

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

**EDUARDO MARIM PISSINATI**

**UMA PROPOSTA DE CORREÇÃO SEMIAUTOMÁTICA DE  
QUESTÕES DISCURSIVAS E DE VISUALIZAÇÃO DE  
ATIVIDADES PARA APOIO À ATUAÇÃO DO DOCENTE**

**VITÓRIA  
2014**

**EDUARDO MARIM PISSINATI**

**Dissertação de MESTRADO - 2014**

**EDUARDO MARIM PISSINATI**

**UMA PROPOSTA DE CORREÇÃO SEMIAUTOMÁTICA DE  
QUESTÕES DISCURSIVAS E DE VISUALIZAÇÃO DE  
ATIVIDADES PARA APOIO À ATUAÇÃO DO DOCENTE**

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

Orientador: Prof. Dr. Elias de Oliveira.

Co-orientador: Profa. Dra. Márcia Gonçalves de Oliveira.

VITÓRIA

2014

**EDUARDO MARIM PISSINATI**

**UMA PROPOSTA DE CORREÇÃO SEMIAUTOMÁTICA DE  
QUESTÕES DISCURSIVAS E DE VISUALIZAÇÃO DE  
ATIVIDADES PARA APOIO À ATUAÇÃO DO DOCENTE**

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 25 de Setembro de 2014.

**COMISSÃO EXAMINADORA**

---

Prof. Dr. Elias de Oliveira  
Universidade Federal do Espírito Santo  
Orientador

---

Profa. Dra. Márcia Gonçalves de Oliveira  
Universidade Federal do Espírito Santo  
Co-orientador

---

Prof. Dr. Edilson de Aguiar  
Universidade Federal do Espírito Santo

---

Profa. Dra. Isaura Alcina Martins Nobre  
Instituto Federal do Espírito Santo

*Dedico este trabalho especialmente ao meu filho Mateus, aos outros que virão e aos meus  
sobrinhos, para que sirva de motivação e inspiração em seus estudos.*

*E também a minha amada esposa que me deu todo apoio nesse período de estudo.*

## **Agradecimentos**

Agradeço a Deus.

Agradeço a minha esposa Fernanda por todo apoio e incentivo nesse período de estudo.

Agradeço aos meus pais Maria Catarina e José Eduardo por terem me incentivado a estudar sempre.

Agradeço à toda minha família pelo apoio e incentivo.

Agradeço ao Prof. Elias de Oliveira por toda orientação e conhecimento adquirido.

Minha gratidão à Prof.<sup>a</sup> Márcia Gonçalves de Oliveira pelas orientações, revisões e dicas sobre o trabalho.

Agradeço aos meus colegas Marcos e Frederico, por toda ajuda nesse período.

Agradeço a todos professores das disciplinas que cursei.

Minha gratidão aos meus companheiros de trabalho do PRODEST e do Tribunal de Justiça, por seu constante apoio e incentivo.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>16</b>
1.1	Motivação . . . . .	17
1.2	Objetivo . . . . .	18
1.3	Metodologia do Trabalho . . . . .	18
1.4	Estrutura desta Dissertação . . . . .	19
<b>2</b>	<b>Revisão da Literatura</b>	<b>21</b>
2.1	Avaliação de questões discursivas apoiadas por computador . . . . .	21
2.2	Visualização de informação no contexto educacional . . . . .	24
<b>3</b>	<b>Conceitos Básicos</b>	<b>26</b>
3.1	Pré-Processamento de Textos . . . . .	26
3.2	Representação de Documentos no Modelo Vetorial . . . . .	28
3.3	Cálculo de Similaridade entre Documentos . . . . .	28
3.4	<i>Clustering</i> . . . . .	28
3.5	O Algoritmo KNN . . . . .	29
3.6	Treemap . . . . .	30

<b>4</b>	<b>Plugin de Integração com Moodle</b>	<b>32</b>
4.1	Instalação . . . . .	33
4.1.1	Etapa 1: Criar usuário do <i>webservice</i> de download de arquivos . . .	33
4.1.2	Etapa 2: Instalação dos <i>webservices</i> de download de respostas e upload de resultados . . . . .	37
4.1.3	Etapa 3: Configuração do servidor remoto - Uso do <i>plugin</i> . . . . .	37
<b>5</b>	<b>Correção Semiautomática</b>	<b>38</b>
5.1	Preditor Semiautomático de notas para questões discursivas . . . . .	38
5.1.1	Funcionamento . . . . .	39
5.2	Predição Semiautomática de notas para códigos de programação . . . . .	46
5.2.1	Funcionamento . . . . .	46
5.2.2	Aplicação e Resultados . . . . .	51
<b>6</b>	<b>Visualização de Informação</b>	<b>53</b>
6.1	Estudo de caso . . . . .	53
6.2	Escolha da Técnica de Visualização de Informação . . . . .	58
6.3	Arquitetura da Ferramenta de Visualização . . . . .	59
<b>7</b>	<b>Metodologia Experimental</b>	<b>62</b>
7.1	Bases de Dados . . . . .	62
7.1.1	Questão 1 . . . . .	63
7.1.2	Questão 2 . . . . .	64
7.1.3	Questão 3 . . . . .	66

7.1.4	Questão 4 . . . . .	66
7.1.5	Questão 5 . . . . .	68
7.2	Métricas . . . . .	69
7.3	Avaliadores Humanos . . . . .	70
7.4	Ponderação de notas . . . . .	71
7.5	Documento representante de um grupo . . . . .	72
<b>8</b>	<b>Experimentos</b>	<b>73</b>
8.1	Experimento 1 . . . . .	73
8.1.1	Discussões . . . . .	75
8.2	Experimento 2 . . . . .	77
8.2.1	Discussões . . . . .	78
8.3	Experimento 3 . . . . .	80
8.4	Experimento 4 . . . . .	81
8.5	Experimento 5 . . . . .	83
8.5.1	Discussões . . . . .	84
8.6	Experimento 6 . . . . .	86
8.7	Experimento 7 . . . . .	89
<b>9</b>	<b>Conclusões e Trabalhos Futuros</b>	<b>94</b>
9.1	Conclusão . . . . .	94
9.2	Trabalhos futuros . . . . .	95
<b>A</b>	<b>Respostas discursivas do grupo 18 do Experimento 1</b>	<b>100</b>

# Lista de Tabelas

6.1	Tabela associando o percentual de acertos às cores. . . . .	55
6.2	Comparativo entre as possíveis técnicas de visualização adequadas a proposta.	59
7.1	Tabela apresentando o resumo da base de dados com respeito a quantidade de palavras nas respostas dadas pelos alunos. . . . .	63
7.2	Resumo das respostas para a Questão 1. . . . .	64
7.3	Resumo das respostas da Questão 2. . . . .	66
7.4	Resumo das respostas da Questão 4. . . . .	68
7.5	Resumo das respostas da Questão 5. . . . .	69
7.6	Comparação entre os avaliadores. . . . .	70
8.1	Resultados do Experimento 1 para o Avaliador1. . . . .	74
8.2	Resultados do Experimento 1 para o Avaliador2. . . . .	75
8.3	Resultados do Experimento 2 para o Avaliador1. . . . .	77
8.4	Resultados do Experimento 2 para o Avaliador2. . . . .	78
8.5	Análise do grupo 15 da questão Q4. . . . .	79
8.6	Análise do grupo 16 da questão Q4. . . . .	80
8.7	Resultados do Experimento 3 para o Avaliador1. . . . .	81

8.8	Resultados Experimento 3 para o Avaliador2. . . . .	81
8.9	Resultado do Experimento 4 para o Avaliador1. . . . .	82
8.10	Resultado do Experimento 4 para o Avaliador2. . . . .	82
8.11	Resultado do Experimento 5 para o Avaliador1, n = 2. . . . .	83
8.12	Resultados do Experimento 5 para o Avaliador2, n = 2. . . . .	83
8.13	Resultado do Experimento 5 para o Avaliador1, n = 3. . . . .	84
8.14	Resultados do Experimento 5 para o Avaliador2, n = 3. . . . .	84
8.15	Resultados da questão Q3 utilizando bigrama, sem a retirada de <i>stopwords</i> e sem o uso de <i>stemming</i> . . . . .	85
8.16	Resultados da questão 3 utilizando trigramas, sem a retirada de <i>stopwords</i> e sem o uso de <i>stemming</i> . . . . .	86
8.17	Resultados do Experimento 6 para o Avaliador1. . . . .	88
8.18	Resultados do Experimento 6 para o Avaliador2. . . . .	88
8.19	Resultados do Experimento 7 para o Avaliador1. . . . .	91
8.20	Resultados do Experimento 7 para o Avaliador2. . . . .	92
8.21	Resultados do Experimento 7 para o Avaliador1 comparado ao objetivo. . .	92
8.22	Resultados do Experimento 7 para o Avaliador2 comparado ao objetivo. . .	93
A.1	Legenda das respostas do grupo 18 formado no Experimento 1. . . . .	100

# Lista de Figuras

3.1	<i>Treemap</i> representando a quantidade de habitantes das Unidades Federativas do Brasil. . . . .	30
4.1	Arquitetura da solução com o <i>plugin</i> . . . . .	32
4.2	Cadastrar usuário para o <i>webservice</i> . . . . .	34
4.3	Cadastrar uma nova <i>role</i> , tela 1/2. . . . .	35
4.4	Cadastrar uma nova <i>role</i> , tela 2/2. . . . .	35
4.5	Vincular Usuário à <i>role</i> . . . . .	36
5.1	Cadastro da questão no <i>Moodle</i> . . . . .	40
5.2	Cadastro da questão no <i>Moodle</i> . . . . .	40
5.3	Cadastro da resposta no <i>Moodle</i> . . . . .	41
5.4	Conteúdo da pasta da questão. . . . .	41
5.5	Arquivo <i>intervalonotas.csv</i> . . . . .	42
5.6	Arquivo <i>notastreino.csv</i> . . . . .	42
5.7	Arquivo <i>resultado.csv</i> . . . . .	44
5.8	Respostas marcadas no <i>Moodle</i> para correção do professor. . . . .	44
5.9	Arquivo <i>notastreino.csv</i> com as notas do professor. . . . .	45

5.10	Arquivo <i>resultado.csv</i> com as notas preditas. . . . .	45
5.11	Cadastro da questão no <i>Moodle</i> . . . . .	47
5.12	Cadastro da questão no <i>Moodle</i> . . . . .	48
5.13	Cadastro da resposta no <i>Moodle</i> . . . . .	49
5.14	Conteúdo da pasta da questão. . . . .	49
5.15	Arquivo <i>intervalonotas.csv</i> . . . . .	50
5.16	Arquivo <i>notastreino.csv</i> . . . . .	50
6.1	<i>Treemap</i> representando a visão das várias turmas de Biblioteconomia da UFES. . . . .	55
6.2	Detalhes que são exibidos ao passar o cursor sobre os retângulos que representam os alunos. . . . .	56
6.3	Visão detalhada da TURMA1. . . . .	57
6.4	Arquitetura da Solução. . . . .	60
6.5	Método <i>definirCor</i> do Módulo de Tratamento de Dados. . . . .	60
8.1	Variação do Fator de Ponderação. . . . .	87
8.2	Variação do número de grupos. . . . .	89
8.3	Variação do Fator de Ponderação. . . . .	91

# Resumo

A atividade de correção de textos discursivos exige dos professores muito esforço e dedicação. Isso se intensifica quanto maior for o número de alunos, até chegar ao ponto em que o professor deixa de aplicar essas atividades por não conseguir avaliá-las, submetendo seus alunos somente a questões objetivas. Além disso, a identificação de alunos que precisam de mais atenção fica comprometida em turmas muito numerosas. Aumentando o número de professores por turma, esses problemas poderiam ser resolvidos, porém a contratação de mais professores representa mais custos para uma instituição. Nesse contexto o uso do computador pode auxiliar nessas atividades sem incluir novos docentes. No entanto essa estratégia deve ser muito bem pensada para não sobrecarregar ainda mais o professor com ferramentas inadequadas. O presente trabalho apresenta uma ferramenta de correção Semiautomática de textos discursivos e uma ferramenta de Visualização que permitirá ao professor acompanhar individualmente os desempenhos de seus alunos mesmo em turmas numerosas. A ferramenta de correção Semiautomática de textos usa técnicas de recuperação inteligente da informação como o agrupamento de textos semelhantes para correção de respostas discursivas. De outro lado, a ferramenta de Visualização usa técnicas de Visualização de Informação para apresentar os resultados ao docente. Ambas foram integradas ao ambiente Virtual de Aprendizagem Moodle de forma que não se alterasse a rotina normal de trabalho daqueles que utilizam esse ambiente.

**Palavras-chave:** Avaliação Semiautomática, Questões Discursivas, Visualização de Informação.

# Abstract

The task of correcting discursive texts requires much effort and dedication from teachers. The larger the number of students, the bigger the work is, until it reaches a point where the teacher fails to apply these activities since he cannot evaluate them, subjecting his students only to objective questions. Furthermore, the identification of students who need more attention is impaired in very large groups. Increasing the number of teachers per class could solve these problems, but more teachers represent more costs to the institution. In this context, the use of the computer can assist in these activities without including new teachers. However this strategy must be well thought not to overload the teacher with inadequate tools. This paper presents a tool for semi-automatic correction of discursive texts and a visualization tool that will allow the teacher to monitor individual performances of their students even in large classes. The semi-automatic correction tool uses intelligent recovering techniques of the information as clustering to correct discursive answers. Besides that, the visualization tool uses information visualization techniques to present the results to the teacher. Both were integrated with Moodle Virtual Learning Environment in order not to change the normal routine of work of those who use it.

Keywords: Semi-automatic Assessment, Discursive Questions, Information Visualization.

# Capítulo 1

## Introdução

O excesso de atividades, associado à cobrança por produtividade, acaba sobrecarregando o educador. Uma pesquisa realizada na Universidade Federal do Espírito Santo (UFES) constatou que, em média, os professores trabalham 18,6 horas semanais em casa, além das 40 regimentais semanais (BORSOI, 2012). Todo esse acúmulo de funções tem causado sérios problemas de saúde aos educadores (CARLOTTO, 2002).

Essa sobrecarga não afeta somente a saúde do docente, mas também compromete a qualidade do ensino. Isso é intensificado quando ele se depara com turmas numerosas, pois fica inviável identificar as necessidades individuais de perfil a perfil de aluno para atuar de forma diferenciada com cada um deles. Dessa forma, o número de atividades desempenhadas pelos alunos fica limitado à quantidade que o professor consegue avaliar.

De modo geral os problemas de sobrecarga poderiam ser resolvidos com o aumento do número de professores por turma, uma vez que suas atividades poderiam ser divididas. Porém, esse aumento gera mais custos para uma instituição de ensino.

O uso do computador pode auxiliar na redução dessa sobrecarga, porém, seu uso inadequado compromete o ensino e desmotiva a utilização desses recursos em novas atividades educacionais (KENSKI, 2003).

Além da sobrecarga nas atividades diárias dos educadores, existe também a sobrecarga em trabalhos específicos, como a avaliação de questões discursivas de concursos e vestibulos.

lares, que possuem muitas correções a serem feitas por avaliador, que geram divergências entre as correções de avaliadores.

No vestibular da Universidade Federal do Espírito Santo (UFES), por exemplo, em uma das questões discursivas de Português, pudemos constatar, através dos experimentos realizados nessa dissertação, que houve uma divergência média entre os avaliadores de 2.33 pontos por resposta, sendo que em algumas respostas houve mais de 10 pontos de divergência, considerando que a escala é de 0 a 10, a divergência foi máxima, conforme levantamento feito na Seção 7.3.

De acordo com o trabalho de Luna (2012) os avaliadores do ENEM entrevistados relataram que, após corrigirem um determinado número de redações, a qualidade no trabalho não era mais a mesma. Isso mostra que podem existir diferenças nos critérios de análise de um mesmo avaliador. Eles apontavam como fator determinante para esse problema o número de redações enviadas por corretor por dia, que no ENEM de 2007 foram 102 textos por dia por avaliador num período de um mês.

## 1.1 Motivação

A grande motivação deste trabalho é poder reduzir o esforço e o tempo gasto pelo professor em determinadas tarefas dando a ele mais disponibilidade que poderiam ser dedicados para melhor qualidade de aulas.

O uso de questões subjetivas na educação dos estudantes é necessária uma vez que os exercícios objetivos não são suficientes para sua formação (SAXENA; GUPTA, 2009). Porém, o esforço de correção dessas atividades subjetivas é muito maior, limitando assim sua aplicação.

O uso do computador se torna adequado nesse tipo de situação pois com ele é possível reduzir o esforço e o tempo que o professor demanda nesse tipo de atividade em quase 50%. Essa redução no esforço será um incentivo à utilização das questões discursivas no ensino.

Não só o tempo e o esforço são melhorados, mas também os critérios de correção, que

são mantidos quando se usa o computador. Além disso, quando a atividade de correção de discursivas é apoiada pela tecnologia é possível analisar mais questões do que quando realizada manualmente, tornando o processo escalável.

Além do auxílio na correção de atividades discursivas, também é necessário auxiliar o docente na visualização e análise dos resultados. Isso deve ser feito pois o docente precisará ter conhecimento dos resultados das correções automáticas. E, quando em turmas numerosas, esse auxílio será necessário para todas atividades.

Em resumo essas motivações se unem para formar uma ferramenta que auxiliará o docente em suas atividades de ensino.

## 1.2 Objetivo

O objetivo deste trabalho é propor uma ferramenta de predição Semiautomática de notas para respostas discursivas e uma ferramenta de visualização de informação capaz de apresentar o desempenho de vários alunos em várias atividades de forma ágil. Essas ferramentas seriam então combinadas e integradas ao ambiente de aprendizagem *Moodle*.

Nos experimentos realizados a ferramenta de predição Semiautomática de notas para respostas discursivas obteve nível de erro semelhante ou inferior quando comparado as diferenças entre humanos.

## 1.3 Metodologia do Trabalho

O primeiro passo da nossa metodologia foi realizar a revisão da literatura com o objetivo de identificar os estudos recentes relacionados ao tema, buscando resgatar o histórico e compreender os avanços realizados ao longo do tempo.

A partir dessa revisão, foram analisadas as técnicas empregadas relacionadas tanto à área de predição de notas quanto à área de visualização de informação. Em seguida, foi criada

uma estratégia para tratamento e processamento dos textos de respostas discursivas do qual foi originado o algoritmo de predição Semiautomática de notas para respostas discursivas de tamanho médio.

Esse algoritmo foi testado com a base de dados de respostas discursivas do vestibular da UFES de 2012 para a disciplina de Português. Essas respostas foram transcritas manualmente e, para cada resposta, foi obtida as notas oficiais dadas pelos avaliadores. Em seguida, essas notas foram comparadas com as notas preditas do algoritmo.

Ao final foi desenvolvido um *plugin* para integrar as ferramentas de correção Semiautomática e de visualização ao ambiente virtual de aprendizagem *Moodle*, que agora funcionam juntamente com as atividades presentes nesse ambiente.

## 1.4 Estrutura desta Dissertação

Neste capítulo, são apresentadas as principais ideias desta dissertação, descrevendo o cenário com seus problemas, suas motivações, e a metodologia proposta. Além desta introdução, este texto é composto pelos seguintes capítulos:

- **Capítulo 2 (Revisão da Literatura):** realiza uma revisão da literatura identificando os avanços obtidos na área até o período recente.
- **Capítulo 3 (Conceitos Básicos):** apresenta os conceitos necessários para o bom entendimento do trabalho.
- **Capítulo 4 (Plugin de Integração com Moodle):** descreve o Plugin desenvolvido para permitir a integração das ferramentas digitais com o *Moodle*.
- **Capítulo 5 (Correção Semiautomática):** descreve o algoritmo proposto para correção Semiautomática de respostas discursivas e apresenta uma proposta para correção Semiautomática de códigos de programação.
- **Capítulo 6 (Visualização de Informação):** detalha a ferramenta de visualização e apresenta um estudo de caso.

- 
- **Capítulo 7 (Metodologia Experimental):** descreve as bases utilizadas nos experimentos, os avaliadores e técnicas utilizadas nos experimentos.
  - **Capítulo 8 (Experimentos e Resultados):** apresenta os experimentos realizados e faz análises sobre os resultados.
  - **Capítulo 9 (Conclusões e Trabalhos Futuros):** apresenta as considerações finais, as contribuições e propostas futuras para aprimoramento do trabalho.

# Capítulo 2

## Revisão da Literatura

Neste capítulo, será realizada uma análise de outros trabalhos similares que serviram de base para o presente trabalho. Para isso, serão apresentadas pesquisas na área de avaliação de questões discursivas apoiadas por computador e pesquisas na área de visualização de informação no contexto educacional.

### 2.1 Avaliação de questões discursivas apoiadas por computador

O estudo da correção de respostas discursivas apoiada por computador vem de meados dos anos 60 com o sistema PEG (*Project Essay Grader*). Proposto por Page (1967), esse sistema foi desenvolvido com o objetivo de avaliar pequenas questões discursivas. Ele baseia sua análise na qualidade de escrita do texto, sem levar em consideração o seu conteúdo. Ou seja, se um aluno fugir ao tema de uma pergunta, por exemplo, mas escrever corretamente, esse sistema consideraria como um texto correto.

Com o surgimento de novas tecnologias de Extração de Informação (EI) e Processamento de Linguagem Natural (PLN) a partir dos anos 90, a área se expandiu e novos trabalhos surgiram, como o *E-Rater* (BURSTEIN et al., 1998; BURSTEIN; LEACOCK; SWARTZ, 2001), que combina técnicas estatísticas e de PLN na análise das respostas discursivas.

Dentre os trabalhos mais recentes, Santos et al. (2007) apresentou acurácia média de 82.10% ao usar a técnica de N-grama associada a retirada de *Stopwords* e *Stemming* (essas técnicas serão abordadas na seção 3.1). Seus testes foram realizados sobre uma base de 4000 respostas discursivas conceituais da disciplina de Introdução a Banco de Dados. Essas respostas estavam divididas em 130 questões diferentes e possuíam de 15 até 30 palavras.

O trabalho de Wang e Brown (2007) em comum com este realiza uma comparação entre a predição humana e a predição da máquina. Ele utiliza como base os dados X, e o indicador Y. Utilizando a técnica Z. Nos testes os resultados dos algoritmos foram superior aos dos humanos atingindo uma taxa de acerto ou ... de XX

O trabalho de Caldas e Favero (2009) usa avaliação automática para mapas conceituais, que são descritos pelos alunos em formato discursivo: *conceito-link-conceito*. Seus experimentos, com o emprego de n-gramas numa base com 400 mapas conceituais, obtiveram uma taxa de acerto de 90% em relação aos avaliadores humanos.

O trabalho de Oliveira et al. (2010) usa técnicas de clusterização para agrupar respostas similares. Isso faz com que o professor corrija somente algumas respostas, reduzindo assim seu esforço. A precisão da proposta foi de 90.4% para seis notas distintas possíveis. A base de dados usada foi de 7000 respostas com até cinco palavras cada uma.

O trabalho de Ávila e Soares (2012) usa algoritmos de comparação e busca textual para correção de questões dissertativas. No trabalho foi usada uma base com 30 respostas divididas em três questões diferentes. Sua análise é realizada a partir das similaridades entre as respostas dos alunos com a resposta modelo ou gabarito. Esse método exige que o professor crie esses modelos, o que acaba restringindo as possibilidades de resposta.

O trabalho de Vilela et al. (2012), por sua vez, apresenta um sistema especialista denominado *SCATeDi*. Esse sistema utiliza Lógica Fuzzy associado com dicionários de sinônimos para comparação dos termos das respostas dos alunos com os termos da resposta modelo criada pelo professor. Essa técnica foi usada para correção automática de questões discursivas da disciplina de geografia da EJA (Educação de Jovens e Adultos). Os resultados de suas predições alcançaram erros de 0.076 a 0.506 pontos por resposta em uma base de quinze respostas divididas em cinco questões diferentes.

A métrica usada no trabalho de Vilela et al. (2012) é baseada na diferença entre a pontuação do avaliador e a pontuação do método proposto. Essa diferença é computada para cada resposta e ao final é gerada a média de erros em pontos por questão. Neste trabalho foi escolhido a mesma métrica para avaliação dos resultados, pois com ela a análise dos resultados ficou mais precisa do que as medidas de Acurácia e Percentual de acertos. Essa discussão foi aprofundada na Seção 7.2.

A pesquisa de Bridgeman, Trapani e Attali (2012), como o presente trabalho, compara o desempenho de avaliadores humanos ao desempenho de um avaliador automático. O estudo mostrou que em média a diferença entre os avaliações humanos e os automáticos foi pequena, mas para determinadas tipos de etnias essa diferença foi maior. Indicando assim o uso do avaliador automático como fiscalizador do processo de correção manual.

Essa pesquisa foi realizada sobre as bases de testes do *TOEFL (Test of English as a Foreign Language)* e do *GRE (Graduate Record Examination)*. O desempenho dos avaliadores humanos foi comparado a uma versão mais atual do software *e-rater* (ATTALI; BURSTEIN, 2006), que possui ferramentas para processamento de linguagem natural, analisando gramática, mecânica, organização, vocabulário e outros aspectos textuais.

O presente trabalho compartilha do mesmo interesse dos trabalhos apresentados nessa seção, que é propor métodos que auxiliem o trabalho docente na correção Semiautomática de respostas discursivas. Utilizando técnicas presentes nos trabalhos aqui apresentados.

Entretanto, nosso trabalho, ao contrário dos demais relatados nessa seção, buscou atuar sobre respostas discursivas com mais palavras, com média de 90 palavras por resposta, sendo que algumas respostas possuem mais de 250 palavras. Com um número maior de palavras é possível validar a proposta sobre uma variabilidade maior nas respostas.

Além disso, os experimentos realizados neste trabalho foram feitos com uma base de respostas do vestibular, com notas dadas por avaliadores oficiais. Ao contrário dos trabalhos apresentados nesta Seção, que usaram respostas de determinados disciplinas de sala de aula. Podendo o presente trabalho apresentar resultados mais consistentes com a realidade.

## 2.2 Visualização de informação no contexto educacional

A área de Visualização de Informação é uma grande aliada na ampliação da percepção do acompanhamento da aprendizagem dos alunos. Ela facilita o caminho entre o usuário e a informação e faz com que o esforço cognitivo na análise das informações seja reduzido, acelerando assim a assimilação das informações. Ela dispõe de uma série de mecanismos voltados para apresentar as informações da melhor maneira possível.

Visualização de Informação pode ser definida como um processo assistido por computador visando revelar informações sobre um fenômeno abstrato, transformando dados abstratos em informações visuais e espaciais. A intenção da Visualização de Informação é otimizar o uso da nossa capacidade de percepção visual e de pensamento para lidar com fenômenos que não observamos facilmente a olho nu (CHEN, 2002).

Nos últimos anos, vários trabalhos têm utilizado as técnicas de Visualização de Informação no contexto educacional. Trabalhando sobre bases de dados de ambientes de Educação a Distância, Silva (2006), por exemplo, propõe a criação de gráficos dinâmicos baseados em combinações e escolhas de atributos. O usuário define o que quer visualizar sem o necessário conhecimento especializado da estrutura da base de dados. Esse modelo se opõe ao de gráficos estáticos, no qual os atributos são predefinidos pelo sistema e o usuário não tem a possibilidade de modificá-los.

A ferramenta *CourseViewer*, proposta por Silva, Inoue e Mendonça (2012), utiliza técnicas de Visualização de Informação para representar com grafos interativos as relações e dependências entre as disciplinas das grades dos cursos da universidade. Ela permite que seus usuários (aluno, professor ou coordenador de curso) tenham a percepção de quanto uma disciplina influenciará no andamento do curso como um todo e apoiará o aluno na escolha de quais disciplinas cursar.

O trabalho de Moissa, Gasparini e Kemczinski (2013) propõe uma ferramenta de Visualização de Informação para facilitar a análise dos dados navegacionais coletados pelo *AdaptWeb* (Ambiente de Ensino-Aprendizagem Adaptativo na Web). A expectativa da ferramenta proposta é que o professor consiga compreender e se adaptar às necessidades de aprendizagem de seus alunos.

---

Um estudo de viabilidade de duas ferramentas de visualização, Ggobi (SWAYNE et al., 2003) e Hoffman (HOFFMAN, 1999), foi realizado por (WEBBER et al., 2013). Nesse estudo, ele utiliza os dados educacionais publicamente disponíveis, coletados através do *software Geometry Tutor*, e mostra a utilização daquelas ferramentas para aumentar a capacidade de percepção do professor durante suas atividades.

Os trabalhos apresentados acima propõem-se a auxiliar o trabalho acadêmico com o uso de recursos visuais. Seguindo a mesma linha, este trabalho apresenta uma alternativa de Visualização de Informação que auxilia o professor em suas atividades diárias. A partir dessa análise, o professor poderá realizar as intervenções necessárias para melhorar a aprendizagem de seus alunos.

# Capítulo 3

## Conceitos Básicos

Neste capítulo, será apresentado o pré-processamento de textos, a representação de documentos, a similaridade entre documentos, a técnica de *clustering*, além de outros conceitos necessários para o bom entendimento do trabalho.

### 3.1 Pré-Processamento de Textos

O pré-processamento de documentos, de acordo com Baeza-Yates, Ribeiro-Neto et al. (2011) pode ser dividido em cinco operações ou transformações:

1. Análise Lexical;
2. Eliminação de *Stopwords*;
3. Remoção de afixos (*Stemming*);
4. Seleção de palavras-chave;
5. Construção de estruturas de categorização de termos.

Um dos principais objetivos da Análise Lexical é a identificação das palavras do texto, que neste trabalho serão chamadas de termos. Os termos, por exemplo, são itens representativos de um documento e não qualquer palavra separada por espaço. Números, por exemplo,

poderão estar separados por espaço mas não serão considerados como termos se não forem representativos o bastante para isso.

Ainda nessa etapa de transformação serão tratados hífen, pontuação e as caixas alta e baixa das palavras. Nos experimentos realizados neste trabalho, os hífen e as pontuações foram retirados e os termos foram convertidos para caixa baixa.

Na etapa de eliminação de *stopwords* são retirados os termos sem relevância. Artigos, preposições e conjunções normalmente são tratados como *stopwords*. Além desses, alguns verbos, advérbios e adjetivos também podem ser retiradas.

Na etapa de remoção de afixos ou *Stemming* são retirados os prefixos e sufixos dos termos. Isso permite que as palavras de uma mesma família (cognatas) sejam tratadas como se fossem a mesma. Por exemplo: pedra, pedreira, pedrada e pedrar são cognatas e possuem o radical "pedr" em comum.

Na etapa de seleção de palavras-chave é feita uma representação do documento através de termos relevantes que o diferenciam dos demais documentos. Uma das técnicas é combinar dois ou três substantivos em um mesmo grupo, sendo que esses substantivos, dentro do texto, estão distantes um do outro por um número preestabelecido de termos. O uso dos substantivos como índices se justifica por esses carregarem mais semântica do que verbos, adjetivos e advérbios (BAEZA-YATES; RIBEIRO-NETO et al., 2011). Neste trabalho todos os termos que não foram eliminados nas etapas anteriores serão usados para representar o documento.

Mais de um termo poderá formar termos complexos quando usada a técnica de n-gramas. A técnica de n-grama consiste em formar termos a partir da junção de dois ou mais termos subsequentes.

Na etapa de construção de estruturas de categorização de termos é criado um dicionário de sinônimos que precisa ser previamente gerado de acordo com um domínio específico. Essa etapa não será usada nos experimentos, pois requer conhecimento prévio da base de dados.

## 3.2 Representação de Documentos no Modelo Vetorial

Dentre as formas de representação de documentos, destacamos a representação por modelo vetorial. Proposto por Salton, Wong e Yang (1975), o modelo vetorial representa os documentos como vetores no espaço  $R^n$  (BAEZA-YATES; RIBEIRO-NETO et al., 2011), onde  $n$  representa o número de termos nos documentos considerados e  $d_j$  representa um documento. Assim, cada documento será representado por um vetor de termos.

Cada posição desse vetor representa um termo  $k_i$  e cada termo terá um peso associado  $w_{i,j}$ , que neste trabalho será a frequência desse termo dentro do documento (*Term Frequency - tf*) (SALTON; WONG; YANG, 1975). Dessa forma, a representação de um documento no modelo vetorial será como a Equação 3.1.

$$d_j = (w_{1,j}, w_{2,j}, \dots, w_{i,j}, \dots, w_{n,j}) \quad (3.1)$$

## 3.3 Cálculo de Similaridade entre Documentos

A representação apresentada na Seção 3.2 permite que operações matemáticas possam ser realizadas entre dois documentos. Como cada documento é representado por um vetor, esses dois documentos formam um ângulo. O cálculo do cosseno desse ângulo indicará o quão próximo um vetor está do outro e, conseqüentemente, representará o quão similar um documento é do outro. A Equação 3.2 apresenta o cálculo dessa medida de similaridade.

$$\cos(d_i, d_j) = \frac{d_i d_j}{\|d_i\| \|d_j\|} = \frac{\sum_{k=1}^n w_{k,i} w_{k,j}}{\sqrt{\sum_{k=1}^n (w_{k,i})^2} \sqrt{\sum_{k=1}^n (w_{k,j})^2}} \quad (3.2)$$

## 3.4 Clustering

*Clustering* é uma técnica que separa elementos com características semelhantes em um mesmo grupo ou *cluster*. No caso de documentos, um método de *clustering* textual separa

automaticamente os documentos de uma coleção qualquer em  $K$  *clusters*, de acordo com algum critério predefinido (BAEZA-YATES; RIBEIRO-NETO et al., 2011).

O *Clustering* pode ser aglomerativo (hierárquico) ou particional. O *Clustering* hierárquico forma uma hierarquia de *clusters*, onde cada *cluster* pode conter outros *clusters*, denominados filhos. Essa estrutura forma uma árvore de *clusters* ou dendograma. Nesse tipo de *clustering* não é necessário informar o número de grupos (BAEZA-YATES; RIBEIRO-NETO et al., 2011).

O *Clustering* por particionamento agrupa os elementos em  $k$  *clusters* diferentes. Onde o valor de  $k$  é informado previamente ao algoritmo. Esse tipo de *clustering* não cria uma estrutura hierárquica associando todos *clusters* como no tipo aglomerativo, o que torna o método particional mais rápido em relação ao hierárquico, principalmente com grandes volumes de dados (BAEZA-YATES; RIBEIRO-NETO et al., 2011).

Neste trabalho usamos o *Clustering* por particionamento, pois além de ser mais rápido, o número de grupos é previamente definido. Como neste trabalho o número de grupos corresponde ao número de respostas que o professor deve corrigir, é possível estimar qual o esforço o professor terá. Ou seja, se o número de respostas que o professor deve corrigir for igual a 40% das respostas existentes, seu esforço de correção também será de 40%, por exemplo.

### 3.5 O Algoritmo KNN

O KNN (*k-nearest neighbor* ou k-vizinhos mais próximos) é um algoritmo de classificação que determina a característica de um determinado elemento baseado nas características dos elementos mais próximos a ele (BAEZA-YATES; RIBEIRO-NETO et al., 2011). Para descobrir os elementos mais próximos, esse algoritmo calcula a distância do elemento referência entre todos os demais. Em seguida, são selecionados os  $k$  elementos que possuem a menor distância.

Neste trabalho, para calcular a distância ou similaridade entre os elementos (ou documentos), foi utilizada a similaridade do cosseno detalhada na Seção 3.3.

Na fase de ponderação de notas, que será detalhada na Seção 7.4, esse algoritmo foi utilizado para encontrar as  $k$  respostas mais próximas da resposta da qual se pretende ter sua nota ponderada.

### 3.6 Treemap

A técnica *Treemap* foi criada por Ben Shneiderman e Brian Johnson, em 1991. Ela é um tipo de visualização baseada em áreas onde o tamanho de cada retângulo representa uma medida (YAU, 2011). A Figura 3.1 apresenta um exemplo do *Treemap*.

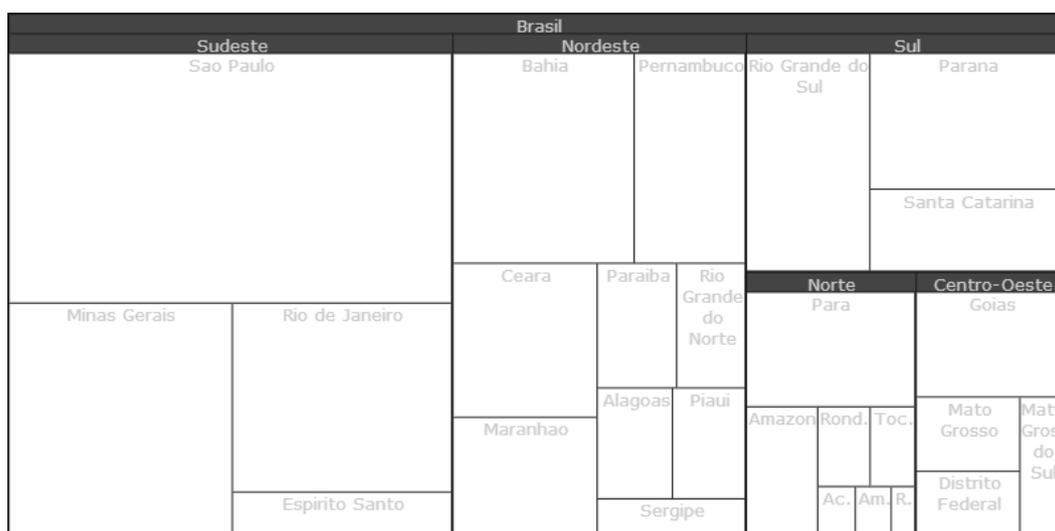


Figura 3.1: *Treemap* representando a quantidade de habitantes das Unidades Federativas do Brasil.

Nessa técnica, cada retângulo representa um objeto distinto, agrupado hierarquicamente. Na Figura 3.1, por exemplo, cada retângulo representa uma unidade federativa do Brasil, agrupada pela sua região e as regiões são agrupadas pelo seu País. Além disso, a área desse retângulo representa uma característica desse objeto, que na Figura 3.1 está representando o número de habitantes de cada unidade federativa. Ou seja, quanto maior o retângulo maior o número de habitantes.

O *Treemap* permite aos usuários comparar os objetos, mesmo em posições hierárquicas diferentes, além de auxiliar na identificação de padrões e exceções de um conjunto de dados. De acordo com a Figura 3.1, por exemplo, o Estado de São Paulo possui mais habitantes

---

que toda a Região Norte, pois a área do retângulo que corresponde ao estado de São Paulo é maior que a área do retângulo que corresponde à região Norte.

## Capítulo 4

### *Plugin de Integração com Moodle*

Um *plugin* funciona como uma extensão do *software* principal e tem como objetivo implementar funcionalidades não existentes nesse *software*. O *plugin* de integração deste trabalho implementa comunicações não existentes originalmente no *Moodle*. Esse *plugin* funciona como intermediário entre as comunicações do *Moodle* e das ferramentas de apoio, conforme Figura 4.1.

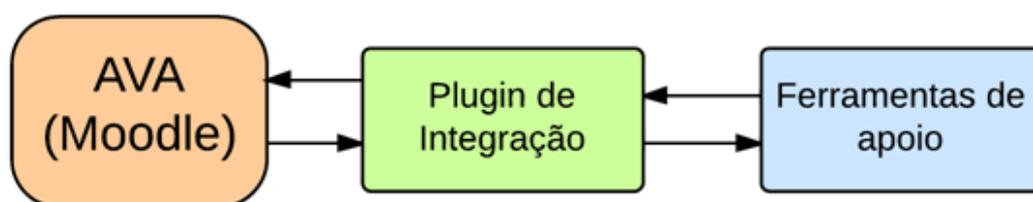


Figura 4.1: Arquitetura da solução com o *plugin*.

O objetivo do *plugin* desenvolvido é integrar as ferramentas de apoio educacional (algumas descritas no Capítulo 5) e o *Moodle*. Essa integração é feita através de *webservices* instalados no mesmo servidor de instalação do *Moodle*. Esses *webservices* possibilitam o *download* e *upload* das informações necessárias à integração.

O *webservice* de *download* busca as respostas de tarefas do tipo *Assignment* do *Moodle*. Tanto as atividades com anexo e sem anexo são baixadas. Para as atividades sem anexo é gerado um arquivo com seu conteúdo.

O *webservice* de *upload* atualiza o Moodle com as notas e *feedbacks* gerados pelas ferramentas de apoio, que são externas ao Moodle.

## 4.1 Instalação

Para utilização do *plugin* é necessário atender os seguintes requisitos:

1. Moodle instalado na versão 2.6.1;
2. Moodle utilizando o Banco de Dados *MySql* na versão 5.1.73-cll;
3. Ter os dados de conexão ao banco de dados (servidor, usuário, senha e nome do banco);
4. Moodle instalado em um servidor Linux;
5. Ter nesse servidor a versão 5.4 do PHP ou superior;
6. Ter o usuário e senha de administrador do Moodle.

A garantia de funcionamento é para as configurações e versões descritas acima. Pois o *Plugin* desenvolvido neste trabalho foi testado e validado com essas configurações, estando fora do escopo a validação em outras configurações. Porém, esse *Plugin* poderá ser adaptado para funcionar com outras versões ou configurações.

Para instalação do *plugin* é necessário passar por três etapas. A primeira é a criação do usuário de *webservice* no Moodle, a segunda é a instalação dos *webservices* no servidor do Moodle e a terceira é a configuração do servidor onde estão as ferramentas de apoio que acessarão o ambiente Moodle.

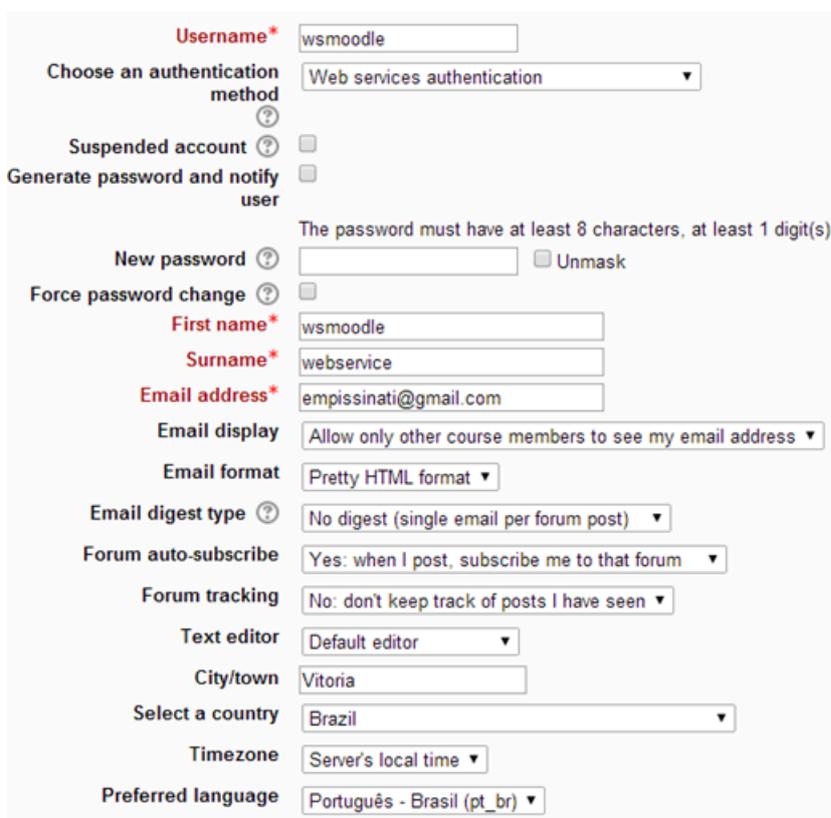
### 4.1.1 Etapa 1: Criar usuário do *webservice* de download de arquivos

Passo 1: Habilitar o uso de *webservices* no Moodle. Acesse o Moodle como administrador e vá em: *Administration*, *Site Administration*, *Advanced Features*, marcar a opção:

*Enable web services*. Depois disso, *role* até o final da página e clique no botão *Save Changes*.

Passo 2: Ative os protocolos que irá trabalhar. Acesse: *Administration, Site Administration, plugins, Web services, Manage protocols*. Clique na coluna *Enable* do item *SOAP Protocol*. Após isso, marque a opção *Web services documentation* e clique no botão *Save Changes*.

Passo 3: Cadastrar usuário para o *webservice*. Acesse: *Administration, Site Administration, Users, Accounts, Add a new user* e preencha o formulário com os dados do novo usuário (conforme exemplo da Figura 4.2). Na opção *Username* informe: *wsmoodle*. Na opção *Choose an authentication method* selecione: *Web services authentication*. Os demais parâmetros devem ser preenchidos como em um cadastro de usuário comum.



The image shows a Moodle user registration form for a webservice user. The form is filled with the following data:

- Username\***: wsmoodle
- Choose an authentication method**: Web services authentication
- Suspended account**:
- Generate password and notify user**:
- New password**: [empty]  Unmask
- Force password change**:
- First name\***: wsmoodle
- Surname\***: webservice
- Email address\***: empissinati@gmail.com
- Email display**: Allow only other course members to see my email address
- Email format**: Pretty HTML format
- Email digest type**: No digest (single email per forum post)
- Forum auto-subscribe**: Yes: when I post, subscribe me to that forum
- Forum tracking**: No: don't keep track of posts I have seen
- Text editor**: Default editor
- City/town**: Vitoria
- Select a country**: Brazil
- Timezone**: Server's local time
- Preferred language**: Português - Brasil (pt\_br)

Figura 4.2: Cadastrar usuário para o *webservice*.

Passo 4: Cadastrar uma nova *role*. Acesse: *Administration, Site Administration, Users, Permissions, Define roles*. Clique no botão *Add a new role*. A próxima tela será conforme a

Figura 4.3, deixe os valores como na Figura e clique em *Continue*.

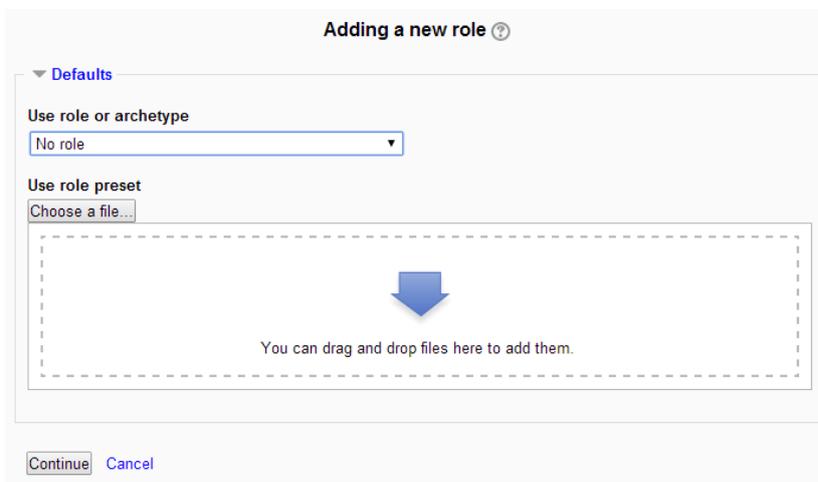


Figura 4.3: Cadastrar uma nova *role*, tela 1/2.

Na tela que segue informe os valores conforme a Figura 4.4.

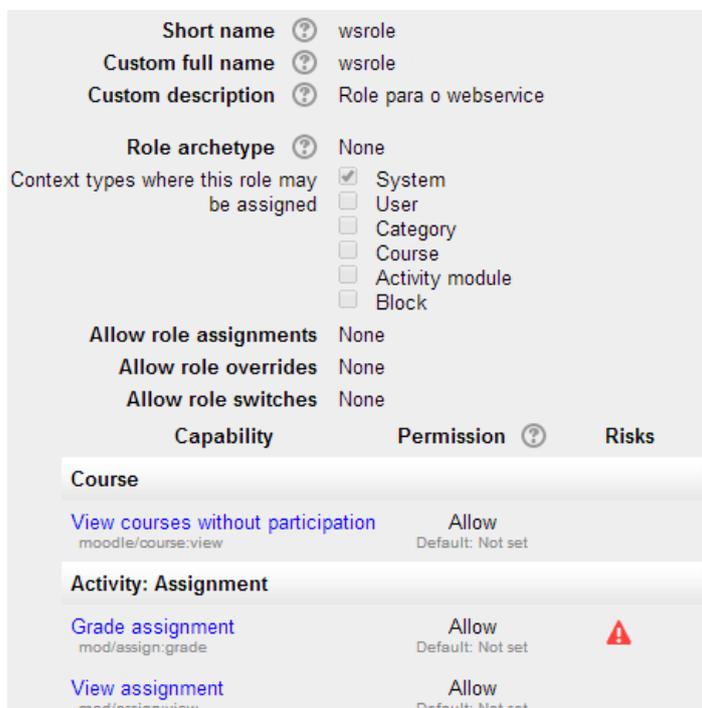


Figura 4.4: Cadastrar uma nova *role*, tela 2/2.

Passo 5: Vincular o usuário à *role* criada. Acesse: *Administration, Site Administration, Users, Permissions, Assign system roles*. Selecione a *role wsrole*. Na tela, conforme a Figura

4.5, encontre o usuário do *webservice* no lado direito e clique em *Add* para vinculá-lo a *role*.

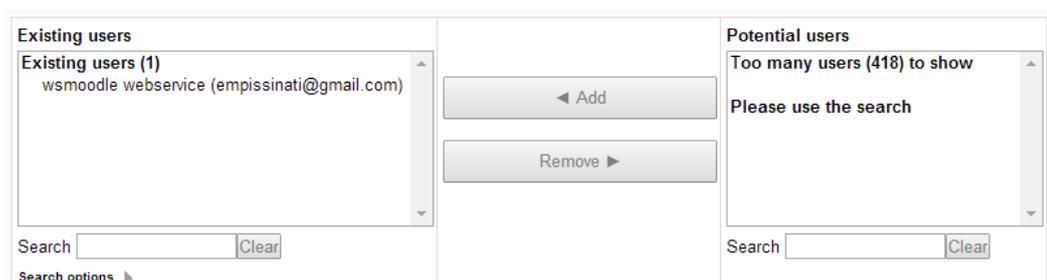


Figura 4.5: Vincular Usuário à *role*.

Passo 6: Cadastrar serviço externo. Acesse: *Administration, Site Administration, plugins, Web services, External services*. Clique no botão *Add*, para o campo *Name* informe *wsfiledownload* e marque as opções: *Enable* e *Can download files*.

Passo 7: Cadastrar um *token* para o *webservice*: *Administration, Site Administration, plugins, Web services, Manage Tokens*. Adicione um token, vincule ao usuário *wsmoodle* e ao serviço *wsfiledownload*.

Passo 8: Testar a Etapa 1:

1. Acesse uma resposta do moodle que possui um anexo;
2. Clique com o botão direito do mouse sobre o nome do arquivo e em seguida copie o endereço do link. (exemplo: `<http://eliasdeoliveira.com.br/moodle2/pluginfile.php/207/assignsubmission_file/submission_files/20/teste.c?forcedownload=1>`);
3. Faça as seguintes modificações no link: Antes do nome `pluginfile`, inclua do texto: `”webservice/”`; Depois do `”forcedownload=1”`, inclua o texto: `&token=[codigotoken]`  
O valor de `[codigotoken]` foi gerado no passo anterior;
4. Exemplo do formato final: `<http://eliasdeoliveira.com.br/moodle2/webservice/pluginfile.php/207/assignsubmission_file/submission_files/20/teste.c?forcedownload=1&token=81a6326f23db9`
5. Informe o link no browser e o arquivo deverá ser baixado;

### 4.1.2 Etapa 2: Instalação dos webservice de download de respostas e upload de resultados

Passo 1: Descompacte o arquivo *webservice.zip* dentro da mesma pasta onde se encontra a pasta do Moodle.

Passo 2: Abra o arquivo *mysql\_Moodle.php* que está dentro da pasta *includes*. Altere os dados de conexão para os de conexão com o banco de dados do Moodle.

### 4.1.3 Etapa 3: Configuração do servidor remoto - Uso do *plugin*

Passo 1: Acesse o servidor remoto que acionará o *plugin*.

Passo 2: Crie o arquivo *clienteMoodle.sh*, como exemplo abaixo:

```
#!/bin/sh
java -jar /home/clienteMoodle.jar g
http://site.com.br/moodle/mod/assign/view.php?id=92
/home/destino/ token

java -jar /home/clienteMoodle.jar s
http://site.com.br/moodle/mod/assign/view.php?id=92
/home/destino/
```

Alterando:

- `"/home/clienteMoodle.jar"` para o caminho onde está o arquivo *clienteMoodle.jar*;
- `"http://site.com.br/moodle/mod/assign/view.php?id=92"` para o link do Moodle;
- `"/home/destino/"` para pasta onde se deseja salvar os dados vindos do Moodle;
- `"token"` para o *token* criado no passo 7 da seção 4.1.1;
- O comando *java* não deve ter quebra de linha.

# Capítulo 5

## Correção Semiautomática

Neste capítulo serão apresentados corretores Semiautomáticos para questões discursivas e para códigos de programação. Ambos integrados com o Ambiente Virtual de Aprendizagem (AVA) *Moodle*.

### 5.1 Preditor Semiautomático de notas para questões discursivas

O Preditor Semiautomático de notas para questões discursivas é uma das propostas desse trabalho para auxiliar o trabalho docente. Seu objetivo é estimar as notas das respostas discursivas com base em respostas já pontuadas por um professor. Esse preditor pode ser previamente treinado para que suas decisões sejam feitas de acordo com o padrão de determinado avaliador ou banca de avaliadores.

A proposta usa técnicas de clusterização para agrupar as respostas discursivas mais similares. Essas respostas são agrupadas em  $n$  grupos diferentes. Esse agrupamento é feito com base na medida de similaridade cosseno entre as respostas, detalhada na Seção 3.3. Para cada grupo de respostas formado é selecionada uma resposta que é submetida ao professor para que ele corrija. A nota que o professor atribuir a essa resposta é replicada às demais respostas do grupo.

Para que seja possível agrupar as respostas semelhantes, cada uma é pré-processada (ver Seção 3.1), vetorizada (ver Seção 3.2) e submetida ao algoritmo de clusterização (ver Seção 3.4). Além dessas técnicas foram utilizadas também a retirada de *stopwords*, *stemming*, n-gramas e ponderação de notas.

Esse avaliador Semiautomático foi integrado ao ambiente virtual de aprendizagem *Moodle*. A integração não modificou as funcionalidades desse ambiente, nem a rotina de trabalho daqueles que o utilizam. Dessa forma é possível utilizar o avaliador mesmo depois que as respostas dos alunos tenham sido carregadas no ambiente. Essa integração foi feita através de um *plugin*, detalhado no Capítulo 4.

### 5.1.1 Funcionamento

Esta subseção descreverá a interação do processador de notas com o ambiente *Moodle* através do *plugin*. Mostrará também a interação do professor e de seus alunos com essa proposta. Esse detalhamento foi dividido em dez passos.

1. Cadastrar nova atividade no *Moodle*;
2. Responder atividade no *Moodle*;
3. Transferir as respostas para o servidor de trabalho;
4. Agrupar as respostas semelhantes;
5. Selecionar as respostas que serão corrigidas manualmente;
6. Encaminhar as respostas selecionadas para o *Moodle*;
7. Corrigir as respostas manualmente;
8. Transferir as notas do *Moodle* para o servidor de trabalho;
9. Corrigir as respostas automaticamente;
10. Atualizar as respostas no *Moodle*.

### Passo 1. Cadastrar nova atividade no Moodle.

O processo inicia quando o professor cadastra uma nova atividade no Moodle do tipo *Assignment*, conforme Figura 5.1. Em seguida ele informa os dados da questão, marcando o tipo de submissão igual a *Online text*, conforme Figura 5.2.

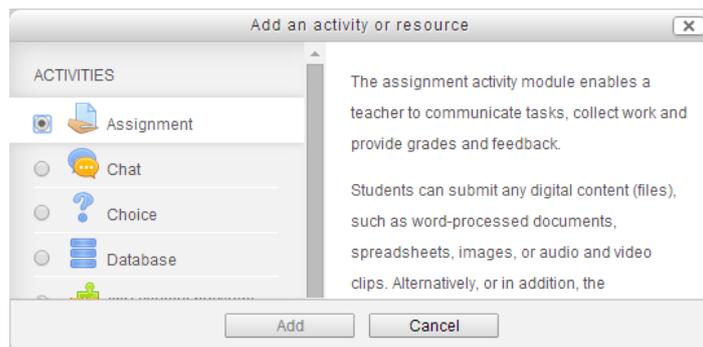


Figura 5.1: Cadastro da questão no Moodle.

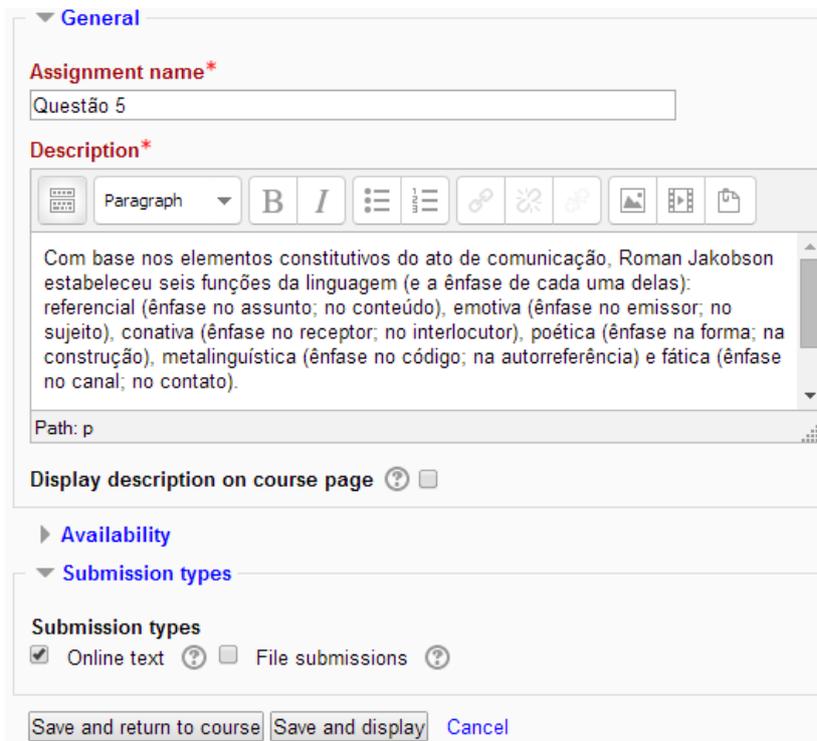


Figura 5.2: Cadastro da questão no Moodle.

### Passo 2. Responder atividade no Moodle.

Depois que o professor disponibilizar uma atividade, os alunos iniciam a submissão das respostas. A Figura 5.3 mostra a visão do aluno no momento da submissão de uma resposta.

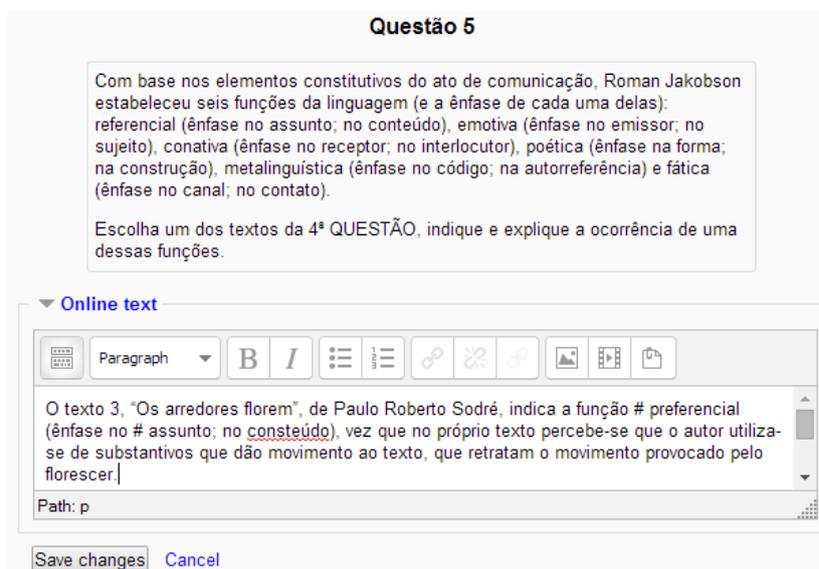


Figura 5.3: Cadastro da resposta no *Moodle*

### Passo 3. Transferir as respostas para o servidor de trabalho.

De três em três horas o *plugin* é acionado e realiza a transferência das repostas informadas pelos alunos no *Moodle* para o servidor de trabalho, atualmente localizado nos computadores do Laboratório de Computação de Alto Desempenho da UFES (LCAD). Nesse servidor, o *plugin* cria uma pasta para cada questão e dentro dessa pasta ele cria uma nova pasta para cada resposta, além de dois arquivos: *intervalonotas.csv* e *notastreino.csv*, conforme Figura 5.4.

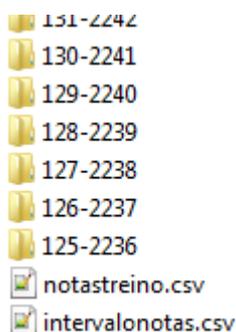


Figura 5.4: Conteúdo da pasta da questão.

O arquivo *intervalonotas.csv* armazena a nota mínima e a nota máxima possível, estabelecendo assim um intervalo. A Figura 5.5 mostra um exemplo cuja nota mínima é 0 e a nota máxima é 100.

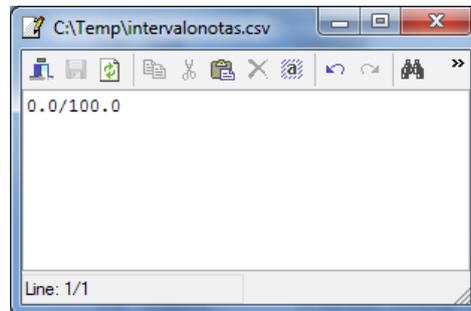


Figura 5.5: Arquivo *intervalonotas.csv*.

O arquivo *notastreino.csv* armazena o identificador das respostas, que corresponde ao nome das pastas de cada resposta, seguido da nota dada pelo professor separados por ponto e vírgula. Caso o professor não tenha atribuído nota, o valor será -1.0. A Figura 5.6 mostra um arquivo que ainda não possui as notas do professor. Esse arquivo será usado pelo preditor de notas para armazenar as notas que o professor pontuou.

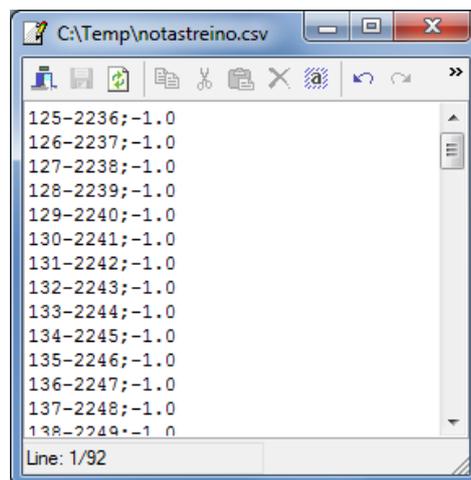


Figura 5.6: Arquivo *notastreino.csv*.

Dentro de cada pasta de resposta criada pelo *plugin* existe um arquivo chamado *resposta.txt* que contém o texto da resposta do aluno.

#### **Passo 4. Agrupar as respostas semelhantes.**

Para que seja possível agrupar as respostas semelhantes, cada resposta é pré-processada (ver Seção 3.1) e vetorizada (Ser seção 3.2). As respostas vetorizadas são unidas para formar uma matriz. Essa matriz é usada para processar o agrupamento das mais similares pelo algoritmo de *clustering*.

O algoritmo de *clustering* usado nesse agrupamento é o *Bisecting K-means*. Esse algoritmo reúne todos elementos em um único grupo, divide esse grupo em dois e depois realiza sucessivas divisões nos grupos existentes até formar  $k$  grupos. Em cada divisão o algoritmo tem como objetivo formar grupos mais similares. A medida de similaridade utilizada nesse processo é o cosseno (Seção 3.3).

Esse algoritmo foi utilizado através de um pacote de ferramentas (*toolkit*) de clusterização chamado *Cluto* (KARYPIS, 2003). O *Cluto* é um *software* que possui funcionalidades encapsuladas de *clustering*. Neste trabalho foi usada a versão 2.1.2.

#### **Passo 5. Selecionar as respostas que serão corrigidas manualmente.**

Para cada grupo, é criado um documento discriminante. Esse documento é gerado a partir da união dos termos de todas as respostas presentes no grupo. Nesse caso, o documento do grupo que for o mais similar ao documento discriminante será selecionado para o professor corrigir.

Essas respostas – selecionadas de cada grupo – serão indicadas no arquivo *resultado.csv*. Esse arquivo será gravado na pasta da questão, criada no Passo 3. O formato do arquivo *resultado.csv* é ilustrado na Figura 5.7. Esse arquivo armazena o identificador das respostas, seguido da nota dada pelo preditor e por um marcador, separados por ponto e vírgula. Esse marcador, que no caso é  $x$ , sinalizará no *Moodle* as respostas que deverão ser corrigidas.

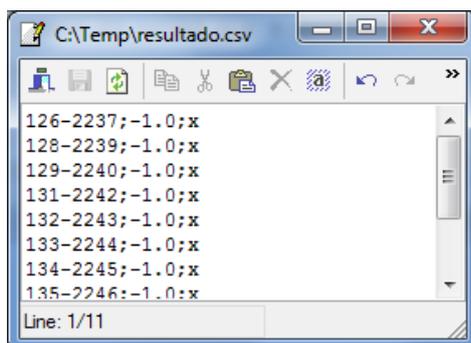


Figura 5.7: Arquivo *resultado.csv*.

### Passo 6. Encaminhar as respostas selecionadas para o Moodle.

O *plugin* fará a transferência das respostas selecionadas para o Moodle. Após a transferência as respostas serão marcadas conforme a Figura 5.8. As respostas marcadas estão com um *x* no campo *Feedback comments*.

First name	Grade	Online text	Feedback comments
Fulano1 Silva1	<input type="text"/> / 100.00	Q "# E vocês, senhoras, esperem da justiça dos homens o castigo deste malvado.". "E vocês, meus filhos, sejam felizes, que eu pedirei para ... (28 words)	x <input type="text"/>
Fulano2 Silva2	<input type="text"/> / 100.00	Q As duas últimas falas do personagem Mestre, com o uso do pronome você, ou vocês ficariam da seguinte forma: E vocês, senhoras, esperem da... (48 words)	<input type="text"/>
Fulano3 Silva3	<input type="text"/> / 100.00	Q Mestre (para Carlos) – Carlos, o Dom Alade julgou mais prudente que você não volte lá, aqui você tem a permissão por ele assinada para ... (66 words)	x <input type="text"/>
Fulano4 Silva4	<input type="text"/> / 100.00	Q Mestre (para Carlos) – Carlos, o Dom Abade julgou mais prudente que você não volte lá. Aqui você tem a permissão assinada por ele para ... (66 words)	<input type="text"/>
Fulano5 Silva5	<input type="text"/> / 100.00	Q – Carlos, o Dom Abade julgou mais prudente que você não volte lá. # Você tem aqui a permissão assinada por ele para que saia # do ... (55 words)	<input type="text"/>

Figura 5.8: Respostas marcadas no Moodle para correção do professor.

### Passo 7. Corrigir as respostas manualmente.

O professor deverá pontuar as respostas que estão marcadas, conforme faz normalmente no ambiente Moodle. Caso o professor pontue alguma resposta que não esteja marcada o processador irá ignorá-la e não fará qualquer alteração nessa nota.

**Passo 8. Transferir as notas do Moodle para o servidor de trabalho.**

Como no Passo 3, o *plugin*, de três em três horas é acionado e realiza a transferência dos dados do Moodle para o servidor de trabalho. As respostas pontuadas pelo professor são registradas no arquivo *notastreino.csv*, onde antes era registrado -1. A Figura 5.9 apresenta esse novo arquivo.

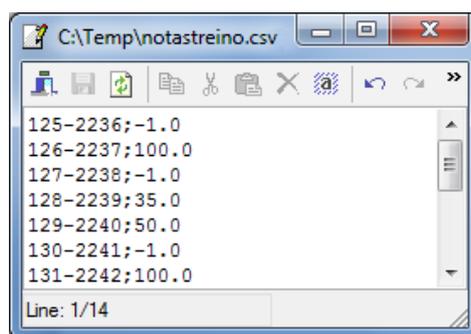


Figura 5.9: Arquivo *notastreino.csv* com as notas do professor.

**Passo 9. Corrigir as respostas automaticamente**

Com base nas respostas pontuadas pelo professor, o Processador faz a pontuação das respostas ainda sem nota. Para cada resposta pontuada pelo professor, o Preditor atribuirá a mesma nota às demais integrantes do mesmo grupo. Depois que todas as respostas estiverem pontuadas é feito o processo de ponderação de notas, somente para as que foram preditas. Esse processo será detalhado na Seção 7.4.

As notas finais preditas serão gravadas no arquivo *resultado.csv*. A Figura 5.10 mostra o arquivo com as notas preditas. Esse arquivo será usado pelo *plugin* para atualizar o Moodle.

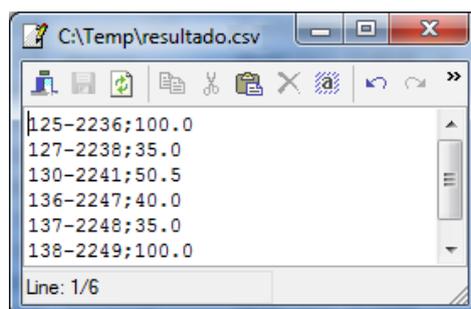


Figura 5.10: Arquivo *resultado.csv* com as notas preditas.

**Passo 10. Atualizar as respostas no Moodle.**

Como no Passo 8, o *plugin*, de três em três horas é acionado e realiza a transferência dos dados do Moodle para o servidor de trabalho. As respostas pontuadas previstas pelo processador são enviadas para o Moodle. Com isso, o processo de predição de notas finaliza e os dados estarão disponíveis para a ferramenta de visualização gerar os gráficos.

## 5.2 Predição Semiautomática de notas para códigos de programação

O desenvolvimento do *plugin* para o Moodle possibilitou o auxílio ao docente em outras atividades, além das respostas discursivas, por exemplo. Nesta subseção será descrito uma forma de apoiar os professores em atividades com códigos de programação que utilizam o Moodle, através de um preditor Semiautomático de notas para códigos de programação.

Esse preditor, denominado Núcleo de Avaliação Diagnóstica (NAD) é um módulo do Sistema Online de Atividades de Programação (SOAP) proposto por Oliveira (2013). O NAD implementa as funcionalidades de avaliação Semiautomática de exercícios de programação e de mapeamento de perfis de aprendizagem por componentes de habilidades (OLIVEIRA, 2013).

A integração do NAD com o Moodle permitiu que usuários desse AVA pudessem ter o apoio na correção de códigos de programação sem ter que acessar ou conhecer um novo sistema e os usuários do SOAP terão acesso às funcionalidades nativas do Moodle. Além disso, com essa integração, será possível utilizar bancos de questões do Moodle com exercícios de programação já corrigidos para reaplicações em novas turmas. Isso fará com que o esforço de correção do professor seja bem menor.

### 5.2.1 Funcionamento

Essa subseção descreverá o funcionamento da integração do NAD com o Moodle, desde a criação da atividade pelo professor até a predição de parte das respostas. Mostrará também

a interação do professor e de seus alunos com essa proposta. Esse detalhamento foi dividido em seis passos.

1. Cadastrar nova atividade no *Moodle*;
2. Responder atividade no *Moodle*;
3. Transferir as respostas para o servidor de trabalho;
4. Selecionar as respostas que serão corrigidas manualmente;
5. Corrigir as respostas manualmente;
6. Corrigir as respostas automaticamente.

**Passo 1. Cadastrar nova atividade no *Moodle*.** O processo inicia quando o professor cadastra uma nova atividade no *Moodle* do tipo *Assignment*, conforme Figura 5.11. Em seguida informa os dados da questão, marcando o tipo de submissão igual a *File submissions*, conforme Figura 5.12.

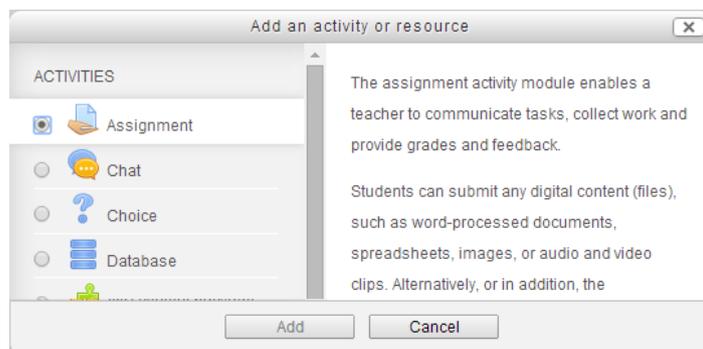
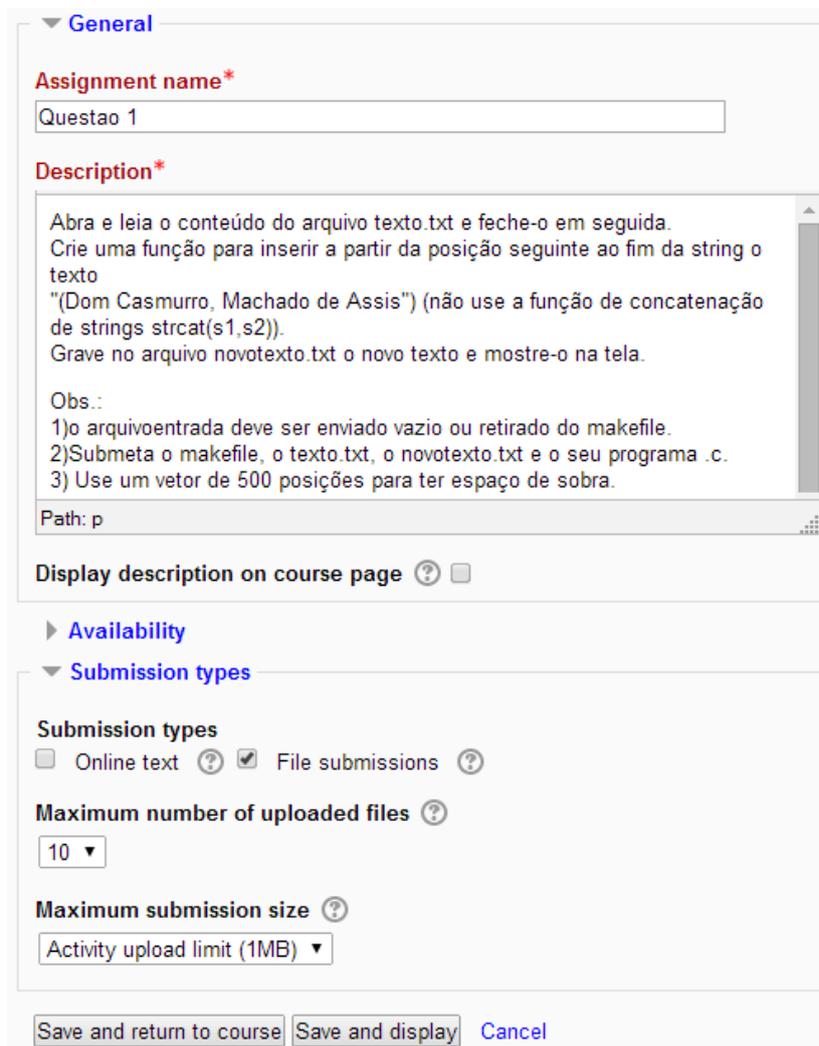


Figura 5.11: Cadastro da questão no *Moodle*.



The screenshot shows the Moodle assignment creation interface. The 'General' section is expanded, showing the 'Assignment name' field with the value 'Questao 1'. The 'Description' field contains the following text: 'Abra e leia o conteúdo do arquivo texto.txt e feche-o em seguida. Crie uma função para inserir a partir da posição seguinte ao fim da string o texto "(Dom Casmurro, Machado de Assis)" (não use a função de concatenação de strings strcat(s1,s2)). Grave no arquivo novotexto.txt o novo texto e mostre-o na tela. Obs.: 1)o arquivoentrada deve ser enviado vazio ou retirado do makefile. 2)Submeta o makefile, o texto.txt, o novotexto.txt e o seu programa .c. 3) Use um vetor de 500 posições para ter espaço de sobra.' The 'Path' field is set to 'p'. There is a checkbox for 'Display description on course page' which is currently unchecked. The 'Availability' section is collapsed. The 'Submission types' section is expanded, showing 'Submission types' with 'Online text' unchecked and 'File submissions' checked. The 'Maximum number of uploaded files' is set to 10, and the 'Maximum submission size' is set to 'Activity upload limit (1MB)'. At the bottom, there are three buttons: 'Save and return to course', 'Save and display', and 'Cancel'.

Figura 5.12: Cadastro da questão no *Moodle*.

### **Passo 2. Responder atividade no *Moodle*.**

Depois que o professor disponibiliza a atividade, os alunos anexam os arquivos necessários para a submissão. Para cada resposta, devem ser anexados os códigos de programação, outros arquivos necessário para execução do algoritmo e um arquivo denominado *makefile*, que conterà a sequência e forma de execução desses arquivos. A Figura 5.13 mostra a visão do aluno no momento da submissão desses arquivos.

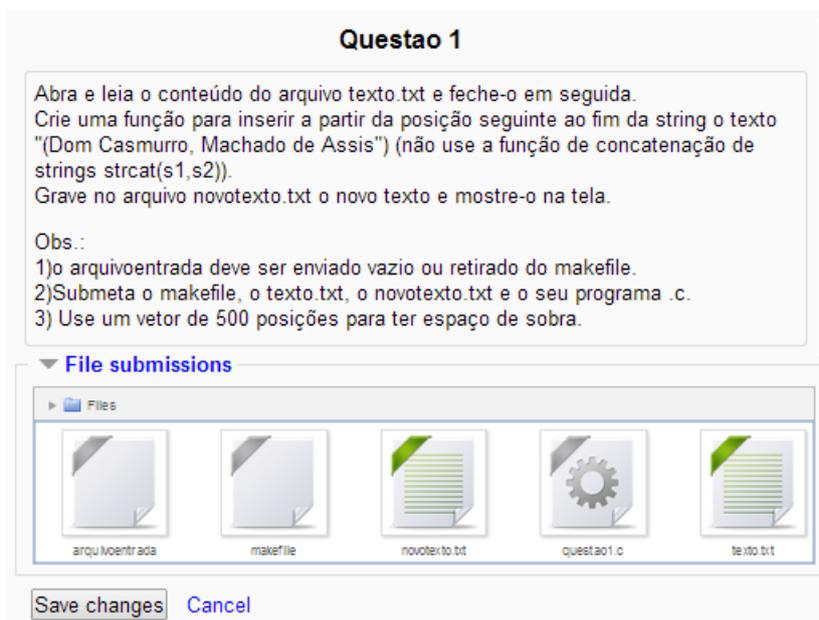


Figura 5.13: Cadastro da resposta no *Moodle*

O padrão de submissão dessas respostas, especificando quais arquivos e seus respectivos conteúdos, é descrito no trabalho de Oliveira (2013).

### **Passo 3. Transferir as respostas para o servidor de trabalho.**

De três em três horas o *plugin* é acionado e realiza o *download* dos arquivos anexados pelos alunos no *Moodle* para o servidor de trabalho. Nesse servidor, o *plugin* cria uma pasta para cada questão, dentro dessa pasta ele cria uma nova pasta para cada resposta e dois arquivos: *intervalonotas.csv* e *notastreino.csv*, conforme Figura 5.14.

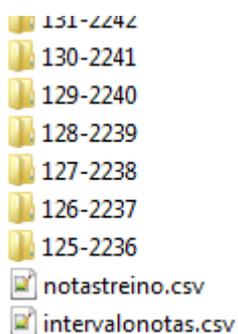


Figura 5.14: Conteúdo da pasta da questão.

O arquivo *intervalonotas.csv* armazena a nota mínima e a nota máxima possível, estabelecendo assim um intervalo. A Figura 5.15 mostra um exemplo cuja nota mínima é 0 e a nota máxima é 100.

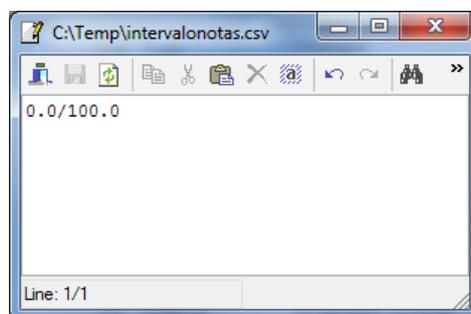


Figura 5.15: Arquivo *intervalonotas.csv*.

O arquivo *notastreino.csv* armazena o identificador das respostas, que corresponde ao nome das pastas de cada resposta, seguido da nota dada pelo professor separados por ponto e vírgula. Caso o professor não tenha atribuído nota, o valor será -1.0. A Figura 5.16 mostra um arquivo que ainda não possui as notas do professor. Esse arquivo será usado pelo processador para armazenar as notas que o professor pontuou.

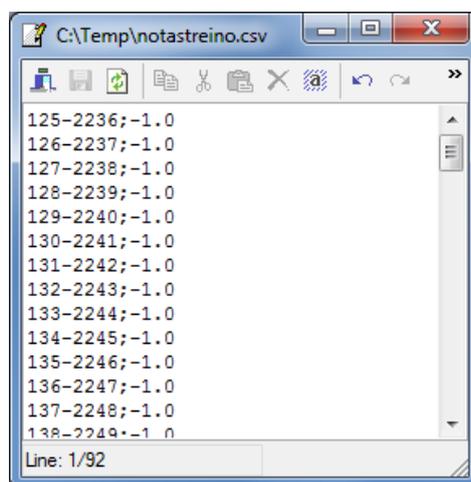


Figura 5.16: Arquivo *notastreino.csv*.

Dentro de cada pasta de resposta criada pelo *plugin* estarão os arquivos anexados pelos alunos em suas submissões.

#### **Passo 4. Selecionar as respostas que serão corrigidas manualmente.**

Dentre a coleção de respostas submetidas, o NAD seleciona uma amostra representativa de respostas para o professor corrigir. Essas respostas selecionadas serão indicadas no arquivo *resultado.csv*. Esse arquivo será gravado na pasta da questão, criada no Passo 3. O formato do arquivo *resultado.csv* é o mesmo descrito no Passo 5 da Subseção 5.1.1.

#### **Passo 5. Corrigir as respostas manualmente.**

O *plugin* fará a transferência das informações de marcação para o *Moodle*. Após a transferência, as respostas serão marcadas com um *x* no campo *Feedback comments*. O professor deverá corrigir essas respostas marcadas. Após isso, o *plugin* fará a transferência dos dados do *Moodle* para o servidor de trabalho.

#### **Passo 6. Corrigir as respostas automaticamente**

Com base nas respostas pontuadas pelo professor, o NAD faz a pontuação das submissões ainda sem nota (OLIVEIRA, 2013). As notas finais serão gravadas no arquivo *resultado.csv*. Esse arquivo será usado pelo *plugin* para atualizar o *Moodle*. Com isso, o processo de predição de notas finaliza e os dados estarão disponíveis para a ferramenta de visualização gerar os gráficos.

### **5.2.2 Aplicação e Resultados**

Um exemplo do uso dessa ferramenta foi na predição Semiautomática das notas de códigos de programação de uma questão de prova da disciplina de Introdução à Programação do curso de Engenharia Elétrica da UFES no período letivo 2014/01. Os alunos submeteram as respostas com os códigos de programação e, após isso, o professor corrigiu parte dessas respostas, conforme descrito na Seção 5.2.1.

Esse processo foi aplicado à uma base com 39 respostas. Dessas 39 respostas, 21 foram selecionadas pelo NAD para o professor corrigir, gerando um esforço de 53,84% na correção. Com base nas correções do professor, o NAD calculou as notas das respostas restantes com um erro médio de 0,51 pontos por questão, para uma escala de 0 a 10 pontos. Sendo que o desvio padrão dessa média foi de 0,39 e o erro máximo foi 1,47 pontos.

---

Como esse processo foi integrado ao *Moodle* através do *plugin*, essa questão pode ser aplicada novamente em outra turma ou em outro período. Com isso, o avaliador poderá usar as respostas já avaliadas para calcular as notas das novas respostas, reduzindo ainda mais o esforço na correção.

Além disso, considerando que o emprego do ensino de programação está presente em várias disciplinas, como, por exemplo, Estrutura de Dados, Inteligência Artificial e Algoritmos, o emprego dessa solução será ainda mais abrangente.

# Capítulo 6

## Visualização de Informação

No intuito de expandir o abrangência deste trabalho, além do auxílio na correção de atividade, é proposta uma ferramenta de visualização que apresenta ao professor informações do desempenho de vários alunos em várias atividades de forma ágil. Isso oferece aos educadores auxílio nas suas análises de como interferir no ensino realizado em suas turmas. Com essa estratégia, foi possível visualizar, em um único gráfico, os desempenhos de aproximadamente 600 alunos em 15 turmas diferentes.

Na proposta de correção automática o docente não avalia todas as respostas, e por isso, não consegue ter noção de como foi o rendimento da turma, a proposta de visualização dará a informação de como foi o resultado desses alunos.

### 6.1 Estudo de caso

Visualizar os perfis de alunos de um curso é uma tarefa difícil se o usuário não possuir as ferramentas mais adequadas para isso. Essa complexidade aumenta quando o contexto é o de turmas com muitos alunos, pois existe bastante informação para ser visualizada e comparada ao mesmo tempo.

Deparamos com essa necessidade nas turmas de Biblioteconomia da UFES, mediadas através do *Moodle*. As turmas possuem muitos alunos e muitas atividades. Para o professor

era difícil identificar quais estudantes estavam se destacando e quais estavam necessitando do seu auxílio.

As turmas de Biblioteconomia também foram estudadas no trabalho de Oliveira et al. (2010), mencionado no Capítulo 2. Nesse trabalho foi proposto uma metodologia computacional melhorada de avaliação formativa para a indexação de documentos textuais.

Diante do cenário dessas turmas, foi proposto o uso de relatórios textuais com os dados desses estudantes. Mas essa abordagem não foi eficaz por impossibilitar a visão geral de suas turmas, o que dificultou a identificação de padrões nessas informações.

Outra abordagem também não eficaz foi o uso de técnicas que davam a visão geral, como gráficos de setores (tradicionalmente chamado gráfico de pizza) e barras, apontando se a turma estaria indo bem ou mal. Entretanto, com essas imagens o professor não possuiria acesso às informações detalhadas dos alunos.

O professor precisa ter a visão geral sem perder os detalhes de cada aluno, visualizando todas essas informações em um único gráfico ou mapa. Esse mapa deverá exibir tanto o desempenho de cada aluno quanto o desempenho de todas as turmas desse professor.

No nosso estudo, o *Treemap* mapeou os resultados das atividades de 597 alunos de 16 turmas diferentes do curso de Biblioteconomia da UFES, durante os anos de 2010 a 2012. Nesse período, eles fizeram 28.017 exercícios das diversas disciplinas existentes no curso. O mapa gerado pode ser visualizado na Figura 6.1.

Na visualização proposta na Figura 6.1, cada retângulo colorido representa um aluno; a área desse retângulo, a quantidade de atividades feitas pelo aluno; e a cor, o percentual de acerto nessas atividades, variando de acordo com a Tabela 6.1 – sendo que cada turma é um agrupamento de alunos, como em uma hierarquia.

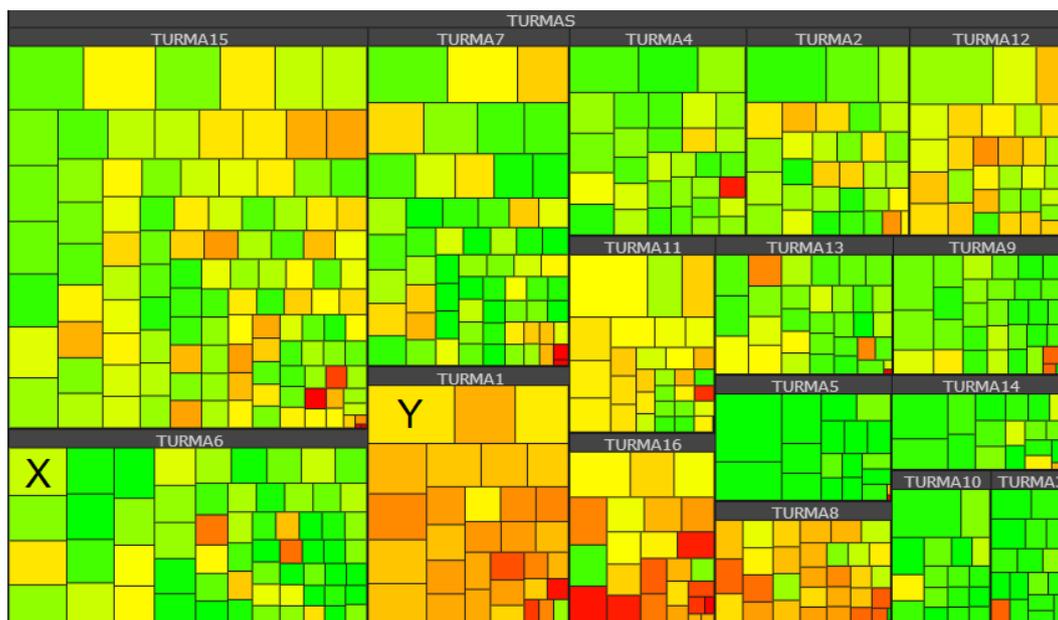


Figura 6.1: *Treemap* representando a visão das várias turmas de Biblioteconomia da UFES.

Cor	% de acertos
Verde RGB (0, 255, 0)	100
Verde tendendo para o Amarelo	
RGB (1, 255, 0) até RGB (254, 255, 0)	99 a 71
Amarelo RGB (255, 255, 0)	70
Amarelo tendendo para o Vermelho	69 a 1
RGB (255, 254, 0) até RGB (255, 1, 0)	
Vermelho RGB (255, 0, 0)	0

Tabela 6.1: Tabela associando o percentual de acertos às cores.

Por exemplo: na Figura 6.1, o quadrado marcado com X representa um aluno da Turma 6 que fez 136 atividades com um percentual de acertos igual a 76,47%; e o marcado com Y, um aluno da Turma 1 que resolveu 241 exercícios com um percentual de acertos igual a 63,07%. A quantidade de atividades e o percentual são exibidos ao passar o cursor sobre os

retângulos, conforme Figura 6.2.

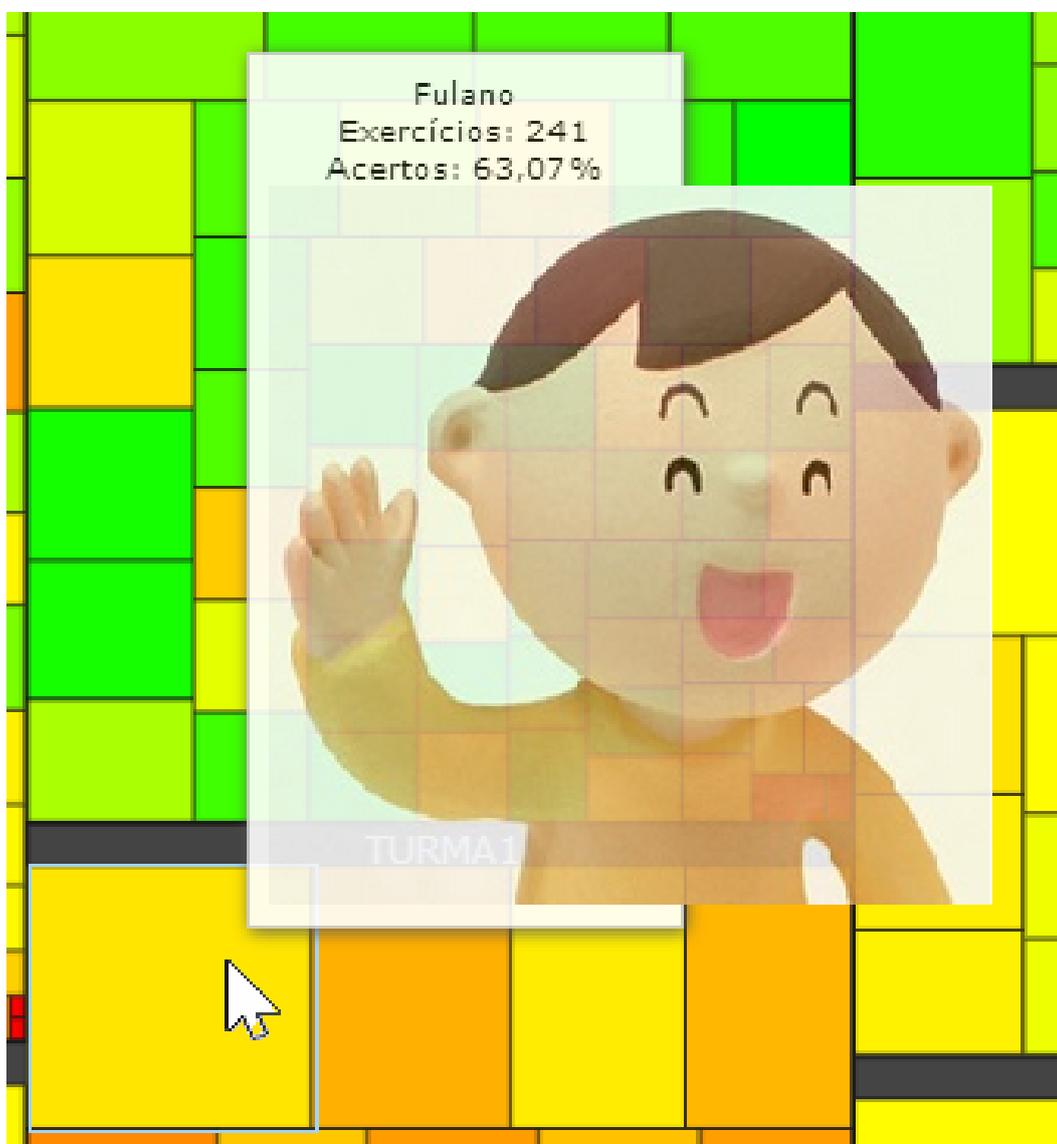


Figura 6.2: Detalhes que são exibidos ao passar o cursor sobre os retângulos que representam os alunos.

Além desses dados, o mapa também apresenta ao professor a foto do aluno, ver Figura 6.2. Essa informação ajudará o professor na identificação do mesmo, caso haja necessidade de uma abordagem pessoal.

Outra interação bastante útil é a visualização dos detalhes, necessária em turmas com muitos alunos. Seu acionamento é feito ao clicar sobre um determinado item, turma ou aluno. A Figura 6.3 ilustra o resultado da expansão da TURMA1. Nessa visão, é possível observar os retângulos menores que podem ficar ocultos na visão geral.

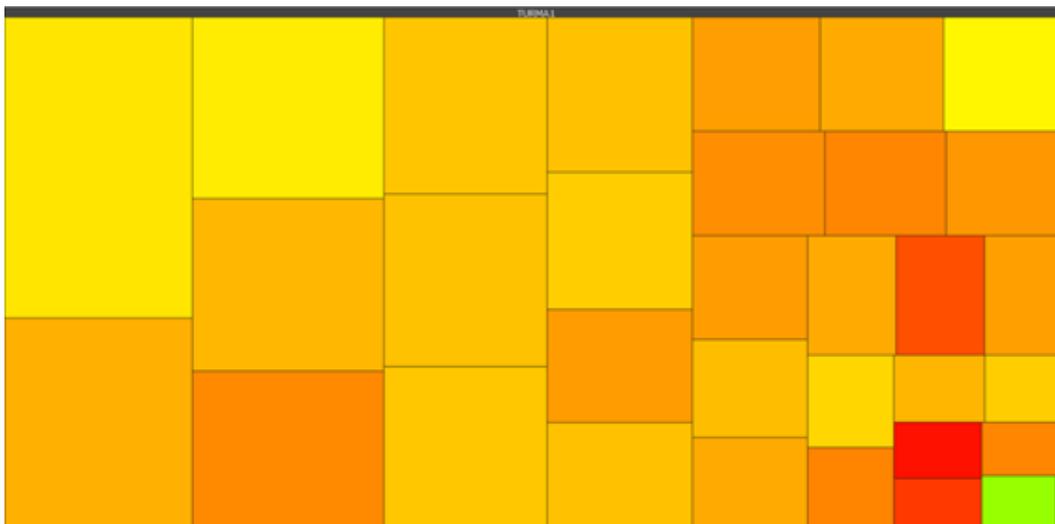


Figura 6.3: Visão detalhada da TURMA1.

Outra informação importante é a ordenação dos retângulos de acordo com a quantidade de atividades: elas diminuem de cima para baixo e da esquerda para direita. Isso faz com que as turmas do mapa fiquem ordenadas dos alunos mais dedicados para os menos dedicados.

Então, comparando os tamanhos dos retângulos, foi possível identificar os alunos que se dedicaram mais às atividades e os que se dedicaram menos. E, de acordo com a cor, foi possível diferenciar os alunos que compreenderam melhor a disciplina dos que tiveram alguma dificuldade.

Com base nessas características de cores e tamanhos, tivemos, rapidamente, a visão geral das situações das turmas. Como o caso das turmas 1, 8, 11 e 16, na Figura 6.1, que apresentam os rendimentos mais baixos, identificados através de sua coloração tendendo para o vermelho.

Da mesma forma, identificamos rapidamente os alunos necessitando do auxílio do docente. Como cada retângulo com a coloração avermelhada representa um aluno que está com um percentual de acertos próximo a zero, identificamos esses alunos e, antes mesmo de consultá-los, concluímos que seu mau desempenho era pela falta de dedicação.

Através dos mapas apresentados, conseguimos comparar os rendimentos dos alunos em uma turma ou em várias, tanto na quantidade de questões realizadas quanto no percentual de acertos de cada um. Foi possível identificar os alunos pelas fotos e detalhar as visualizações

de turmas específicas.

Conforme foi mostrado, podemos observar uma maneira de visualizar um enorme volume de dados de forma clara, objetiva e eficiente. Com uma ferramenta apresentando os dados dos alunos dessa forma, o professor terá em mãos um mapa de sua turma e um coordenador terá em mãos um mapa de toda sua coordenação. Assim, saberão onde atuar de forma rápida, com poucos segundos de análise do gráfico.

## **6.2 Escolha da Técnica de Visualização de Informação**

A escolha da técnica de visualização de informação foi baseada na necessidade de apresentar ao usuário muitos dados de um ponto de vista mais abrangente sem perder o ponto de vista mais específico. Ou seja, possibilitar a visão de todos os dados de uma só vez e também de um único item nesse mesmo gráfico. Um exemplo disso seria mostrar os desempenhos de várias turmas e ao mesmo tempo ser possível visualizar o desempenho de um único aluno.

Outro fator considerado foi a necessidade de organização dos dados hierarquicamente.

Baseado nessas necessidades, foram analisadas algumas das técnicas presentes na literatura. A Tabela 6.2 apresenta o resultado:

Técnica	Visão geral	Visão individual	Muitos dados	Hierarquia	Uso total do espaço
Coordenadas Paralelas (INSELBERG, 1985)	Sim	Sim	Não	Não	Não
<i>Stick Figures</i> (PICKETT; GRINSTEIN, 1988)	Sim	Não	Sim	Não	Não
<i>Treemap</i> (JOHNSON; SHNEIDERMAN, 1991)	<b>Sim</b>	<b>Sim</b>	<b>Sim</b>	<b>Sim</b>	<b>Sim</b>
<i>Cone Trees</i> (ROBERTSON; MACKINLAY; CARD, 1991)	Sim	Sim	Sim	Sim	Não
<i>Scatterplots</i> (FURNAS; BUJA, 1994)	Sim	Sim	Não	Não	Não
Orientadas a Pixels (KEIM; KRIGEL, 1996)	Sim	Não	Sim	Não	Sim
Baseadas em Grafos	Sim	Sim	Sim	Sim	Não

Tabela 6.2: Comparativo entre as possíveis técnicas de visualização adequadas a proposta.

Dentre as técnicas Grafos, *Cone Tree* e *Treemap*, que foram as mais pontuadas, *Treemap* se destaca por utilizar 100% da área disponível para visualização, suportando um número maior de dados de forma organizada hierarquicamente (JOHNSON; SHNEIDERMAN, 1991). O uso total da área disponível para visualização permite ao usuário obter muita informação usando pouco espaço.

### 6.3 Arquitetura da Ferramenta de Visualização

A arquitetura proposta consiste nos seguintes módulos (vide Figura 6.4):

(i) Módulo de Extração de Dados: responsável por buscar os dados no banco de dados do *Moodle* e tem como saída um arquivo no formato CSV;

(ii) Módulo de Tratamento de Dados: responsável por transformar os dados em informação,

fazer agrupamentos, calcular os percentuais de acertos, determinar as cores e as áreas dos objetos que serão representados no gráfico, e fazer a conversão dos arquivos CSV para o formato JSON;

(iii) Módulo de Visualização Gráfica: responsável por transformar as informações do arquivo JSON em visualizações gráficas. Esse módulo faz uso da biblioteca *JavaScript JavaScrit InfoVis Toolkit* para exibição dos gráficos.

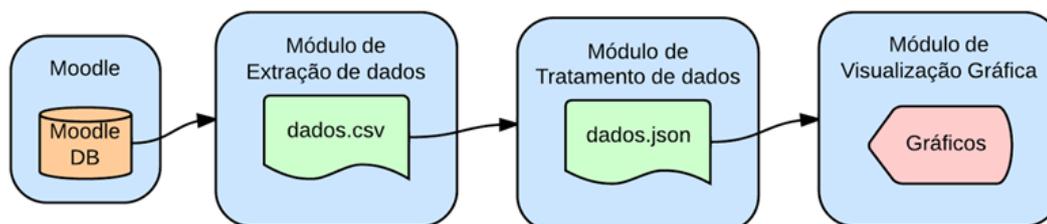


Figura 6.4: Arquitetura da Solução.

Uma das funcionalidades do Módulo de Tratamento de Dados é a definição da cor de cada objeto. Essa cor é representada no modelo RGB, que possui uma combinação da intensidade de vermelho, verde e azul para cada cor. O método *definirCor*, representado na Figura 6.5, define essas intensidades baseado no percentual de acerto passado por parâmetro. Esse percentual varia de 0 a 100 e as cores variam do vermelho ao verde.

```

public static String definirCor(Double percentualAcerto)
{
    int vermelho = 0, verde = 0, azul = 0;

    if (percentualAcerto > 70) {
        vermelho = (int) ((100 - percentualAcerto) * 255)/30;
        verde = 255;
    }
    else {
        vermelho = 255;
        verde = (int) (percentualAcerto * 255)/70;
    }

    return "rgb(" + vermelho + "," + verde + "," + azul + ")";
}
  
```

Figura 6.5: Método *definirCor* do Módulo de Tratamento de Dados.

Outra funcionalidade desse módulo é a definição da área do retângulo associado a cada

---

aluno. A área de um retângulo é definida relativa às áreas dos demais retângulos, ou seja, terá uma área maior o retângulo do aluno que realizou mais exercícios. Com isso, a relação sempre se manterá, mas o tamanho absoluto irá variar de acordo com a área disponível para exibição do gráfico.

# Capítulo 7

## Metodologia Experimental

Neste capítulo, são detalhadas cada uma das bases de dados utilizadas nos experimentos. Também é apresentada uma comparação entre dois avaliadores do vestibular, cuja comparação será usada para avaliar a qualidade do algoritmo de predição de notas.

### 7.1 Bases de Dados

Os testes foram realizados sobre uma amostra das respostas discursivas de Português do vestibular da UFES. No total são 460 respostas de 5 questões diferentes, representando um total de 92 alunos.

Cada questão é composta por 92 respostas descritas em língua portuguesa e as notas de seus avaliadores. A quantidade de palavras por resposta varia de acordo com a questão e com o aluno. A Tabela 7.1 apresenta um resumo da base de dados.

As colunas Q1, Q2, Q3, Q4 e Q5 da Tabela 7.1 se referem, respectivamente às cinco questões. A primeira linha mostra a quantidade de respostas por questão, a segunda linha mostra a média de palavras por resposta para cada questão, a terceira linha apresenta o desvio padrão e a última linha mostra a quantidade de palavras da maior resposta.

	Q1	Q2	Q3	Q4	Q5
Quantidade de respostas	92	92	92	92	92
Média de palavras por resposta	127,64	131,57	57,45	80,21	60,51
Desvio padrão	62,95	55,06	16,52	34,29	35,63
Maior resposta	254	268	115	218	194

Tabela 7.1: Tabela apresentando o resumo da base de dados com respeito a quantidade de palavras nas respostas dadas pelos alunos.

Adiante serão apresentadas essas cinco questões.

### 7.1.1 Questão 1

Para essa primeira questão, formamos uma base de dados contendo 92 respostas discursivas para a seguinte pergunta:

”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.”.

A questão acima dá ao aluno a opção de escolher dentre os três romances para elaborar sua resposta o que, conseqüentemente, amplia a diversidade de respostas. A Tabela 7.2 apresenta quantidade de respostas que escreveram sobre determinado romance e o seu percentual. As respostas em branco não selecionaram romances.

Romance escolhido	Quantidade de respostas	Percentual das respostas
Boca do inferno	7	7,61%
Ensaio sobre a cegueira	29	31,52%
Kitty aos 22: divertimento	52	56,52%
Nenhum	4	4,35%

Tabela 7.2: Resumo das respostas para a Questão 1.

### 7.1.2 Questão 2

Para essa segunda questão, formamos uma base de dados contendo 92 respostas discursivas para a seguinte pergunta:

”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:

A) Justifique o título da obra que você escolheu.

B) 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)”.

A questão acima dá ao aluno a opção de escolher dentre três obras para elaborar sua resposta o que, conseqüentemente, amplia a diversidade de respostas. A Tabela 7.3 apresenta a quantidade de respostas que os alunos escreveram sobre determinada obra e o seu percentual. As respostas em branco não selecionaram romances.

Obra escolhida	Quantidade de respostas	Percentual das respostas
Romanceiro da Inconfidência	15	16,30%
Vidas secas	60	65,22%
O noviço	14	15,22%
Nenhum	3	3,26%

Tabela 7.3: Resumo das respostas da Questão 2.

### 7.1.3 Questão 3

Para essa terceira questão, formamos uma base de dados contendo 92 respostas discursivas para a seguinte pergunta:

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.

### 7.1.4 Questão 4

Para essa quarta questão, formamos uma base de dados contendo 92 respostas discursivas para a seguinte pergunta:

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)

A) 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.).

B) 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.

A letra "A" da questão acima dá ao aluno a opção de escolher dentre os três textos para elaborar sua resposta ampliando assim a diversidade de respostas. A Tabela 7.4 apresenta a quantidade de respostas que escreveram sobre determinado texto e o seu percentual. As respostas em branco não selecionaram romances.

Texto escolhido	Quantidade de respostas	Percentual das respostas
TEXTO 1 - O navio negreiro	64	69,57%
TEXTO 2 - 7	22	23,91%
TEXTO 3 - Os arredores florem	4	4,35%
Nenhum	2	2,17%

Tabela 7.4: Resumo das respostas da Questão 4.

### 7.1.5 Questão 5

Para essa quinta questão, formamos uma base de dados contendo 92 respostas discursivas para a seguinte pergunta:

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 4ª QUESTÃO, indique e explique a ocorrência de uma dessas funções.

A questão acima dá ao aluno a opção de escolher dentre os três textos para elaborar sua resposta o que, conseqüentemente, amplia a diversidade de respostas. A Tabela 7.5 apresenta

a quantidade de respostas que escreveram sobre determinado romance e o seu percentual. As respostas em branco não selecionaram romances.

Texto escolhido	Quantidade de respostas	Percentual das respostas
TEXTO 1 - O navio negreiro	37	40,22%
TEXTO 2 - 7	47	51,09%
TEXTO 3 - Os arredores florem	6	6,52%
Nenhum	2	2,17%

Tabela 7.5: Resumo das respostas da Questão 5.

## 7.2 Métricas

Uma métrica muito comum no contexto estudado é a Precisão. A Precisão é a fração das notas preditas corretamente. Essa métrica permite somente julgamentos binários (não graduais) (BAEZA-YATES; RIBEIRO-NETO et al., 2011), ou seja, uma predição de nota é igual ou é diferente a de um avaliador. Ela não difere uma predição com diferença de 0,5 ponto de uma predição com diferença de 10 pontos.

No caso da comparação entre os avaliadores, se utilizássemos a precisão, teríamos um valor de 46,09% de precisão entre o Avaliador1 e o Avaliador2. Uma alternativa para o uso dessa medida seria utilizar um valor para margem de erro, ou seja, considerar o desvio de 1,0 ponto como estando correto. Mas isso retorna ao problema anterior de julgamentos binários, pois uma nota com 1,1 ponto de diferença seria tratada da mesma forma que uma nota de 9 pontos de diferença. Por esse motivos não utilizamos essa métrica neste trabalho.

Então, visando contornar esses problemas, como feito no trabalho de Vilela et al. (2012), usamos o Erro Médio de Pontuação por Resposta. Essa métrica será utilizada para comparar quão semelhante é a predição de nota entre dois avaliadores. Ela indica, em pontos, a média das diferenças entre as notas atribuídas à uma resposta pelos avaliadores. Essa métrica foi utilizada por ser gradual, ou seja, é possível identificar qualquer variação na medida, pois qualquer valor decimal será considerado em seu cálculo.

### 7.3 Avaliadores Humanos

Para a realização dos experimentos, foram considerados dois avaliadores do vestibular da UFES para as questões discursivas de português detalhadas na Seção 7.1. Neste trabalho esses avaliadores serão denominados de Avaliador1 e Avaliador2.

Como um dos objetivos desse trabalho é apresentar um algoritmo que consiga avaliar as respostas discursivas com um grau semelhante a um avaliador humano, nesta seção é feita a comparação entre esses dois avaliadores com o objetivo de medir o quão distante estão suas correções.

Essa comparação será feita a partir do erro médio de pontuação por resposta (Seção 7.2) entre os dois avaliadores. Ela é feita a partir do cálculo do módulo da diferença da nota dada pelo Avaliador1 pela nota atribuída pelo Avaliador2 para uma mesma resposta.

A Tabela 7.6 apresenta para cada questão a média dessas diferenças e seu desvio padrão. A última linha representa a medida final que é a média dessas médias e a média dos desvios padrões. Essa medida será usada para saber se o algoritmo conseguirá se comportar como um dos avaliadores.

A escala de pontuação das questões do vestibular varia de 0 a 2 pontos. Neste trabalho, essa escala foi alterada para 0 a 10.

Base	Média das Diferenças	Desvio padrão das Diferenças
Q1	1,46	1,75
Q2	2,33	1,95
Q3	0,26	0,72
Q4	2,16	1,76
Q5	0,68	1,85
Média	1,38	1,75

Tabela 7.6: Comparação entre os avaliadores.

Para avaliar o desempenho do algoritmo em relação ao Avaliador1, por exemplo, as notas preditas desse algoritmo serão comparadas com as respectivas notas do Avaliador1. Depois

disso, serão geradas as médias de suas diferenças para cada questão e uma média final da Média das Diferenças e do Desvio padrão das diferenças.

Nossa esperança é de que esses valores finais sejam menores ou iguais aos valores encontrados na comparação entre os dois avaliadores humanos, que são os valores da última linha da Tabela 7.6. Com isso, será possível dizer se a proposta é semelhante ou não à de um avaliador humano.

Somente as notas previstas pelo algoritmo serão consideradas em sua análise. As notas que serão informadas pelo professor não serão consideradas no cálculo das médias.

## 7.4 Ponderação de notas

A ponderação de notas permite que a pontuação de uma respostas predita seja ajustada com base nas notas das respostas mais semelhantes a ela. A seleção das respostas mais semelhantes é feita através do algoritmo KNN (Seção 3.5), que seleciona as  $k$  respostas mais similares baseada na medida do cosseno (Seção 3.3).

O valor de  $k$  será denominado de fator de ponderação e nos experimentos serão feitos testes com esses fatores variando de 1 a 30. Isso significa que para  $k=3$  a ponderação da nota será feita com base nas três repostas mais próximas.

O cálculo da nota ponderada será feito das seguinte forma, seja  $r$  a resposta que se deseja ponderar,  $n$  sua nota,  $X_k = \{x_1, x_2, \dots, x_k\}$  o conjunto das  $k$  respostas mais próximas a  $r$  e  $N_k = \{n_1, n_2, \dots, n_k\}$  o conjunto das notas das  $k$  respostas mais próximas a  $r$ . Então, a nota ponderada de  $r$ ,  $NP_r$ , será dada por:

$$NP_r = n + \frac{\sum_{i=1}^k (n_i - n) * sim(r, x_i)}{k} \quad (7.1)$$

A ponderação poderá ser tanto para mais quanto para menos pontos. Essa variação será determinada de acordo com as notas dos vizinhos mais próximos. Caso as  $k$  notas desses vizinhos mais próximos sejam iguais a nota a ser ponderada, essa última não sofrerá ponderação.

## **7.5 Documento representante de um grupo**

O documento representante de um grupo é aquele que possui a maioria das características em comum com os demais.

Para encontrar esse documento representante é gerado um documento discriminante (OLIVEIRA et al., 2010) desse grupo. Nesse trabalho o documento discriminante é um documento artificial que possui todos os termos dos demais documentos do grupo. Então para cada documento do grupo é calculada sua similaridade com o documento discriminante. O documento mais similar ao documento discriminante é eleito o documento representante do grupo.

# Capítulo 8

## Experimentos

Nos capítulos anteriores, foram apresentadas as ferramentas desenvolvidas e as técnicas utilizadas para realizar a preparação das bases de dados e os experimentos.

Neste capítulo estão descritos os experimentos realizados para identificar os melhores parâmetros para correção Semiautomática das questões de Língua Portuguesa do vestibular da UFES com as pontuações dos avaliadores oficiais. Também são apresentados os resultados e discussões sobre esses resultados.

Nos experimentos buscamos melhorar a predição das notas de forma a obter uma aproximação com as notas atribuídas pelos dois avaliadores humanos. Não foi buscado melhorar o desempenho para um avaliador ou uma questão específica, pois o objetivo desses experimentos foi buscar uma proposta de predição que se adeque independente de questão ou avaliador.

Porém, caso se deseja medir o potencial da proposta para uma questão específica com um perfil de avaliador, isso poderia ser feito.

### 8.1 Experimento 1

O primeiro experimento com as bases de respostas, detalhadas na Seção 7.1, foi realizado através do agrupamento das respostas em 22 grupos diferentes, conforme procedimento descrito na Seção 5.1. A escolha dos 22 grupos segue a ideia do sistema NADID proposto por

(OLIVEIRA et al., 2010), onde é sugerida a quantidade de grupos igual à quantidade de notas diferentes + 1. No caso do vestibular as notas para questões discursivas variam de 0 a 2 totalizando 21 notas diferentes, acrescentando mais um grupo chegamos ao total de 22 grupos. Conforme descrito na Seção 7.3, a escala de notas dos avaliadores foi alterada de 0 a 2 para 0 a 10.

Após o agrupamento, para cada grupo formado foi selecionada a primeira resposta desse grupo e encaminhada ao professor para sua correção. Então, a nota dada pelo professor é reproduzida aos demais elementos do grupo.

As Tabelas 8.1 e 8.2 apresentam os resultados obtidos para cada base de dados e cada nota de avaliador. A primeira e a segunda coluna indicam, respectivamente, a base utilizada e os avaliadores analisados, a terceira coluna apresenta a média das diferenças de nota entre o avaliador Semiautomático e o avaliador humano. A quarta coluna apresenta o desvio padrão das diferenças de nota entre o avaliador Semiautomático e o avaliador humano. A última linha da tabela apresenta a média dos valores das coluna 3 e 4.

Base	Gabarito	Média das Diferenças	Desvio padrão das Diferenças
Q1	Avaliador1	3,16	3,05
Q2	Avaliador1	1,64	1,60
Q3	Avaliador1	1,92	1,85
Q4	Avaliador1	2,55	2,26
Q5	Avaliador1	3,05	4,24
Média		2,46	2,60

Tabela 8.1: Resultados do Experimento 1 para o Avaliador1.

Base	Gabarito	Média das Diferenças	Desvio padrão das Diferenças
Q1	Avaliador2	2,30	2,29
Q2	Avaliador2	2,88	2,21
Q3	Avaliador2	2,18	2,15
Q4	Avaliador2	1,77	1,48
Q5	Avaliador2	3,53	4,39
Média		2,53	2,50

Tabela 8.2: Resultados do Experimento 1 para o Avaliador2.

O resultado final de cada experimento será dado pela média dos resultados do Avaliador1 e do Avaliador2 apresentados na última linha de cada tabela.

Então, para o Experimento 1, o resultado final será a média dos resultados apresentados na última linha das Tabelas 8.1 e 8.2. Sendo a média das diferenças igual a 2,50 e o desvio padrão das diferenças igual a 2,55. Esses resultados serão usados para comparar os desempenhos dos experimentos e o objetivo em cada experimento será reduzir esses valores. Além disso, esses resultados finais de cada experimento indicam o quão próximo cada experimento está do desempenho dos avaliadores humanos (Seção 7.3).

### 8.1.1 Discussões

Ao analisar a Tabela 8.1, é possível identificar que os valores da coluna 3 para as questões Q1 e Q5 são os mais elevados. Analisando o motivo do valor apresentado para Questão 1, foi identificado que o grupo 18, que é um dos 22 grupos formados pelo processo de clusterização, é constituído por respostas com as notas: 0; 10; 10; 10; 10; 10; 9; 10; 2,5; 9; 10; 10; 8. Como o algoritmo seleciona a primeira nota de cada grupo para o avaliador corrigir e replicar aos demais elementos do grupo, nesse grupo foi selecionada uma resposta cuja nota é zero, levando as demais notas do grupo a serem pontuadas com zero também.

Se o elemento selecionado pelo algoritmo fosse a segunda nota do grupo 18, por exemplo, o erro médio seria de 1,92 e não 3,16 apresentado na Tabela 8.1. Isso mostra que a seleção

de um elemento representativo no grupo, para ser corrigido pelo professor, pode melhorar o resultado. A identificação do elemento representante de um grupo foi abordada na Seção 7.5 e será utilizada pelo Experimento 2.

Nesse mesmo exemplo de agrupamento, cabe uma análise mais aprofundada, pois existe uma resposta com nota zero que foi agrupada junto às outras respostas com notas bem maiores. Então, foi verificado que o Avaliador2 atribuiu cinco pontos à essa resposta que recebeu zero pelo Avaliador1. E, como ocorreu uma diferença considerável entre as notas dos avaliadores, essa resposta foi submetida a um novo avaliador, que por sua vez atribuiu nota sete. Isso mostra que há discordância entre os avaliadores humanos.

Além disso, se considerarmos que o Avaliador1 pontuou incorretamente essa resposta, pois dois avaliadores divergiram dele e o algoritmo classificou essa resposta como similar às outras respostas do grupo. Podemos sugerir que o Avaliador1 apresentou discordância nos seus critérios de correção usado nas respostas que avaliou desse grupo.

Esse caso indicou que a ferramenta proposta pode auxiliar na fiscalização dos avaliadores humanos, sugerindo respostas a serem analisadas ou reavaliadas. Esse auxílio pode ser feito através da identificação de grupos de respostas semelhantes que receberam notas divergentes. Essa fiscalização automática pode auxiliar no combate à fraude em concursos, por exemplo.

As respostas do grupo 18 podem ser visualizadas na sua forma original no Apêndice A. Ao analisar essas respostas é possível observar que o algoritmo agrupou corretamente as que falam do mesmo tema: "Kitty aos 22: divertimento". Se avaliadores diferentes avaliassem temas diferentes, o algoritmo poderia auxiliar nessa separação de temas de acordo com os avaliadores. Outra consideração é que não foi necessário separar as respostas de temas diferentes para processá-las.

Com relação à questão Q5, foi observado que as notas do Avaliador1, para essa questão, estavam nos extremos, ou seja, 75% de suas notas são maiores ou iguais a 9 e 20,5% são zero, então qualquer erro tem um impacto grande no resultado. Apesar da média das diferenças das notas dessa questão ser pior em relação às outras questões, o número de notas iguais entre o avaliador automático e o Avaliador1 foi superior ao das outras questões, ou seja, ele foi mais preciso em sua avaliação.

O mesmo cenário ocorreu na questão Q5 do Avaliador2, onde mais de 90% das notas são máximas (10) ou são mínima (0) apresentando uma média das diferenças ruim embora tenha apresentado boa precisão.

Outro ponto interessante desse experimento foi que a média das diferenças alcançada na questão Q2 é a que mais diverge entre as tabelas. Na Tabela 8.1 seu valor foi 1,64 e na Tabela 8.2 foi 2,88. Ao analisar o motivo dessa divergência foi verificado que as notas dadas pelos avaliadores 1 e 2 divergem mais nessa questão, conforme analisado na Seção 7.3. Acontece o mesmo com a questão Q4, onde a diferença entre as notas dos avaliadores também é grande.

## 8.2 Experimento 2

O segundo experimento modificou a etapa de seleção de respostas. Nessa etapa uma resposta de cada grupo é escolhida para o professor avaliar. Neste experimento, ao invés de selecionar a primeira resposta do grupo, foi selecionada a resposta que representa as demais do grupo, ou a mais representativa. Isso faz com que respostas com padrão menos semelhante ao grupo não sejam selecionadas. A resposta que o professor avalia tem sua nota replicada às demais presentes no grupo, conforme detalhado na Seção 5.1.

As Tabelas 8.3 e 8.4 apresentam os resultados. Os valores entre parênteses presentes nas tabelas mostram o valor da redução ou aumento do erro em relação as tabelas do experimento anterior.

Base	Gabarito	Média das Diferenças	Desvio padrão das Diferenças
Q1	Avaliador1	1,98 (-1,18)	2,11 (-0,94)
Q2	Avaliador1	1,88 (0,24)	1,77 (0,17)
Q3	Avaliador1	1,21 (-0,71)	1,59 (-0,26)
Q4	Avaliador1	2,75 (0,20)	2,29 (0,03)
Q5	Avaliador1	3,03 (-0,02)	4,25 (0,01)
Média		2,17 (-0,29)	2,40 (-0,20)

Tabela 8.3: Resultados do Experimento 2 para o Avaliador1.

Base	Gabarito	Média das Diferenças	Desvio padrão das Diferenças
Q1	Avaliador2	1,63 (-0,67)	2,24 (-0,05)
Q2	Avaliador2	2,34 (-0,54)	2,06 (-0,15)
Q3	Avaliador2	1,25 (-0,93)	1,73 (-0,42)
Q4	Avaliador2	2,04 (0,27)	1,93 (0,45)
Q5	Avaliador2	2,67 (-0,86)	3,84 (-0,55)
Média		1,99 (-0,54)	2,36 (-0,14)

Tabela 8.4: Resultados do Experimento 2 para o Avaliador2.

O resultado final de cada experimento será dado pela média dos resultados do Avaliador1 e do Avaliador2 apresentados na última linha de cada tabela.

Então, para o Experimento 2, o resultado final será a média dos resultados apresentados na última linha das Tabelas 8.3 e 8.4. Sendo então a média das diferenças igual a 2,08 e o desvio padrão das diferenças igual a 2,38. Com relação ao Experimento 1 a redução foi 16,63% para a média das diferenças e 6,67% para o desvio padrão das diferenças.

### 8.2.1 Discussões

A questão levantada no Experimento 1, de que a questão Q1 do Avaliador1 poderia ter seu desempenho melhorado com a seleção do melhor elemento do grupo, é comprovada. Pois, houve uma redução da média das diferenças de 3,16, no Experimento 1, para 1,98, no Experimento 2, o que corresponde a uma variação de 37,34%. E com relação ao desvio padrão, a variação foi de 30,82%, indo de 3,05, para 2,11, no Experimento 2.

Com relação aos resultados da questão Q4, cabe analisar por que não houve melhora, tanto para o Avaliador1 quanto para o Avaliador2.

Buscando entender esse fato que aconteceu com a avaliação da questão Q4 usando o gabarito do Avaliador1, foi identificado que o erro ocorreu nos grupos 15 e 16. Esses grupos foram representados pelas tabelas 8.5 e 8.6.

Nas Tabelas 8.5 e 8.6, a primeira coluna corresponde ao código do grupo, a segunda coluna possui as notas dadas pelo avaliador 1 nas respostas desse grupo, a terceira e a quinta coluna possuem um X para representar a resposta selecionada para correção do avaliador humano nos experimentos 1 e 2 respectivamente e a quarta e a sexta coluna representam a nota dada pelo avaliador automático nos experimentos 1 e 2, respectivamente.

A última linha das Tabelas 8.5 e 8.6 apresenta a média das diferenças para o grupo 15 e para o grupo 16 no Experimento 1 e no Experimento 2. Ressaltando que a média das diferenças só considera notas dadas automaticamente, não sendo considerados nas notas marcadas com X.

Grupo	Avaliador1	Seleção Exp1	Nota Exp1	Seleção Exp2	Nota Exp2
15	4,5	X	4,5		9
15	4		4,5		9
15	9		4,5		9
15	9		4,5	X	9
15	5		4,5		9
15	2,5		4,5		9
15	5		4,5		9
Média das diferenças			2,08		4

Tabela 8.5: Análise do grupo 15 da questão Q4.

Grupo	Avaliador1	Seleção Exp1	Nota Exp1	Seleção Exp2	Nota Exp2
16	6	X	6		10
16	2,5		6		10
16	3,5		6		10
16	10		6	X	10
16	3		6		10
16	10		6		10
16	9		6		10
16	3,5		6		10
Média das diferenças			3,21		4,64

Tabela 8.6: Análise do grupo 16 da questão Q4.

Nas Tabelas 8.5 e 8.6 é possível observar que a escolha do primeiro elemento acabou ocasionando um resultado melhor que a escolha de um elemento mais representativo. O mesmo cenário foi observado na questão Q4 para o Avaliador2, mas com outros grupos.

Isso pode acontecer quando a formação de um grupo possui respostas com grande variação entre as notas, mas devido a alguma similaridade entre elas, essas respostas foram agrupadas no mesmo grupo. Espera-se que com o uso de técnicas de Recuperação de Informação seja possível eliminar a similaridade que não condiz com a pontuação e assim melhorar o resultado final.

### 8.3 Experimento 3

Esse experimento apresenta os resultados com o uso da técnica de retirada de *stopwords*, detalhada na Seção 3.1, juntamente com a técnica usada no Experimento 2. As Tabelas 8.7 e 8.8 apresentam esses resultados.

Base	Gabarito	Média das Diferenças	Desvio padrão das Diferenças
Q1	Avaliador1	2,00 (0,02)	1,98 (-0,13)
Q2	Avaliador1	1,61 (-0,27)	1,72 (-0,05)
Q3	Avaliador1	1,34 (0,13)	1,71 (0,12)
Q4	Avaliador1	3,01 (0,26)	2,50 (0,21)
Q5	Avaliador1	0,96 (-2,07)	2,62 (-1,63)
Média		1,78 (-0,39)	2,11 (-0,29)

Tabela 8.7: Resultados do Experimento 3 para o Avaliador1.

Base	Gabarito	Média das Diferenças	Desvio padrão das Diferenças
Q1	Avaliador2	1,63 (0,00)	2,12 (-0,12)
Q2	Avaliador2	2,84 (0,50)	2,65 (0,59)
Q3	Avaliador2	1,39 (0,14)	1,92 (0,19)
Q4	Avaliador2	1,89 (-0,15)	1,94 (0,01)
Q5	Avaliador2	1,89 (-0,78)	3,56 (-0,28)
Média		1,93 (-0,06)	2,44 (0,08)

Tabela 8.8: Resultados Experimento 3 para o Avaliador2.

O resultado final de cada experimento será dado pela média dos resultados do Avaliador1 e do Avaliador2 apresentados na última linha de cada tabela.

Então, para o Experimento 3, o resultado final será a média dos resultados apresentados na última linha das Tabelas 8.7 e 8.8. Sendo então a média das diferenças igual a 1,86 e o desvio padrão das diferenças igual a 2,28. Com relação ao Experimento 2 a redução foi 10,82% para a média das diferenças e 4,41% para o desvio padrão das diferenças.

## 8.4 Experimento 4

Esse experimento apresenta os resultados com o uso da técnica de *stemming*, detalhada na Seção 3.1, juntamente com as técnicas usadas nos experimentos 2 e 3. As Tabelas 8.9 e

8.10 apresentam esses resultados.

Base	Gabarito	Média das Diferenças	Desvio padrão das Diferenças
Q1	Avaliador1	1,86 (-0,14)	1,68 (-0,30)
Q2	Avaliador1	1,63 (0,02)	1,69 (-0,03)
Q3	Avaliador1	1,40 (0,06)	1,65 (-0,06)
Q4	Avaliador1	2,31 (-0,70)	1,98 (-0,52)
Q5	Avaliador1	1,38 (0,42)	3,18 (0,56)
Média		1,72 (-0,06)	2,04 (-0,07)

Tabela 8.9: Resultado do Experimento 4 para o Avaliador1.

Base	Gabarito	Média das Diferenças	Desvio padrão das Diferenças
Q1	Avaliador2	1,44 (-0,19)	1,85 (-0,27)
Q2	Avaliador2	2,64 (-0,20)	1,97 (-0,68)
Q3	Avaliador2	1,43 (0,04)	1,87 (-0,05)
Q4	Avaliador2	1,55 (-0,34)	1,60 (-0,34)
Q5	Avaliador2	0,90 (-0,99)	2,35 (-1,21)
Média		1,59 (-0,34)	1,93 (-0,51)

Tabela 8.10: Resultado do Experimento 4 para o Avaliador2.

O resultado final de cada experimento será dado pela média dos resultados do Avaliador1 e do Avaliador2 apresentados na última linha de cada tabela.

Então, para o Experimento 4, o resultado final será a média dos resultados apresentados na última linha das Tabelas 8.9 e 8.10. Sendo então a média das diferenças igual a 1,66 e o desvio padrão das diferenças igual a 1,99. Com relação ao Experimento 3 a redução foi 10,78% para a média das diferenças e 12,75% para o desvio padrão das diferenças.

## 8.5 Experimento 5

Esse experimento apresenta os resultados com o uso da técnica de n-grama, mencionada na Seção 3.1. Essa técnica foi utilizada para  $n = 2$  e  $n = 3$ , juntamente com as técnicas dos experimentos 2, 3 e 4. As Tabelas 8.11 e 8.12 apresentam esses resultados para bigrama ( $n = 2$ ).

Base	Gabarito	Média das Diferenças	Desvio padrão das Diferenças
Q1	Avaliador1	2,64 (0,78)	2,40 (0,72)
Q2	Avaliador1	1,83 (0,20)	2,27 (0,58)
Q3	Avaliador1	1,30 (-0,10)	1,75 (0,10)
Q4	Avaliador1	2,76 (0,45)	2,05 (0,07)
Q5	Avaliador1	2,86 (1,48)	4,29 (1,11)
Média		2,28 (0,56)	2,55 (0,51)

Tabela 8.11: Resultado do Experimento 5 para o Avaliador1,  $n = 2$ .

Base	Gabarito	Média das Diferenças	Desvio padrão das Diferenças
Q1	Avaliador2	2,49 (1,05)	3,02 (1,17)
Q2	Avaliador2	2,79 (0,15)	2,34 (0,37)
Q3	Avaliador2	1,36 (-0,07)	1,83 (-0,04)
Q4	Avaliador2	1,74 (0,19)	1,83 (0,23)
Q5	Avaliador2	2,43 (1,53)	3,92 (1,57)
Média		2,16 (0,57)	2,59 (0,66)

Tabela 8.12: Resultados do Experimento 5 para o Avaliador2,  $n = 2$ .

Esse experimento terá dois resultados finais, uma para bigrama e outro para trigrama.

O resultado final para a utilização de bigrama será dado pela média dos resultados apresentados na última linha das Tabelas 8.11 e 8.12. Sendo então a média das diferenças igual a 2,22 e o desvio padrão das diferenças igual a 2,57. Com relação ao Experimento 4 o aumento no erro foi de 34,14% para a média das diferenças e 29,47% para o desvio padrão das diferenças. Esse experimento não melhorou os resultados alcançados no Experimento 4.

As Tabelas 8.13 e 8.14 apresentam os resultados para trigrama ( $n = 3$ ).

Base	Gabarito	Média das Diferenças	Desvio padrão das Diferenças
Q1	Avaliador1	2,64 (0,78)	2,40 (0,72)
Q2	Avaliador1	1,83 (0,20)	2,27 (0,58)
Q3	Avaliador1	1,30 (-0,10)	1,75 (0,10)
Q4	Avaliador1	2,76 (0,45)	2,05 (0,07)
Q5	Avaliador1	2,86 (1,48)	4,29 (1,11)
Média		2,28 (0,56)	2,55 (0,51)

Tabela 8.13: Resultado do Experimento 5 para o Avaliador1,  $n = 3$ .

Base	Gabarito	Média das Diferenças	Desvio padrão das Diferenças
Q1	Avaliador2	3,08 (1,64)	3,16 (1,31)
Q2	Avaliador2	2,84 (0,20)	2,51 (0,54)
Q3	Avaliador2	1,29 (-0,14)	1,73 (-0,14)
Q4	Avaliador2	1,94 (0,39)	1,83 (0,23)
Q5	Avaliador2	3,10 (2,20)	4,30 (1,95)
Média		2,45 (0,86)	2,71 (0,78)

Tabela 8.14: Resultados do Experimento 5 para o Avaliador2,  $n = 3$ .

O resultado final para a utilização de trigrama será dado pela média dos resultados apresentados na última linha das Tabelas 8.13 e 8.14. Sendo então a média das diferenças igual a 2,37 e o desvio padrão das diferenças igual a 2,63. Com relação ao Experimento 4 o aumento no erro foi de 42,90% para a média das diferenças e 32,49% para o desvio padrão das diferenças. Esse experimento também não melhorou os resultados alcançados no Experimento 4.

### 8.5.1 Discussões

Nos experimentos 3, 4 e 5 foi observado que a questão Q3 apresentou um resultado ruim nos experimentos 3 e 4 e um resultado melhor no experimento 5, para os dois avaliadores,

ou seja, o desempenho da questão Q3 foi contrário ao desempenho médio.

Ao buscar as diferenças entre a questão Q3 e as outras questões, foi identificado que o enunciado da questão Q3 (seção 7.1.3) solicita ao aluno a reescrita de um trecho de um texto específico, enquanto as demais questões solicitam informações conceituais que não consideram a forma da escrita.

Considerando o fato de que a retirada de *stopwords* e a técnica de *stemming* alteram a forma como o texto é escrito, pois uma elimina algumas palavras, e a outra reduz as palavras aos seus radicais, retirando o plural de algumas dessas palavras, por exemplo.

Considerando também que a técnica de n-gramas, ao contrário das outras, mantém a forma de escrita do texto, pois une as palavras adjacentes. É possível sugerir que a retirada de *stopwords* e a técnica de *stemming* prejudicaram o resultado da questão Q3, fazendo com que as respostas de questão Q3 fossem agrupadas independente da forma que foram escritas.

Visando analisar essa hipótese, foi realizado um novo experimento com a base de respostas da questão Q3. Nesse novo experimento foi utilizado n-gramas para  $n=2$  e  $n=3$ , sem a retirada de *stopwords* e sem aplicação de *stemming*. A Tabela 8.15 apresenta os resultados para bigrama e a Tabela 8.16 apresenta os resultados para trigrama.

A primeira coluna da tabela apresenta a questão, a segunda o avaliador, a terceira e quarta coluna apresentam os resultados.

Base	Gabarito	Média das Diferenças	Desvio padrão das Diferenças
Q3	Avaliador1	1,32	1,55
Q3	Avaliador2	1,29	1,68
Média		1,31	1,62

Tabela 8.15: Resultados da questão Q3 utilizando bigrama, sem a retirada de *stopwords* e sem o uso de *stemming*.

Base	Gabarito	Média das Diferenças	Desvio padrão das Diferenças
Q3	Avaliador1	1,17	1,51
Q3	Avaliador2	1,18	1,51
Média		1,17	1,51

Tabela 8.16: Resultados da questão 3 utilizando trigrama, sem a retirada de *stopwords* e sem o uso de *stemming*.

É possível observar, através das Tabelas 8.15 e 8.16, que a questão Q3 obteve uma melhora em seus resultados para ambos avaliadores, comparado a todos os experimentos anteriores. Esse melhor resultado foi com a utilização de trigrama, sem a retirada de *stopwords* e sem o uso de *stemming* (Tabela 8.16).

Esse experimento mostrou que o uso de técnicas que modificam a forma que o texto é escrito pode influenciar negativamente os resultados onde a forma de escrita do texto é levada em consideração pelos avaliadores. Ou seja, em alguns casos, a retirada de *stopwords* e a utilização de *stemming* não trarão os melhores resultados. Já a utilização de n-gramas pode ser uma boa estratégia para esses cenários.

Porém, como o uso de n-gramas não obteve melhora no desempenho médio do algoritmo, ele não será acumulado para os próximos experimentos.

## 8.6 Experimento 6

Esse experimento apresenta os resultados com o uso de uma técnica específica da área de predição de notas, que é a ponderação de notas, descrita na Seção 7.4. Ela foi aplicada juntamente com as técnicas usadas nos experimentos anteriores, com exceção da técnica de n-grama empregada no Experimento 5.

O primeiro passo foi experimentar qual fator de ponderação apresentará o melhor resultado. Assim será possível verificar se a ponderação apresentará bons resultados e qual será o melhor fator. Os fatores, definidos na Seção 7.4 variam de 1 a 30. O Gráfico 8.1 apresenta o resultado desses experimentos.

O eixo x do Gráfico 8.1 representa os fatores de ponderação usados. O valor zero no eixo x representa a não utilização da ponderação de notas. O eixo y representa a Média das diferenças de notas das 5 questões. As curvas do gráfico representam o Avaliador1, o Avaliador2 e a média dos dois avaliadores.

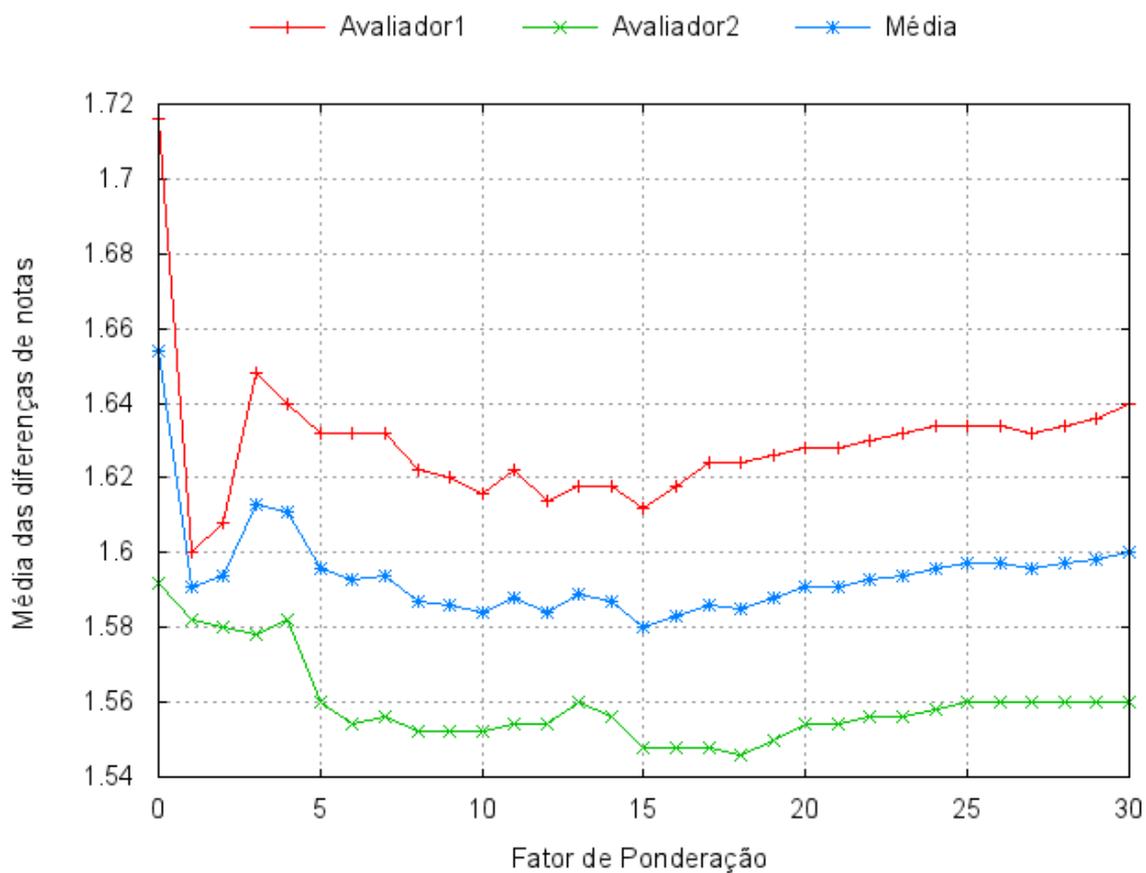


Figura 8.1: Variação do Fator de Ponderação.

É possível observar no Gráfico 8.1 que o fator 15 obteve o melhor resultado médio. As Tabelas 8.17 e 8.18 apresentam os resultados para esse fator de ponderação.

Base	Gabarito	Média das Diferenças	Desvio padrão das Diferenças
Q1	Avaliador1	1,83 (-0,03)	1,57 (-0,11)
Q2	Avaliador1	1,48 (-0,15)	1,54 (-0,15)
Q3	Avaliador1	1,27 (-0,13)	1,62 (-0,03)
Q4	Avaliador1	1,82 (-0,49)	1,50 (-0,48)
Q5	Avaliador1	1,66 (0,28)	2,76 (-0,42)
Média		1,61 (-0,11)	1,80 (-0,24)

Tabela 8.17: Resultados do Experimento 6 para o Avaliador1.

Base	Gabarito	Média das Diferenças	Desvio padrão das Diferenças
Q1	Avaliador2	1,44 (0,00)	1,64 (-0,21)
Q2	Avaliador2	2,24 (-0,40)	1,68 (-0,29)
Q3	Avaliador2	1,35 (-0,08)	1,85 (-0,02)
Q4	Avaliador2	1,44 (-0,11)	1,30 (-0,30)
Q5	Avaliador2	1,27 (0,37)	2,16 (-0,19)
Média		1,55 (-0,04)	1,73 (-0,20)

Tabela 8.18: Resultados do Experimento 6 para o Avaliador2.

O resultado final de cada experimento será dado pela média dos resultados do Avaliador1 e do Avaliador2 apresentados na última linha de cada tabela.

Então, para o Experimento 6, o resultado final será a média dos resultados apresentados na última linha das Tabelas 8.17 e 8.18. Sendo então a média das diferenças igual a 1,58 e o desvio padrão das diferenças igual a 1,77. Com relação ao Experimento 4 a redução foi 4,53% para a média das diferenças e 11,08% para o desvio padrão das diferenças. A comparação foi feita em relação ao Experimento 4, pois a técnica utilizada no Experimento 5 não foi acumulada nos experimentos.

## 8.7 Experimento 7

Esse experimento busca atingir a meta estabelecida de semelhança com os avaliadores humanos aumentando o número de grupos que são formados. Esse aumento tem influencia direta sobre o esforço desse professor. Pois ele deverá corrigir uma questão de cada grupo. Dessa forma, quanto maior o número de grupos mais questões deverão ser corrigidas manualmente.

O gráfico da Figura 8.2 mostra a variação do número de grupos para o experimento realizado na seção 8.6. O eixo x desse gráfico representa as quantidades de grupos consideradas, que varia de 22 a 60. O eixo y representa a Média das diferenças de notas das 5 questões. As curvas do gráfico representam o Avaliador1, o Avaliador2, a Média dos dois avaliadores e a curva objetivo que representa a média das diferenças entre os dois avaliadores humanos.

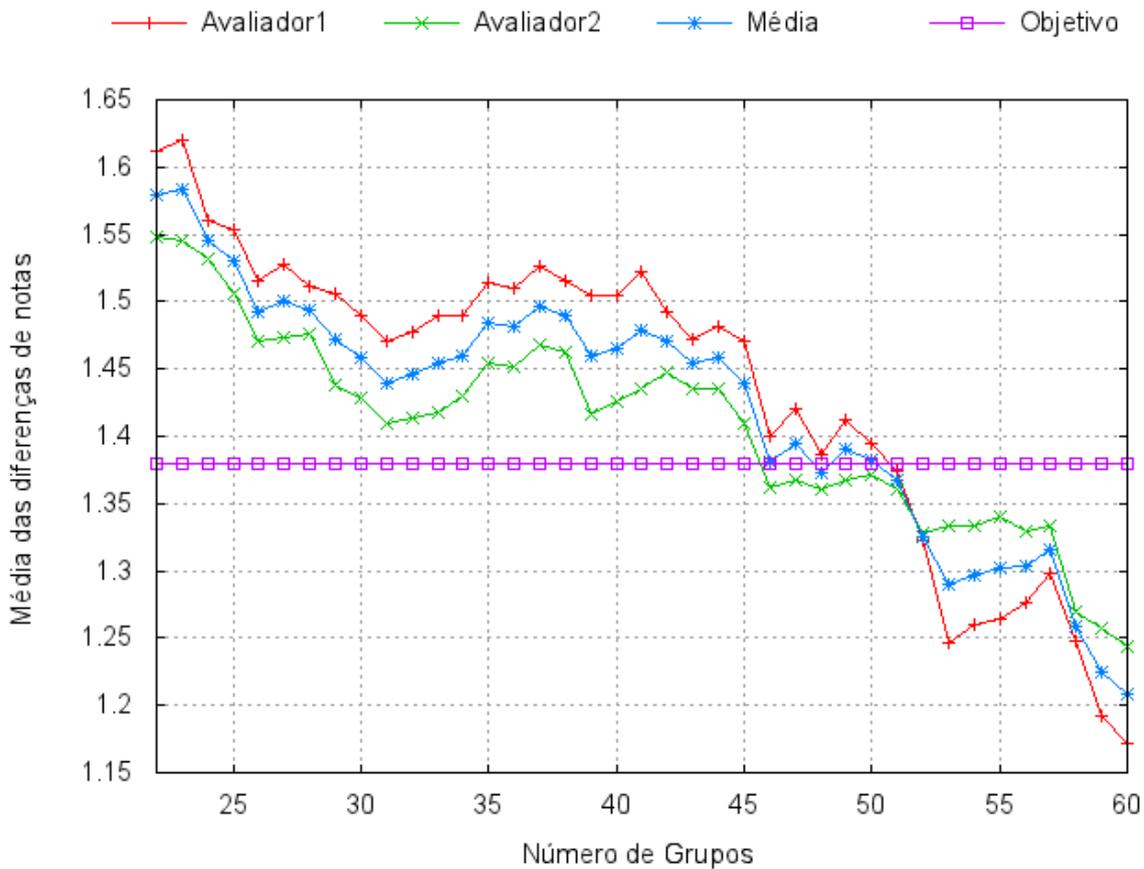


Figura 8.2: Variação do número de grupos.

Nesse gráfico é possível observar que a curva Objetivo é ultrapassada por todas as outras curvas quando o número de grupos é igual a 51. Isso significa que nesse ponto a Média das diferenças de notas entre o algoritmo e os avaliadores humanos é menor que a Média das diferenças entre os dois avaliadores humanos. Ou seja, o preditor de notas Semiautomático teve o erro médio de pontuação por questão inferior ao de um avaliador humano.

Para isso, foi necessário uma amostra de treino de 55,43% das respostas para atingir essa semelhança de um avaliador humano. E o esforço na correção de todas as respostas foi reduzido em 44,56%.

Além disso, é possível observar no gráfico que o par Média das diferenças de notas versus quantidade de grupo nem sempre caminham juntos, ou seja, nem sempre um aumento na amostra de treino representará um resultado melhor. Isso pode ser observado no intervalo de 31 a 37 grupos ou no intervalo de 53 a 58 grupos.

Outra observação é que com 46 grupos, 50% das respostas o algoritmo se tornou semelhante ao Avaliador2.

Além da média das diferenças ter sido alcançada, a Figura 8.3 mostra que o desvio padrão, com 51 grupos, também ficou abaixo dos avaliadores humanos.

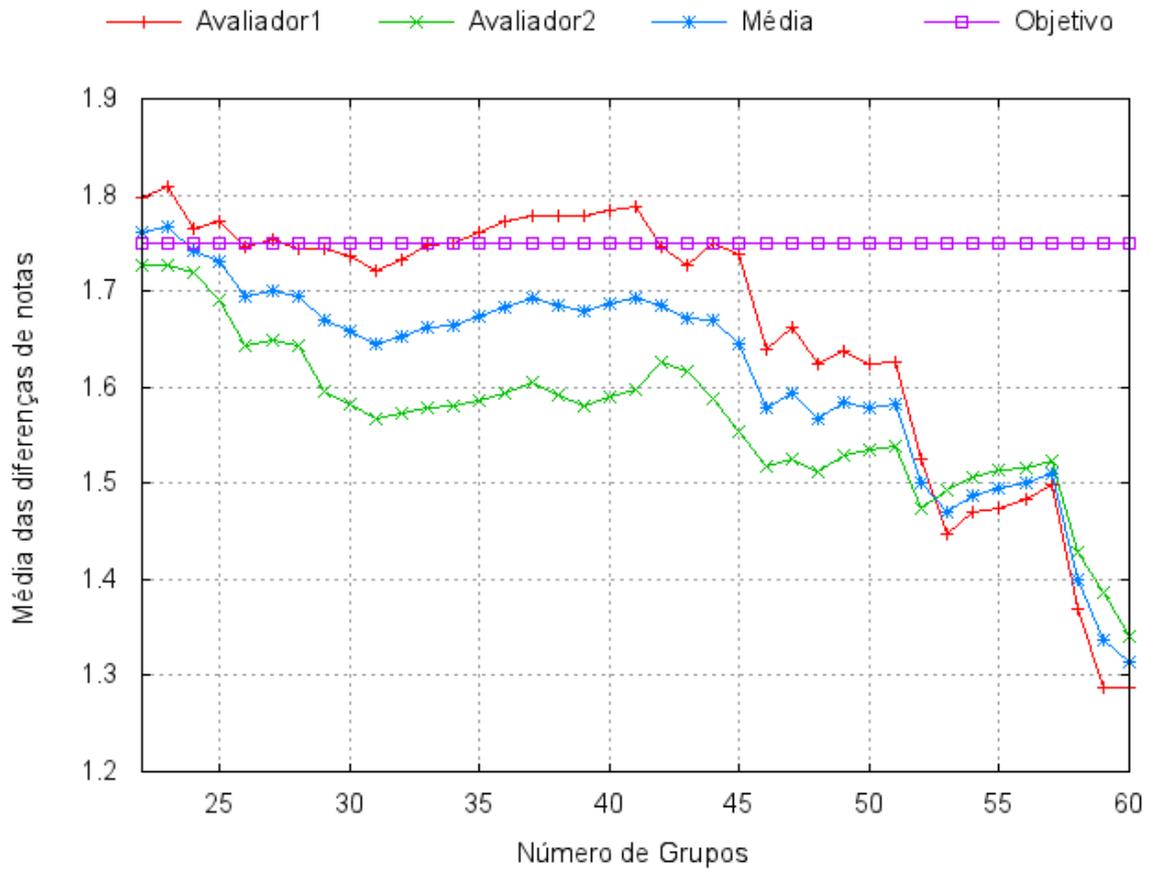


Figura 8.3: Variação do Fator de Ponderação.

Base	Gabarito	Média das Diferenças	Desvio padrão das Diferenças
Q1	Avaliador1	1,38 (-0,45)	1,32 (-0,25)
Q2	Avaliador1	1,26 (-0,22)	1,47 (-0,07)
Q3	Avaliador1	0,92 (-0,35)	1,21 (-0,41)
Q4	Avaliador1	1,60 (-0,22)	1,50 (0,00)
Q5	Avaliador1	1,71 (0,05)	2,63 (-0,13)
Média		1,37 (-0,24)	1,63 (-0,17)

Tabela 8.19: Resultados do Experimento 7 para o Avaliador1.

Base	Gabarito	Média das Diferenças	Desvio padrão das Diferenças
Q1	Avaliador2	0,99 (-0,45)	1,05 (-0,59)
Q2	Avaliador2	2,17 (-0,07)	1,67 (-0,01)
Q3	Avaliador2	0,85 (-0,50)	1,23 (-0,62)
Q4	Avaliador2	1,37 (-0,07)	1,21 (-0,09)
Q5	Avaliador2	1,42 (0,15)	2,53 (0,37)
Média		1,36 (-0,19)	1,54 (-0,19)

Tabela 8.20: Resultados do Experimento 7 para o Avaliador2.

O resultado final para o Experimento 7 será a média dos resultados apresentados na última linha das Tabelas 8.19 e 8.20. Sendo então a média das diferenças igual a 1,37 e o desvio padrão das diferenças igual a 1,59. Com relação ao Experimento 6 a redução foi 13,29% para a média das diferenças e 10,17% para o desvio padrão das diferenças.

Outra comparação a ser apresentada é a do Experimento 7 com o objetivo almejado. Esse objetivo é a variação das notas entre os avaliadores apresentada na Seção 7.3. A Tabela 8.21 apresenta o resultado para o Avaliador1 e a Tabela 8.22 apresenta o resultado para o Avaliador2.

Os valores entre parênteses são os resultados do objetivo. E os valores em negrito são os resultados alcançados pelo preditor Semiautomático de notas que foram melhores que os avaliadores humanos.

Base	Gabarito	Média das Diferenças	Desvio padrão das Diferenças
Q1	Avaliador1	<b>1,38 (1,46)</b>	<b>1,32 (1,75)</b>
Q2	Avaliador1	<b>1,26 (2,33)</b>	<b>1,47 (1,95)</b>
Q3	Avaliador1	0,92 (0,26)	1,21 (0,72)
Q4	Avaliador1	<b>1,60 (2,16)</b>	<b>1,50 (1,76)</b>
Q5	Avaliador1	1,71 (0,68)	2,63 (1,85)
Média		<b>1,37 (1,38)</b>	<b>1,63 (1,75)</b>

Tabela 8.21: Resultados do Experimento 7 para o Avaliador1 comparado ao objetivo.

Base	Gabarito	Média das Diferenças	Desvio padrão das Diferenças
Q1	Avaliador2	<b>0,99 (1,46)</b>	<b>1,05 (1,75)</b>
Q2	Avaliador2	<b>2,17 (2,33)</b>	<b>1,67 (1,95)</b>
Q3	Avaliador2	0,85 (0,26)	1,23 (0,72)
Q4	Avaliador2	<b>1,37 (2,16)</b>	<b>1,21 (1,76)</b>
Q5	Avaliador2	1,42 (0,68)	2,53 (1,85)
Média		<b>1,36 (1,38)</b>	<b>1,54 (1,75)</b>

Tabela 8.22: Resultados do Experimento 7 para o Avaliador2 comparado ao objetivo.

É possível observar nas Tabelas 8.21 e 8.22 que em média o desempenho do preditor Semiautomático de notas, dado um número de amostras, consegue ser semelhante aos avaliadores humanos. E, nas questões Q1, Q2 e Q3 o preditor apresentou melhores resultados que os avaliadores humanos, tanto na média das diferenças quanto no desvio padrão das diferenças. Ou seja, foi possível reduzir o número de erros por resposta e reduzir a variação desse erro.

# Capítulo 9

## Conclusões e Trabalhos Futuros

### 9.1 Conclusão

Este trabalho apresentou um algoritmo de predição de notas Semiautomática para respostas discursivas. Ele foi testado com as questões de Língua Portuguesa do vestibular da UFES e comparado à variação de pontuação apresentada pelos avaliadores oficiais. Nos experimentos realizados o algoritmo apresentou desempenho semelhante ao dos avaliadores. Se fosse aplicada, a redução no esforço da atividade de avaliação seria de 44,56%.

Os resultados apresentados mostram que essa proposta pode ser comparada ao desempenho de um avaliador humano, o que torna possível seu emprego no auxílio a correções de respostas discursivas.

Esse algoritmo de predição é um passo inicial no desenvolvimento de um sistema de predição de notas Semiautomático para respostas discursivas que poderá ser usado no dia-a-dia do docente ou no auxílio do trabalho de correção de respostas discursivas de concursos e vestibulares, tanto na predição da nota oficial quanto na validação de notas de revisores. Além disso, poderá ser aplicado também no treinamento de avaliadores.

No trabalho foi observado que as técnicas de *stemming* e retirada de *stopwords* obtiveram bons resultados quando usadas na análise de respostas de questões conceituais, onde não importa a forma de escrita da resposta e sim seu conteúdo. Sendo que, em respostas de

questões cuja forma de escrita era importante, como uso de plural, por exemplo, os resultados do uso dessas técnicas foram ruins.

Diferente das propostas dessa área, foi apresentada uma ferramenta de Visualização de Informação que auxiliará o professor a conhecer o desempenho dos alunos cuja respostas não foram avaliadas por ele. Essa ferramenta ajudará à ampliar nosso entendimento sobre determinado cenário. Mostramos um dos gráficos gerados por ela com os dados do ambiente virtual de aprendizagem *Moodle* para visualizar as informações dos alunos.

Com o uso dos mapas construídos com o apoio de técnicas de Visualização de Informação, observamos que o professor consegue identificar diferentes situações de suas turmas e, a partir daí, interferir mais especificamente sobre elas. Isto é possível porque a análise dos resultados ficou mais rápida e a comparação de resultados, mais simples.

Foi desenvolvido também um *plugin* que integrou essas propostas ao *Moodle* tornando possível o uso dessas propostas por docentes que já utilizam esse ambiente e possuem cursos em andamento. Além dessas, o *plugin* também integrou outras soluções de apoio ao trabalho docente já existentes, como o NAD proposto por Oliveira (2013).

## 9.2 Trabalhos futuros

Sugerimos como trabalhos futuros, a partir deste trabalho, os seguintes:

- Aplicar a proposta de predição Semiautomática de notas deste trabalho com outras bases de dados e com outras disciplinas para comparar os resultados alcançados entre bases diferentes visando verificar a aplicabilidade e acerto da predição com maior significância estatística;
- Como este trabalho não considera o conteúdo das frases ou significado das palavras é possível utilizar novas propostas que contemple a análise semântica sobre a base de dados utilizada neste trabalho, com objetivo de melhorar os resultados alcançados, ou seja, reduzir o erro na predição ou o esforço do professor.

- 
- Sobre a proposta de visualização: estudar novas formas de visualização para que seja possível apresentar mais dimensões dos dados, apresentar a evolução e a tendência dos discentes bem como dar a possibilidade do aluno ter uma visão de seu rendimento para que ele possa se autoavaliar.

# Referências Bibliográficas

ATTALI, Y.; BURSTEIN, J. Automated essay scoring with e-rater® v. 2. *The Journal of Technology, Learning and Assessment*, v. 4, n. 3, 2006.

ÁVILA, R. L. de; SOARES, J. M. Concepção de ferramenta de apoio à correção de questões dissertativas com base na adaptação de algoritmos de comparação e busca textual combinados com técnicas de pré-processamento de textos. *RENOTE*, v. 10, n. 3, 2012.

BAEZA-YATES, R.; RIBEIRO-NETO, B. et al. *Modern information retrieval*. New York: ACM Press Books, 2011.

BORSOI, I. C. F. Trabalho e produtividade: saúde e modo de vida de docentes de instituições públicas de ensino superior. *Cadernos de Psicologia Social do Trabalho*, Universidade de São Paulo, v. 15, n. 1, p. 81–100, 2012.

BRIDGEMAN, B.; TRAPANI, C.; ATTALI, Y. Comparison of human and machine scoring of essays: Differences by gender, ethnicity, and country. *Applied Measurement in Education*, Taylor & Francis Group, v. 25, n. 1, p. 27–40, 2012.

BURSTEIN, J. et al. Automated scoring using a hybrid feature identification technique. *Proceedings of the 17th international conference on Computational linguistics-Volume 1*, Association for Computational Linguistics, p. 206–210, 1998.

BURSTEIN, J.; LEACOCK, C.; SWARTZ, R. Automated evaluation of essays and short answers. *Proceedings of the 5th CAA Conference*, © Loughborough University, 2001.

CALDAS, V. M.; FAVERO, E. L. Uma ferramenta de avaliação automática para mapas conceituais como auxílio ao ensino em ambientes de educação a distância. *Anais do Simpósio Brasileiro de Informática na Educação*, v. 1, n. 1, 2009.

CARLOTTO, M. S. A síndrome de burnout e o trabalho docente. *Psicologia em estudo*, v. 7, n. 1, p. 21–29, 2002.

CHEN, C. *Information Visualization*. 2002. 1–4 p. (<http://www.pages.drexel.edu/~cc345/papers/ivs2002v1n1.pdf>).

FURNAS, G. W.; BUJA, A. Projection views: Dimensional inference through sections and projections. *Journal of Computational and Graphical Statistics*, Taylor & Francis Group, v. 3, n. 4, p. 323–353, 1994.

- HOFFMAN, P. E. *Table visualizations: a formal model and its applications*. Tese (Doutorado) — University of Massachusetts Lowell, 1999.
- INSELBERG, A. The plane with parallel coordinates. *The Visual Computer*, Springer-Verlag, v. 1, n. 2, p. 69–91, 1985.
- JOHNSON, B.; SHNEIDERMAN, B. Tree-maps: A space-filling approach to the visualization of hierarchical information structures. *Visualization, 1991. Visualization '91, Proceedings., IEEE Conference on*, p. 284–291, 1991.
- KARYPIS, G. *CLUTO - A clustering toolkit*. 2003. Dept. of Computer Science, University of Minnesota.
- KEIM, D. A.; KRIEGEL, H.-P. Visualization techniques for mining large databases: A comparison. *Knowledge and Data Engineering, IEEE Transactions on, IEEE*, v. 8, n. 6, p. 923–938, 1996.
- KENSKI, V. M. Aprendizagem mediada pela tecnologia. *Revista Diálogo Educacional, Curitiba*, v. 4, n. 10, p. 47–56, 2003.
- LUNA, E. Á. dos A. Avaliação em larga escala e produção textual: reflexões sobre o enem. *ESTUDOS EM EDUCAÇÃO E LINGUAGEM-ISSN: 2237-5880*, v. 1, n. 1, p. 48–56, 2012.
- MOISSA, B.; GASPARINI, I.; KEMCZINSKI, A. Análise do comportamento do aluno através de técnicas de web analytics e sua proposta de representação interativa em um ambiente e-learning. *Nuevas Ideas en Informática Educativa TISE*, 2013.
- OLIVEIRA, E. S. d. 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. 2010.
- OLIVEIRA, M. G. d. *Núcleos de Avaliações Diagnóstica e Formativa para Regulação da Aprendizagem de Programação*. Tese (Doutorado) — Tese de Doutorado. Centro Tecnológico, Universidade Federal do Espírito Santo, 2013.
- PAGE, E. B. Grading essays by computer: Progress report. *Proceedings of the Invitational Conference on Testing Problems*, Princeton, p. 87–100, 1967.
- PICKETT, R.; GRINSTEIN, G. Iconographic displays for visualizing multidimensional data. *Systems, Man, and Cybernetics, 1988. Proceedings of the 1988 IEEE International Conference on*, v. 1, p. 514–519, 1988.
- ROBERTSON, G. G.; MACKINLAY, J. D.; CARD, S. K. Cone trees: animated 3d visualizations of hierarchical information. *Proceedings of the SIGCHI conference on Human factors in computing systems*, p. 189–194, 1991.
- SALTON, G.; WONG, A.; YANG, C. A vector space model for automatic indexing. *Communications of the ACM, ACM*, v. 18 (11), p. 613–620, 1975.
- SANTOS, T. et al. Avaliação automática de questões conceituais discursivas. *IX Argentine Symposium on Artificial Intelligence-ASAI*, p. 128–138, 2007.

- SAXENA, S.; GUPTA, P. R. Automatic assessment of short text answers from computer science domain through pattern based information extraction. *Proceeding of ASCNT*, p. 109–118, 2009.
- SILVA, C. d. *Exploração de bases de dados de ambientes de Educação a Distância por meio de ferramentas de consulta apoiadas por Visualização de Informação*. Tese (Doutorado) — Tese de Doutorado. Instituto de Computação, Universidade Estadual de Campinas. <http://www.ic.unicamp.br/~celmar/tese> (20/Abril/2007), 2006.
- SILVA, C. G. da; INOUE, M. T.; MENDONÇA, P. J. C. de. CourseViewer – Ferramenta para visualização de catálogos de cursos universitários e históricos escolares. *Anais do VIII Simpósio Brasileiro de Sistemas de Informação*, p. 492–503, 2012.
- SWAYNE, D. F. et al. Ggobi: evolving from xgobi into an extensible framework for interactive data visualization. *Computational Statistics & Data Analysis*, Elsevier, v. 43, n. 4, p. 423–444, 2003.
- VILELA, R. F. et al. Scatedi: Sistema inteligente para avaliação de desempenho escolar em avaliações discursivas. *Anais do Workshop de Informática na Escola*, v. 1, n. 1, 2012.
- WANG, J.; BROWN, M. S. Automated essay scoring versus human scoring: A comparative study. *The Journal of Technology, Learning and Assessment*, v. 6, n. 2, 2007.
- WEBBER, C. G. et al. Facilitando a análise de dados educacionais através de ferramentas de visualização. *RENOTE*, v. 11, n. 3, p. 492–503, 2013.
- YAU, N. *Visualize this: the FlowingData guide to design, visualization, and statistics*. Indianapolis: John Wiley & Sons, 2011.

# Apêndice A

## Respostas discursivas do grupo 18 do Experimento 1

Abaixo seguem os textos das respostas discursivas agrupadas no grupo 18 no Experimento 1. Cada resposta é constituída de um código seguido do texto da resposta. Esse grupo 18 é formado pelas respostas descritas neste capítulo de acordo com a legenda na Tabela A.1. As respostas estão na mesma sequência apresentada no Experimento 1.

	Nota	Identificador
1	0	863.146-1
2	10	863.184-1
3	10	863.404-1
4	10	863.453-1
5	10	863.623-1
6	10	863.692-1
7	9	863.734-1
8	10	863.777-1
9	2,5	863.875-1
10	9	863.898-1
11	10	863.954-1
12	10	864.022-1
13	8	864.085-1

Tabela A.1: Legenda das respostas do grupo 18 formado no Experimento 1.

863.146-1

Em "Kitty aos 22: divertimento", de Reinaldo Santos Neves (2006), ocorrem os seguintes traços: a problematização das relações amorosas, em especial as femininas e homoeróticas, o ressurgimento do narrador clássico, cujo relato beira a oralidade; a multiplicação de textos de autoria e temática femininas, influencia do mercado editorial. Ao longo do texto de caráter retroco e um tanto quanto singelo é explanado de forma viva as relações amorosas, todos os intempéries, se questionamentos e inseguranças das relações interpessoais. O narrador declara todo o texto, busca a oralidade, envolvendo o leitor de uma forma simples e clara. Temática feminina, voltada para o público atual, cujos valores e discernimento sobre a vida estão sobrepujados na ênfase do bem estar e da satisfação pessoal, na busca de se encontrar, na busca da dependência de outros para se obter a própria felicidade. É bem certo que há influência do mercado editorial onde se escreve aquilo que o público gostaria de ler, aquilo que será aceito, de acordo com as pesquisas de mercado, de acordo com as estatísticas, voltadas aos valores dessa vida e a eterna insatisfação humana.

863.184-1

Na narrativa pós-moderna "kitty aos 22: divertimento", de Reinaldo Santos Neves, verifica-se a ocorrência, dentre outros dos traços citados, do cruzamento da literatura com a linguagem da mídia e da internet; da problematização das relações amorosas, em especial as femininas e homoeróticas; do ressurgimento do narrador clássico, cujo relato beira a oralidade; da hegemonia de personagens anônimos, comuns, anti-heróicos e da mistura de gêneros; literários ou não. Com o objetivo de atribuir verossimilhança à composição de sua obra, que narra o universo de Marica Catarina Leme, uma jovem do século XXI, o autor utiliza diversos recursos linguísticos, tais como as expressões "orkut" e "blog", gírias e observações, representativas do universo midiático e da internet; termos de baixo calão, diminutivos e apelidos, que conferem aspecto de oralidade ao texto. Há também a presença de um casal homoafetivo na trama, os jovens Bob e Benjy, trazendo a questão das relações amorosas homoerótica. Além disso, em oposição ao maniqueísmo, às divagações e idealizações de outrora, em "Kitty", o autor cria personagens comuns, com atitudes éticas e anti-éticas, como qualquer ser humano. A intertextualidade e a mistura de gênero também podem ser verificadas na obra, pois o narrador faz referência ao "Conto da Cinderela", a obras cinematográficas

e a estilos musicais modernos, como rock.

863.404-1

1) O romance *Kitty aos 22: divertimento*, de Reinaldo Santos Neves, narra o cotidiano de Kitty, uma jovem de vinte e dois anos e aborda alguns de seus relacionamentos. Há, por exemplo, a abordagem que é feita a seu relacionamento com a sua amiga Lu, que é moradora do bairro Jardim [inint]. Kitty, moradora do bairro Mata da Praia, deixa transparecer na narrativa que se considera superior a Lu, por ela morar em um local de condições econômicas inferiores ao seu. Isso se evidencia, quando ambas se interessam pelo mesmo rapaz, e Kitty não entende o suposto interesse dele por Lu. Desse triângulo amoroso, acontece o rompimento das relações de amizade entre ambas. Além disso, aborda-se também a relação homoafetiva de seu pai, a quem ela chama de Daddy e de seus dois amigos que morrem, vítimas de um assassinato, no início do enredo. Dessas abordagens feitas pelo autor, nota-se a problematização das relações amorosas, em especial as femininas e homoeróticas. Outra característica presente no livro é o cruzamento da literatura com a linguagem da mídia e da internet. O próprio nome da personagem central, Kitty, é característico de apelidos que se encontra na internet, em redes sociais, em substituição dos nomes dos usuários. A cidade de Vitória, onde se desenrola a trama, é chamada de Mic, a personagem Lu-ciana é chamada de Lu, caracterizando assim a linguagem jovem utilizada largamente nas redes sociais e transmitida na mídia através de novelas, propagandas e filmes.

863.453-1

”*Kitty aos 22: divertimento*”, de Reinaldo Santos Neves é uma obra de autor capixaba que reflete bem os hábitos sociais e a maneira como se relacionam os jovens ao seu tempo. A obra apresenta claros elementos de cruzamento da literatura com a linguagem da mídia e da internet. Esses cruzamentos ocorrem desde a apresentação e utilização de novas tecnologias no desenvolver da trama até, de forma mais objetiva, quando a protagonista se queixa de não ter recebido mensagem de determinado ”e-mail”, fazendo assim a troca do nome pelo endereço eletrônico sem prejuízo na identificação. Outro ponto importante é o cunho social envolvido na ficção. A obra faz um retrato social de um comportamento comum em maior ou menos grau. O descompromisso inicial da protagonista, sua postura, sua libertinagem e

anceios são os mesmos de muitas jovens com o mesmo perfil. O amadurecimento que passa a ser exigido é nada mais que o caminho social propôs-to na contemporaneidade. Percebe-se então que o texto tem sim cunho social no sentido de expor o comportamento cotidiano de forma absolutamente transparente.

863.623-1

01) Na obra *Kitty aos 22: divertimento*, de Reinaldo Santos Neves há a ocorrência de fatores que caracterizam a produção literária contemporânea. Dentre eles, podemos elencar o cruzamento da literatura com a linguagem da mídia e da internet, perceptível no uso de palavras como "e-mail" e "blog", próprias da linguagem de internet. A personagem de Kitty mantém um "blog", página na internet em que publica fotos e comentários sobre seu cotidiano – uma espécie de diário virtual. Há ainda o uso de "e-mails" e de outras ferramentas de comunicação por meio do computador, denotando a grande presença do uso de ferramentas tecnológicas na vida dos personagens. Podemos citar ainda, como elemento caracterizador da literatura contemporânea, a utilização hegemônica de personagens anônimas, comuns, e anti-heróicas. Kitty, personagem principal do romance, é uma jovem bonita, rica, fútil, materialista, que deseja ser sempre o centro de todas as atenções. Mummy, sua mãe, passa o dia ébria ou de ressaca. Phil é o padrasto que tem desejo sexual pela enteada, Kitty. Os amigos de Kitty – Deb, Jujuba – só querem se divertir e às vezes fazem uso de entorpecentes. Todas as personagens são apresentadas com suas qualidades e defeitos, seus medos, seus anseios, seus erros. São personagens que muitas vezes podem até ser comparadas com pessoas que conhecemos.

863.692-1

Português 1) Escolho a obra de Reinaldo Santos Neves, *Kitty aos 22: divertimento*; para indicar os dois traços sugeridos entre os elencados de nossa Literatura contemporânea por Flávio Carneiro. 1º - O cruzamento da literatura com a linguagem da mídia e da internet. O autor tem como sua personagem principal uma jovem de 22 anos e em seu texto busca a aproximação do leitor com o mundo retratado e vivido por sua personagem, o autor teve o cuidado de pesquisar como falam, vivem, curtem, os jovens. O texto vem escrito com uma linguagem que vemos em microblogs, aplicativos de mensagens instantâneas, linguagem de

mídia televisiva também – na parte em que aborda o evento de moda realizado na cidade onde se passa a história, até mesmo o apelido da personagem principal foi retirado da mania das jovens de apelidos em inglês. 2º - A problematização das relações amorosas No livro duas jovens se interessam pelo mesmo rapaz, o que gera conflito na amizade delas, além disso tem a personagem principal que atrai com sua beleza o desejo dos homens, entre eles um ex colega de faculdade que vive se declarando em busca de ser correspondido, mas o máximo que consegue é o olhar piedoso de Kitty, e por fim o trágico desfecho da história com a mocinha em perigo por causa de sua relação amorosa.

863.734-1

Português 1) Escolho a obra de Reinaldo Santos Neves, Kitty aos 22: divertimento; para indicar os dois traços sugeridos entre os elencados de nossa Literatura contemporânea por Flávio Carneiro. 1º - O cruzamento da literatura com a linguagem da mídia e da internet. O autor tem como sua personagem principal uma jovem de 22 anos e em seu texto busca a aproximação do leitor com o mundo retratado e vivido por sua personagem, o autor teve o cuidado de pesquisar como falam, vivem, curtem, os jovens. O texto vem escrito com uma linguagem que vemos em microblogs, aplicativos de mensagens instan-tâneas, linguagem de mídia televisiva também – na parte em que aborda o evento de moda realizado na cidade onde se passa a história, até mesmo o apelido da personagem principal foi retirado da mania das jovens de apelidos em inglês. 2º - A problematização das relações amorosas No livro duas jovens se interessam pelo mesmo rapaz, o que gera conflito na amizade delas, além disso tem a personagem principal que atrai com sua beleza o desejo dos homens, entre eles um ex colega de faculdade que vive se declarando em busca de ser correspondido, mas o máximo que consegue é o olhar piedoso de Kitty, e por fim o trágico desfecho da história com a mocinha em perigo por causa de sua relação amorosa.

863.777-1

1) Kitty aos 22: divertimento que representa bem, em muitos aspectos essa tendência do pós-modernismo: um panorama funcional contemporâneo, ou seja, que busca aproximar a personagem, o cenário e a linguagem do que se está escrevendo. Nesse sentido, Kitty aos 22 traz um cruzamento da literatura com a linguagem da mídia e da internet, pois o autor busca

informações para compor suas personagens em "blogs" e páginas da internet a fim de caracterizar o mundo jovial e moderno do livro, portanto é comum no romance deparar-se com expressões chulas e palavras de baixo calão, além de técnicas do meio "cibernético" como abreviação das palavras e uso do estrangeirismo, como exemplo temos "mic" para Vitória (onde se passa o romance) e "daddy" para papai. Em tempo, percebe-se também na obra a problematização das relações amorosas, em especial as femininas e homoeróticas, tendo vista que Reinaldo traz atona uma discussão filosófica sobre a relação de Kitty (a personagem) com Bruno, por exemplo, ou com Phil, não uma relação propriamente dita, mas o desejo que este alimenta. Já em se tratando de relação homoeróticas, observa-se que o romance inicia-se com um velório, de dois homossexuais amigos de Kittu, deixando transparecer uma discussão de temas polêmicos da sociedade.

863.875-1

Romance analisado: *Kitty aos 22: divertimento*. Traço I: O ressurgimento do narrador clássico, cujo o relato beira a oralidade. O autor, Reinaldo Santos Neves, usa de uma linguagem simples, acessível e direta. Isso provoca no leitor a impressão de estar em contato direto com o autor, recebendo deste, a mensagem, mantendo este leitor ligado nessa relação tão direta estabelecida entre ambos. Traço II: A mistura de gêneros, literários ou não. O autor Reinaldo Santos Neves, desprende-se de formalizar em um gênero, literário ou não, sua escrita. A multiplicidade genérica observada em sua obra nota-se na intenção do autor em expor de forma objetiva e simples as histórias da vida de uma jovem de vinte e dois anos de idade, para tanto, ele abre mão de formalizar o conteúdo, e dá a sua obra um tom generalesco, objetivando seu entendimento.

863.898-1

A análise da obra "*Kitty aos 22: divertimento*", de Reinaldo Santos Neves, permite identificar traços presentes no panorama da ficção contemporânea de Flávio Carneiro. Desta forma, o primeiro traço observado é o cruzamento da literatura com a linguagem da mídia e da internet, que se faz presente em toda a rotina de Kitty. A internet é um dos principais meios usados pela personagem para se comunicar com os seus amigos e expor as suas ideias, inclusive, a obra utilizou diversos termos utilizados na linguagem internética. Além disso,

observa-se outro traço que esta- relacionado a problematização das relações amo-rosas, em especial as femininas e homoeróticas. A protagonista tem dificuldade em firmar os seus relacionamentos e chegar a acusar um de seus namorados, Breno, de gay, pelo fato de transarem pouco e há tempos isso não ocorrer. Ainda, em outra situação, ao sair com Bruno, o mancha negra, Kitty está vestida de forma provocativa, cede à carícias íntimas, porém não permite que o sexo se consume, o que gera grande indignação em Bruno, que passa a ofendê-lo e humilhá-la, demonstrando a banalização do sexo.

863.954-1

O romance "Kitty aos 22: divertimento" de Reinaldo Santos Neves conta a história de Maria Catarina, uma jovem de 22 anos, re-latando amizades, romances, baladas e os demais acontecimentos do universo jovial atual. Cabe ressaltar a notoriedade da problematização das relações amo-rosas, em especial as femininas e homoeróticas. É inegável esse destaque vez que o livro é centrado em Kitty que namora Bruno, mas já namorou Guto, além de Nênis ter se declarado para ela, assim como Bruno. Isso demonstra os problemas das relações amorosas de Kitty que mantém um relação com casa um deles, mesmo que não de maneira física, mas devido a ligação pelo sentimento amor. Além disso, cabe citar a amiga Lau que está imersa em desilusões amorosas constantemente. Há ainda o relato de problemas homoeróticos com a presença de um casal homossexual na trama, que é brutalmente assassinado. Sem contar as cantadas que Kitty recebe de lésbicas e em alguns momentos o desejo pelas amigas. Apesar do livro ser ficcional, nota-se claramente o cunho social do mês-mo, inclusive, relatando com precisão os fatos da cidade de Vitória, ou melhor da Grande Vitória – talvez por ser capixaba. Numa análise da juventude capixaba não é surpresa encontrar-se com Kitty, Lau, Deb e os demais, pois o livro representa a materialização, em palavras, do universo da juventude capixaba. Portanto, é inquestionável seu cunho social.

864.022-1

Podemos observar a ocorrência dos seguintes traços citados no texto, na obra Kitty aos 22: divertimento, de Reinaldo Santos Neves, quais sejam: 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. O primeiro traço citado está presente em toda obra, vez que

a linguagem utilizada pelo autor foi baseada na linguagem utilizada pelos jovens nas redes sociais. O uso de abreviaturas, como exemplo Vitória é tido por Mic, caracteriza esse traço presente na obra, podemos citar ainda que durante a narrativa, digo, os personagens fazem uso de blogs, fotologs, e-mail e outros recursos tecnológicos, bem como a participação de Kitty em um evento de moda ocorrido em Vitória que é amplamente divulgado pela mídia. Quanto ao segundo traço, ele pode ser percebido no envolvimento de Kitty com o mancha negra, que inicialmente se apresenta como o bom moço, o príncipe encantado, mas que na verdade só queria manter relações sexuais com Kitty, da mesma forma como estava em um vídeo que ele, o mancha negra, tinha de Kitty e seu ex-namorado em momentos de intimidade corrobora com esse traço o fato de que inicialmente o Bruno, o mancha negra, ter se envolvido com a amiga de Kitty, e depois de "dar um fora" nela, teria investido em Kitty, sendo o pivô do fim da amizade entre as duas.

864.085-1

1 - Na obra, Kitty aos 22: divertimento, de Reinaldo Santos Neves, já a ocorrência de traço, como o cruzamento da literatura com a linguagem da mídia e da internet. Percebe-se que durante a história, a personagem Kitty, utiliza uma linguagem característica do seu universo jovem, ou seja, palavras contraídas, expressões virtuais e expressões inventadas por ela e suas amigas, como a palavra "Mic", que diz respeito ao nome da cidade de Vitória. Outro traço que é encontrado na obra, Kitty aos 22: divertimento, de Reinaldo Santos Neves, é a problematização das relações amorosas, em relação às femininas e homoeróticas. Nota-se que a protagonista mantém relacionamentos superficiais, de forma que, ela leva uma vida voltada para libertinagem, sem se aprofundar em um relacionamento amoroso. Com relação à questão homoerótica, há um fato na história, de dois amigos de Kitty, que são homossexuais, porém eles morrem no início da obra.