446	0,2651	42	0,05686
50	1 Item de Classe 50,0			
0	0,5	0,00483	1	0,25
mtpc-13-362	12		16	281
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
100	0,4360	0,0561	12	0,0781
90	1 Item de Classe 90,0			
0	1 Item de Classe 0,0			
Tec-II-5-159	336		29	669
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
100	0,4359	0,8705	280	0,2775
70	0,2376	0,0015	1	0,2140
0	0,5573	0,3019	101	0,5902
Tec-II-5-117	23		19	461

Continua na próxima página



Dataset	C. Final		Documentos	C. Inicial
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
100	0,5746	0,0998	23	0,25
50	0,2712	0,0022	1	0,1914
0	0,1558	0,0024	1	0,1730
mtpc-13-281	347		19	927
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
100	0,3851	0,0081	5	0,3685
90	0,3334	0,5284	288	0,2197
50	0,1911	0,1476	88	0,1630
Filosofia-20-328	409		49	1830
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
100	0,2469	0,1624	185	0,2056
90	0,3136	0,2142	248	0,2622
80		1 Item de Classe 80,0		
60		1 Item de Classe 60,0		
50		1 Item de Classe 50,0		
Tec-II-5-185	34		19	180
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
100	0,4606	0,0056	1	0,1662
80	0,5746	0,2675	32	0,125
0	0,3528	0,0083	1	0,3333
Filosofia-20-387	635		49	1382
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
100	0,2463	0,2002	146	0,2250
90	0,4042	0,2049	130	0,3480
70	0,6223	0,7044	317	0,5677
60	0,2429	0,0101	14	0,2281
50	0,5856	0,2985	171	0,5899
Tec-II-5-3	2234		44	4672
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
100	0,3362	0,6384	1658	0,2141
77		1 Item de Classe 77,0		

Continua na próxima página

Dataset	C. Final		Documentos	C. Inicial
0	0,3060	0,2901	865	0,1885
Tec-II-5-4	2352		44	4672
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
100	0,3362	0,6869	1860	0,2090
77		1 Item de Classe 77,0		
0	0,3060	0,2827	827	0,1793
Tec-II-5-131	53		20	540
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
100	0,5450	0,2528	55	0,5
0	0,1982	0,0022	1	0,1905
Tec-II-5-187	559		21	1063
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
100	0,2678	0,7655	457	0,2423
0	0,3504	0,3355	171	0,3024
Tec-II-5-186	98		20	854
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
100	0,6088	0,1546	66	0,5
90	0,2853	0,0531	34	0,2886
0	0,1826	0,00132	1	0,1111
Tec-II-5-8	872		45	1772
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
100	0,7339	1,2928	572	0,6841
93	0,5014	0,1378	176	0,4462
90		1 Item de Classe 90,0		
0	0,4190	0,3700	285	0,3677
Tec-II-5-82	707		45	1772
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
100	0,7339	1,4243	335	0,7614
93	0,5014	0,3388	235	0,4688
90		1 Item de Classe 90,0		
0	0,4190	0,34737	260	0,4790
Tec-II-5-189	160		25	976

Continua na próxima página

Dataset	C. Final		Documentos	C. Inicial
Classe	SI Inicial	Coeficiente	N. Características	SI Final
100	0,2018	0,1021	68	0,1453
50	0,5102	0,1291	63	0,5
0	0,3036	0,0478	35	0,2632
Filosofia-20-301	273		50	1632
Classe	SI Inicial	Coeficiente	N. Características	SI Final
100	0,2480	0,0911	109	0,1901
90	0,5891	0,2169	177	0,5
80		1 Item de Classe 80,0		
60		1 Item de Classe 60,0		
Tec-II-5-6	1411		44	4343
Classe	SI Inicial	Coeficiente	N. Características	SI Final
100	0,4358	0,3571	837	0,2903
0	0,3629	0,3333	716	0,29078
Tec-II-5-77	1667		44	4343
Classe	SI Inicial	Coeficiente	N. Características	SI Final
100	0,4358	0,4823	1098	0,2949
0	0,3629	0,3475	747	0,2954
Tec-II-5-11	1422		44	4343
Classe	SI Inicial	Coeficiente	N. Características	SI Final
100	0,4358	0,3374	782	0,2947
0	0,3629	0,3623	750	0,2926

Última página

Tabela C.6: Similaridade interna anterior e posterior à seleção de características em atividades da base de dados com mais de uma classe de nota.

Dada a expectativa de aumento de similaridade interna para melhoria do processo de avaliação, visualizamos nas figuras à seguir o aumento de similaridade interna da classe para cada nota representada na Tabela C.6.

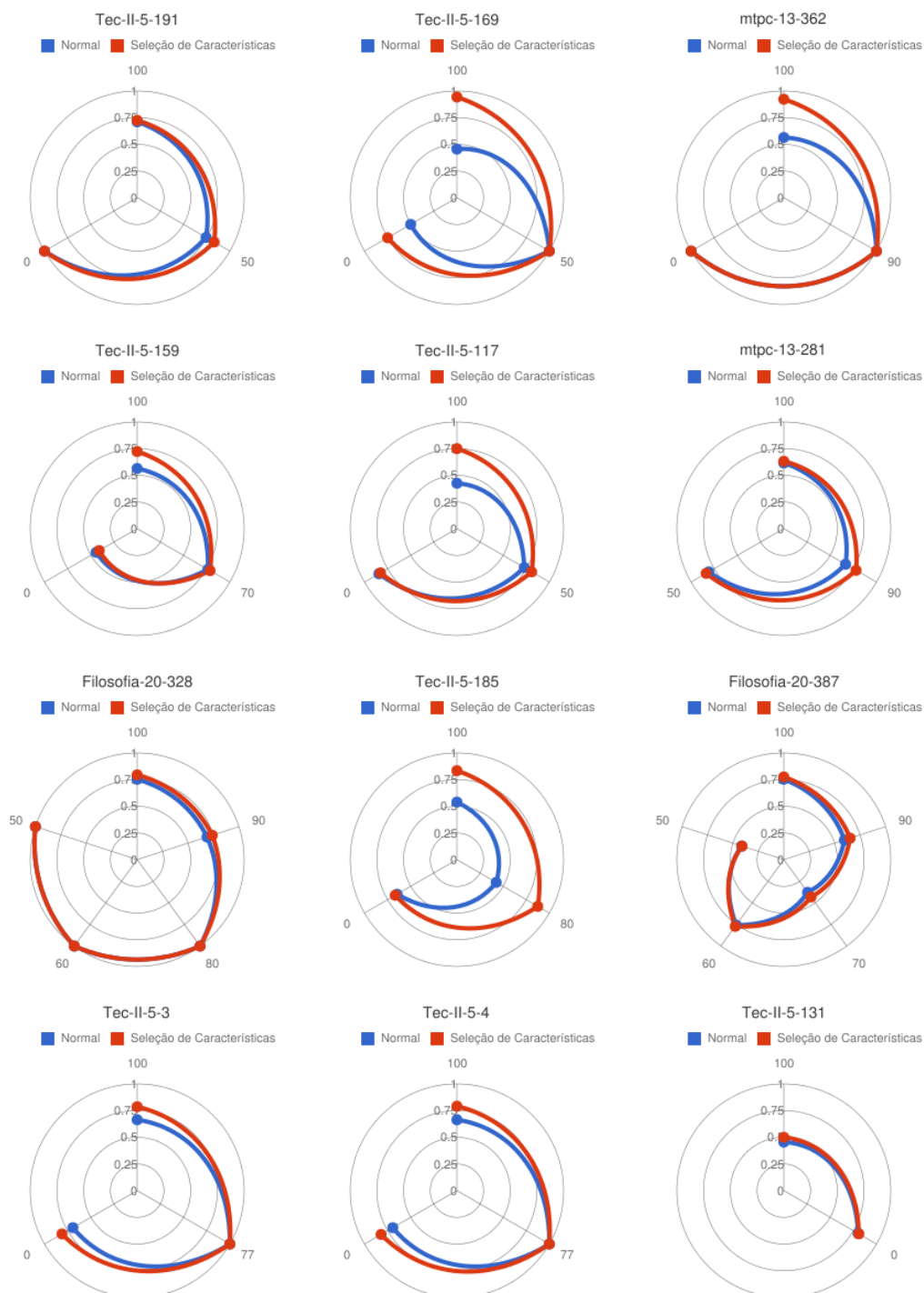


Figura C.1: Gráficos para aumento de similaridade por nota.

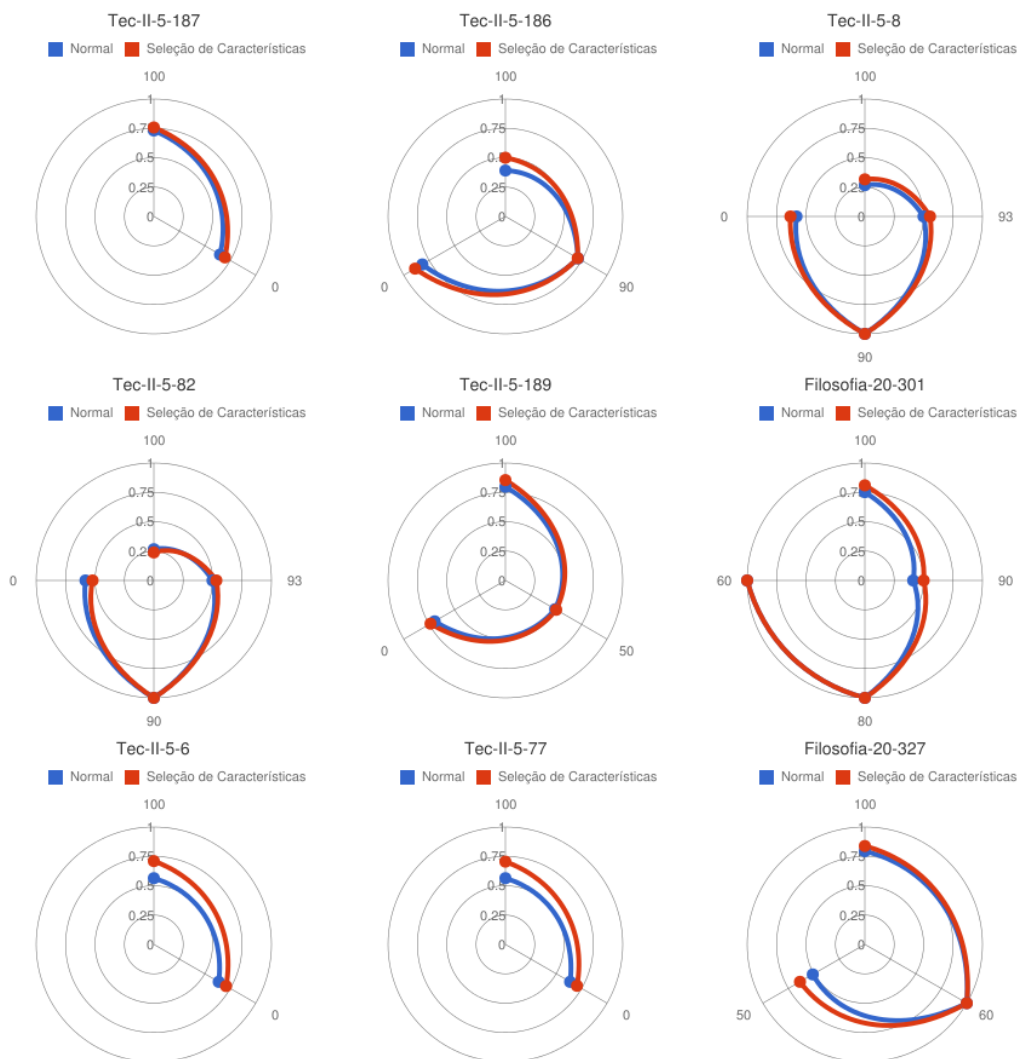


Figura C.2: Gráficos para aumento de similaridade por nota das atividades da base de dados de Disciplinas da UFES.

### C.3 Base de Dados Ciência da Computação UNT

Base de dados extraída de atividades da disciplina de Estrutura de Dados para a Ciência da Computação da Universidade do Norte do Texas. Foram aplicadas dez listas de exercícios com sete questões cada e duas provas com dez questões. Os dados foram coletados por (MOHLER; BUNESCU; MIHALCEA, 2011) e contou com 30 alunos na disciplina. A Tabela C.2 mostra para cada atividade o identificador, o nome, a quantidade e o percentual referente da base de treinamento e total de submissões, total de características selecionadas pelo mapa de características, percentual referente à seleção e o total de características da base de dados.

#	Dataset	Submissões			Características		
		Treino	(em %)	Total	Selecionadas	(em %)	Total
1	DS-CC UNT 1.1	16	55,17	29	101	14,41	701
2	DS-CC UNT 1.2	13	44,83	29	3	0,61	490
3	DS-CC UNT 1.3	14	48,28	29	407	43,21	942
4	DS-CC UNT 1.4	22	75,86	29	9	13,04	69
5	DS-CC UNT 1.5	13	44,83	29	4	1,01	395
6	DS-CC UNT 1.6	17	58,62	29	154	34,30	449
7	DS-CC UNT 1.7	14	48,28	29	7	1,02	689
8	DS-CC UNT 2.1	13	43,33	30	2	0,52	384
9	DS-CC UNT 2.2	22	73,33	30	94	13,53	695
10	DS-CC UNT 2.3	16	53,33	30	7	0,99	704
11	DS-CC UNT 2.4	18	60,00	30	2	1,02	197
12	DS-CC UNT 2.5	21	70,00	30	54	21,86	247
13	DS-CC UNT 2.6	22	73,33	30	288	34,16	843
14	DS-CC UNT 2.7	19	63,33	30	4	0,75	536
15	DS-CC UNT 3.1	15	48,39	31	26	10,61	245
16	DS-CC UNT 3.2	22	70,97	31	11	3,91	281
17	DS-CC UNT 3.3	21	67,74	31	4	0,84	475
18	DS-CC UNT 3.4	29	93,55	31	3	0,54	556
19	DS-CC UNT 3.5	18	58,06	31	3	0,92	327
20	DS-CC UNT 3.6	16	51,61	31	185	37,37	495
21	DS-CC UNT 3.7	12	38,71	31	3	0,46	659
22	DS-CC UNT 4.1	22	73,33	30	4	0,65	619
23	DS-CC UNT 4.2	14	46,67	30	474	52,73	899
24	DS-CC UNT 4.3	24	80,00	30	71	14,14	502
25	DS-CC UNT 4.4	18	60,00	30	8	1,13	707
26	DS-CC UNT 4.5	14	46,67	30	92	26,29	350
27	DS-CC UNT 4.6	18	60,00	30	2	1,87	107

Continua na próxima página

#	<i>Dataset</i>	Submissões			Características		
		Treino	(em %)	Total	Selecionadas	(em %)	Total
28	DS-CC UNT 4.7	22	73,33	30	2	1,28	156
29	DS-CC UNT 5.1	21	75,00	28	40	6,28	637
30	DS-CC UNT 5.2	16	57,14	28	5	0,81	618
31	DS-CC UNT 5.3	15	53,57	28	147	34,35	428
32	DS-CC UNT 5.4	21	75,00	28	4	1,77	226
33	DS-CC UNT 6.1	15	57,69	26	2	0,59	337
34	DS-CC UNT 6.2	16	61,54	26	2	1,45	138
35	DS-CC UNT 6.3	12	46,15	26	14	8,05	174
36	DS-CC UNT 6.4	20	76,92	26	2	0,60	331
37	DS-CC UNT 6.5	18	69,23	26	3	1,70	176
38	DS-CC UNT 6.6	14	53,85	26	30	11,76	255
39	DS-CC UNT 6.7	19	73,08	26	83	29,33	283
40	DS-CC UNT 7.1	18	69,23	26	1	0,19	528
41	DS-CC UNT 7.2	12	46,15	26	2	0,41	488
42	DS-CC UNT 7.3	16	61,54	26	6	0,93	642
43	DS-CC UNT 7.4	14	53,85	26	4	1,82	220
44	DS-CC UNT 7.5	17	65,38	26	28	5,91	474
45	DS-CC UNT 7.6	13	50,00	26	2	0,50	397
46	DS-CC UNT 7.7	17	65,38	26	138	33,91	407
47	DS-CC UNT 8.1	23	85,19	27	3	0,63	475
48	DS-CC UNT 8.3	13	48,15	27	251	39,28	639
49	DS-CC UNT 8.4	16	59,26	27	3	0,42	718
50	DS-CC UNT 8.5	17	62,96	27	8	1,43	560
51	DS-CC UNT 8.6	21	77,78	27	38	6,25	608
52	DS-CC UNT 8.7	19	70,37	27	156	22,77	685
53	DS-CC UNT 9.1	20	74,07	27	2	0,37	538
54	DS-CC UNT 9.2	15	55,56	27	1	0,76	131
55	DS-CC UNT 9.3	13	48,15	27	24	3,10	775
56	DS-CC UNT 9.4	17	62,96	27	285	31,60	902
57	DS-CC UNT 9.5	18	66,67	27	32	6,08	526
58	DS-CC UNT 9.6	16	59,26	27	4	2,37	169
59	DS-CC UNT 9.7	14	51,85	27	61	21,79	280
60	DS-CC UNT 10.1	15	62,50	24	3	0,57	527
61	DS-CC UNT 10.2	17	70,83	24	3	1,06	282
62	DS-CC UNT 10.3	21	87,50	24	3	2,80	107
63	DS-CC UNT 10.4	13	54,17	24	1	0,34	291
64	DS-CC UNT 10.5	18	75,00	24	253	39,04	648
65	DS-CC UNT 10.6	18	75,00	24	4	0,80	503
66	DS-CC UNT 10.7	20	83,33	24	3	2,05	146
67	DS-CC UNT 11.1	19	63,33	30	1	0,12	814
68	DS-CC UNT 11.1	15	50,00	30	2	0,59	340
69	DS-CC UNT 11.2	14	46,67	30	18	10,65	169
70	DS-CC UNT 11.3	16	53,33	30	5	1,01	494
71	DS-CC UNT 11.4	18	60,00	30	2	0,53	375

Continua na próxima página

#	<i>Dataset</i>	Submissões			Características		
		Treino	(em %)	Total	Selecionadas	(em %)	Total
72	DS-CC UNT 11.5	17	56,67	30	1	0,19	538
73	DS-CC UNT 11.6	19	63,33	30	1	0,34	298
74	DS-CC UNT 11.7	17	56,67	30	102	15,18	672
75	DS-CC UNT 11.8	15	50,00	30	121	16,49	734
76	DS-CC UNT 11.9	14	46,67	30	238	30,91	770
77	DS-CC UNT 12.1	17	60,71	28	154	48,28	319
78	DS-CC UNT 12.1	17	60,71	28	1	0,35	289
79	DS-CC UNT 12.2	21	75,00	28	225	37,38	602
80	DS-CC UNT 12.3	12	42,86	28	22	29,33	75
81	DS-CC UNT 12.4	15	53,57	28	28	3,12	898
82	DS-CC UNT 12.5	19	67,86	28	1	0,18	547
83	DS-CC UNT 12.6	15	53,57	28	138	22,48	614
84	DS-CC UNT 12.7	13	46,43	28	47	23,50	200
85	DS-CC UNT 12.8	15	53,57	28	176	26,39	667
86	DS-CC UNT 12.9	18	64,29	28	617	63,15	977

Última página

Tabela C.8: *Datasets* coletados por (MOHLER; BUNESCU; MIHALCEA, 2011) em turmas de Ciência da Computação durante as aulas de Estrutura de Dados na Universidade do Norte do Texas.

A Tabela C.10 apresenta para todas as atividades dessa base de dados o desempenho de avaliação antes e depois do processo de seleção de características com as métricas de Erro Quadrático Médio - MSE e Erro Médio Absoluto - MAE, *Accuracy* - ACC, *Precision*, *Recall* e F1 para os dois classificadores testados KNN e CBC.



<i>Dataset</i>	KNN						CBC					
	MSE	MAE	ACC	<i>Precision</i>	<i>Recall</i>	F1	MSE	MAE	ACC	<i>Precision</i>	<i>Recall</i>	F1
DS-CC UNT 6.1												
Inicial	0,003	0,015	0,923	0,954	0,923	0,931	0,003	0,015	0,923	0,954	0,923	0,931
Apos G.A.	0,003	0,015	0,923	0,954	0,923	0,931	0,137	0,238	0,577	0,763	0,577	0,458
DS-CC UNT 3.7												
Inicial	0,222	0,258	0,677	0,833	0,677	0,603	0,190	0,226	0,710	0,850	0,710	0,646
Apos G.A.	0,317	0,426	0,387	0,833	0,387	0,313	0,512	0,548	0,387	0,833	0,387	0,289
DS-CC UNT 1.3												
Inicial	0,026	0,076	0,759	0,895	0,759	0,788	0,026	0,076	0,759	0,895	0,759	0,788
Apos G.A.	0,054	0,117	0,690	0,835	0,690	0,651	0,026	0,076	0,759	0,895	0,759	0,788
DS-CC UNT 8.6												
Inicial	0,025	0,052	0,889	0,916	0,889	0,879	0,037	0,037	0,963	0,981	0,963	0,969
Apos G.A.	0,043	0,096	0,778	0,848	0,778	0,751	0,043	0,096	0,778	0,848	0,778	0,751
DS-CC UNT 9.4												
Inicial	0,213	0,267	0,630	0,780	0,630	0,579	0,015	0,059	0,741	0,756	0,741	0,722
Apos G.A.	0,136	0,207	0,630	0,762	0,630	0,554	0,058	0,111	0,741	0,786	0,741	0,722
DS-CC UNT 4.3												
Inicial	0,003	0,013	0,933	0,952	0,933	0,936	0,001	0,007	0,967	0,971	0,967	0,967
Apos G.A.	0,003	0,013	0,933	0,952	0,933	0,936	0,001	0,007	0,967	0,971	0,967	0,967
DS-CC UNT 10.4												
Inicial	0,005	0,025	0,875	0,938	0,875	0,891	0,005	0,025	0,875	0,938	0,875	0,891
Apos G.A.	0,013	0,067	0,667	0,824	0,667	0,651	0,258	0,342	0,542	0,761	0,542	0,42
DS-CC UNT 3.4												
Inicial	0,001	0,006	0,968	0,970	0,968	0,967	0,000	0,000	1,000	1,000	1,000	1
Apos G.A.	0,000	0,000	1,000	1,000	1,000	1,000	0,003	0,013	0,935	0,945	0,935	0,933

Continua na próxima página

<i>Dataset</i>	KNN						CBC					
	MSE	MAE	ACC	Precision	Recall	F1	MSE	MAE	ACC	Precision	Recall	F1
DS-CC UNT 12.3												
Inicial	0,149	0,204	0,643	0,773	0,643	0,660	0,143	0,189	0,679	0,802	0,679	0,699
Apos G.A.	0,118	0,161	0,679	0,829	0,679	0,702	0,238	0,339	0,464	0,814	0,464	0,417
DS-CC UNT 3.6												
Inicial	0,013	0,052	0,774	0,840	0,774	0,741	0,009	0,045	0,774	0,791	0,774	0,745
Apos G.A.	0,013	0,026	0,935	0,954	0,935	0,940	0,006	0,032	0,839	0,874	0,839	0,824
DS-CC UNT 2.5												
Inicial	0,167	0,167	0,833	0,877	0,833	0,826	0,067	0,067	0,933	0,940	0,933	0,931
Apos G.A.	0,188	0,193	0,800	0,863	0,800	0,796	0,092	0,140	0,700	0,801	0,700	0,620
DS-CC UNT 10.2												
Inicial	0,017	0,050	0,833	0,907	0,833	0,843	0,005	0,025	0,875	0,917	0,875	0,879
Apos G.A.	0,000	0,000	1,000	1,000	1,000	1,000	0,027	0,033	0,958	0,963	0,958	0,952
DS-CC UNT 1.7												
Inicial	0,215	0,248	0,586	0,785	0,586	0,522	0,189	0,214	0,690	0,758	0,690	0,631
Apos G.A.	0,007	0,034	0,828	0,819	0,828	0,808	0,372	0,400	0,483	0,719	0,483	0,388
DS-CC UNT 6.3												
Inicial	0,003	0,015	0,923	0,936	0,923	0,924	0,003	0,015	0,923	0,936	0,923	0,924
Apos G.A.	0,009	0,046	0,769	0,885	0,769	0,785	0,012	0,062	0,692	0,897	0,692	0,731
DS-CC UNT 7.5												
Inicial	0,012	0,031	0,923	0,923	0,923	0,923	0,012	0,031	0,923	0,923	0,923	0,923
Apos G.A.	0,006	0,015	0,962	0,981	0,962	0,968	0,031	0,077	0,808	0,772	0,808	0,774
DS-CC UNT 7.1												
Inicial	0,000	0,000	1,000	1,000	1,000	1,000	0,000	0,000	1,000	1,000	1,000	1,000
Apos G.A.	0,000	0,000	1,000	1,000	1,000	1,000	0,086	0,154	0,692	0,799	0,692	0,619

Continua na próxima página

<i>Dataset</i>	KNN						CBC					
	MSE	MAE	ACC	Precision	Recall	F1	MSE	MAE	ACC	Precision	Recall	F1
DS-CC UNT 12.2												
Inicial	0,198	0,211	0,750	0,815	0,750	0,723	0,169	0,186	0,786	0,830	0,786	0,766
Apos G.A.	0,205	0,221	0,750	0,823	0,750	0,710	0,205	0,221	0,750	0,811	0,750	0,722
DS-CC UNT 3.5												
Inicial	0,035	0,071	0,806	0,917	0,806	0,819	0,031	0,065	0,806	0,920	0,806	0,808
Apos G.A.	0,190	0,265	0,581	0,802	0,581	0,521	0,190	0,265	0,581	0,802	0,581	0,521
DS-CC UNT 12.6												
Inicial	0,017	0,086	0,571	0,786	0,571	0,459	0,044	0,079	0,750	0,823	0,750	0,687
Apos G.A.	0,049	0,100	0,643	0,792	0,643	0,555	0,083	0,129	0,643	0,792	0,643	0,564
DS-CC UNT 9.5												
Inicial	0,001	0,007	0,963	0,981	0,963	0,969	0,001	0,007	0,963	0,981	0,963	0,969
Apos G.A.	0,059	0,104	0,815	0,863	0,815	0,778	0,172	0,178	0,815	0,863	0,815	0,763
DS-CC UNT 11.8												
Inicial	0,092	0,120	0,800	0,839	0,800	0,792	0,117	0,137	0,800	0,889	0,800	0,809
Apos G.A.	0,192	0,217	0,733	0,802	0,733	0,745	0,092	0,183	0,633	0,688	0,633	0,578
DS-CC UNT 2.6												
Inicial	0,013	0,040	0,867	0,878	0,867	0,866	0,001	0,007	0,967	0,973	0,967	0,968
Apos G.A.	0,007	0,020	0,933	0,946	0,933	0,935	0,001	0,007	0,967	0,973	0,967	0,968
DS-CC UNT 7.6												
Inicial	0,018	0,062	0,769	0,843	0,769	0,788	0,017	0,054	0,808	0,908	0,808	0,829
Apos G.A.	0,017	0,054	0,808	0,908	0,808	0,829	0,045	0,131	0,577	0,758	0,577	0,536
DS-CC UNT 8.4												
Inicial	0,218	0,274	0,593	0,811	0,593	0,540	0,055	0,111	0,704	0,892	0,704	0,738
Apos G.A.	0,077	0,133	0,704	0,892	0,704	0,735	0,218	0,274	0,593	0,811	0,593	0,540

Continua na próxima página

<i>Dataset</i>	KNN						CBC					
	MSE	MAE	ACC	<i>Precision</i>	<i>Recall</i>	F1	MSE	MAE	ACC	<i>Precision</i>	<i>Recall</i>	F1
DS-CC UNT 6.5												
Inicial	0,017	0,054	0,808	0,867	0,808	0,819	0,017	0,038	0,885	0,924	0,885	0,900
Apos G.A.	0,018	0,046	0,846	0,919	0,846	0,865	0,031	0,092	0,692	0,817	0,692	0,649
DS-CC UNT 12.5												
Inicial	0,001	0,007	0,964	1,000	0,964	0,982	0,001	0,007	0,964	1,000	0,964	0,982
Apos G.A.	0,001	0,007	0,964	1,000	0,964	0,982	0,003	0,032	0,679	0,802	0,679	0,591
DS-CC UNT 4.7												
Inicial	0,021	0,027	0,967	0,983	0,967	0,972	0,021	0,027	0,967	0,983	0,967	0,972
Apos G.A.	0,001	0,007	0,967	0,983	0,967	0,970	0,001	0,007	0,967	0,983	0,967	0,970
DS-CC UNT 10.7												
Inicial	0,043	0,050	0,917	0,958	0,917	0,936	0,023	0,050	0,875	0,924	0,875	0,884
Apos G.A.	0,002	0,008	0,958	1,000	0,958	0,978	0,037	0,067	0,875	0,929	0,875	0,872
DS-CC UNT 2.7												
Inicial	0,212	0,247	0,700	0,826	0,700	0,662	0,036	0,087	0,767	0,878	0,767	0,779
Apos G.A.	0,036	0,087	0,767	0,878	0,767	0,779	0,224	0,267	0,667	0,811	0,667	0,621
DS-CC UNT 2.4												
Inicial	0,072	0,120	0,800	0,848	0,800	0,754	0,036	0,060	0,900	0,912	0,900	0,882
Apos G.A.	0,064	0,160	0,600	0,792	0,600	0,492	0,400	0,400	0,600	0,792	0,600	0,492
DS-CC UNT 7.2												
Inicial	0,002	0,008	0,962	0,981	0,962	0,968	0,002	0,008	0,962	0,981	0,962	0,968
Apos G.A.	0,002	0,008	0,962	0,981	0,962	0,968	0,020	0,100	0,500	0,795	0,500	0,414
DS-CC UNT 9.1												
Inicial	0,004	0,022	0,889	0,903	0,889	0,869	0,000	0,000	1,000	1,000	1,000	1,000
Apos G.A.	0,000	0,000	1,000	1,000	1,000	1,000	0,093	0,156	0,741	0,816	0,741	0,662

Continua na próxima página

<i>Dataset</i>	KNN						CBC					
	MSE	MAE	ACC	Precision	Recall	F1	MSE	MAE	ACC	Precision	Recall	F1
DS-CC UNT 11.1												
Inicial	0,180	0,257	0,633	0,783	0,633	0,525	0,000	0,000	1,000	1,000	1,000	1,000
Apos G.A.	0,000	0,000	1,000	1,000	1,000	1,000	0,180	0,257	0,633	0,783	0,633	0,525
DS-CC UNT 5.3												
Inicial	0,029	0,071	0,786	0,841	0,786	0,792	0,040	0,100	0,714	0,754	0,714	0,721
Apos G.A.	0,031	0,086	0,714	0,742	0,714	0,714	0,040	0,100	0,714	0,754	0,714	0,721
DS-CC UNT 2.2												
Inicial	0,007	0,020	0,933	0,947	0,933	0,935	0,017	0,060	0,767	0,857	0,767	0,760
Apos G.A.	0,017	0,060	0,767	0,850	0,767	0,767	0,012	0,047	0,800	0,830	0,800	0,800
DS-CC UNT 12.4												
Inicial	0,174	0,214	0,714	0,845	0,714	0,728	0,117	0,129	0,857	0,925	0,857	0,889
Apos G.A.	0,073	0,150	0,536	0,811	0,536	0,440	0,081	0,150	0,571	0,791	0,571	0,493
DS-CC UNT 8.3												
Inicial	0,031	0,096	0,630	0,836	0,630	0,618	0,027	0,089	0,630	0,719	0,630	0,600
Apos G.A.	0,058	0,156	0,556	0,789	0,556	0,468	0,040	0,096	0,741	0,843	0,741	0,723
DS-CC UNT 1.2												
Inicial	0,011	0,041	0,828	0,900	0,828	0,836	0,019	0,069	0,724	0,874	0,724	0,742
Apos G.A.	0,148	0,269	0,483	0,810	0,483	0,435	0,070	0,172	0,483	0,801	0,483	0,396
DS-CC UNT 8.1												
Inicial	0,039	0,044	0,926	0,946	0,926	0,923	0,001	0,007	0,963	0,981	0,963	0,966
Apos G.A.	0,061	0,067	0,926	0,957	0,926	0,934	0,076	0,081	0,889	0,915	0,889	0,878
DS-CC UNT 7.3												
Inicial	0,194	0,262	0,615	0,789	0,615	0,552	0,120	0,169	0,731	0,820	0,731	0,686
Apos G.A.	0,018	0,077	0,654	0,808	0,654	0,584	0,194	0,262	0,615	0,789	0,615	0,552

Continua na próxima página

<i>Dataset</i>	KNN						CBC					
	MSE	MAE	ACC	Precision	Recall	F1	MSE	MAE	ACC	Precision	Recall	F1
DS-CC UNT 10.6												
Inicial	0,148	0,175	0,750	0,843	0,750	0,739	0,013	0,033	0,917	0,958	0,917	0,937
Apos G.A.	0,063	0,117	0,750	0,857	0,750	0,715	0,183	0,200	0,750	0,857	0,750	0,728
DS-CC UNT 4.4												
Inicial	0,216	0,267	0,633	0,766	0,633	0,592	0,100	0,140	0,767	0,855	0,767	0,751
Apos G.A.	0,047	0,087	0,800	0,863	0,800	0,817	0,253	0,293	0,600	0,789	0,600	0,513
DS-CC UNT 9.3												
Inicial	0,247	0,319	0,519	0,813	0,519	0,498	0,067	0,141	0,593	0,827	0,593	0,644
Apos G.A.	0,081	0,170	0,593	0,752	0,593	0,597	0,139	0,222	0,556	0,773	0,556	0,581
DS-CC UNT 5.1												
Inicial	0,031	0,086	0,750	0,769	0,750	0,721	0,007	0,021	0,929	0,943	0,929	0,930
Apos G.A.	0,030	0,079	0,786	0,855	0,786	0,761	0,071	0,129	0,750	0,829	0,750	0,676
DS-CC UNT 7.4												
Inicial	0,072	0,115	0,769	0,803	0,769	0,752	0,014	0,038	0,885	0,906	0,885	0,882
Apos G.A.	0,045	0,115	0,577	0,804	0,577	0,518	0,377	0,392	0,577	0,797	0,577	0,497
DS-CC UNT 6.2												
Inicial	0,002	0,008	0,962	0,963	0,962	0,960	0,000	0,000	1,000	1,000	1,000	1,000
Apos G.A.	0,037	0,000	1,000	1,000	1,000	1,000	0,037	0,092	0,769	0,838	0,769	0,703
DS-CC UNT 10.3												
Inicial	0,000	0,000	1,000	1,000	1,000	1,000	0,000	0,000	1,000	1,000	1,000	1,000
Apos G.A.	0,000	0,000	1,000	1,000	1,000	1,000	0,125	0,125	0,875	0,897	0,875	0,844
DS-CC UNT 11.1												
Inicial	0,035	0,087	0,733	0,930	0,733	0,800	0,060	0,117	0,700	0,885	0,700	0,769
Apos G.A.	0,068	0,123	0,733	0,943	0,733	0,805	0,068	0,123	0,733	0,943	0,733	0,805

Continua na próxima página

<i>Dataset</i>	KNN						CBC					
	MSE	MAE	ACC	Precision	Recall	F1	MSE	MAE	ACC	Precision	Recall	F1
DS-CC UNT 10.1												
Inicial	0,097	0,150	0,667	0,726	0,667	0,644	0,042	0,108	0,667	0,727	0,667	0,654
Apos G.A.	0,038	0,108	0,667	0,828	0,667	0,648	0,203	0,250	0,625	0,795	0,625	0,529
DS-CC UNT 7.7												
Inicial	0,223	0,269	0,654	0,832	0,654	0,577	0,055	0,092	0,846	0,934	0,846	0,826
Apos G.A.	0,014	0,069	0,654	0,832	0,654	0,599	0,134	0,177	0,731	0,878	0,731	0,697
DS-CC UNT 6.4												
Inicial	0,022	0,046	0,885	0,942	0,885	0,903	0,045	0,069	0,885	0,940	0,885	0,906
Apos G.A.	0,063	0,115	0,769	0,854	0,769	0,767	0,155	0,162	0,808	0,882	0,808	0,768
DS-CC UNT 11.2												
Inicial	0,000	0,003	0,967	1,000	0,967	0,983	0,000	0,003	0,967	1,000	0,967	0,983
Apos G.A.	0,000	0,003	0,967	1,000	0,967	0,983	0,000	0,003	0,967	1,000	0,967	0,983
DS-CC UNT 6.6												
Inicial	0,003	0,015	0,923	0,937	0,923	0,908	0,003	0,015	0,923	0,937	0,923	0,908
Apos G.A.	0,000	0,000	1,000	1,000	1,000	1,000	0,002	0,008	0,962	0,965	0,962	0,956
DS-CC UNT 4.2												
Inicial	0,053	0,133	0,633	0,845	0,633	0,640	0,049	0,140	0,567	0,869	0,567	0,582
Apos G.A.	0,059	0,133	0,633	0,649	0,633	0,626	0,053	0,147	0,567	0,869	0,567	0,571
DS-CC UNT 12.7												
Inicial	0,000	0,000	1,000	1,000	1,000	1,000	0,000	0,000	1,000	1,000	1,000	1,000
Apos G.A.	0,000	0,000	1,000	1,000	1,000	1,000	0,016	0,079	0,607	0,780	0,607	0,496
DS-CC UNT 11.5												
Inicial	0,117	0,133	0,833	0,849	0,833	0,808	0,042	0,050	0,933	0,950	0,933	0,933
Apos G.A.	0,408	0,417	0,567	0,768	0,567	0,444	0,408	0,417	0,567	0,768	0,567	0,444

Continua na próxima página

<i>Dataset</i>	KNN						CBC					
	MSE	MAE	ACC	Precision	Recall	F1	MSE	MAE	ACC	Precision	Recall	F1
DS-CC UNT 3.3												
Inicial	0,092	0,123	0,806	0,835	0,806	0,783	0,004	0,019	0,903	0,940	0,903	0,910
Apos G.A.	0,021	0,065	0,774	0,853	0,774	0,764	0,206	0,245	0,677	0,792	0,677	0,597
DS-CC UNT 3.1												
Inicial	0,023	0,039	0,903	0,952	0,903	0,918	0,023	0,039	0,903	0,952	0,903	0,918
Apos G.A.	0,116	0,155	0,774	0,807	0,774	0,771	0,116	0,155	0,774	0,807	0,774	0,771
DS-CC UNT 12.8												
Inicial	0,285	0,350	0,536	0,859	0,536	0,541	0,264	0,329	0,536	0,841	0,536	0,552
Apos G.A.	0,249	0,314	0,536	0,845	0,536	0,560	0,249	0,314	0,536	0,845	0,536	0,560
DS-CC UNT 11.7												
Inicial	0,112	0,197	0,567	0,608	0,567	0,527	0,105	0,163	0,633	0,754	0,633	0,627
Apos G.A.	0,064	0,117	0,667	0,837	0,667	0,643	0,070	0,117	0,700	0,864	0,700	0,699
DS-CC UNT 2.3												
Inicial	0,017	0,073	0,667	0,822	0,667	0,676	0,023	0,073	0,733	0,828	0,733	0,752
Apos G.A.	0,047	0,127	0,600	0,683	0,600	0,588	0,021	0,080	0,667	0,778	0,667	0,653
DS-CC UNT 5.2												
Inicial	0,019	0,036	0,929	1,000	0,929	0,963	0,019	0,036	0,929	1,000	0,929	0,963
Apos G.A.	0,020	0,043	0,893	0,963	0,893	0,924	0,023	0,057	0,821	0,897	0,821	0,836
DS-CC UNT 11.9												
Inicial	0,023	0,080	0,667	0,800	0,667	0,690	0,015	0,053	0,767	0,933	0,767	0,816
Apos G.A.	0,015	0,053	0,767	0,831	0,767	0,792	0,015	0,053	0,767	0,912	0,767	0,815
DS-CC UNT 12.1												
Inicial	0,163	0,189	0,750	0,819	0,750	0,748	0,106	0,136	0,750	0,782	0,750	0,743
Apos G.A.	0,208	0,250	0,643	0,641	0,643	0,631	0,079	0,118	0,750	0,790	0,750	0,733

Continua na próxima página



<i>Dataset</i>	KNN						CBC					
	MSE	MAE	ACC	<i>Precision</i>	<i>Recall</i>	F1	MSE	MAE	ACC	<i>Precision</i>	<i>Recall</i>	F1
DS-CC UNT 12.9												
Inicial	0,164	0,186	0,714	0,885	0,714	0,732	0,072	0,118	0,714	0,841	0,714	0,748
Apos G.A.	0,164	0,186	0,714	0,885	0,714	0,732	0,147	0,164	0,714	0,876	0,714	0,739
DS-CC UNT 1.5												
Inicial	0,030	0,069	0,828	0,864	0,828	0,837	0,023	0,048	0,897	0,940	0,897	0,914
Apos G.A.	0,114	0,214	0,552	0,617	0,552	0,493	0,110	0,207	0,552	0,624	0,552	0,494
DS-CC UNT 11.3												
Inicial	0,116	0,153	0,700	0,744	0,700	0,698	0,104	0,160	0,600	0,725	0,600	0,608
Apos G.A.	0,120	0,187	0,567	0,812	0,567	0,524	0,064	0,107	0,767	0,872	0,767	0,778
DS-CC UNT 2.1												
Inicial	0,073	0,153	0,633	0,684	0,633	0,620	0,073	0,113	0,800	0,928	0,800	0,840
Apos G.A.	0,255	0,340	0,500	0,807	0,500	0,454	0,255	0,340	0,500	0,807	0,500	0,454
DS-CC UNT 8.7												
Inicial	0,039	0,089	0,778	0,869	0,778	0,785	0,024	0,074	0,741	0,795	0,741	0,735
Apos G.A.	0,147	0,185	0,741	0,840	0,741	0,707	0,067	0,111	0,778	0,864	0,778	0,788
DS-CC UNT 1.6												
Inicial	0,044	0,110	0,655	0,811	0,655	0,644	0,043	0,103	0,690	0,807	0,690	0,674
Apos G.A.	0,149	0,221	0,621	0,661	0,621	0,599	0,048	0,090	0,793	0,823	0,793	0,799
DS-CC UNT 8.5												
Inicial	0,086	0,104	0,852	0,922	0,852	0,885	0,086	0,104	0,852	0,926	0,852	0,881
Apos G.A.	0,062	0,074	0,889	0,972	0,889	0,924	0,138	0,215	0,630	0,786	0,630	0,582
DS-CC UNT 12.1												
Inicial	0,021	0,043	0,857	0,961	0,857	0,902	0,004	0,018	0,893	1,000	0,893	0,943
Apos G.A.	0,004	0,018	0,893	1,000	0,893	0,943	0,227	0,296	0,607	0,814	0,607	0,525

Continua na próxima página

<i>Dataset</i>	KNN						CBC					
	MSE	MAE	ACC	<i>Precision</i>	<i>Recall</i>	F1	MSE	MAE	ACC	<i>Precision</i>	<i>Recall</i>	F1
DS-CC UNT 3.2												
Inicial	0,001	0,006	0,968	0,970	0,968	0,968	0,003	0,013	0,935	0,945	0,935	0,935
Apos G.A.	0,034	0,039	0,935	0,938	0,935	0,933	0,008	0,039	0,806	0,889	0,806	0,813
DS-CC UNT 4.1												
Inicial	0,017	0,047	0,833	0,872	0,833	0,820	0,017	0,047	0,867	0,871	0,867	0,866
Apos G.A.	0,032	0,080	0,800	0,875	0,800	0,790	0,075	0,133	0,733	0,831	0,733	0,696
DS-CC UNT 1.1												
Inicial	0,022	0,069	0,759	0,849	0,759	0,738	0,017	0,041	0,862	0,908	0,862	0,867
Apos G.A.	0,033	0,069	0,828	0,880	0,828	0,821	0,028	0,055	0,862	0,908	0,862	0,866
DS-CC UNT 4.5												
Inicial	0,053	0,093	0,767	0,898	0,767	0,790	0,196	0,260	0,533	0,729	0,533	0,451
Apos G.A.	0,101	0,200	0,533	0,828	0,533	0,517	0,232	0,293	0,500	0,689	0,500	0,433
DS-CC UNT 10.5												
Inicial	0,007	0,033	0,833	0,879	0,833	0,822	0,018	0,042	0,875	0,910	0,875	0,872
Apos G.A.	0,007	0,033	0,833	0,879	0,833	0,822	0,003	0,017	0,917	0,951	0,917	0,923
DS-CC UNT 5.4												
Inicial	0,070	0,093	0,857	0,889	0,857	0,850	0,047	0,064	0,893	0,911	0,893	0,888
Apos G.A.	0,000	0,000	1,000	1,000	1,000	1,000	0,160	0,200	0,750	0,847	0,750	0,729
DS-CC UNT 9.7												
Inicial	0,000	0,000	1,000	1,000	1,000	1,000	0,000	0,000	1,000	1,000	1,000	1,000
Apos G.A.	0,000	0,000	1,000	1,000	1,000	1,000	0,000	0,000	1,000	1,000	1,000	1,000
DS-CC UNT 11.4												
Inicial	0,093	0,163	0,667	0,832	0,667	0,591	0,041	0,077	0,833	0,913	0,833	0,8535
Apos G.A.	0,031	0,110	0,600	0,818	0,600	0,511	0,344	0,363	0,600	0,818	0,600	0,5110

Continua na próxima página

<i>Dataset</i>	KNN						CBC					
	MSE	MAE	ACC	Precision	Recall	F1	MSE	MAE	ACC	Precision	Recall	F1
DS-CC UNT 9.6												
Inicial	0,000	0,000	1,000	1,000	1,000	1,000	0,000	0,000	1,000	1,000	1,000	1,000
Apos G.A.	0,000	0,000	1,000	1,000	1,000	1,000	0,000	0,000	1,000	1,000	1,000	1,000
DS-CC UNT 11.6												
Inicial	0,033	0,033	0,967	1,000	0,967	0,983						
DS-CC UNT 4.6												
Inicial	0,033	0,033	0,967	0,952	1,000	0,976	0,033	0,033	0,967	0,952	1,000	0,975
Apos G.A.	0,267	0,267	0,733	0,619	1,000	0,765	0,133	0,133	0,867	1,000	0,840	0,913
DS-CC UNT 6.7												
Inicial	0,039	0,039	0,962	1,000	0,958	0,979	0,039	0,039	0,962	1,000	0,958	0,978
Apos G.A.	0,154	0,154	0,846	0,870	0,952	0,909	0,039	0,039	0,962	1,000	0,958	0,978
DS-CC UNT 1.4												
Inicial	0,000	0,000	1,000	1,000	1,000	1,000	0,000	0,000	1,000	1,000	1,000	1,000
Apos G.A.	0,000	0,000	1,000	1,000	1,000	1,000	0,000	0,000	1,000	1,000	1,000	1,000
DS-CC UNT 9.2												
Inicial	0,000	0,000	1,000	1,000	1,000	1,000	0,000	0,000	1,000	1,000	1,000	1,000
Apos G.A.	0,000	0,000	1,000	1,000	1,000	1,000	0,444	0,444	0,556	0,539	1,000	0,700

Última página

Tabela C.10: Avaliação da classificação para cada atividade da base de dados conforme 6 métricas e dois classificadores.

A Tabela C.12 apresenta para todas as atividades com mais de uma classe dessa base de dados o processo de seleção de características com número de documentos e características - inicial e final. Para cada classe dessas atividades com mais de um item é analisada a similaridade interna - SI inicial e final, o coeficiente de redução e o número de características de classe.

<i>Dataset</i>	C. Final	Documentos		C. Inicial
DS-CC UNT 6,1	2	26		337
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,2074	0,003	1	0
4	0,4022	0,0045	1	0,1111
2	1 Item de Classe 2,0			
DS-CC UNT 3,7	3	31		659
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,3699	0,0019	1	0,08
3	1 Item de Classe 3,0			
2	0,4074	0,0015	1	0
1	0,5	0,0015	1	0
0	1 Item de Classe 0,0			
DS-CC UNT 1,3	407	29		942
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,2749	0,0311	22	0,273
4	1 Item de Classe 4,0			
3	0,2699	0,0012	1	0,2769
2	0,9007	0,8386	395	0,25
DS-CC UNT 8,6	38	27		608
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,2279	0,0018	1	0,5252
4	0,5618	0,1316	40	0,5
3	0,4074	0,0016	1	0,5556
2	1 Item de Classe 2,0			
1	1 Item de Classe 1,0			
0	1 Item de Classe 0,0			

Continua na próxima página

<i>Dataset</i>	C. Final		Documentos	C. Inicial
DS-CC UNT 9,4	285		27	902
Classe	SI Inicial	Coeficiente	N. Características	SI Final
5	0,2462	0,0938	54	0,2295
4	0,5	0,0011	1	0,5
3	0,3876	0,1314	71	0,3333
2	0,5625	0,7005	192	0,5
1	1 Item de Classe 1,0			
0	0,5422	0,0011	1	0,5426
DS-CC UNT 4,3	71		30	502
Classe	SI Inicial	Coeficiente	N. Características	SI Final
5	0,3072	0,002	1	0,0481
4	0,2795	0,0025	1	0,2
3	0,3881	0,003	1	0,3333
1	0,5806	0,2709	68	0,25
DS-CC UNT 10,4	1		24	291
Classe	SI Inicial	Coeficiente	N. Características	SI Final
5	0,3402	0,0044	1	0,0494
4	0,4504	0	0	0
1	1 Item de Classe 1,0			
DS-CC UNT 3,4	3		31	556
Classe	SI Inicial	Coeficiente	N. Características	SI Final
5	0,1809	0,002	1	0,0165
4	0,2348	0,0018	1	0
3	0,2014	0,0019	1	0,0059
DS-CC UNT 12,3	22		28	75
Classe	SI Inicial	Coeficiente	N. Características	SI Final
10	0,9176	0,1067	8	0,0625
8	0,8669	0,1867	7	0,25
6	0,9183	0,2133	8	0,25
4	1 Item de Classe 4,0			
2	0,6417	0,0267	2	0
0	1 Item de Classe 0,0			

Continua na próxima página

<i>Dataset</i>	C. Final		Documentos	C. Inicial
DS-CC UNT 3,6	185		31	495
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,4885	0,6056	185	0,1711
4	0,4914	0,002	1	0,3818
3	0,5445	0,002	1	0,5
2	0,5	0,004	1	0,5
1	1 Item de Classe 1,0			
DS-CC UNT 2,5	54		30	247
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,1517	0,004	1	0,065
4	1 Item de Classe 4,0			
1	0,5884	0,0081	1	0,6768
0	0,3147	0,2632	52	0,04
DS-CC UNT 10,2	3		24	282
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,6155	0,0035	1	1
4	0,2484	0,004	1	0,0123
3	0,4161	0,0053	1	0,1111
2	1 Item de Classe 2,0			
0	1 Item de Classe 0,0			
DS-CC UNT 1,7	7		29	689
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,2465	0,0097	6	0,02
4	0,5	0,0015	1	0,25
3	1 Item de Classe 3,0			
0	1 Item de Classe 0,0			
DS-CC UNT 6,3	14		26	174
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,6752	0,092	12	0,0625
4	0,2846	0,0057	1	0,0556
1	0,5	0,0115	1	0,5
DS-CC UNT 7,5	28		26	474

Continua na próxima página

<i>Dataset</i>	C. Final		Documentos	C. Inicial
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,3484	0,0025	1	0,5184
3			1 Item de Classe 3,0	
2	0,5446	0,1139	27	0,25
0			1 Item de Classe 0,0	
DS-CC UNT 7,1	1		26	528
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5			1 Item de Classe 5,0	
4	0,1776	0,0022	1	0,0178
2			1 Item de Classe 2,0	
1			1 Item de Classe 1,0	
DS-CC UNT 12,2	225		28	602
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
10	0,1909	0,0017	1	0,2196
6	0,3701	0,0075	3	0,3653
5			1 Item de Classe 5,0	
4			1 Item de Classe 4,0	
3	0,3161	0,0017	1	0,1875
0	1	0,3704	223	0
DS-CC UNT 3,5	3		31	327
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,3116	0,0033	1	0,0069
4	0,5	0,0031	1	0
3			1 Item de Classe 3,0	
1	0,5277	0,0031	1	0
DS-CC UNT 12,6	138		28	614
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
10	0,2636	0,3281	139	0,1022
8			1 Item de Classe 8,0	
3			1 Item de Classe 3,0	
0	0,8	0,0016	1	0,5
DS-CC UNT 9,5	32		27	526

Continua na próxima página

<i>Dataset</i>	C. Final		Documentos	C. Inicial
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,2301	0,0021	1	0,0473
4	1 Item de Classe 4,0			
3	0,5875	0,0019	1	0,25
2	0,5766	0,1141	30	0,25
0	1 Item de Classe 0,0			
DS-CC UNT 11,8	121		30	734
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
10	0,3842	0,0018	1	0,25
9	1 Item de Classe 9,0			
5	1 Item de Classe 5,0			
3	0,3507	0,0777	38	0,3333
0	0,2578	0,1848	90	0,2039
DS-CC UNT 2,6	288		30	843
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,2858	0,0153	10	0,1988
4	0,4528	0,3081	140	0,3533
3	0,2882	0,0012	1	0,2433
1	0,61	0,5914	180	0,5
DS-CC UNT 7,6	2		26	397
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	1 Item de Classe 5,0			
4	0,3754	0,0025	1	0,0625
3	0,2321	0,0029	1	0,0204
0	1 Item de Classe 0,0			
DS-CC UNT 8,4	3		27	718
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,2582	0,0016	1	0,0313
4	1 Item de Classe 4,0			
3	0,5	0,0014	1	0
2	1 Item de Classe 2,0			
1	1 Item de Classe 1,0			

Continua na próxima página



<i>Dataset</i>	C. Final		Documentos	C. Inicial
0	0,3648	0,0014	1	0
DS-CC UNT 6,5	3		26	176
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,5229	0,0057	1	0
4	0,223	0,0065	1	0,0313
3	0,4259	0,0057	1	0
DS-CC UNT 12,5	1		28	547
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
10	0,1529	0,0019	1	0,0030864198
9	1 Item de Classe 9,0			
DS-CC UNT 4,7	2		30	156
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,4835	0,0513	1	0,7623
4	1 Item de Classe 4,0			
1	1 Item de Classe 1,0			
0	0,2728	0,0128	1	0,2134
DS-CC UNT 10,7	3		24	146
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,3063	0,0068	1	0,0069
3	1 Item de Classe 3,0			
2	0	0,0068	1	0
0	0,2297	0,0086	1	0,04
DS-CC UNT 2,7	4		30	536
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,3569	0,0019	1	0
4	0,3662	0,0019	1	0,1111
3	0,1817	0,0021	1	0,0123
2	1 Item de Classe 2,0			
0	0,2222	0,0019	1	0
DS-CC UNT 2,4	2		30	197
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,2138	0,006	1	0,0296

Continua na próxima página

<i>Dataset</i>	C. Final		Documentos	C. Inicial
3	1 Item de Classe 3,0			
2	0,5	0,0102	1	0,25
1	1 Item de Classe 1,0			
0	1 Item de Classe 0,0			
DS-CC UNT 7,2	2		26	488
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,207	0,002	1	0
4	1 Item de Classe 4,0			
3	0,5083	0,002	1	0,25
DS-CC UNT 9,1	2		27	538
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,2099	0,002	1	0,0069
4	0,5698	0,0037	1	0,25
2	1 Item de Classe 2,0			
DS-CC UNT 11,1	1		30	814
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
10	0,1874	0,0013	1	0,0039
5	1 Item de Classe 5,0			
4	1 Item de Classe 4,0			
3	1 Item de Classe 3,0			
DS-CC UNT 5,3	147		28	428
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,3796	0,0023	1	0,2774
4	0,5172	0,0023	1	0,5
3	0	0,3364	144	0
2	0,3905	0,0023	1	0,356
DS-CC UNT 2,2	94		30	695
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,3349	0,0017	1	0,3355
4	0,3862	0,0014	1	0,3981
3	0,3323	0,0018	1	0,382
2	1 Item de Classe 2,0			

Continua na próxima página

<i>Dataset</i>	C. Final		Documentos	C. Inicial
1	0,539	0,0547	19	0,5
0	0,6591	0,2129	74	0,6581
DS-CC UNT 12,4	28		28	898
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
10	0,231	0,0344	27	0,0156
8	1 Item de Classe 8,0			
2	0,1856	0,0011	1	0,1308
DS-CC UNT 8,3	251		27	639
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,4002	0,815	247	0,2823
4	0,5557	0,0016	1	0,5
3	0,5167	0,0016	1	0,5
2	0,5288	0,0188	6	0,5
1	1 Item de Classe 1,0			
DS-CC UNT 1,2	3		29	490
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,5333	0,002	1	0
4	1 Item de Classe 4,0			
3	0,2566	0,002	1	0
2	0,5556	0,002	1	0
1	1 Item de Classe 1,0			
DS-CC UNT 8,1	3		27	475
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,1891	0,0022	1	0,0185
4	0,5497	0,0021	1	0
1	1 Item de Classe 1,0			
0	0,5417	0,0021	1	0
DS-CC UNT 7,3	6		26	642
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,2416	0,0018	1	0,0781
4	1 Item de Classe 4,0			
3	0,5431	0,0093	4	0,1111

Continua na próxima página

<i>Dataset</i>	C. Final		Documentos	C. Inicial
2	1 Item de Classe 2,0			
1	0,3523	0,0016	1	0
DS-CC UNT 10,6	4		24	503
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,2375	0,002	1	0,0123
4	0,5	0	0	0,25
3	0,5	0,002	1	0
2	0,7339	0,004	1	0,25
0	0,356	0,002	1	0
DS-CC UNT 4,4	8		30	707
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,2972	0,0017	1	0,0612
4	0,4741	0,0196	6	0,4444
3	0,5589	0,0014	1	0,25
1	0,3134	0,0018	1	0,04
0	1 Item de Classe 0,0			
DS-CC UNT 9,3	24		27	775
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,2537	0	0	0,494
4	1 Item de Classe 4,0			
3	1 Item de Classe 3,0			
2	0,4276	0,055	23	0,2222
1	0,5959	0,0013	1	0,7041
0	1 Item de Classe 0,0			
DS-CC UNT 5,1	40		28	637
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,2287	0,0718	37	0,054
4	0,2663	0,002	1	0,08
3	0,351	0,0016	1	0,2222
2	1 Item de Classe 2,0			
1	0,3708	0,0016	1	0,3333
DS-CC UNT 7,4	4		26	220

Continua na próxima página

<i>Dataset</i>	C. Final		Documentos	C. Inicial
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,3008	0,0053	1	0,0204
4	0,5861	0,0091	1	0,25
3	0,5	0,0091	1	0,5
2	0,5	0,0045	1	0
0	1 Item de Classe 0,0			
DS-CC UNT 6,2	2		26	138
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,3518	0,0081	1	0,1
4	0,3267	0,0072	1	0
3	1 Item de Classe 3,0			
DS-CC UNT 10,3	3		24	107
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,4702	0,0093	1	0,0051
4	0,5	0,0187	1	0,25
3	0,7696	0,0093	1	0
2	1 Item de Classe 2,0			
0	1 Item de Classe 0,0			
DS-CC UNT 11,1	2		30	340
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
10	0,3058	0,0029	1	0
7	1 Item de Classe 7,0			
6	1 Item de Classe 6,0			
3	0,4768	0,0044	1	0,1111
DS-CC UNT 10,1	3		24	527
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,4415	0,0019	1	0
4	0,3607	0,0028	1	0,1111
3	0,2948	0,0025	1	0,125
2	0,3961	0	0	0
1	1 Item de Classe 1,0			
0	1 Item de Classe 0,0			

Continua na próxima página

<i>Dataset</i>	C. Final		Documentos	C. Inicial
DS-CC UNT 7,7	138		26	407
Classe	SI Inicial	Coeficiente	N. Características	SI Final
5	0,1624	0,0027	1	0,1249
4	0,7588	0,6732	137	0,25
3	0,5	0,0025	1	0,5
1		1 Item de Classe 1,0		
0		1 Item de Classe 0,0		
DS-CC UNT 6,4	2		26	331
Classe	SI Inicial	Coeficiente	N. Características	SI Final
5	0,1611	0,003	1	0,0059
4		1 Item de Classe 4,0		
3		1 Item de Classe 3,0		
2	0,294	0,003	1	0
0		1 Item de Classe 0,0		
DS-CC UNT 11,2	18		30	169
Classe	SI Inicial	Coeficiente	N. Características	SI Final
10	0,3857	0,127	17	0,06
8		1 Item de Classe 8,0		
3		1 Item de Classe 3,0		
0	0,5	0,0118	1	0,5
DS-CC UNT 6,6	30		26	255
Classe	SI Inicial	Coeficiente	N. Características	SI Final
5	0,4432	0,0049	1	0,4055
2	0,4375	0,13	29	0,0156
1		1 Item de Classe 1,0		
DS-CC UNT 4,2	474		30	899
Classe	SI Inicial	Coeficiente	N. Características	SI Final
5	0,5335	0,7248	277	0,4435
4		1 Item de Classe 4,0		
3	0,4421	0,0017	1	0,3531
2	0,4502	0,6452	270	0,3445
1	0,5506	0	0	0,5

Continua na próxima página

<i>Dataset</i>	C. Final		Documentos	C. Inicial
0	0,559	0,1201	54	0,5
DS-CC UNT 12,7	47		28	200
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
10	0,3958	0,0974	14	0,1332
9	0,6081	0,34	34	0,25
8	1 Item de Classe 8,0			
DS-CC UNT 11,5	1		30	538
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
10	0,2887	0,002	1	0,0051
8	1 Item de Classe 8,0			
5	1 Item de Classe 5,0			
0	1 Item de Classe 0,0			
DS-CC UNT 3,3	4		31	475
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,2844	0,0023	1	0,26
4	0,371	0,0021	1	0
2	0,4536	0,0028	1	0,0625
0	0,5	0,0021	1	0
DS-CC UNT 3,1	26		31	245
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,4291	0,1376	26	0,0556
4	1 Item de Classe 4,0			
1	0,4009	0	0	0,8194
DS-CC UNT 12,8	176		28	667
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
10	0,3931	0,4601	175	0,25
9	1 Item de Classe 9,0			
8	1 Item de Classe 8,0			
5	0,3766	0,0015	1	0,4651
4	1 Item de Classe 4,0			
0	0,2322	0,0015	1	0,12
DS-CC UNT 11,7	102		30	672

Continua na próxima página

<i>Dataset</i>	C. Final		Documentos	C. Inicial
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
10	0,3129	0,0019	1	0,2555
7			1 Item de Classe 7,0	
6			1 Item de Classe 6,0	
5			1 Item de Classe 5,0	
4	0,4005	0,006	3	0,3917
2	0,5945	0,2946	99	0,25
1	0,5158	0,0022	1	0,3333
DS-CC UNT 2,3	7		30	704
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5			1 Item de Classe 5,0	
4	0,2715	0,0036	2	0,08
3	0,4462	0,0014	1	0,04
2	0,2904	0,0076	4	0,0625
1			1 Item de Classe 1,0	
DS-CC UNT 5,2	5		28	618
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,16	0,0017	1	0,0994
4	0,5	0,0129	4	0,25
DS-CC UNT 11,9	238		30	770
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
10	0,5846	0,7735	228	0,5
8	0,2183	0,0072	4	0,2398
6	0,5579	0,0182	7	0,5435
4			1 Item de Classe 4,0	
2	0,5	0,0013	1	0,5
DS-CC UNT 12,1	154		28	319
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
10	0,3795	0,0042	1	0,4269
9	1	0,2602	83	0
5			1 Item de Classe 5,0	
0	0,1723	0,3651	91	0,1619

Continua na próxima página



<i>Dataset</i>	C. Final		Documentos	C. Inicial
DS-CC UNT 12,9	617		28	977
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
10	0,4437	0,0937	61	0,4556
9	1 Item de Classe 9,0			
8	0,6581	1,4234	427	0,5157
7	0,4239	0,0875	57	0,3795
6	0,469	0,6309	282	0,3403
5	1 Item de Classe 5,0			
4	1 Item de Classe 4,0			
0	0,376	0,001	1	0,2942
DS-CC UNT 1,5	4		29	395
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,3082	0,0038	1	0,1389
4	0,3758	0,0038	1	0,1111
3	0,5	0,0025	1	0,25
2	0,543	0,0025	1	0
DS-CC UNT 11,3	5		30	494
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
10	0,3443	0,002	1	0,04
6	0,5	0,002	1	0,25
4	0,5	0,002	1	0
2	0,3029	0,002	1	0
0	0,3525	0,002	1	0,1111
DS-CC UNT 2,1	2		30	384
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,2664	0	0	0
4	1 Item de Classe 4,0			
3	0,5	0,0026	1	0
1	0,5344	0,0026	1	0
DS-CC UNT 8,7	156		27	685
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,4536	0,3829	139	0,3439

Continua na próxima página

<i>Dataset</i>	C. Final		Documentos	C. Inicial
4	0,5	0,0467	16	0,5
3	0,376	0,0015	1	0,3354
2	0,2331	0,0015	1	0,24
1	0,5	0,0015	1	0,5
0	0,5	0,0015	1	0,5
DS-CC UNT 1,6	154		29	449
Classe	SI Inicial	Coeficiente	N. Características	SI Final
5	0,7788	0,6815	153	0,25
4	0,2796	0,003	1	0,2917
3	0,4432	0,0022	1	0,5351
2	0,2439	0,0022	1	0,2208
1	0,3476	0,0033	1	0,3333
DS-CC UNT 8,5	8		27	560
Classe	SI Inicial	Coeficiente	N. Características	SI Final
5	0,2235	0,002	1	0,0496
3	0,5166	0,0214	6	0,25
2		1 Item de Classe 2,0		
1	0,3948	0,0018	1	0,1111
DS-CC UNT 12,1	1		28	289
Classe	SI Inicial	Coeficiente	N. Características	SI Final
10	0,2272	0,0037	1	0,0044
3		1 Item de Classe 3,0		
2		1 Item de Classe 2,0		
DS-CC UNT 3,2	11		31	281
Classe	SI Inicial	Coeficiente	N. Características	SI Final
5	0,3717	0,0036	1	0,122
4	0,2131	0,0036	1	0,08
3		1 Item de Classe 3,0		
1	0,7107	0,0569	8	0,25
0	0,5225	0,0071	1	0,5
DS-CC UNT 4,1	4	30	619	
Classe	SI Inicial	Coeficiente	N. Características	SI Final

Continua na próxima página

<i>Dataset</i>	C. Final		Documentos	C. Inicial
5	0,2133	0,0018	1	0,02
4	0,3141	0,0016	1	0
3	0,3209	0,0016	1	0
2	0,2566	0,0016	1	0
DS-CC UNT 1,1	101		29	701
Classe	SI Inicial	Coeficiente	N. Características	SI Final
5	0,4806	0,0014	1	0,6808
4	0,312	0,0014	1	0,25
3	0,5118	0,2853	100	0,25
2	0,2665	0,0019	1	0,1875
1	1 Item de Classe 1,0			
DS-CC UNT 4,5	92		30	350
Classe	SI Inicial	Coeficiente	N. Características	SI Final
5	0,5833	0,3814	89	0,1111
4	0,4367	0,0029	1	0,3333
2	0,125	0,0029	1	0,125
1	0,3333	0,0029	1	0,2222
0	1 Item de Classe 0,0			
DS-CC UNT 10,5	253		24	648
Classe	SI Inicial	Coeficiente	N. Características	SI Final
5	0,3199	0,5319	221	0,1511
4	0,6089	0,0015	1	0,5
3	1 Item de Classe 3,0			
2	0,5584	0,0015	1	0,5
1	0,6195	0,1204	39	0,5
DS-CC UNT 5,4	4		28	226
Classe	SI Inicial	Coeficiente	N. Características	SI Final
5	0,2299	0,0054	1	0,0496
4	0,5	0,0044	1	0
3	1 Item de Classe 3,0			
2	0,5	0,0044	1	0
1	0,3934	0,0044	1	0

Continua na próxima página

<i>Dataset</i>	C. Final		Documentos	C. Inicial
DS-CC UNT 9.7	61		27	280
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,2285	0,2832	61	0,0608
1	1 Item de Classe 1,0			
DS-CC UNT 11.4	2		30	375
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
10	0,2945	0,0029	1	0,0083
7	1 Item de Classe 7,0			
4	1 Item de Classe 4,0			
9	1 Item de Classe 9,0			
6	1 Item de Classe 6,0			
2	0,5373	0,0053	1	0,25
0	1 Item de Classe 0,0			
DS-CC UNT 9.6	4		27	169
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,2794	0,0203	3	0,0153
0	0,5	0,0118	1	0,25
DS-CC UNT 11.6	1		30	298
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
10	0,1425	0,0035	1	0,0028
DS-CC UNT 4.6	2		30	107
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,6742	0,0093	1	0
0	0,5249	0,0117	1	0,08
DS-CC UNT 6.7	83		26	283
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,3232	0	0	0,2153
4	0,7	0,5866	83	0,25
DS-CC UNT 1.4	9		29	69
Classe	SI Inicial	Coeficiente	N. Caracteristicas	SI Final
5	0,4254	0,1341	8	0,02
0	0,5	0,029	1	0,25

Continua na próxima página

<i>Dataset</i>	C. Final		Documentos	C. Inicial
DS-CC UNT 9.2	1		27	131
Classe	SI Inicial	Coeficiente	N. Características	SI Final
5	0,3095	0,0082	1	0,0051
2	1 Item de Classe 2,0			

Última página

Tabela C.12: Similaridade interna anterior e posterior à seleção de características em atividades da base de dados com mais de uma classe de nota.

## C.4 Caracterização das Questões Curtas

O tamanho das respostas representa o conteúdo e adiciona dificuldades na avaliação sem conhecimento prévio da base da atividade. Assim, para todas as bases de dados é realizada a descrição das atividades quanto ao tamanho das respostas e da base de dados. Na Tabela C.14 é apresentada a dimensionalidade (número de características x número de documentos) e o tamanho das respostas segundo a média ( $\bar{x}$ ), desvio padrão ( $\sigma$ ) e a maior (máx) e menor (min) norma dos vetores de resposta.

<i>Dataset</i>	N. Documentos	N. Características	Submissões (em palavras)			
			$\bar{x}$	$\sigma$	máx.	min.
VEST Q1	90	2117	76	35	142	1
VEST Q2	89	2101	77	30	162	23
VEST Q3	91	185	35	9	70	14
VEST Q4	90	1051	48	19	134	16
VEST Q5	90	1135	36	20	118	11
FILOSOFIA-20-387	49	195	12	4	19	3
FILOSOFIA-20-388	49	186	12	4	18	4
FILOSOFIA-20-389	46	315	15	3	21	5
FILOSOFIA-20-390	49	344	14	3	18	6
MTPC-13-220	17	65	10	4	15	5
MTPC-13-221	17	87	9	3	13	2
MTPC-13-258	18	262	24	10	50	11
MTPC-13-281	19	158	16	5	22	2
MTPC-13-282	14	434	49	25	112	17
MTPC-13-283	13	540	62	24	111	26
MTPC-13-304	7	47	14	5	20	5
MTPC-13-307	9	407	95	69	211	10
MTPC-13-309	7	286	66	31	135	31
MTPC-13-321	9	367	69	30	121	18
MTPC-13-338	13	80	14	4	20	8
MTPC-13-339	12	96	14	3	19	9
MTPC-13-340	11	86	13	3	17	9

Continua na próxima página

<i>Dataset</i>	N. Documentos	N. Características	Submissões (em palavras)			
			$\bar{x}$	$\sigma$	máx.	min.
MTPC-13-341	10	61	10	2	15	8
MTPC-13-342	9	79	14	3	19	9
MTPC-13-343	10	90	13	4	17	6
MTPC-13-350	14	230	30	25	88	13
MTPC-13-351	13	94	15	5	19	3
MTPC-13-352	13	213	48	28	115	11
MTPC-13-353	17	57	5	6	19	0
MTPC-13-354	17	70	12	5	18	4
MTPC-13-355	16	140	15	4	20	7
MTPC-13-356	17	99	15	2	18	11
MTPC-13-357	17	72	12	4	17	5
MTPC-13-358	16	140	17	2	21	11
MTPC-13-359	15	101	16	4	19	4
MTPC-13-360	17	138	23	10	43	12
MTPC-13-361	14	81	11	5	19	4
MTPC-13-362	16	51	10	3	15	7
MTPC-13-363	15	66	13	4	17	7
MTPC-13-364	15	81	13	5	19	5
MTPC-13-365	15	59	11	4	18	4
MTPC-13-366	14	74	15	2	18	12
MTPC-13-367	15	59	19	4	24	12
MTPC-13-368	14	57	10	4	19	4
MTPC-13-369	15	62	14	4	19	7
MTPC-13-378	17	114	13	4	19	7
MTPC-13-379	17	119	14	3	19	8
MTPC-13-380	17	111	15	3	19	9
MTPC-13-381	16	112	13	5	21	5
MTPC-13-382	15	129	15	4	20	7
MTPC-13-383	17	103	16	4	21	5
MTPC-13-384	16	125	15	4	23	5
MTPC-13-392	13	18	4	2	9	3
MTPC-13-395	17	62	10	3	16	5

Continua na próxima página

<i>Dataset</i>	N. Documentos	N. Características	Submissões (em palavras)			
			$\bar{x}$	$\sigma$	máx.	min.
MTPC-13-397	16	170	16	6	30	10
MTPC-13-405	14	159	17	7	30	10
MTPC-13-407	16	206	19	8	33	10
MTPC-13-410	17	8	2	1	3	1
MTPC-13-411	17	123	16	4	22	3
MTPC-13-412	17	185	33	19	90	14
MTPC-13-413	17	130	13	6	22	3
MTPC-13-414	16	21	4	2	10	2
MTPC-13-415	16	29	5	3	11	2
MTPC-13-416	17	83	14	3	18	5
MTPC-13-417	16	25	5	5	18	2
MTPC-13-418	17	139	17	2	20	11
MTPC-13-419	17	114	35	19	63	7
MTPC-13-420	17	31	3	3	15	2
MTPC-13-468	14	58	8	5	17	1
TEC-II-5-114	22	163	12	5	21	1
TEC-II-5-115	26	148	27	15	58	4
TEC-II-5-116	18	53	25	7	43	3
TEC-II-5-117	19	90	8	4	16	1
TEC-II-5-118	15	48	4	5	15	0
TEC-II-5-129	22	143	14	5	25	4
TEC-II-5-130	22	84	13	4	20	6
TEC-II-5-131	20	105	13	5	20	3
TEC-II-5-132	21	99	7	4	15	2
TEC-II-5-153	29	159	15	4	20	0
TEC-II-5-154	28	148	14	3	18	0
TEC-II-5-155	27	233	14	4	21	4
TEC-II-5-156	28	99	15	4	19	4
TEC-II-5-157	27	132	13	5	19	4
TEC-II-5-158	29	80	12	6	22	2
TEC-II-5-159	29	105	13	3	18	3
TEC-II-5-160	29	124	12	4	19	5

Continua na próxima página



<i>Dataset</i>	N. Documentos	N. Características	Submissões (em palavras)			
			$\bar{x}$	$\sigma$	máx.	min.
TEC-II-5-161	29	92	7	6	19	1
TEC-II-5-162	27	126	9	6	24	1
TEC-II-5-166	25	122	10	5	19	2
TEC-II-5-167	25	116	13	4	18	5
TEC-II-5-168	25	80	9	3	16	3
TEC-II-5-169	25	38	3	4	15	1
TEC-II-5-170	25	57	7	5	17	2
TEC-II-5-177	27	37	6	6	19	2
TEC-II-5-178	27	122	9	6	19	2
TEC-II-5-179	27	111	13	3	18	6
TEC-II-5-180	26	72	11	3	16	6
TEC-II-5-181	25	77	10	7	18	1
TEC-II-5-185	19	46	5	2	9	2
TEC-II-5-186	20	161	12	5	20	3
TEC-II-5-187	21	188	18	4	25	7
TEC-II-5-188	24	125	12	5	19	1
TEC-II-5-189	25	192	12	4	19	5
TEC-II-5-190	26	216	15	4	21	6
TEC-II-5-191	24	154	14	4	21	5
TEC-II-5-192	25	143	13	6	27	3
TEC-II-5-4	44	660	54	49	272	9
TEC-II-5-5	98	857	45	24	162	4
TEC-II-5-6	44	637	54	44	179	6
TEC-II-5-7	48	604	48	35	155	10
TEC-II-5-8	45	243	42	30	115	5
DS-CC UNT 10.1	24	150	22	11	44	4
DS-CC UNT 10.2	24	96	16	9	44	7
DS-CC UNT 10.3	24	58	7	5	21	3
DS-CC UNT 10.4	24	89	18	11	41	7
DS-CC UNT 10.5	24	149	32	12	59	9
DS-CC UNT 10.6	24	130	19	14	58	1
DS-CC UNT 10.7	24	66	7	7	23	1

Continua na próxima página

<i>Dataset</i>	N. Documentos	N. Características	Submissões (em palavras)			
			$\bar{x}$	$\sigma$	máx.	min.
DS-CC UNT 1.1	29	226	21	12	58	8
DS-CC UNT 11.1	30	91	10	7	27	2
DS-CC UNT 11.10	30	196	26	10	47	4
DS-CC UNT 11.2	30	58	7	7	27	1
DS-CC UNT 11.3	30	155	15	10	46	2
DS-CC UNT 11.5	30	159	18	11	60	2
DS-CC UNT 11.6	30	102	9	8	29	1
DS-CC UNT 11.7	30	185	21	9	42	7
DS-CC UNT 11.8	30	188	22	10	50	10
DS-CC UNT 11.9	30	208	23	9	50	11
DS-CC UNT 1.2	29	162	14	14	55	1
DS-CC UNT 12.1	28	76	12	5	27	5
DS-CC UNT 12.10	28	115	13	17	79	1
DS-CC UNT 12.2	28	195	18	12	52	1
DS-CC UNT 12.3	28	45	5	2	10	1
DS-CC UNT 12.4	28	217	31	12	66	14
DS-CC UNT 12.5	28	172	18	14	76	3
DS-CC UNT 12.6	28	169	24	10	41	4
DS-CC UNT 12.7	28	65	9	9	27	1
DS-CC UNT 12.8	28	181	23	15	80	2
DS-CC UNT 12.9	28	220	38	21	84	2
DS-CC UNT 1.3	29	282	26	26	114	4
DS-CC UNT 1.4	29	33	5	3	14	1
DS-CC UNT 1.5	29	122	15	8	35	5
DS-CC UNT 1.6	29	136	18	12	55	3
DS-CC UNT 1.7	29	164	29	15	60	1
DS-CC UNT 2.1	30	132	12	10	46	2
DS-CC UNT 2.2	30	156	26	10	51	8
DS-CC UNT 2.3	30	189	21	12	54	4
DS-CC UNT 2.4	30	88	9	5	28	2
DS-CC UNT 2.5	30	100	8	9	48	0
DS-CC UNT 2.6	30	223	27	26	164	7

Continua na próxima página

<i>Dataset</i>	N. Documentos	N. Características	Submissões (em palavras)			
			$\bar{x}$	$\sigma$	máx.	min.
DS-CC UNT 2.7	30	177	15	9	36	0
DS-CC UNT 3.1	31	76	13	7	30	5
DS-CC UNT 3.2	31	89	12	10	46	2
DS-CC UNT 3.3	31	156	19	11	50	2
DS-CC UNT 3.4	31	159	17	9	41	7
DS-CC UNT 3.5	31	110	14	11	60	1
DS-CC UNT 3.6	31	152	19	10	40	5
DS-CC UNT 3.7	31	201	21	14	65	2
DS-CC UNT 4.1	30	184	19	14	61	3
DS-CC UNT 4.2	30	227	27	16	79	11
DS-CC UNT 4.3	30	157	15	17	79	2
DS-CC UNT 4.4	30	198	24	14	55	2
DS-CC UNT 4.5	30	126	12	11	51	1
DS-CC UNT 4.6	30	56	4	6	26	1
DS-CC UNT 4.7	30	68	6	11	47	1
DS-CC UNT 5.1	28	170	24	8	45	11
DS-CC UNT 5.2	28	133	25	7	39	14
DS-CC UNT 5.3	28	144	18	13	50	1
DS-CC UNT 5.4	28	85	10	7	23	1
DS-CC UNT 6.1	26	103	14	9	47	4
DS-CC UNT 6.2	26	60	9	5	26	2
DS-CC UNT 6.3	26	56	11	4	22	3
DS-CC UNT 6.4	26	119	13	11	37	1
DS-CC UNT 6.5	26	65	10	3	20	5
DS-CC UNT 6.6	26	85	17	8	36	3
DS-CC UNT 6.7	26	86	14	8	33	3
DS-CC UNT 7.1	26	127	22	9	42	8
DS-CC UNT 7.2	26	153	18	14	80	2
DS-CC UNT 7.3	26	187	23	15	76	7
DS-CC UNT 7.4	26	80	11	6	22	1
DS-CC UNT 7.5	26	119	22	13	76	7
DS-CC UNT 7.6	26	133	15	7	32	5

Continua na próxima página

<i>Dataset</i>	N. Documentos	N. Características	Submissões (em palavras)			
			$\bar{x}$	$\sigma$	máx.	min.
DS-CC UNT 7.7	26	133	16	7	35	2
DS-CC UNT 8.1	27	129	19	9	39	7
DS-CC UNT 8.2	27	57	8	7	27	3
DS-CC UNT 8.3	27	182	26	13	61	11
DS-CC UNT 8.4	27	178	26	20	111	2
DS-CC UNT 8.5	27	172	18	7	38	8
DS-CC UNT 8.6	27	185	21	28	142	3
DS-CC UNT 8.7	27	178	25	21	86	1
DS-CC UNT 9.1	27	155	20	11	50	5
DS-CC UNT 9.2	27	42	7	7	28	1
DS-CC UNT 9.3	27	205	26	12	55	4
DS-CC UNT 9.4	27	210	29	25	138	10
DS-CC UNT 9.5	27	178	17	9	43	5
DS-CC UNT 9.6	27	68	6	7	26	1
DS-CC UNT 9.7	27	102	10	10	38	1
KAGGLE 1	1672	2152	24	11	71	0
KAGGLE 10	1640	2187	22	16	127	1
KAGGLE 2	1278	1966	29	11	80	0
KAGGLE 3	1891	2110	26	8	71	2
KAGGLE 4	1738	2679	21	8	59	3
KAGGLE 5	1795	2591	14	11	142	1
KAGGLE 6	1797	2725	14	12	89	1
KAGGLE 7	1799	2601	20	12	79	1
KAGGLE 8	1799	2689	28	18	147	1
KAGGLE 9	1798	3070	26	19	160	1

Última página

Tabela C.14: Tamanho dos documentos para cada atividade de todas as bases de dados após a etapa de pré-processamento.