

VITA - UMA FERRAMENTA PEDAGÓGICA PARA ACOMPANHAMENTO E AUXÍLIO NA RELAÇÃO ALUNO-PROFESSOR

Leonardo de Souza Preuss

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Ciência da Computação, Centro Federal de Educação Tecnológica Celso Suckow da Fonseca CEFET/RJ, como parte dos requisitos necessários à obtenção do título de mestre.

Orientadores:

Joel dos Santos

Gustavo Paiva Guedes e Silva

Rio de Janeiro,

Janeiro de 2019

VITA - Uma ferramenta pedagógica para acompanhamento e auxílio na relação aluno-professor

Dissertação de Mestrado em Ciência da Computação, Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, CEFET/RJ.

Leonardo de Souza Preuss

Aprovada por:

Presidente, Prof. Joel André Ferreira dos Santos, D.Sc. (orientador)

Prof. Gustavo Paiva Guedes e Silva, D.Sc. (co-orientador)

Prof. Eduardo Soares Ogasawara, D.Sc.

Prof. Alessandro Rosa Lopes Zachy, D.Sc. (CEFET/RJ)

Prof. Diego Gimenez Passos, D.Sc. (UFF)

Rio de Janeiro,
Janeiro de 2019

Ficha catalográfica elaborada pela Biblioteca Central do CEFET/RJ

S168 Salles, Rebecca Pontes
Avaliação de agregação temporal na previsão da temperatura
de superfície do mar do Oceano Atlântico / Rebecca Pontes
Salles.—2016.
xiii, 58f. : il. (algumas color.) , grafs. , tabs. ; enc.

Projeto Final (Graduação) Centro Federal de Educação
Tecnológica Celso Suckow da Fonseca , 2016.
Bibliografia : f. 51-58
Orientador : Eduardo Soares Ogasawara

1. Computação. 2. Precipitação (Meteorologia) – Previsão. 3.
Análise de séries temporais. 4. Temperatura atmosférica. 5.
Atlântico, Oceano, I. Ogasawara, Eduardo Soares (Orient.). II.
Título.

CDD 004

AGRADECIMENTOS

Agradece-se à CAPES e ao CNPq pelo financiamento parcial desta pesquisa.

A Deus por ter me dado saúde e força para superar as dificuldades. A toda minha família que não mediram esforços para que eu chegasse até esta etapa.

A todos os professores do Programa de Pós-Graduação em Ciência da Computação, que foram tão importantes na minha vida acadêmica.

Agradece-se também as contribuições dos alunos que se voluntariaram à participação da pesquisa.

Aos orientadores Joel dos Santos e Gustavo Guedes em especial pela paciência na orientação e o grande incentivo que tornaram possível a realização desse trabalho.

RESUMO

VITA - Uma ferramenta pedagógica para acompanhamento e auxílio na relação
aluno-professor

Leonardo de Souza Preuss

Orientadores:

Joel dos Santos

Gustavo Paiva Guedes e Silva

Resumo da Dissertação submetida ao Programa de Pós-graduação em Ciência da Computação do Centro Federal de Educação Tecnológica Celso Suckow da Fonseca CEFET/RJ como parte dos requisitos necessários à obtenção do título de mestre.

Um grande problema do sistema de educação atual é o alto índice de evasão em cursos de graduação. Pesquisas na área da educação indicam que o nível de satisfação do estudante é um importante indicador na medida do sucesso de uma universidade. Estudantes com altos níveis de satisfação são mais propensos a continuar no curso, em contraste àqueles que estão insatisfeitos. Nas universidades brasileiras, a qualidade docente e o método de ensino são considerados fatores cruciais e altamente responsáveis na satisfação dos alunos.

Nesse contexto, é de grande valia um mecanismo capaz de melhorar a percepção do professor a respeito dos alunos. Isso permite que o professor antecipe a insatisfação dos alunos em uma disciplina e tenha maior agilidade na compreensão de suas necessidades e na adequação das aulas. Reduzindo, assim, a taxa de evasão nas universidades brasileiras. O presente trabalho propõe um sistema que auxilia o professor na percepção do andamento dos alunos em suas aulas, revelando seu grau de satisfação ou insatisfação durante o período letivo. A avaliação é feita tendo como base comentários escritos sobre cada aula, fornecidos pelos alunos.

A ferramenta proposta utiliza técnicas da Análise de Sentimentos para processar as avaliações e apresentar ao professor um resultado em formato de gráfico, individual ou por turma. Este gráfico indica apenas os sentimentos predominantes na avaliação do aluno. Assim, é mantida a privacidade do mesmo (com respeito à avaliação fornecida), bem como evita interpretações diferenciadas por parte do professor. Experimentos realizados com 4 turmas de graduação apontam para uma boa precisão da estimativa da satisfação feita por Virtual Therapist (VITA) e sua utilidade para professores.

Palavras-chave:

Sistema de Apoio ao Ensino; Análise de Sentimentos; LIWC.

Rio de Janeiro,

Janeiro de 2019

ABSTRACT

VITA - Uma ferramenta pedagógica para acompanhamento e auxílio na relação
aluno-professor

Leonardo de Souza Preuss

Advisors:

Joel dos Santos

Gustavo Paiva Guedes e Silva

Abstract of dissertation submitted to Programa de Pós-graduação em Ciência da Computação - Centro Federal de Educação Tecnológica Celso Suckow da Fonseca CEFET/RJ as partial fulfillment of the requirements for the degree of master.

A big problem with the current educational system is the high dropout rate in undergraduate courses. Research in the educational area indicates that student satisfaction is an important factor in the success of a university. Students with high satisfaction levels are more prone to stay in the course in contrast to dissatisfied ones. In Brazilian universities, the quality of the teaching staff and the teaching methodology are considered crucial factors and highly responsible for student satisfaction.

In this context, it is imperative a mechanism capable of improving teachers perception about students. Such a mechanism allows the teacher to anticipate student dissatisfaction about a class and to have more agility in to comprehend their needs and to adapt his class accordingly. Thus, it would reduce the risk of dropout in Brazilian universities. The present work proposes a system to ease the teacher identification of the student's progress in class, indicating their satisfaction or dissatisfaction level along the semester. The evaluation is performed on written comments about each class provided by students.

The proposed tool uses Sentiment Analysis techniques to process the evaluations and presents to the teacher its result in a chart, either individual or for the whole class. This chart indicates the prevailing feelings in the student evaluation. Thus maintaining the student privacy (concerning the evaluation they provide), as well as different interpretations by the teacher. Experiments conducted with 4 classes of an undergraduate course points to a good precision of the satisfaction estimation provided by VITA and its usefulness for teachers.

Key-words:

Educational Support System; Sentiment Analysis; LIWC.

Rio de Janeiro,

Janeiro de 2019

Sumário

I	Introdução	1
II	Computação Afetiva	5
II.1	Processamento de Linguagem Natural – NLP	5
II.2	Análise de Sentimentos	8
II.3	Léxicos Afetivos	8
II.3.1	LIWC – Linguistic Inquiry and Word Count	8
II.3.2	ANEW – Affective Norms for English Words	11
II.4	Observações Finais	11
III	Trabalhos relacionados	13
IV	VITA - Virthual Therapist	16
IV.1	Arquitetura do VITA	17
IV.2	Processo de Mineração de Dados	18
IV.3	Modelo	19
IV.3.1	Cálculo da satisfação	21
IV.3.2	Deteccção de negação e intensificadores/atenuadores	23
IV.4	Implementação	24
V	Avaliação do VITA	26
V.1	Discussão e considerações finais	31
VI	Conclusão	33
	Referências Bibliográficas	35

Lista de Figuras

II.1	Exemplos de n-gramas.	6
II.2	Modelo valência-excitação de James Russel. Cada ponto do gráfico apresenta um sentimento estimado a partir dos valores de valência e excitação. Adaptado de Russell [1980]	12
IV.1	Visão geral da arquitetura da ferramenta VITA	17
IV.2	Etapas envolvidas no processamento dos comentários.	18
IV.3	Vetor contendo os pares categoria-frequência	20
IV.4	Exemplo de vetor ponderado com suas respectivas categorias	22
IV.5	Telas principais das versões aluno e professor do VITA.	24
IV.6	Telas de seleção da versão aluno do VITA.	25
V.1	Gráfico de dispersão da correlação de Spearman entre a satisfação estimada por VITA e a indicada pelo aluno. A linha em azul mostra a correlação entre os dois eixos. Quanto mais próxima de um ângulo de 45° , maior a correlação linear entre as variáveis e quanto mais próxima do centro, mais os valores de um eixo coincidem com os do outro. A faixa destacada em cinza mostra o intervalo de confiança. Quanto mais estreito, maior o nível de confiança na estimativa.	29
V.2	Gráfico de dispersão da correlação de Spearman entre a satisfação estimada pelo professor e a indicada pelo aluno. A linha em azul mostra a correlação entre os dois eixos. Quanto mais próxima de um ângulo de 45° , maior a correlação linear entre as variáveis e quanto mais próxima do centro, mais os valores de um eixo coincidem com os do outro. A faixa destacada em cinza mostra o intervalo de confiança. Quanto mais estreito, maior o nível de confiança na estimativa.	30
V.3	Gráfico de dispersão da correlação de Sperman entre a satisfação indicada pelo professor e a estimada pelo VITA. A linha em azul mostra a correlação entre os dois eixos. Quanto mais próxima de um ângulo de 45° , maior a correlação linear entre as variáveis e quanto mais próxima do centro, mais os valores de um eixo coincidem com os do outro. A faixa destacada em cinza mostra o intervalo de confiança. Quanto mais estreito, maior o nível de confiança na estimativa.	30

V.4 Gráfico comparativo da média aritmética das notas preenchidas pelos alunos da disciplina MC durante um período de 7 aulas e da média calculada por VITA. O gráfico em azul foi gerado pelo sistema proposto enquanto o gráfico em vermelho é relativo à nota dos alunos.

Lista de Tabelas

II.1	Tabela comparativa entre as formas da palavra antes e após a lematização e estematização	7
II.2	Tabela parcial das categorias do LIWC para dimensão psicológica (adaptado de [Pennebaker et al., 2015])	9
II.3	Tabela contendo exemplos de palavras do ANEW-br com suas respectivas valências e excitação	11
IV.1	Tabela mostrando as novas categorias criadas e suas respectivas palavras	20
V.1	Matriz de confusão para valores estimados por VITA (n_{vita}) e àqueles informados pelo aluno (n_{aluno}).	28
V.2	Comparação entre os verdadeiro-positivos (tp), falso-positivos (fp) e falso-negativos (fn) para as diferentes notas (classes) estimadas por VITA. A precisão e abrangência também foram calculadas para cada classe. O melhor resultado encontrado se deu para a Nota 5	28

Lista de Abreviações

ABSA	Aspect-Based Sentiment Analysis	35
ANEW-BR	Affective Norms For English Words	2, 11, 26, 35
API	Application Programming Interface	18
ATC	Automatic Text Categorization	1
BOW	Bag-of-Words	6
CPA	Comissão Própria De Avaliação	3
HTTP	Hypertext Transfer Protocol	17
JAZZY	Java Spell Checker API	18
JSON	Javascript Object Notation	17, 19
KDD	Knowledge Discovery In Databases	18
LIWC	Linguistic Inquiry And Word Count	2, 3, 4, 5, 8, 9, 11, 15, 17, 18, 19, 20, 21, 23, 26, 34, 35
NLP	Natural Language Processing	3, 5, 6, 7
OCR	Optical Chracter Recognition	8
POS	Part-of-Speech Tagging	6
SBD	Sentence Boundary Disambiguation	34
VITA	Virtual Therapist	v, vi, viii, ix, x, 4, 28, 29, 30, 31, 32, 34
WSD	Word Sense Disambiguation	7

Capítulo I Introdução

Pesquisas na área da educação indicam que o nível de satisfação do estudante é um importante indicador na medida do sucesso de uma universidade [Weerasinghe et al., 2017; Elliott and Shin, 2002]. Estudantes com altos níveis de satisfação são mais propensos a continuar no curso [Seldin, 1993], em contraste àqueles que estão insatisfeitos. Alguns fatores são responsáveis pela satisfação do aluno, como a infraestrutura do campus, a relação interpessoal dos alunos, bem como a didática docente envolvida. Além disso, a qualidade docente e o método de ensino são considerados fatores cruciais e altamente responsáveis na satisfação dos alunos [de Souza and Reinert, 2010].

Diversos problemas do ensino público no Brasil e no mundo estão relacionados ao baixo desempenho acadêmico em disciplinas iniciais, muitas vezes causados pela falta de informação sobre as dificuldades e necessidades dos alunos perante a disciplina [Manhães et al., 2011]. Isto gera um impacto direto na satisfação dos alunos diante do curso, que acabam evadindo-se das instituições de ensino. Esse é um grave problema e traz grandes desperdícios de recursos às universidades [Vallejos and Steel, 2014], já que uma grande quantidade de verbas necessárias para sustentar um aluno na universidade são aplicadas sem retorno na preparação desse.

Tal cenário se agrava em turmas com um grande número alunos ou que não permitam um contato mais próximo entre aluno e professor (como no ensino à distância). Dessa forma, torna-se mais complicado para o professor acompanhá-los individualmente. Por este motivo, é de grande valia a criação de um mecanismo capaz de melhorar a percepção do professor a respeito dos alunos. Tal mecanismo deve permitir que o professor identifique a necessidade de mudanças na didática de uma disciplina, bem como levantar a necessidade de um acompanhamento mais próximo de algum aluno. Dessa forma o professor tem mais agilidade na compreensão das necessidades dos alunos e na realização de mudanças necessárias para atendê-las.

Conforme o crescente aumento das páginas *Web*, passou a ser necessária a tarefa de se organizar documentos de acordo com seu assunto, também conhecida como Categorização Automática de Textos - do inglês Automatic Text Categorization (ATC). Ao mesmo tempo, surgiram tendências de visitantes das páginas poderem tecer comentários sobre de matérias, seja de jornais, blogs, produtos, etc. Dessa forma, seria importante classificar documentos de acordo com a relevância, que poderia variar, por exemplo, entre bom, ruim ou neutro [Pang et al., 2002].

No entanto, a tarefa de ATC não era suficiente para se classificar os comentários fornecidos pelos

visitantes das páginas nestas categorias. Deste modo, Análise de Sentimentos é uma área que surge da necessidade de se classificar o sentimento predominante dos autores de textos quando fazem a avaliação de um objeto em questão, explorando e analisando grandes quantidades de comentários textuais sobre determinado assunto. Muitas empresas utilizam da Análise de Sentimentos para avaliar a opinião dos usuários sobre um produto lançado no mercado. Por este motivo, muitas vezes essa área é chamada de *Opinion Mining* [Liu, 2012]. Além disso, a Análise de Sentimentos tem um grande potencial para extrair informações complementares como o sentimento do autor do documento (e.g., raiva, tristeza, felicidade, etc.), ao invés de uma simples orientação semântica de um documento (polaridade).

Este trabalho busca responder a seguinte pergunta de pesquisa: *O acompanhamento do nível de satisfação dos alunos no decorrer do semestre ajuda na percepção do professor a respeito das dificuldades enfrentadas por esses alunos?* Para isso utiliza-se do potencial da Análise de Sentimentos em extrair informações de um comentário como mecanismo capaz de melhorar a percepção do professor a respeito de necessidades existentes na turma. Nesse trabalho, usamos o termo satisfação para referir à percepção do aluno tanto da didática do professor, quanto de sua dificuldade em um determinado assunto abordado em aula.

O objetivo desse trabalho é propor um sistema que apresenta ao professor o grau de satisfação ou insatisfação dos alunos de uma turma durante o período letivo. A avaliação é feita tendo como base comentários escritos sobre cada aula, fornecidos pelos alunos, usando técnicas da Análise de Sentimentos. O resultado da análise de tais comentários são gráficos que apresentam o grau de satisfação dos alunos individualmente, ou da turma como um todo.

As informações utilizadas para produção dos gráficos têm como base um módulo que avalia os comentários fornecidos pelos alunos a cada aula, extraindo o grau de satisfação de cada um, formando uma análise global e individual, numa abordagem similar a de trabalhos de *User eXperience (UX)* [Hassenzahl, 2008]. A avaliação é feita com base em dicionários em que verbetes de um idioma são categorizados com base em aspectos afetivos, sociais, psicológicos e linguísticos. A proposta aqui apresentada utiliza os dicionários de léxicos afetivos Linguistic Inquiry and Word Count (LIWC) [Tausczik and Pennebaker, 2010], na versão português do Brasil, e Affective Norms for English Words (ANEW-br) [Bradley et al., 1999] para o cruzamento dos dados e maior precisão na estimativa dos sentimentos e grau de satisfação encontrados [Preuss et al., 2018]. Por fim, os tipos de gráficos disponibilizados no sistema foram escolhidos com a aplicação do método de Condorcet em pesquisa com os professores a fim de se conhecer as melhores formas de visualização.

Uma vantagem da abordagem proposta é que o professor não toma ciência dos comentários fornecidos pelos alunos, apenas recebe informações da situação desses em forma de gráficos. Dessa forma, os alunos possuem maior liberdade para discorrerem sobre sua percepção das aulas. Acresce-

se a isso o potencial da Análise de Sentimentos em extrair informações adicionais, como o sentimento do aluno, podendo-se futuramente estender a proposta para colher tais informações. Tal melhoria seria inviável caso a proposta se restringisse à obter a nota do aluno apenas em um simples valor numérico.

Outra vantagem oferecida está na utilização de uma abordagem computacional para avaliação dos comentários, que não se encontra sujeita a diferenças na interpretação, tal como nas avaliações tradicionais de fim de curso, em que são escritos comentários manualmente, que poderiam ser compreendidos de forma diferenciada pelo professor. Ainda, a abordagem proposta prevê a avaliação continuada por parte dos alunos, a cada aula. Assim, permite-se que o professor observe tendências ao longo do curso. A Comissão própria de Avaliação (CPA)¹, comissão cuja tarefa é coordenar a autoavaliação interna para toda a universidade, produz relatórios anuais e recomendações de ajustes para o ano letivo seguinte. Essa frequência de avaliação pode não ser suficiente para alunos que resolvem abandonar o curso neste ínterim. Assim, a ferramenta aqui proposta também pode ser de grande utilidade fornecendo informações mais detalhadas e em uma frequência maior para a CPA.

É também importante ressaltar que essa abordagem permite uma análise mais fiel em relação a abordagem tradicional, em que os comentários se destinam a uma visão geral do curso. Além disso, a ferramenta proposta torna-se bastante adequada em turmas de grande porte, já que a percepção da satisfação, do ponto de vista do professor, é degradada pela quantidade de alunos. Outra aplicação possível é em turmas de ensino a distância nas quais o baixo contato do aluno com o professor impacta nessa percepção.

As principais contribuições deste trabalho são: (i) propor uma adaptação do léxico afetivo LIWC para alterar a polaridade de uma palavra com base no uso de palavras de negação e intensificadores, (ii) fornecer um modelo de Análise de Sentimentos para a extração do grau de satisfação de alunos a partir de textos escritos e (iii) disponibilizar um aplicativo para dispositivos móveis que permite o uso deste modelo em sala de aula. Resultados dos experimentos apontam que a avaliação feita por VITA apresenta uma boa precisão quando comparada com avaliações feitas por professores dos comentários escritos. Ainda, a ferramenta apresentou boa precisão em relação a nota atribuída pelos alunos das aulas, quando desconsideradas discrepâncias entre o texto escrito e a nota atribuída pelo aluno. Por fim, a ferramenta se apresentou útil para os casos em que o professor não tinha um bom conhecimento do aluno sendo avaliado.

O restante dessa dissertação está estruturado como apresentado nos parágrafos a seguir. O Capítulo II fornece o referencial teórico para compreensão do trabalho proposto. O capítulo apresenta em mais detalhes a Computação Afetiva e áreas relacionadas, como a Análise de sentimentos e a Natural Language Processing (NLP). São discutidos o uso dos dicionários de léxicos afetivos

¹<http://www.cefet-rj.br/index.php/cpa>

LIWC e ANEW-br, que são base para a construção da abordagem proposta nesse trabalho. O Capítulo III apresenta os trabalhos relacionados, revelando o estado da arte de trabalhos mais relevantes para o mesmo segmento, utilizados ao redor do mundo. O capítulo apresenta ainda uma comparação entre a proposta aqui apresentada e trabalhos relacionados. O Capítulo IV discute o modelo de avaliação proposto neste trabalho. É discutida ainda a arquitetura da ferramenta VITA, que implementa o modelo proposto. O Capítulo V apresenta a avaliação realizada com o VITA, de forma a verificar sua eficácia para uso no apoio ao ensino. O Capítulo VI faz as considerações finais apontando possíveis trabalhos futuros.

Capítulo II Computação Afetiva

A Computação Afetiva foi uma área originada em 1995 por Roselind Picard [Picard, 1995], que observou e estudou a necessidade de se incluir um elemento considerado ausente nas máquinas: a emoção. Segundo Picard [Picard, 2010], apenas o lado racional era visto como útil, e não as emoções, que eram até então ignoradas. A capacidade de perceber, interpretar, reagir à, ou expressar emoções foi considerada uma grande lacuna na computação – um elemento chave para que se pudesse evoluir no desenvolvimento de máquinas inteligentes.

Sua motivação inicial era fornecer às máquinas a capacidade de simular a empatia. Tal qualidade é muito valorizada no ser humano, sendo identificada como a capacidade de um indivíduo de se imaginar e se colocar nas circunstâncias de outro [Leiberg and Anders, 2006]. A criação de sistemas capazes de apresentar tais qualidades promoveriam um grande salto na evolução da computação como um todo, estabelecendo relação com as Ciências Cognitivas [Picard, 1995], fornecendo suporte à Psicologia, e dando origem a futuras áreas, como a Análise de Sentimentos.

Assim, este capítulo está estruturado como segue. A Seção II.1 apresenta os principais conceitos de uma área fortemente relacionada à Análise de Sentimentos – o Processamento de Linguagem Natural. A seguir, a Seção II.2 aborda os principais conceitos sobre a Análise de Sentimentos. A Seção II.3 apresenta dois dicionários de léxicos afetivos, que são ferramentas muito utilizadas para descoberta de padrões e perfis linguísticos no campo da escrita, são eles o LIWC e o ANEW-br. Por fim, a Seção II.4 finaliza o capítulo realçando a importância dos conceitos abordados.

II.1 Processamento de Linguagem Natural – NLP

Uma área muito promissora e em destaque atualmente envolve a análise de grandes quantidades de textos escritos em linguagem natural, que são textos escritos pelo ser humano. Tais linguagens sofrem uma evolução orgânica, em contraste às linguagens artificiais, onde sua construção é integralmente planejada. Assim, esta área permite a realização de inúmeras tarefas de processamento referentes à questões gramaticais, sintáticas e semânticas do modelo linguístico para um determinado idioma [Liddy, 2001]. A esta área dá-se o nome de Processamento de Linguagem Natural (do inglês *Natural Language Processing – NLP*).

NLP é um ramo da Ciência da Computação que se intersecta com a Linguística Computacional

e a Inteligência Artificial. NLP agrega várias tarefas de processamento, que consistem em fazer com que o computador seja capaz de interpretar um texto de modo semelhante ao ser humano. É também considerado a base da Mineração de Texto [Hotho et al., 2005]), uma área que estende a Mineração de Dados, tendo o diferencial de trabalhar com dados não-estruturados [Tan, 1999]. Desta forma, esta área realiza desde tarefas textuais mais básicas que partem da análise de sintaxe, até tarefas mais complexas, como a análise semântica e a transformação de fala em texto. Um sistema NLP implementado em plena capacidade deve ser capaz de parafrasear, responder perguntas, traduzir e elaborar conclusões sobre um texto [Liddy, 2001].

A tarefa mais simples, a de análise sintática, contém vários métodos para tratamento do texto, de modo análogo à etapa de pré-processamento da Mineração de Dados. Em primeiro lugar temos a tokenização, uma tarefa básica de identificação e separação das palavras (*tokens* ou termos) de um texto. A partir dessa tarefa, é construído o chamado Bag-of-Words (BOW), ou seja, o conjunto completo e único de *tokens*, com suas respectivas frequências de repetição. O BOW tem grande utilidade para cálculos e informações estatísticas acerca de uma palavra, como a frequência de um termo em um ou mais documentos [Matsubara et al., 2003]. A seguir, é feita a correção ortográfica das palavras, removendo possíveis erros de digitação que tenham ocorrido na escrita do texto. Como resultado essa etapa gera um novo BOW com suas palavras corrigidas e, potencialmente, novas frequências de repetição. Após essas etapas, as palavras do *BOW* estão no formato necessário para seu uso nas técnicas apresentadas a seguir.

Outro conceito de grande relevância, é conhecido como n-grama, onde se analisam grupos de palavras, de acordo com o valor de “n”. Por exemplo, um uni-grama, é equivalente ao *token*, onde uma palavra é avaliada por vez. Já o bi-grama são duas palavras consecutivas, e assim por diante. Esta técnica facilita a análise onde os termos possuem dependência um do outro, e a análise individual perde o sentido correto das palavras naquele contexto [Hotho et al., 2005]. A Figura II.1 apresenta alguns exemplos de n-gramas obtidos a partir de um texto.

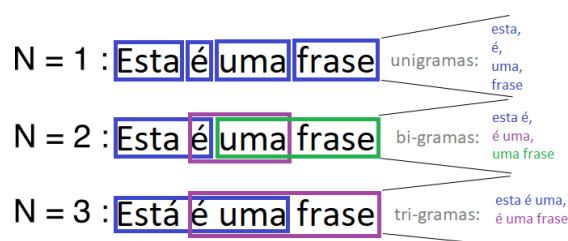


Figura II.1: Exemplos de n-gramas.

A partir daí, outras tarefa mais complexas são possíveis, como a associação de rótulos a partes do discurso (Part-of-Speech Tagging (POS)), onde são identificadas as funções sintáticas de palavras, como verbos, adjetivos, substantivos, etc. Então, é adicionado um rótulo à estas palavras, de modo que esta seja, numa tarefa posterior, facilmente distinguível de outra idêntica, mas que exerce outra

função na frase [Collobert et al., 2011]. Outras duas tarefas bastante comuns são a lematização e a estematização.

A lematização consiste em transformar uma palavra para seu lema, que é uma forma deflexionada de uma palavra. Isso reduz a dimensionalidade, devido ao fato de variantes de uma palavra serem, na realidade, apenas uma. Por exemplo, o lema de um verbo é seu infinitivo. Uma vantagem da lematização é que verbos ou palavras irregulares, podem ser deflexionadas para sua forma raiz corretamente. Já na estematização, o objetivo também é obter a forma raiz de uma palavra, porém sua desvantagem é não deflexionar corretamente palavras quando sua forma raiz é irregular. Em contrapartida, é um procedimento mais rápido e muitas vezes substitui a lematização por questões de performance.

Tabela II.1: Tabela comparativa entre as formas da palavra antes e após a lematização e estematização

Palavra	Lematização	Estematização
Viagem	Viajar	Viag
Viajando	Viajar	Viaj
Viajei	Viajar	Viaj
Cantores	Cantar	Cant
Canto	Cantar	Cant
Pescando	Pescar	Pesc
Pesquei	Pescar	Pesq

Já na análise semântica, conforme o próprio nome sugere, as tarefas executadas levam em conta o significado das palavras. Sendo assim, existe uma tarefa voltada para isolar o significado de palavras idênticas, mas que possuem significados diferentes. Esta tarefa recebe o nome de Word Sense Disambiguation (WSD), que permite a desambiguação do sentido das palavras quando em contextos diferentes [Hotho et al., 2005].

Com a desambiguação, passa a ser possível o reconhecimento ou extração de entidades nomeadas, que tem por objetivo identificar e fazer a distinção entre nomes próprios, cidades e outros. Por exemplo, o nome de um artista pode aparecer de diferentes formas em um texto. Enquanto Bill Clinton ou William Jefferson Clinton se referem à mesma pessoa, o método de reconhecimento de entidades deve ser capaz de identificar quando homônimos se referem a entidades diferentes [Witten, 2005]. Da mesma forma que Roberto Carlos pode se referir ao cantor, poderia também se referir ao ex-jogador de futebol.

Há também tarefas mais complexas, que se utilizam dos resultados das tarefas mais simples, como a Extração de Relacionamentos, onde é necessário saber quais os relacionamentos existentes naquele texto. Com a extração de relacionamentos, dada uma pergunta, o sistema NLP deve ser capaz de responder questões sobre o grau de parentesco entre membros de uma família, por exemplo. Ademais, esta área engloba outras tarefas semânticas como, identificar o idioma de um

texto, realizar a tradução para outro idioma, produzir um resumo, elaborar um argumento para responder uma pergunta sobre o conteúdo do texto, reconhecer caracteres do texto a partir de imagens (Optical Character Recognition (OCR)), bem como a análise de sentimentos.

II.2 Análise de Sentimentos

A Análise de Sentimentos é o ramo da Ciência da Computação que estende a Mineração de Dados, e surge da tarefa de se automatizar a classificação de textos, em geral associados à páginas Web, de acordo com seu assunto. Por exemplo, as tarefas de identificar se um determinado *e-mail* é *spam*, reconhecer o idioma de um texto, classificá-lo em ficção científica, romance, drama, etc; identificar o gênero do autor de um texto, e a própria análise de sentimentos são tarefas da Classificação Automática de Textos [Pang et al., 2008].

A Análise de Sentimentos tem como objetivo explorar e identificar o sentimento ou opinião dos autores por meio de seus textos. Exemplos de tais textos são comentários em fóruns, blogs, redes sociais, dentre outros. De acordo com [Gupta and Lehal, 2009] é estimado que 80% dos dados estejam armazenados em formato textual. Assim, muitas áreas surgem na tentativa de adquirir novos conhecimentos a partir de dados textuais [Pang et al., 2008]. Busca-se com a Análise de Sentimentos categorizar estes textos, a fim de que se possa organizá-los com base em seu conteúdo. Por este motivo, a Análise de Sentimentos é muitas vezes chamada de Mineração de Opinião [Liu, 2012].

Uma forma comum de categorização é por meio da extração da polaridade da opinião, que indica se um comentário remete a opinião de ordem positiva (e.g., felicidade, saúde, amor, etc.) ou negativa (e.g., raiva, tristeza, depressão) [Wiegand et al., 2010]. A identificação e classificação dos sentimentos na Análise de Sentimentos é comumente realizada de duas formas: (i) por meio do aprendizado de máquina ou (ii) com o uso de dicionário de léxicos afetivos [Gonçalves et al., 2013]. No caso do aprendizado de máquina, os algoritmos mais utilizados são Naive Bayes, SVM, dentre outros. Já os dicionários de léxicos afetivos contém palavras associadas a emoções, sentimentos, processos psicológicos, dentre outros [Filho et al., 2013]. O presente trabalho faz uso dos dicionários de léxicos afetivos LIWC e ANEW, detalhados na seção a seguir.

II.3 Léxicos Afetivos

II.3.1 LIWC – Linguistic Inquiry and Word Count

O LIWC (*Linguistic Inquiry And Word Count*) [Tausczik and Pennebaker, 2010] é um dicionário de léxicos afetivos em que os verbetes estão associados à categorias psicológicas, afetivas, gramaticais, dentre outras. O dicionário original, em inglês, foi desenvolvido pelo psicólogo James

Pennebaker, que afirma que o estilo de escrita contém informações implícitas, e a forma de se estruturar e organizar uma frase é mais informativo que o conteúdo em si.

Como o próprio nome diz, o LIWC é uma ferramenta para contagem de palavras e categorias. Cada palavra está contida em uma ou mais categorias, que por sua vez estão distribuídas nas 5 dimensões do LIWC. Assim, dentro da dimensão psicológica, por exemplo, existem várias categorias, como ansiedade, tristeza, emoção positiva, dentre outras [Tausczik and Pennebaker, 2010]. Dado um texto, realiza-se a contagem de todas as categorias identificadas, formando-se um vetor com a categoria e a frequência.

A única versão do LIWC em português do Brasil (de 2007) [Filho et al., 2013] conta com mais de 127 mil verbetes e 64 categorias. A Tabela II.2 apresenta algumas dimensões e categorias do LIWC. É interessante destacar que um verboete pode estar associado a mais de uma categoria. Por exemplo, a palavra “bom” está associada às categorias de emoção positiva, afeto, sentimento, etc. Por outro lado, a palavra “mau” está associada às categorias afeto, emoção negativa, raiva, dentre outras.

Tabela II.2: Tabela parcial das categorias do LIWC para dimensão psicológica (adaptado de [Pennebaker et al., 2015])

Categoria	Abreviação	Exemplos	Palavras na categoria
processos afetivos	afeto	felicidade, choro	1393
emoção positiva	posemo	amor, bom, doce	620
emoção negativa	negemo	machucar, feio, desagradável	744
ansiedade	anx	preocupado, temeroso	116
raiva	raiva	ódio, matar, irritado	230
tristeza	sad	chorando, tristeza, triste	116
processos sociais	social	colega, conversa, eles	756
família	family	filho, pai, tia	118
amigos	friends	camarada, vizinho	95
referência feminina	female	garota, ela, mãe	124
referência masculina	male	garoto, ele, pai	116

Na tabela II.2 são apresentadas algumas categorias para a dimensão psicológica do dicionário. As categorias de processos afetivos e processos sociais são as categorias mais genéricas para a dimensão psicológica, por isto estão demarcadas em tom de cinza mais escuro. Estão contidas na categoria de processos afetivos outras cinco categorias: emoção positiva, emoção negativa, ansiedade, raiva e tristeza. Dentre estas, as que estão em tom de cinza mais claro, são na realidade subcategorias da emoção negativa. Estão contidas na categoria de processos sociais outras quatro categorias: família, amigos, referência feminina e referência masculina. A tabela apresenta ainda exemplos de palavras e o total de palavras contidas em cada categoria.

Com os estudos e o avanço de Pennebaker e sua equipe, foi possível destacar duas classes de

palavras muito distintas: as palavras de conteúdo, e as palavras de estilo. As palavras de conteúdo são palavras que se referem a substantivos, verbos, adjetivos, dentre outras. Já as palavras de estilo (também chamadas de palavras funcionais) são palavras que não possuem conteúdo em si, mas exercem papel de conectivos, pronomes, preposições, etc.

Apesar disso, mesmo que tais palavras funcionais possuam proporção muito menor em comparação as palavras de conteúdo, foi observado que estas respondem, para a língua inglesa, por 55% de tudo o que é falado, ouvido ou escrito. Além disso, o cérebro trabalha as duas categorias de palavras de maneiras muito distintas. Assim, a partir de uma ótica psicológica, a categoria de palavras de conteúdo tratam do que a pessoa está dizendo enquanto a categoria de palavras de estilo refletem a forma em que a pessoa está dizendo [Tausczik and Pennebaker, 2010].

É importante lembrar que o LIWC é uma ferramenta de análise que não leva em consideração o contexto no qual a palavra está inserida. Logo, para toda a análise feita, uma palavra pode conter uma grande amplitude de significados e, portanto, pertencerá a várias categorias, mesmo que divergentes entre si. Ainda assim, o LIWC pode ser muito utilizado para descobrir a polaridade de um texto. Para tal, um procedimento comumente realizado é o de contar o número de palavras que pertencem às categorias positivas e subtraí-lo do número de palavras que pertencem às categorias negativas. O próprio resultado desta operação informará a polaridade da frase (se maior que zero positivo, caso contrário negativo). Assim, mesmo sendo o LIWC independente de contexto, bons resultados foram obtidos para estimar a polaridade da frase em comparação com outros algoritmos léxicos que não são isentos de contexto [Gonçalves et al., 2013].

Por este motivo, este dicionário é muito utilizado para estudos no campo linguístico, onde o perfil de escrita de um indivíduo, ou seja, a maneira particular de cada um escrever revela traços da própria personalidade, informações implícitas, ou mesmo uma mentira, que não está escrita diretamente no texto. Assim, Pennebaker considera que o estilo linguístico pode ser mais informativo do que o conteúdo em si.

De forma a se ilustrar tal informação, em um de seus estudos, Pennebaker cita um caso famoso ocorrido nos Estados Unidos no ano de 1994 com Susan Smith, onde alegava-se que seus filhos teriam sido sequestrados. Ao ser entrevistada pelo repórter, ela disse: "Meus filhos me queriam. Eles precisavam de mim. E agora não posso ajudá-los". Neste trecho, apesar de sua fala indicar uma suposta preocupação com seus filhos, a frase está escrita no passado, sugerindo que ela já considerava seus filhos mortos e, portanto, já sabia do desfecho da situação. Normalmente um parente ainda mantém esperança sobre o desaparecimento de alguém, e por isso sua fala é no tempo presente. Ao término da investigação as autoridades descobriram que ela tinha afogado seus próprios filhos na piscina [Newman et al., 2003].

II.3.2 ANEW – Affective Norms for English Words

O ANEW (*Affective Norms for English Words*) [Bradley et al., 1999] é outro dicionário de léxicos afetivos. No entanto, seus verbetes estão associados aos conceitos de *valência* (polaridade), *excitação* (ativação) e *dominância*. Estas componentes provêm de conceitos diretamente relacionados à psicologia e dizem respeito, respectivamente, a polaridade de uma emoção (positiva ou negativa), ao grau de ativação desta emoção, e o quanto a emoção é de submissão ou dominância. No ANEW, as 3 dimensões estão compreendidas em uma escala de valores decimais com seu mínimo em 1 e máximo em 9, contendo também seus respectivos desvios-padrão para cada palavra. Na Tabela II.3 é possível ver exemplos de algumas palavras. Na versão utilizada do ANEW-br não são encontrados valores de dominância. Todavia, os valores de dominância não são utilizados neste trabalho.

Palavra	Valência	Excitação
Amor	8,75	4,39
Agradável	8,15	3,32
Feliz	8,69	4,13
Triste	1,73	5,28
Angustiado	2,22	6,65
Desamparado	2,48	5,00

Tabela II.3: Tabela contendo exemplos de palavras do ANEW-br com suas respectivas valências e excitação

A tabela II.3 mostra os níveis de valência e excitação para algumas palavras do ANEW-br. As tabela está separada por palavras positivas e negativas, na ordem. É possível constatar que as palavras negativas possuem uma valência abaixo de cinco (neutro). Acima deste valor, estão as palavras positivas. Já a excitação varia de acordo com a intensidade da emoção relacionada à palavra.

A partir das duas dimensões mencionadas, é possível gerar um gráfico indicando a emoção sentida por um indivíduo. Uma forma comum de tal representação exhibe a valência no eixo horizontal e a excitação no eixo vertical. Com estas informações podemos estimar, mediante o ângulo do gráfico gerado, a emoção sentida pelo usuário. Por exemplo, uma baixa excitação e uma polaridade bastante positiva, indicam um estado de calma. Já uma grande excitação e valência negativa, indicam um estado de estresse ou raiva. A Figura II.2 apresenta exemplos de sentimentos que podem ser obtidos a partir dessa representação.

II.4 Observações Finais

Este capítulo apresentou diversos conceitos e técnicas da NLP, que embasam a Análise de Sentimentos. Foram apresentados também os léxicos afetivos LIWC e ANEW-br, seu funcionamento

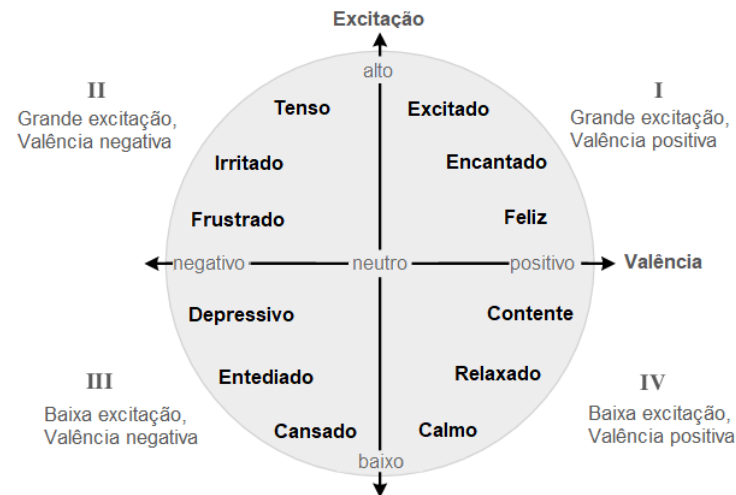


Figura II.2: Modelo valência-excitação de James Russel. Cada ponto do gráfico apresenta um sentimento estimado a partir dos valores de valência e excitação. Adaptado de Russell [1980]

e suas potenciais aplicações nesta área.

A abordagem seguida por esse trabalho será de utilizar os conceitos mencionados da Análise de Sentimentos para o apoio a educação. Estudos correlacionam problemas emocionais ou distúrbios no comportamento de alunos, com uma maior dificuldade de aprendizado, baixo desempenho acadêmico e altas taxas de abandono em cursos [Vaez and Laflamme, 2008; Pritchard and Wilson, 2003]. O trabalho de [Andrews and Wilding, 2004] mostra que, de forma específica, a ansiedade e a dificuldade de aprendizado estão intimamente relacionadas. Estados de ansiedade podem ser definidos como um comportamento evasivo, em que há uma tendência de fuga diante de uma ameaça iminente [Heimberg et al., 2004]. Portanto, aqueles que se encontram nesta situação de fuga, apresentam maior dificuldade para o aprendizado. Ademais, de acordo com [Baylor and Rosenberg-Kima, 2006], o indivíduo que possui dificuldades para aprender sobre determinado assunto, apresenta tendência de reduzir sua motivação para conclusão da tarefa.

A ideia é permitir que professores possam conhecer a satisfação dos alunos a cerca de uma disciplina e, com base nesse resultado, optar por mudanças na metodologia ou planejamento das aulas visando melhorar o desempenho dos alunos e, conseqüentemente, reduzir a evasão. Os sentimentos dos alunos serão identificados com o uso combinado do dicionário LIWC e ANEW-br sobre comentários dos alunos a respeito das aulas. O capítulo a seguir apresenta a ferramenta proposta nesse trabalho, sua arquitetura e uso.

Capítulo III Trabalhos relacionados

No contexto da Análise de Sentimentos, existem alguns trabalhos publicados na literatura voltados ao auxílio na área da educação. Dentre eles, podem-se destacar [Aung and Myo, 2017; Barrón-Estrada et al., 2017; Pong-Inwong and Kaewmak, 2016; Balahadia et al., 2016; Rani and Kumar, 2017; Ortigosa et al., 2014; Robinson et al., 2013; Rude et al., 2004]. Os artigos apresentados nesse capítulo foram obtidos com o uso da *string de busca sentimentanalysis ∧ teaching* na base Scopus, tendo como resultado 60 artigos. Dentre estes, 8 trabalhos foram os que mais se assemelharam ao presente trabalho.

Em [Aung and Myo, 2017], os autores apresentam um sistema de análise de sentimentos em comentários de alunos sobre os professores. Esse trabalho utilizou uma abordagem léxica para a análise de sentimentos, de forma semelhante ao presente trabalho. No entanto, o dicionário utilizado para classificação das palavras é diferente do dicionário utilizado nesse trabalho. No trabalho mencionado, o resultado da avaliação do professor é expressado em 7 níveis, sendo eles: fortemente positivo, moderadamente positivo, fracamente positivo, neutro, fracamente negativo, moderadamente negativo e fortemente negativo. Um grande diferencial em relação o trabalho aqui proposto é que [Aung and Myo, 2017] não classifica os comentários de acordo com as emoções presentes (ex.: felicidade, raiva, tristeza).

Em [Barrón-Estrada et al., 2017], os autores apresentam um módulo para análise de sentimentos de comentários de alunos para a *API (Application Programming Interface)* Sensei. Sensei é um ambiente Web que funciona como uma espécie de sistema de recomendação, fornecendo instruções para os estudantes a partir dos comentários feitos. A ideia do módulo é obter as informações dos alunos estritamente sobre os exercícios de aulas. Dessa forma, o aluno resolve os exercícios e fornece ao módulo seu comentário sobre os exercícios realizados para que os professores adaptem os exercícios conforme a necessidade e, assim, o ciclo se repete. Em [Barrón-Estrada et al., 2017], o resultado é expressado em formato de polaridade (positivo, negativo ou neutro) e também não utiliza o LIWC para classificação, apesar de utilizar uma abordagem léxica.

Em [Pong-Inwong and Kaewmak, 2016], os autores desenvolvem um trabalho de análise de sentimentos para comentários de alunos sobre os professores em mídias sociais. Cada aluno faz a postagem de sua opinião sobre o professor uma vez por semana. Além disso, também é proposta a melhoria de métodos de análise de sentimentos por meio da comparação entre métodos

tradicionais de aprendizado de máquina, com os chamados *Ensemble Learning*, no qual múltiplos métodos de aprendizado de máquina são utilizados. Os resultados indicaram que os métodos de *Ensemble Learning* são significativamente mais eficientes que os métodos de aprendizado de máquina tradicionais.

Em [Balahadia et al., 2016], os autores apresentam uma ferramenta para avaliação de professores por estudantes, colocando a avaliação como parte da escala de desempenho de professores. Ademais, faz também a medição do quanto uma avaliação qualitativa pode complementar uma avaliação quantitativa. Neste trabalho os resultados da análise de sentimentos só assumem dois estados: positivo ou negativo. O trabalho utiliza uma abordagem léxica para a análise de sentimentos e fornece módulos para análises estatísticas e compilação de resultado, dentre outros. Além disso, o trabalho mostra que a necessidade da tarefa de análise de sentimentos é grande no que diz respeito ao trabalho tedioso do professor em ler os comentários dos alunos. Este trabalho se diferencia ao atual na análise da polaridade, já que o trabalho aqui proposto leva em conta algumas emoções negativas.

Em [Rani and Kumar, 2017], os autores desenvolvem um trabalho de Análise de Sentimentos em páginas de pesquisa de avaliação de professores para comentários de alunos. O sistema desenvolvido fornece a polaridade dos sentimentos (positivo, negativo ou neutro) além de 8 categorias de emoções (tristeza, ódio, medo, desgosto, surpresa, prazer, expectativa e confiança). A partir daí o sistema ainda calcula o grau de satisfação e insatisfação do estudante. O trabalho utiliza uma abordagem léxica, com base no dicionário NRC *Emotion Lexicon*, para calcular a polaridade da frase, bem como classificá-la dentre as 8 categorias de emoções. É importante ressaltar, entretanto, que o léxico utilizado neste trabalho não é voltado para o idioma português do Brasil.

Em [Ortigosa et al., 2014], os autores descrevem a criação de uma aplicação para o *Facebook*, chamada *SentBuk*. A aplicação executa o processo de análise de sentimentos sobre os comentários escritos na linha do tempo de um usuário. A ferramenta é capaz de verificar a polaridade do sentimento do autor do texto, bem como a alteração na emoção daquele usuário em textos subsequentes. Ela permite também a busca de emoções por amigo, classificação de usuários de acordo com as mensagens escritas, estatísticas, etc. A ferramenta *SentBuk* conta tanto com uma abordagem de análise léxica quanto por aprendizado de máquina. Os resultados da pesquisa demonstraram que foi possível atingir 83% de precisão na avaliação dos sentimentos do usuário. Além disso, uma das aplicações sugeridas pelos autores da pesquisa se encaixam no contexto de *e-learning* adaptativo, onde o sistema que provê o conteúdo informativo se ajusta de acordo com o sentimento percebido no usuário. Outra possibilidade é seu uso em aulas e cursos a distância, servindo de *feedback* para os professores sobre o andamento das aulas, já que o contato interpessoal entre aluno e professor é menor nestes casos.

Em [Robinson et al., 2013], os autores apresentam um trabalho capaz de prever o desempenho de estudantes ao final do semestre a partir de breves textos introdutórios sobre si mesmos, escritos no início do semestre. Para isto, a pesquisa utilizou o dicionário LIWC para fazer a contagem das palavras e das categorias (gramaticais e psicológicas) às quais pertencem, em textos autobiográficos que variavam entre uma e duas páginas. A pesquisa foi feita na Universidade do Texas, contendo amostras de textos totais fornecidas por 363 participantes, dentre eles, 275 mulheres e 88 homens durante seis semestres. Esta pesquisa revelou que o uso de palavras de determinadas categorias linguísticas, como o uso demorado de marcas de pontuação, simplicidade de palavras, uso de pronomes pessoais no singular, verbos no passado, preocupações relacionadas a questões domésticas, vida social, e mais especificamente palavras relacionadas a comida, bebida e sexo podem ser decisivas no processo de predição do desempenho destes estudantes ao final do semestre.

Em [Rude et al., 2004], os autores tratam do estilo linguístico de escrita por parte de estudantes com depressão. Os estudantes são classificados em três grupos: um dos que atualmente estão passando por problemas relacionados à depressão, um segundo dos que já tiveram depressão alguma vez na vida, e o terceiro dos que nunca tiveram depressão. A amostra de participantes era composta de 31 estudantes atualmente com depressão, 26 estudantes com depressão alguma vez na vida e 67 que nunca tiveram depressão alguma vez na vida. Foi observado através de redações, que o uso de pronomes pessoais no singular, assim como, palavras com polaridade negativa, é significativamente maior em pessoas com depressão do que com pessoas que já tiveram depressão alguma vez na vida. Além disso, a quantidade de pronomes pessoais usados por estudantes que já tiveram depressão alguma vez na vida é expressivamente maior do que daqueles que nunca tiveram.

Capítulo IV VITA - Virtual Therapist

VITA é o nome dado ao sistema aqui proposto por ter sido anteriormente designado para funcionar como uma espécie de terapeuta virtual para textos escritos. No entanto, a ideia teve sua melhor aplicação quando direcionada a educação, para avaliação de estudantes com base em suas percepções sobre a aula. Assim, a proposta deste trabalho visa permitir a avaliação, em formato de escrita livre, das aulas lecionadas. Dessa maneira, o aluno pode discorrer sem restrições sobre cada aula ministrada. Esta proposta foi realizada na forma da ferramenta VITA. VITA permite que o professor tenha acesso ao sistema, verificando o nível de satisfação geral da turma, bem como a satisfação individual dos alunos.

Um ponto positivo ao se utilizar esta abordagem é o fato do professor não precisar ler todos os comentários dos alunos para saber como está indo a turma – uma tarefa trabalhosa quando se considera turmas de muitos alunos. A própria ferramenta faz essa mediação entre alunos e professor, fornecendo uma avaliação dos comentários recebidos. Além disso, é importante mencionar que a avaliação realizada pela VITA não está sujeita à interpretações diferenciadas motivadas por questões emocionais, psicológicas ou seletivas por parte do docente. Toda a análise é feita levando-se em conta os mesmos critérios, independente do estado emocional do avaliador, ou do autor do comentário. Por outro lado, o fato de o professor não ter acesso ao conteúdo dos comentários dos alunos, fornece a eles mais conforto na hora de prestarem seu depoimento sobre a disciplina. Espera-se com isso que a qualidade dos depoimentos sejam melhores e, conseqüentemente, a avaliação provida pela ferramenta.

No estágio atual, o sistema está desenvolvido para plataforma Web e aplicativo móvel, restrito aos dispositivos Android. VITA também conta com um servidor encarregado de armazenar e produzir o resultado das avaliações fornecidas pelos alunos e um banco de dados. As seções a seguir apresentam a arquitetura da ferramenta (Seção IV.1), o processo de mineração de dados utilizado (Seção IV.2), o modelo de Análise de Sentimentos empregado (Seção IV.3) e sua implementação (Seção IV.4).

IV.1 Arquitetura do VITA

VITA segue uma arquitetura de três camadas, sendo elas: cliente, servidor e base de dados. A parte cliente é onde ocorre a interação do usuário com o sistema, em que o usuário pode ser tanto o professor quanto o aluno. A versão aluno do cliente envia, via protocolo Hypertext Transfer Protocol (HTTP), o texto a ser analisado ao servidor. O servidor, por sua vez, recebe a solicitação, processa-a e armazena o resultado em um banco de dados. Em um segundo momento, caso o professor tenha a intenção de visualizar o desempenho da turma, ou de um aluno específico, a versão professor do VITA solicitará ao servidor a avaliação de sua disciplina, para determinada aula. O servidor obtém as informações da aula escolhida no banco de dados, remonta-as e as retorna ao cliente em formato JSON. As informações enviadas ao cliente permitem a montagem dos gráficos de satisfação apresentados ao professor.

Atualmente parte cliente de VITA está implementado para a plataforma Android e Web, de modo que possa atingir a maior parte dos dispositivos móveis. A escrita do código foi feita em Java 7/8, tanto para o cliente quanto para o servidor. Foi utilizado o Apache Tomcat 9.0 como o servidor e banco de dados MySql 5.6.

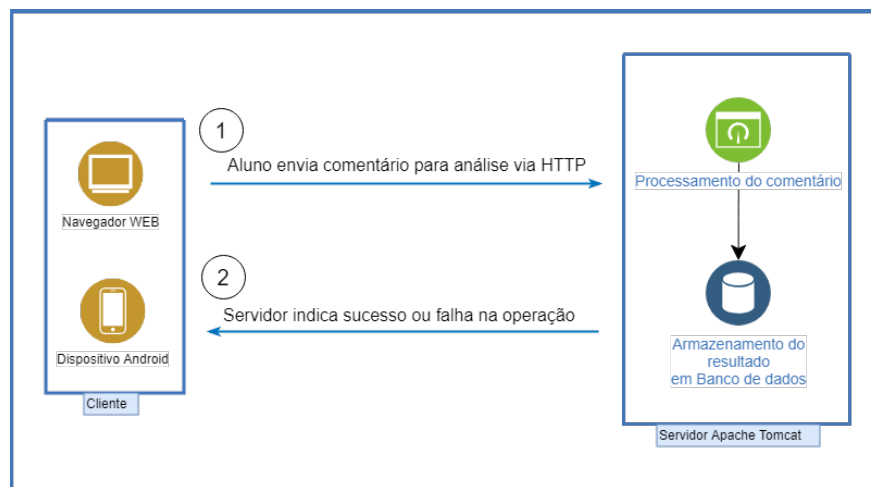


Figura IV.1: Visão geral da arquitetura da ferramenta VITA

De acordo com a Figura IV.1 a parte que cabe ao aplicativo é o envio do texto para análise ao servidor Web. A requisição ao servidor é realizada de mesma forma que uma requisição HTTP a um servidor Web, que pode ser tanto GET quanto POST. De modo geral, após receber uma requisição, o servidor a repassa ao módulo de processamento de análise de sentimentos, que é atualmente composto apenas pelo LIWC. Ao final, o resultado é convertido no formato Javascript Object Notation (JSON) e retornado para o cliente. O aplicativo então lê o resultado e o converte para um formato de gráfico.

IV.2 Processo de Mineração de Dados

A Análise de Sentimentos segue a mesma abordagem do processo de descoberta de conhecimento – Knowledge Discovery in Databases (KDD) que envolvem a seleção (coleta) da informação, o pré-processamento, para tratamento de possíveis erros nas informações obtidas, transformação dos dados para que sejam compatíveis com os métodos de processamento a serem aplicados, a mineração de dados propriamente dita e a avaliação dos dados para produção de conhecimento.

Como geralmente é feito na Mineração de Texto, o processamento realizado com o comentário do aluno para análise de sentimentos compreende várias etapas que envolvem a coleta dos textos (seleção), remoção de textos inválidos e correção ortográfica (pré-processamento), a Análise de Sentimentos (Mineração) e a produção do resultado. A diferença aqui é de que a coleta dos dados é para a produção do conhecimento da satisfação do aluno. A etapa de transformação dos dados está ausente no trabalho atual, podendo ser futuramente implementada para extração de informações adicionais. Uma ilustração destas etapas pode ser observada na Figura IV.2.

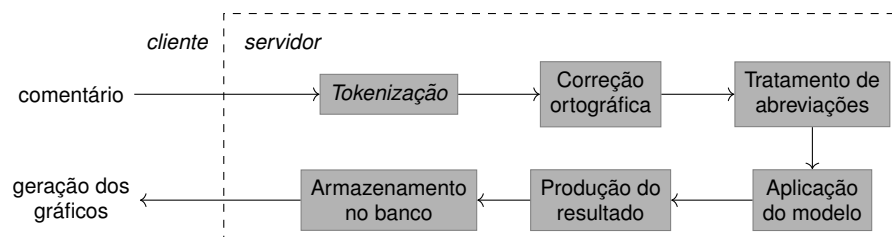


Figura IV.2: Etapas envolvidas no processamento dos comentários.

A etapa de *tokenização* identifica as palavras de um documento, por meio de espaços em branco, quebras de linha, marcas de tabulação, etc. É neste momento em que é produzido o *BOW* (*Bag-of-Words*), ou seja, o conjunto completo e único de *tokens*, com suas respectivas frequências de repetição. A seguir, é feita a correção ortográfica das palavras, removendo possíveis erros de digitação que tenham ocorrido na escrita do texto com a Application Programming Interface (API) Java Java Spell Checker API (Jazzy). Tal biblioteca é capaz de trabalhar com dicionários personalizados, desde que as palavras do dicionário estejam contidas exclusivamente em cada linha. Assim, tal dicionário foi adaptado para o português através de um *script* automatizado, tendo como base o LIWC, obtendo assim um dicionário com aproximadamente 118 mil palavras.

Na terceira etapa ocorre a transformação de formas abreviadas comuns de escrita (e.g., vc, tbm, blz) para a forma padrão da Língua Portuguesa. Após essas etapas, as palavras do *BOW* estão no formato necessário para sua busca no léxico utilizado.

O resultado da aplicação do modelo é um vetor com a contagem de palavras para cada categoria do léxico utilizado. No caso deste trabalho, para o cálculo da satisfação do aluno, é realizada uma filtragem das categorias de emoção positiva, emoção negativa e as subcategorias desta última, são

elas: ansiedade, raiva e tristeza. O vetor produzido é finalmente armazenado no banco de dados, de modo que todas as informações sejam recuperadas posteriormente.

Na última etapa, ocorre a conversão deste vetor para o formato JSON e o retorno ao cliente. Um exemplo deste objeto retornado é apresentado na Listagem 1.

```

1  {"satisfaction": 0.28,
2   "dissatisfaction": 0.72}
```

Listing 1: Exemplo de objeto JSON retornado pelo servidor

Quando o docente solicita a avaliação da aula, são resgatadas do banco as informações necessárias para remontar o vetor, retorná-los ao cliente, e a partir deste, produzir os gráficos a serem visualizados pelo professor. Os gráficos disponíveis são do tipo barras, setores e radar. Essa escolha foi baseada em uma pesquisa com 11 professores do curso de Ciência da Computação, utilizando o método de Condorcet [Young, 1988] para seleção dos três melhores gráficos.

IV.3 Modelo

Conforme mencionado anteriormente, o modelo proposto aqui visa estimar o nível de satisfação do aluno diante das aulas de uma disciplina. Para isso, é tomada como base a polaridade do texto escrito. Assim, este trabalho propõe um modelo de avaliação dos comentários dos alunos com base em léxicos afetivos. O modelo aqui proposto se utiliza de uma versão adaptada do LIWC_{AN}, uma abordagem apresentada em [Rodrigues and Guedes, 2017], na qual é feita a combinação dos dicionários LIWC com o ANEW, tendo como objetivo melhorar a precisão da análise [Preuss et al., 2018]. Esta adaptação leva em conta a inversão da polaridade das palavras pelo uso de negações, bem como a intensificação ou atenuação causada pelo uso de advérbios de intensidade.

Adaptações do LIWC

Uma grande contribuição da proposta apresentada é a adaptação feita sobre o dicionário LIWC. Como mostrado na Seção II.3.1 LIWC também associa palavras às categorias gramaticais. Uma categoria de grande utilidade para este trabalho é a categoria de negação (“*negate*”) já existente no LIWC. Nela estão contidas palavras como “não”, “nem”, “nunca”, “nada”, “ninguém”, etc. A adaptação realizada do LIWC é a inclusão de outras palavras no dicionário que tenham a função de inverter o sentido de uma frase, como as palavras “sem”, “tirar”, etc.

Outra importante adaptação do LIWC diz respeito aos advérbios de intensidade. LIWC, por padrão, não faz distinção de categorias para palavras que intensificam ou atenuam um discurso. Todos os diferentes tipos de advérbios, inclusive os de intensidade estão inseridos na categoria “*adverb*”. Por este motivo, uma adaptação do LIWC é criar duas subcategorias de modo a con-

templar distintamente os advérbios de intensidade quando tem função de amplificar ou amenizar uma qualidade. As categorias criadas receberam o nome de “*intens*” e “*atenuat*” para palavras que intensificam ou atenuam um discurso, respectivamente. A Tabela IV.1 mostra as novas categorias criadas e as palavras inseridas.

intens	atenuat
muito	pouco
muitíssimo	pouquíssimo
mais	menos
bastante	insuficiente
tanto	apenas
tão	quase
extremamente	-
deveras	-
demais	-

Tabela IV.1: Tabela mostrando as novas categorias criadas e suas respectivas palavras

Como visto anteriormente, ao utilizar-se o LIWC para análise da polaridade de uma sentença, é acrescentado ao vetor V de categorias uma unidade para cada categoria na qual uma palavra está classificada. A abordagem comumente utilizada quando encontra-se uma palavra amplificadora, é dar um percentual de acréscimo às categorias da palavra qualificadora. Em contrapartida, ao se encontrar uma palavra atenuadora, a palavra qualificadora passa a ter apenas um determinado percentual do valor padrão (1). A utilização das categorias de negação, intensificação e negação será apresentada em mais detalhes na Seção IV.3.1.

Execução do modelo

De posse do vetor das palavras contidas no texto submetido para análise, o qual é obtido após os passos descritos na Seção IV.1, é criado um vetor vazio para guardar as categorias encontradas. Nesse vetor cada índice é associado a uma das categorias do LIWC. Para cada palavra no vetor de palavras, é pesquisado no LIWC quais as categorias estão associadas a esta palavra. As categorias encontradas são então inseridas em um vetor específico para armazená-las de forma única, guardando também a frequência de uso para cada uma delas. Este processo é repetido para todas as palavras até que não existam mais palavras no dicionário. O vetor que contém os pares categoria-frequência pode ser visto na Figura IV.3.



Figura IV.3: Vetor contendo os pares categoria-frequência

A Figura IV.3 apresenta um exemplo do vetor que contém as categorias e frequência de aparição de cada categoria no texto. Os rótulos abaixo dos quadrados azuis designam algumas categorias do LIWC, como **funct** (palavras funcionais), **cogmech** (mecanismo cognitivo), **verb** (verbo), **preps** (preposições), **negemo** (emoção negativa), **anx** (ansiedade) e por último **sad** (tristeza). Em cada retângulo azul, está contida a frequência de uma determinada categoria no texto.

IV.3.1 Cálculo da satisfação

A estratégia de contagem das categorias do LIWC foi baseada no trabalho de [Guedes et al., 2016]. A aplicação do LIWC_{AN} parte desse vetor que relaciona as categorias encontradas em um texto e suas respectivas contagens. Sua diferença para a pura aplicação do LIWC é que a união entre os dicionários acrescenta um peso aos verbetes do LIWC, de acordo com a excitação e valência de uma palavra encontrada no ANEW. Assim garante-se maior correspondência entre a polaridade das palavras. A adaptação aqui proposta altera a forma como o pesos advindos do ANEW são calculados em [Rodrigues and Guedes, 2017]. A fórmula a seguir determina o peso $p(w)$ a ser atribuído para uma palavra w , com base nos valores de valência ($V(w)$) e excitação ($A(w)$) obtidos do ANEW.

$$p(w) = \frac{|V(w) - 5| \times A(w)}{36} \quad (\text{IV.1})$$

Na fórmula apresentada na Equação IV.1 o valor de valência é subtraído de 5, deslocando sua escala original $[1, 9]$ para o intervalo $[-4, 4]$. Assim, a neutralidade da valência fica representada pelo valor central 0. O valor resultante é multiplicado pelo valor de excitação (A , do inglês *arousal*). Como a escala de excitação não possui um valor central, permanece no mesmo intervalo $[1, 9]$. Ao final, é feita a divisão por 36, de forma que o resultado seja um valor no intervalo $[-1, 1]$.

A partir do modelo *bag of words* (*BOW*) [Kao and Poteet, 2007] obtido nas fases anteriores à utilização do modelo proposto nesse trabalho, a aplicação do LIWC_{AN} é feita seguindo a abordagem descrita no Algoritmo 1.

V é o vetor que se espera ter como resultado, encarregado de guardar as categorias e suas respectivas frequências. Este é inicializado com um vetor vazio e contendo índices para todas as categorias do LIWC por meio da função *initialize*, aguardando ser preenchido nos passos seguintes. Para cada palavra (t) existente no *BOW*, é feita a busca ao LIWC por meio da função *findWord*, que retorna um vetor C com as categorias associadas àquela palavra. Após, é feito o cálculo do peso obtido no ANEW usando a Equação IV.1. Caso a palavra solicitada não exista no ANEW, o peso retornado é 0, sendo apenas o peso do LIWC (w_L) acrescentado. Cabe ressaltar que esta divisão foi realizada de forma a permitir o ajuste das contribuições do ANEW e do LIWC. Assim, a parcela de peso para o LIWC é parametrizada por um parâmetro $\alpha_L \in [0, 1]$ e do ANEW por um parâmetro $\alpha_A = 1 - \alpha_L$.

Algorithm 1 Algoritmo de contagem de categorias por meio da aplicação do LIWC_{AN}.

```

 $T = BOW, w_L = \text{peso LIWC}$ 
Require:  $T \neq \emptyset$ 
 $initialize(V)$ 
for all token  $t \in T$  do
   $C = findWord(t)$ 
   $w_A = p(t)$ 
  if  $C \cap \{posemo, negemo, anx, anger, sad\} \neq \emptyset$  then
    if  $searchNegation(t)$  then
       $invert(C)$ 
    end if
     $w_I \leftarrow searchIntensifier(t)$ 
  end if
  for all  $c \in C$  do
     $V[c] \leftarrow V[c] + w_I \cdot (\alpha_L \cdot w_L + \alpha_A \cdot w_A)$ 
  end for
end for
return  $V$ 

```

Após a obtenção dos valores de peso, o valor das categorias em C é armazenado na posição correspondente de V , considerando o peso do ANEW encontrado. O procedimento se repete até que todas as palavras do BOW tenham sido visitadas. Ao final do procedimento, temos o vetor V preenchido. Uma imagem deste vetor pode ser vista na Figura IV.4.

negemo	posemo	anx	anger	sad
35,6	40,1	10,5	22,7	18,5

Figura IV.4: Exemplo de vetor ponderado com suas respectivas categorias

A satisfação é então calculada a partir da razão entre o valor da categoria de emoção positiva com a soma dos valores encontrados para categorias de emoção negativa. As categorias levadas em conta neste trabalho foram de emoção positiva (*posemo*), emoção negativa (*negemo*), ansiedade (*anx*), raiva (*anger*) e tristeza (*sad*), sendo as demais categorias desconsideradas.

É importante ressaltar que, apesar da categoria *negemo* contemplar as subcategorias *anx*, *sad* e *anger*, nem toda palavra negativa pertence a uma das três últimas categorias citadas. De forma a se estabelecer um certo rigor, o modelo verifica se toda palavra contida em uma das subcategorias de emoção negativa (*anx*, *anger* e *sad*) está também contida na categoria de emoção negativa (*negemo*). Caso positivo, a categoria *negemo* não é contabilizada. Assim, a categoria de emoção negativa é contada levando em conta apenas as emoções negativas que não se encontram em suas três subcategorias. Para o exemplo da Figura IV.4, a satisfação encontrada é de $35,6 / (35,6 + 40,1 + 10,5 + 22,7 + 18,5)$ resultando em 0,28 ou 28%.

IV.3.2 Detecção de negação e intensificadores/atenuadores

As palavras de negação (e.g., não, nem, nunca, nada, ninguém, etc.), conhecidas como *shifter words*, tem a função de inverter a polaridade de uma palavra qualificadora. Deste modo, quando há a presença de tais palavras, ocorre uma inversão das categorias: as palavras positivas são acrescentadas à categorias negativas e vice-versa.

A regra para detecção das palavras de negação, primeiro é avaliado se a palavra sendo classificada é uma palavra de qualidade. É considerada uma palavra de qualidade toda aquela que se encontra em uma das 5 categorias do LIWC utilizadas (*posemo*, *negemo*, *anx*, *anger* e *sad*). Caso seja, é feita a busca por uma palavra de negação (função *searchNegation*) imediatamente próxima da palavra de qualidade. Encontrada a negação, é feita a inversão da polaridade da palavra (função *invert*). Caso contrário é mantida a polaridade encontrada.

No modelo aqui proposto a busca pela negação é feita de duas maneiras. A primeira forma é buscar a negação imediatamente antes da palavra de qualidade. Por exemplo, na sentença “**Não** gostei da aula”, a palavra de negação é encontrada imediatamente antes da palavra de qualidade, invertendo sua polaridade. A segunda forma é buscar pela negação imediatamente antes do verbo mais próximo da palavra qualificadora. Como o LIWC contém categoria de verbo, o algoritmo realiza uma busca pelo verbo mais próximo no sentido contrário das palavras da sentença. Ao encontrar o verbo, verifica-se a existência de uma negação que o anteceda. Por exemplo, na sentença “A aula **não** foi boa”, a palavra de negação é encontrada imediatamente antes do verbo, invertendo a polaridade da palavra de qualidade *boa*.

As palavras de intensidade, conhecidas como *intensifier words*, nada mais são que advérbios de intensidade, como *muito*, *pouco*, *mais*, *menos*, dentre outros. Deste modo, quando há a presença de tais palavras, estas exercem uma função de intensificar ou atenuar uma palavra positiva ou negativa.

A regra para detecção de palavras intensificadoras segue uma abordagem similar à da negação, no entanto, a procura é feita para palavras que pertençam à categoria de intensidade. A busca por uma palavra de intensidade (função *searchIntensifier*) é feita imediatamente antes da palavra qualificadora. Caso não encontre, é feita a busca imediatamente após a palavra qualificadora. Por exemplo, na sentença “A aula foi **muito** boa”, a palavra intensificadora é encontrada imediatamente antes da palavra qualificadora. Por outro lado, na sentença “Gostei **muito** da aula” a palavra intensificadora é encontrada imediatamente após a palavra qualificadora. Em ambos os exemplos as palavras qualificadoras são intensificadas. Na sentença “Aula foi **pouco** explicativa” é possível observar um exemplo de sentença onde a palavra qualificadora é atenuada.

O uso combinado da negação com advérbios de intensidade é avaliado de forma diferenciada. Por exemplo, na sentença “**nem tão** boa”, não se pode considerar a negação (*nem*) como inversora

da polaridade de “boa” nem a palavra intensificadora (*tão*) no sentido de amplificar essa negação. Neste caso, o uso do intensificador combinado com a negação está fazendo papel de um leve amenizador. Ou seja, o aluno quer dizer que a aula foi boa, porém não em sua totalidade, funcionando como um leve amenizador.

Uma outra situação acontece quando há o uso de uma palavra de negação juntamente com uma palavra atenuadora. Por exemplo, na sentença “**nem** um **pouco** boa”, não se pode considerar o atenuador combinado com a negação como uma atenuação da palavra qualificadora *boa*. Neste caso, a interpretação mais comum é de que a palavra atenuadora pode ser desconsiderada e o sentido da frase diretamente invertido pela negação.

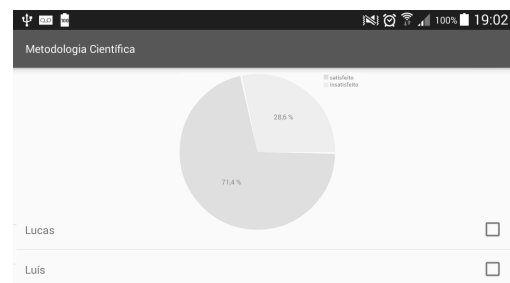
No modelo proposto nesse trabalho, a detecção da negação e de palavras intensificadora e atenuadoras é feita após o cálculo da polaridade da palavra, conforme apresentado na Seção IV.3.1. O resultado, portanto, é um valor no intervalo $[-1, 1]$, de acordo com a Equação IV.1. O uso de um intensificador multiplica esse peso por um fator β_I e o uso de um atenuador multiplica esse peso por β_A , conforme discutido na Seção IV.3. Já no caso de um leve amenizador pela combinação da negação com um intensificador, o peso é multiplicado por β_N .

A Seção a seguir apresenta detalhes da implementação do VITA.

IV.4 Implementação

Conforme já apresentado anteriormente, a ferramenta VITA está implementada para dispositivos Android, contando também com uma versão Web. Ambos os perfis de alunos e professores podem acessar o sistema, considerando que as opções disponíveis são distintas de acordo com o perfil de acesso. Enquanto o aluno apenas envia a avaliação, o professor apenas as visualiza. A Figura IV.5a mostra o *layout* de envio por parte do aluno. Já a Figura IV.5b apresenta a visão da turma pelo professor para dispositivos Android.

(a) Versão Web para perfil aluno.



(b) Versão para perfil professor.

Figura IV.5: Telas principais das versões aluno e professor do VITA.

Na versão aluno, a tela principal apresenta a lista de disciplinas disponíveis e permite ao aluno escolher para qual disciplina será fornecida a avaliação, selecionando-a na lista. À direita do nome

da disciplina, uma caixa de marcação (*checkbox*) indica, caso marcada, se foram concluídas todas as avaliações daquela disciplina. Caso haja alguma pendência, a caixa aparece desmarcada. A Figura IV.6a apresenta um exemplo da tela em questão.



Figura IV.6: Telas de seleção da versão aluno do VITA.

Ao selecionar uma disciplina, o aplicativo apresenta uma segunda tela onde são listadas todas as aulas já lecionadas da disciplina escolhida, juntamente com a data da aula. À direita de cada aula, novamente aparece a *checkbox*, que neste caso, serve para identificar se já foi realizada a avaliação daquela aula específica. A Figura IV.6b apresenta um exemplo dessa tela. Quando o usuário escolhe uma aula para fazer a avaliação, o aplicativo passa para a tela de avaliação similar àquela apresentada na Figura IV.5a, fornecendo um campo específico para a escrita do texto da avaliação.

Já a versão do aplicativo para o professor apresenta algumas visualizações diferentes. Conforme esperado, não há a tela para enviar a avaliação, já que não cabe ao professor. Ao invés disso, é apresentada uma tela mostrando a avaliação geral da turma, a lista de alunos, e uma caixa de seleção informando qual aluno fez a avaliação e qual está pendente, como apresentado na Figura IV.5b. Caso um aluno seja selecionado, sua avaliação individual aparecerá numa nova tela que se abre.

Capítulo V Avaliação do VITA

Apesar de a aplicação da análise de sentimentos feita por este trabalho estar amparado por vários estudos, é necessário verificar em termos práticos se o modelo aqui proposto funciona de maneira eficiente. Por isso, os experimentos conduzidos buscaram avaliar a ferramenta de três pontos. O primeiro visando verificar se os resultados obtidos pela ferramenta são fiéis à realidade da satisfação dos alunos. O segundo visando verificar o quão próximo é a avaliação da ferramenta em relação ao entendimento do aluno a respeito da sua própria satisfação. O terceiro visando verificar se a avaliação produzida pela ferramenta é realmente útil para o professor. Para isso, foram conduzidos dois experimentos: o primeiro foi realizado de forma bastante simples envolvendo três turmas, sendo o segundo um experimento mais amplo.

Os valores dos parâmetros apresentados na Seção IV.3 para realização dos experimentos foram pré-definidos da seguinte maneira: a distribuição de pesos do LIWC e ANEW foram estabelecidas de forma que o ANEW-br contribua mais com o LIWC. Logo, α_L foi definido em 0.3 e α_A foi definido em 0.7. Já os intensificadores β_I e β_A foram definidos para valores de 1.75 e 0.25, respectivamente. Já o leve amenizador β_N foi definido para 0.75.

No primeiro experimento foram recolhidos 14 comentários textuais, avaliados por 4 professores do curso. Foram escolhidos professores com bastante experiência em docência. Cada professor classificou o nível de satisfação presente nos comentários numa escala de 0 a 10. O experimento realizado visou estimar a proximidade da avaliação do VITA com aquela feita pelos professores. O grau de satisfação fornecido pela ferramenta foi ajustado para uma escala de satisfação VITA (n_{vita}) de 0 a 10. Para efeito de comparação, foi considerado o nível de satisfação identificado pelos professores (n_{prof}) como a média das classificações feitas pelos professores.

Por fim, foi calculada a diferença entre as duas avaliações (n_{vita} e n_{prof}) utilizando a correlação de Spearman [Myers and Sirois, 2004]. A correlação de Spearman mostra o quão correlacionado estão duas variáveis, apresentando um valor no intervalo $[-1, 1]$. Uma variável próxima de 1 possui perfeita correlação com outra, enquanto o valor -1 indica uma perfeita correlação inversa. O valor 0 mostra que as variáveis não se relacionam de forma alguma. O valor de correlação de Spearman obtido para este cenário foi de 0,58. Assim, concluímos que a ferramenta apresentou resultados preliminares positivos.

Já no segundo experimento foram convidados 65 alunos de quatro disciplinas do curso de Ba-

charelado em Ciência da Computação. De forma efetiva, participaram do experimento 24 alunos de quatro disciplinas, sendo elas: Metodologia Científica (MC), Estrutura de Dados e Algoritmos (EDA), Gerência de Projetos de Tecnologia da Informação (GPTI), e Programação Orientada a Objetos (POO). O período de avaliação se deu durante 10 semanas, coletando um total de 68 avaliações de alunos distribuídas durante o período contemplado, resultando em uma média de 2 a 3 comentários por aluno.

Durante o experimento foi solicitado aos alunos que informassem uma nota de 0 a 5, para fins de estabelecer uma correlação entre a nota informada e a satisfação avaliada pela ferramenta. As avaliações foram enviadas para diferentes dias de aula. Essa avaliação teve como principal meta estimar a proximidade da avaliação do VITA com àquela feita pelos alunos.

Dentre os comentários coletados, foi realizada uma seleção de modo a extrair apenas os comentários considerados válidos para análise. Ficou estabelecido que para o comentário ser considerado válido, deveria seguir os seguintes critérios:

1. Possuir pelo menos 70 palavras (legíveis);
2. Conter a opinião do aluno - Conteúdo deve se restringir à opinião sobre a aula (e.g., não comentar sobre fatos ou assuntos diversos);
3. Fornecer nota consistente - A nota informada pelo aluno não pode ser incompatível com o que foi escrito (e.g., fornecer a nota mínima para uma aula muito bem avaliada na escrita).

Comentários que não se enquadraram nestes quesitos foram considerados *outliers* e foram excluídos da análise. Por este motivo, dos 68 comentários, 7 não seguiram os critérios estabelecidos, sendo portanto removidos da análise, restando o total de 61 amostras.

A partir destes dados coletados, foi possível gerar uma matriz de confusão que compara valores de satisfação estimados por VITA e valores indicados pelos alunos para cada classificação fornecida. Desta vez, a escala para comparação com a nota dos alunos foi transformada para o intervalo entre 0 e 5. O valor mais alto da escala (5) indica plena satisfação do aluno, enquanto o menor valor da escala (0) indica que o aluno não está satisfeito de forma alguma.

Foi também aplicada, nesta matriz, uma discretização seguindo um critério de arredondamento, de forma a se ter valores inteiros. Por exemplo, se o valor encontrado por VITA fosse de 4.5, seria arredondado para 5. Caso houvesse um valor com casa decimal inferior a 0.5 (e.g., 3.4), este seria arredondado para 3. Nesta matriz é possível perceber quantas vezes VITA classificou de forma semelhante ou diferenciada em relação a avaliação feita pelo aluno. Os valores em negrito equivalem à quantidade de comentários que VITA classificou de forma semelhante. A matriz de confusão foi gerada diretamente pelo pacote *caret* do *RStudio*. A Tabela V.1 mostra a matriz de confusão.

		Aluno				
		Nota 5	Nota 4	Nota 3	Nota 2	Total
VITA	Nota 5	20	7	0	0	27
	Nota 4	13	13	0	0	26
	Nota 3	0	4	3	1	8
	Nota 2	0	0	0	0	0
	Total	33	24	3	1	61

Tabela V.1: Matriz de confusão para valores estimados por VITA (n_{vita}) e àqueles informados pelo aluno (n_{aluno}).

A partir desta matriz, se torna fácil encontrar medidas bastante úteis como a precisão e a abrangência. O cálculo da precisão é uma forma de se verificar o quão bom um classificador é em relação à classificação positiva. Ou seja, qual o percentual de amostras classificadas corretamente (verdadeiro-positivos - tp) dentre o universo total de amostras classificadas. Deste modo, a precisão de VITA pode ser encontrada por meio do cálculo entre o número de amostras corretamente classificadas por VITA (tp) e o total de amostras, que é a soma entre os tp e os falso-positivos (fp). Já a abrangência é a relação entre tp e a soma entre fp e os falso-negativos (fn).

$$\text{Precisão} = \frac{tp}{tp + fp} \quad (\text{V.1})$$

$$\text{Abrangência} = \frac{tp}{tp + fn} \quad (\text{V.2})$$

A partir destas fórmulas, é possível gerar uma tabela contendo as notas (classes) e seus respectivos valores de tp , fp , fn , bem como a precisão e abrangência para tais classes. A Tabela V.2 mostra estes valores. Nesta tabela é possível verificar que os melhores resultados classificados por VITA se deram na classe 5. A precisão e a abrangência gerais encontradas foram 0,56 e 0,57 respectivamente.

Classe	tp	fp	fn	Precisão	Abrangência
Nota 5	20	7	13	0,74	0,60
Nota 4	13	13	11	0,50	0,54
Nota 3	3	5	0	0,38	1,00
Total	33	25	24	0,56	0,57

Tabela V.2: Comparação entre os verdadeiro-positivos (tp), falso-positivos (fp) e falso-negativos (fn) para as diferentes notas (classes) estimadas por VITA. A precisão e abrangência também foram calculadas para cada classe. O melhor resultado encontrado se deu para a Nota 5

Além disso, foi também calculada a correlação de Spearman para os valores obtidos entre VITA (n_{vita}) e a nota do aluno (n_{aluno}), obtendo um valor de 0.46 para 61 amostras avaliadas e um p -value de 1.9×10^{-4} . Um gráfico de dispersão gerado para esta correlação pode ser visualizado na Figura V.1.

Foi também importante conhecer o quanto as notas de VITA se diferenciaram da nota fornecida

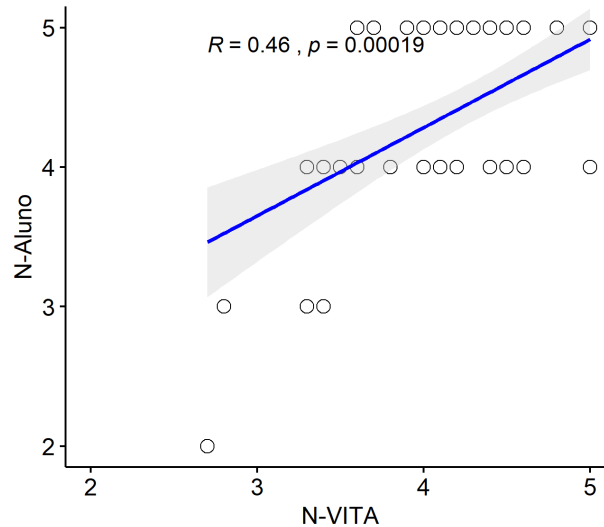


Figura V.1: Gráfico de dispersão da correlação de Spearman entre a satisfação estimada por VITA e a indicada pelo aluno. A linha em azul mostra a correlação entre os dois eixos. Quanto mais próxima de um ângulo de 45° , maior a correlação linear entre as variáveis e quanto mais próxima do centro, mais os valores de um eixo coincidem com os do outro. A faixa destacada em cinza mostra o intervalo de confiança. Quanto mais estreito, maior o nível de confiança na estimativa.

pelo aluno. Para isso foi calculada a média do módulo da diferença entre n_{vita} e n_{aluno} para todas as 61 amostras de comentários. Além disso, também foi calculado o desvio-padrão. Os valores encontrados foram de 0.48 para a média e 0.39 para o desvio-padrão. Sendo assim, 68% das notas estimadas por VITA diferiram da nota do aluno por, no mínimo 0,09 e no máximo 0,87.

Além de coletarmos a nota que os alunos forneceram e compará-las com a nota estimada por VITA, dentre os 61 comentários dos alunos, em 28 destes obtivemos a nota em que o professor opinava sobre qual seria a satisfação do aluno para determinada aula. Em outras palavras, foi comparada a percepção do professor com a percepção do aluno relacionados à satisfação. Os 33 comentários faltantes se devem ao fato do professor não conseguir identificar o aluno pelo nome, não tendo sido possível atribuir uma nota. A correlação de Pearson foi calculada com valor de 0.52 e p -value de 0.005 para 28 amostras. A Figura V.2 mostra o gráfico desta correlação.

Por fim, comparamos também a relação entre as notas atribuída pelo VITA (n_{vita}) e a opinião do professor (n_{prof}). A correlação foi de 0,07 para 28 amostras e p -value de 0,71. Tal correlação foi considerada desprezível, indicando que não foi possível estabelecer correlação alguma entre as notas estimadas por VITA e a opinião do professor (n_{vita}). A Figura V.3 mostra o gráfico desta correlação.

De forma a se aferir a proximidade da satisfação calculada por VITA para toda uma turma ao longo de um determinado intervalo de aulas, apresentamos mais um comparativo, desta vez em formato de gráfico temporal. Neste comparativo temos dois gráficos: um foi obtido por meio da nota dos alunos e o outro gerado pelo sistema aqui proposto. Cada aula desenhada no gráfico é a

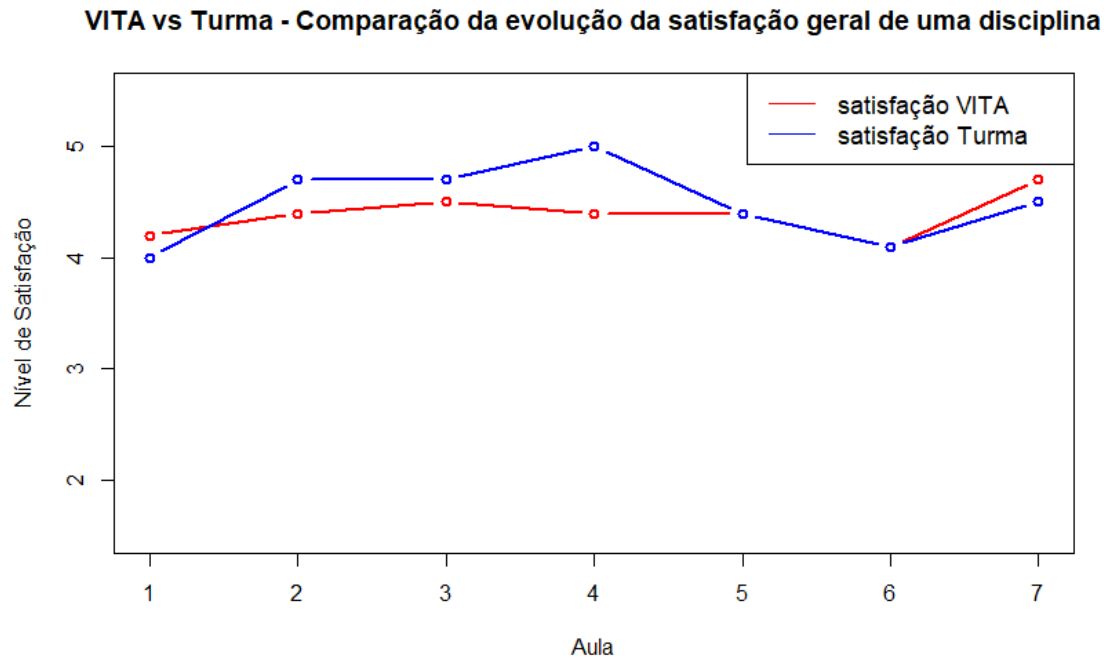


Figura V.4: Gráfico comparativo da média aritmética das notas preenchidas pelos alunos da disciplina MC durante um período de 7 aulas e da média calculada por VITA. O gráfico em azul foi gerado pelo sistema proposto enquanto o gráfico em vermelho é relativo à nota dos alunos.

V.1 Discussão e considerações finais

Levando-se em conta que o sistema proposto apresentou uma correlação de Pearson de 0,46 quando confrontada com a nota do aluno, esta se caracteriza uma correlação fraca. Entretanto, algumas considerações devem ser feitas para melhor interpretação e profundidade na análise dos resultados obtidos. Em primeiro lugar, o comentário do aluno nem sempre é consistente com a nota informada à aula. Observemos o comentário fornecido a uma aula:

“Aula boa, consegui pegar bem o conteúdo e sanar minhas dúvidas. O maior problema é a falta de criação de programas durante a aula, sendo os exercícios copiados do quadro e várias partes são quase que obrigatórias a serem perguntadas, pela falta de entendimento do professor de que os alunos não entendem ou não tiveram essa parte da matéria. E outro problema é o horário, que torna a aula muito cansativa.”

Analisando este comentário, é perfeitamente visível que o aluno não está plenamente satisfeito com aula, pois cita alguns problemas, como a falta de uma parte prática, a grande necessidade de tirar dúvidas em conteúdos dos quais o professor considera que o aluno já o conhece, bem como o horário que torna a aula cansativa. No entanto, a nota fornecida pelo aluno a esta mesma aula foi 5, o que representa um certo contraste com o comentário escrito. Em contrapartida, VITA classificou este comentário como nota 3, o que aparenta estar mais próximo da realidade expressa no comentário.

Já em relação a correlação de Pearson entre a opinião do professor e do aluno, esta obteve um resultado inferior quando comparada com a correlação entre VITA e a nota do aluno. Este ponto corrobora com a hipótese de que a percepção do professor não é suficiente para saber como está a qualidade de aula segundo os alunos. No entanto, há de se considerar que a quantidade de amostras, no primeiro caso, foi significativamente inferior (28) ao segundo caso (61), impactando diretamente no resultado. Logo, testes com mais amostras seriam necessários para que, de fato, fosse possível formar uma conclusão a respeito deste comparativo.

Além disso, o gráfico produzido na correlação de Pearson para n_{prof} e n_{aluno} possui um deslocamento para a parte superior, em relação ao centro do gráfico, indicando a tendência do professor subestimar a satisfação do aluno. Ou seja, neste experimento o professor esteve mais propenso a considerar o aluno menos satisfeito do que verdadeiramente estava. No caso de VITA o gráfico se apresenta mais próximo do centro, indicando maior simetria na estimativa da satisfação.

Por meio de uma entrevista com o professor sobre a utilidade da estimativa apresentada pela ferramenta, pode-se perceber que boa parte das avaliações foram aderentes a percepção do professor. Alunos que tinham sido percebidos pelo professor como não estando empolgados com a matéria apresentaram uma estimativa de satisfação próxima a esperada. Um ponto destacado pelo professor foi que, para os alunos cuja percepção era de que gostavam da matéria, as estimativas de satisfação do VITA superaram a expectativa do professor.

É também importante mencionar que a ferramenta desenvolvida apresentará uma maior precisão conforme se aumenta a quantidade de comentários. Assim, a satisfação mostrada para uma turma não deve ser considerada válida no caso de haver apenas uma fração de comentários para determinada disciplina.

Capítulo VI Conclusão

Pesquisas na área da educação indicam que o nível de satisfação do estudante é um importante indicador na medida do sucesso de uma universidade. Estudantes com altos níveis de satisfação são mais propensos a continuar no curso, em contraste àqueles que estão insatisfeitos. Tal cenário se agrava em turmas com um grande número alunos ou que não permitam um contato mais próximo entre aluno e professor (como no ensino à distância). Dessa forma, torna-se mais complicado para o professor acompanhá-los individualmente. Sendo de grande valia a criação de um mecanismo capaz de melhorar a percepção do professor a respeito dos alunos.

O presente trabalho apresentou um método alternativo de avaliação do ensino por parte dos alunos. A avaliação é feita por meio da escrita livre de comentários a respeito das aulas lecionadas. Dessa maneira, o aluno pode discorrer sem restrições sobre cada aula ministrada, tendo maior liberdade para expressar sua opinião. O resultado apresentado ao professor é na forma de gráficos que indicam a satisfação dos alunos de uma turma individualmente ou para a classe como um todo. Assim o aluno pode se sentir mais a vontade para expressar sua opinião dado que o texto escrito não é acessado pelo professor.

Ressalta-se que por permitir que a avaliação seja feita continuamente, a proposta apresentada nesse trabalho permite uma maior agilidade na avaliação de turmas grandes. Dada a possibilidade de acompanhar a evolução da satisfação dos alunos ao longo do semestre, o professor pode realizar alterações na condução de suas aulas visando evitar altos índices de insatisfação e consequente abandono do curso. Além disso, em cenários de turmas com grande número de alunos ou que não permitam um contato mais próximo entre aluno e professor, pretende-se estimular uma maior aproximação entre alunos e professores com o uso da abordagem aqui proposta.

A proposta desse trabalho se baseia no uso de técnicas de Análise de Sentimentos para identificar as emoções contidas em comentários escritos de alunos, estimando assim sua satisfação. O modelo proposto estima o nível de satisfação do com base em léxicos afetivos. É feita a combinação dos dicionários LIWC com o ANEW, ambos em sua versão em português, tendo como objetivo melhorar a precisão do cálculo da polaridade do comentário escrito. O modelo proposto leva também em consideração possíveis inversões de polaridade do texto devido ao uso de palavras de negação. São considerados também efeitos de intensificação ou atenuação da polaridade do texto causada pelo uso de advérbios de intensidade.

O modelo de estimação da satisfação do aluno serviu de base para a implementação da ferramenta VITA (*Virtual Therapist*). VITA é implementado em uma abordagem cliente-servidor. O servidor VITA recebe os comentários dos alunos e estima a satisfação a ele associado. Já o cliente VITA, permite o envio do comentários dos alunos, bem como o acesso do professor ao sistema, verificando o nível de satisfação geral da turma, bem como a satisfação individual dos alunos. Duas implementações do cliente VITA (versão aluno) estão disponíveis: como um aplicativo para Android ou como uma página web.

Com base na ferramenta VITA foram feitos experimentos com turmas do curso de Bacharelado em Ciência da Computação. Resultados dos experimentos apontam que a avaliação feita por VITA apresenta uma boa precisão quando comparada com avaliações feitas por professores dos comentários escritos. Ainda, a ferramenta apresentou boa precisão em relação a nota atribuída pelos alunos das aulas, quando desconsideradas discrepâncias entre o texto escrito e a nota atribuída pelo aluno. Por fim, a ferramenta se apresentou útil para os casos em que o professor não tinha um bom conhecimento do aluno sendo avaliado.

Dessa forma, ressalta-se as principais contribuições deste trabalho como: (i) ter proposto uma adaptação do léxico afetivo LIWC para alterar a polaridade de uma palavra com base no uso de palavras de negação e intensificadores, (i) ter apresentado um modelo de Análise de Sentimentos para a extração do grau de satisfação de alunos a partir de textos escritos e (ii) ter disponibilizado um aplicativo para dispositivos móveis e uma página web que permite o uso deste modelo em sala de aula. Além disso, ressalta-se que o trabalho aqui apresentado foi publicado no SBIE (Simpósio Brasileiro de Informática na Educação) [Preuss et al., 2018].

Um limiar para a precisão alcançada pela ferramenta está relacionado ao modelo de Análise de Sentimentos adotado. Como a proposta apresentada nesse trabalho foca no uso dos léxico afetivos LIWC e ANEW, a análise da polaridade do texto depende da capacidade de tais léxicos de identificar essa polaridade. Um trabalho futuro nesse sentido é integrar outros léxicos ao modelo proposto, visando uma maior precisão na polaridade estimada. Outro possível trabalho futuro se encontra na adição do uso da voz (*Speech-To-Text*) tanto por questões de acessibilidade, quanto para tornar possível capturar outras características emocionais, não presentes em texto, como a entonação da voz. Ainda com relação ao aplicativo, um trabalho futuro é a adição de um fotômetro, de forma a facilitar a identificação de um aluno pelo professor.

Acresce-se a isso que o modelo de VITA ainda é um modelo implementado de forma bastante simples. Sendo assim, não estão presentes etapas da Análise de Sentimentos como a Desambiguação de Limite de Sentença – do inglês Sentence Boundary Disambiguation (SBD). A SBD tem por objetivo separar as frases e sentenças de um texto. Outra melhoria possível concerne a implementação da detecção de características individuais de forma independente, detalhando a satisfação

por subitens, como a pontualidade do professor, didática, andamento da aula, e outras, que seriam possíveis ao implementar a Análise de Sentimentos baseada em Aspectos – do inglês Aspect-Based Sentiment Analysis (ABSA). Ademais, diferentemente do LIWC, o ANEW-br é sensível à ausência das etapas de lematização ou estemização, que não foram implementadas já que não era uma necessidade inicial, visto que LIWC já possui tais variações de forma das palavras. Ainda, um outro trabalho futuro considerando o modelo aqui proposto é o uso de expressões idiomáticas, bem como do uso de duplo intensificadores (e.g., muito pouco) como forma de calcular os pesos de palavras de qualidade.

Atualmente o modelo aqui proposto foi implementado para seu uso por dispositivos móveis, requerendo que os alunos escrevam comentários para cada aula. Uma possível extensão de VITA é seu uso integrado a sistemas já existentes de acompanhamento das aulas, como, por exemplo, o Moodle. Dessa forma podem ser utilizados comentários escritos pelos alunos nesses sistemas para a identificação do seu nível de satisfação.

Por fim, os experimentos realizados e aqui retratados, visaram avaliar a precisão da estimativa realizada por VITA, bem como sua utilidade para o professor. Um trabalho em andamento é a utilização do VITA em quatro turmas regulares do curso de Ciência da Computação, abrangendo alunos do segundo ao quinto períodos do curso. Este uso contínuo da ferramenta permitirá avaliar sua influência na condução das aulas. Um trabalho futuro é avaliar como o uso de uma ferramenta como o VITA pode influenciar na redução da evasão escolar.

Referências Bibliográficas

- Andrews, B. and Wilding, J. M. (2004). The relation of depression and anxiety to life-stress and achievement in students. *British Journal of Psychology*, 95(4):509–521. 12
- Aung, K. Z. and Myo, N. N. (2017). Sentiment analysis of students’ comment using lexicon based approach. In *2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS)*, pages 149–154. 13
- Balahadia, F. F., Fernando, M. C. G., and Juanatas, I. C. (2016). Teacher’s performance evaluation tool using opinion mining with sentiment analysis. In *2016 IEEE Region 10 Symposium (TENSYP)*, pages 95–98. 13, 14
- Barrón-Estrada, M. L., Zatarain-Cabada, R., Oramas-Bustillos, R., and González-Hernández, F. (2017). Sentiment analysis in an affective intelligent tutoring system. In *2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT)*, pages 394–397. 13
- Baylor, A. L. and Rosenberg-Kima, R. B. (2006). Interface agents to alleviate online frustration. In *Proceedings of the 7th international conference on Learning sciences*, pages 30–36. International Society of the Learning Sciences. 12
- Bradley, M. M., Lang, P. J., Bradley, M. M., and Lang, P. J. (1999). Affective norms for english words (anew): Instruction manual and affective ratings. 2, 11
- Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., and Kuksa, P. (2011). Natural language processing (almost) from scratch. *Journal of Machine Learning Research*, 12(Aug):2493–2537. 7
- de Souza, S. A. and Reinert, J. N. (2010). Avaliação de um curso de ensino superior através da satisfação/insatisfação discente. *Avaliação: Revista da Avaliação da Educação Superior*, 15(1). 1
- Elliott, K. M. and Shin, D. (2002). Student satisfaction: An alternative approach to assessing this important concept. *Journal of Higher Education Policy and Management*, 24(2):197–209. 1

- Filho, P. P. B., Pardo, T. A. S., and Aluísio, R. M. (2013). An evaluation of the brazilian portuguese liwc dictionary for sentiment analysis. In *Proceedings of the 9th Brazilian Symposium in Information and Human Language Technology*, pages 215–219, Fortaleza, CE, Brazil. Sociedade Brasileira de Computação. 8, 9
- Gonçalves, P., Araújo, M., Benevenuto, F., and Cha, M. (2013). Comparing and combining sentiment analysis methods. In *Proceedings of the first ACM conference on Online social networks*, pages 27–38. ACM. 10
- Gonçalves, P., Araújo, M., Benevenuto, F., and Cha, M. (2013). Comparing and combining sentiment analysis methods. In *Proceedings of the First ACM Conference on Online Social Networks*, pages 27–38. 8
- Guedes, G. P., Bezerra, E., Ferrari, L., and Duarte, F. (2016). Diferenças de gênero no uso de português em redes sociais: evidências a partir do liwc. In *WebMedia 2016 Short Papers ()*. 21
- Gupta, V. and Lehal, G. S. (2009). A survey of text mining techniques and applications. *Journal of Emerging Technologies in Web Intelligence*, pages 60 – 76. 8
- Hassenzahl, M. (2008). User experience (ux): Towards an experiential perspective on product quality. In *Proceedings of the 20th Conference on L’Interaction Homme-Machine*, pages 11–15. 2
- Heimberg, R. G., Turk, C. L., and Mennin, D. S. (2004). *Generalized anxiety disorder: Advances in research and practice*. Guilford Press. 12
- Hotho, A., Nürnberger, A., and Paaß, G. (2005). A brief survey of text mining. *LDV Forum - GLDV Journal for Computational Linguistics and Language Technology*. 6, 7
- Kao, A. and Poteet, S. R. (2007). *Natural language processing and text mining*. Springer Science & Business Media. 21
- Leiberg, S. and Anders, S. (2006). The multiple facets of empathy: a survey of theory and evidence. In Anders, S., Ende, G., Junghofer, M., Kissler, J., and Wildgruber, D., editors, *Understanding Emotions*, volume 156 of *Progress in Brain Research*, pages 419 – 440. Elsevier. 5
- Liddy, E. D. (2001). Natural language processing. 5, 6
- Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1):1–167. 2, 8

- Manhães, L. M. B., da Cruz, S. M. S., Costa, R. J. M., Zavaleta, J., and Zimbrão, G. (2011). Previsão de estudantes com risco de evasão utilizando técnicas de mineração de dados. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*, volume 1. 1
- Matsubara, E. T., Martins, C. A., and Monard, M. C. (2003). Pretext: Uma ferramenta para pré-processamento de textos utilizando a abordagem bag-of-words. *Technical Report*, 209. 6
- Myers, L. and Sirois, M. J. (2004). Spearman correlation coefficients, differences between. *Encyclopedia of statistical sciences*, 12. 26
- Newman, M. L., Pennebaker, J. W., Berry, D. S., and Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin*, 29(5):665–675. PMID: 15272998. 10
- Ortigosa, A., Martin, J., and Carro, R. (2014). Sentiment analysis in facebook and its application to e-learning. 31:527–541. 13, 14
- Pang, B., Lee, L., et al. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, 2(1–2):1–135. 8
- Pang, B., Lee, L., and Vaithyanathan, S. (2002). Thumbs up?: sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*, pages 79–86. Association for Computational Linguistics. 1
- Pennebaker, J. W., Boyd, R. L., Jordan, K., and Blackburn, K. (2015). The development and psychometric properties of liwc2015. Technical report. x, 9
- Picard, R. W. (1995). Affective computing. 5
- Picard, R. W. (2010). Affective computing: From laughter to iccc. *IEEE Transactions on Affective Computing*, 1(1):11–17. 5
- Pong-Inwong, C. and Kaewmak, K. (2016). Improved sentiment analysis for teaching evaluation using feature selection and voting ensemble learning integration. In *2016 2nd IEEE International Conference on Computer and Communications (ICCC)*, pages 1222–1225. 13
- Preuss, L., dos Santos, J., and Guedes, G. P. (2018). Vita-estimando a satisfação de estudantes por meio da análise de sentimentos. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*, volume 29, page 1143. 2, 19, 34

- Pritchard, M. E. and Wilson, G. S. (2003). Using emotional and social factors to predict student success. *Journal of college student development*, 44(1):18–28. 12
- Rani, S. and Kumar, P. (2017). A sentiment analysis system to improve teaching and learning. *Computer*, 50(5):36–43. 13, 14
- Robinson, R. L., Navea, R., and Ickes, W. (2013). Predicting final course performance from students’ written self-introductions: A liwc analysis. *Journal of Language and Social Psychology*, 32(4):469–479. 13, 15
- Rodrigues, R. G. a. and Guedes, G. P. (2017). A hybrid affective lexicon for brazilian portuguese. In *Proceedings of the XVI Brazilian Symposium on Human Factors in Computing Systems, IHC 2017*, pages 52:1–52:4, New York, NY, USA. ACM. 19, 21
- Rude, S., Gortner, E.-M., and Pennebaker, J. (2004). Language use of depressed and depression-vulnerable college students. 18:1121–1133. 13, 15
- Russell, J. (1980). A circumplex model of affect. 39:1161–1178. viii, 12
- Seldin, P. (1993). The use and abuse of student ratings of professors. *The Chronicle of Higher Education*. 1
- Tan, A.-H. (1999). Text mining: The state of the art and the challenges. In *In Proceedings of the PAKDD 1999 Workshop on Knowledge Discovery from Advanced Databases*, pages 65–70. 6
- Tausczik, Y. R. and Pennebaker, J. W. (2010). The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of language and social psychology*, 29(1):24–54. 2, 8, 9, 10
- Vaez, M. and Laflamme, L. (2008). Experienced stress, psychological symptoms, self-rated health and academic achievement: A longitudinal study of swedish university students. *Social Behavior and Personality: an international journal*, 36(2):183–196. 12
- Vallejos, C. and Steel, M. (2014). Bayesian survival modelling of university outcomes. 1
- Weerasinghe, I. S., Lalitha, R., and Fernando, S. (2017). Students satisfaction in higher education literature review. *American Journal of Educational Research*, 5(5):533–539. 1
- Wiegand, M., Balahur, A., Roth, B., Klakow, D., and Montoyo, A. (2010). A survey on the role of negation in sentiment analysis. In *Proceedings of the Workshop on Negation and Speculation in Natural Language Processing, NeSp-NLP ’10*, pages 60–68, Stroudsburg, PA, USA. Association for Computational Linguistics. 8

- Witten, I. H. (2005). Text mining. *Practical handbook of Internet computing*, pages 14–1. 7
- Young, H. P. (1988). Condorcet’s theory of voting. *American Political Science Review*, 82(4):1231–1244. 19