CEFET/RJ
DIRETORIA DE PESQUISA E PÓS-GRADUAÇÃO

ABORDAGENS BASEADAS EM ONTOLOGIAS PARA ANÁLISE DE SENTIMENTOS EM PORTUGUÊS DO BRASIL

Érica Carneiro Queiroz da Silva

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 grau de mestre.

Orientadores:
Kele Teixeira Belloze, D.Sc.
Gustavo Paiva Guedes e Silva, D.Sc.

Rio de Janeiro,
Agosto de 2024

Abordagens Baseadas em Ontologias para Análise de Sentimentos em Português do Brasil

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

Érica Carneiro Queiroz da Silva

Aprovada por:

Presidente, Prof. Kele Teixeira Belloze, D.Sc. (orientadora)

Gustavo Paiva Guedes e Silva, D.Sc. (coorientador)

Eduardo Bezerra da Silva, D.Sc.

Ronaldo Ribeiro Goldschmidt, D.Sc.

Rio de Janeiro,

Agosto de 2024

Ficha catalográfica elaborada pela Biblioteca Central do CEFET/RJ

S586 Silva, Érica Carneiro Queiroz da
Abordagens baseadas em ontologias para análise de sentimentos em português do Brasil / Érica Carneiro Queiroz da Silva – 2024.
97f : il. (algumas color.) , enc.

Dissertação (Mestrado). Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, 2024.
Bibliografia : f. 86-97.
Orientadora: Kele Teixeira Belloze.
Orientador: Gustavo Paiva Guedes e Silva.

1. Processamento da linguagem natural (Computação). 2. Análise de Sentimento. 3. Aprendizado de máquina. I. Belloze, Kele Teixeira (Orient.). II. Silva, Gustavo Paiva Guedes e (Orient.). III. Título.

CDD 004.0151

Elaborada pelo bibliotecário Leandro Mota de Menezes – CRB/7 nº 5281

DEDICATÓRIA

Dedico esta pesquisa à Divindade, a meu marido, que me apoiou em todos os momentos da elaboração deste trabalho, sempre me motivando a ser uma pessoa melhor. A meus pais e meu irmão, cujo apoio e dedicação pelo desenvolvimento científico me inspiraram a buscar novas aspirações. Esta pesquisa é também dedicada a todas as mulheres da comunidade científica e tem a modesta pretensão de inspirar mais meninas a buscarem novos conhecimentos que as permitam alçar voos cada vez mais altos em nossa sociedade.

AGRADECIMENTOS

O Mestrado é um período cujo devotamento, empenho e renúncia nos leva muitas vezes a refletir sobre nossas escolhas. É um ciclo razoavelmente longo em que abandonamos nosso hedonismo e, em alguns momentos, negligenciamos pessoas que amamos em prol de conquistar o título de mestre. "Parabéns, Mestre em Ciências!", dirão eles, porém poucos realmente sabem os percalços que passamos até a chegada do título. Perdemos pessoas queridas, que gostaríamos que compartilhassem desse momento conosco, mas também ganhamos novos amigos e renovamos nossa própria essência.

São altos e baixos, erros e acertos que forçam-nos ao amadurecimento, tornam-nos mais resilientes e, acima de tudo, focados... compenetrados em alcançar um objetivo maior que qualquer outra proposta ou evento, por mais trágico, cômico ou irrecusável que seja. Particularmente, sempre almejei seguir a formação, contudo, eventualmente os planos não saem exatamente como imaginávamos e a vida toma rumos inesperados que nos trazem gratas surpresas.

Da Linguística clássica de Saussure à Linguística Computacional vão-se aí algum par de séculos e uma curva de aprendizagem razoavelmente íngreme, e, nessa escalada sinuosa, em meio às tantas páginas desta pesquisa, reflito e pauso para dedicar algumas delas ao agradecimento de pessoas que foram fundamentais para a realização deste sonho.

Agradeço, primeiramente, a meu irmão, Guilherme, por ser uma grande inspiração para mim, tanto como ser humano, quanto no campo científico e sempre me apoiar a seguir em frente, lembrando-me constantemente que: *"academia é isso, ou você encara ou não encara"*. A meus pais, Rosa e Léo (*in Memoriam*), principalmente minha mãe, por continuamente me incentivar a buscar mais do que só do mesmo. Dedico este trabalho a meu pai, que partiu durante o desenvolvimento desta pesquisa, mas que sempre acreditou em mim e me incentivou a ver o mundo de uma maneira peculiar e a sempre questionar o porquê das coisas serem.

Agradeço a meu companheiro, Erick, por seu apoio incondicional em todos os momentos, por sempre ver o quê tenho de melhor e, principalmente, por acreditar em mim, muitas vezes mais do que eu mesma. Há também o agradecimento "fofura do dia", aos meus gatinhos, Ghost e Amora por estarem ao meu lado a todo o tempo, nas longas e solitárias madrugadas de estudo, pelas brincadeiras e momentos de descontração tão importantes nessa fase.

Meu agradecimento especial a meus orientadores, Prof.a Dr.a Kele T. Belloze e Prof. Dr. Gus-

tavo P. Guedes, por sempre me incentivarem e acreditarem em mim e, sobretudo, pelo profissionalismo e dedicação de ambos, doando seu tempo de descanso para me ajudar, orientar e ouvir. Agradeço-lhes, ainda, por serem uma inspiração na continuidade do meu desenvolvimento acadêmico e pessoal por perseguir novas descobertas.

Agradeço, também a todo o corpo docente do PPCIC, em especial aos Prof.s Dr.s Diego Brandão e Eduardo Bezerra por terem acreditado em mim e me motivado a ultrapassar barreiras que nunca pensei ser capaz de romper e por me encorajarem e proporcionarem o conhecimento de alçar novas perspectivas e mesmo uma nova profissão. Aos meus colegas discentes, principalmente, Luís, Viny, Thiago, Matheus e Yukio pelo compartilhamento de conhecimento, ombros amigos, sessões de "terapia"e desabafos.

Agradeço aos meus companheiros e amigos e a todo corpo docente da UERJ, que me capacitaram em minha primeira formação, que foram, em seu momento, fundamentais para minha maturidade acadêmica e analítica. Um agradecimento especial a minha amiga Luciana, por me aturar e ser parceira desde os tempos da escola e do vestibular. Também não posso deixar de mencionar a turminha dos "*Supervised Clusters*". Vocês podem não saber, mas me incentivaram a seguir em meus objetivos e se tornaram exemplos de coleguismo e profissionalismo em Ciência de Dados para mim. Agradeço e reconheço a mim mesma por ter tido a coragem de superar diversas barreiras e concluir esse sonho e mais essa etapa da minha vida.

Mais uma, vez, agradeço ao CEFET-RJ, ao PPCIC e a todo o seu colegiado pelo apoio, incentivo, motivação e inspiração. Profs., vocês são incríveis! Por fim, agradeço ao órgão de fomento, responsável por fomentar esta pesquisa.

"It's worth remembering that with regard to cognitive science, we're kind of pre-Galilean, just beginning to open up the subject."
(Noam Chomsky)

RESUMO

Abordagens Baseadas em Ontologias para Análise de Sentimentos em Português do Brasil

Érica Carneiro Queiroz da Silva

Orientadores:
Kele Teixeira Belloze, D.Sc.
Gustavo Paiva Guedes e Silva, D.Sc.

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 grau de mestre.

Com os avanços do Aprendizado de Máquina (AM), muitas soluções estatísticas foram desenvolvidas para resolver problemas de Análise de Sentimentos (AS) e Processamento de Linguagem Natural (PLN). No entanto, até o momento, nenhum processo de classificação estatística seria capaz de resolver relações semânticas e linguísticas da mesma forma que o cérebro humano.

De acordo com a literatura, somente ferramentas extremamente robustas seriam capazes de gerar correlações entre os sentidos das palavras, trazendo o real significado simbólico aos textos. Ontologias são capazes de representar estruturas semânticas de modo a correlacionar conceitos em um domínio específico, assim como seres humanos associam palavras e recebem mensagens de acordo com sua realidade cultural, histórica e social. Apesar de recorrentes na área de Ciência da Computação, ontologias são raramente aplicadas à AS, com exceção de alguns exemplos na língua inglesa, cujos estudos apresentaram resultados promissores ao classificar sentimentos. O Brasil, atualmente não possui nenhum *framework* robusto que execute tarefas de AS com aspectos de abordagem ontológica ou base de conhecimento e poucas são as pesquisas utilizando a metodologia na língua portuguesa.

Este trabalho apresenta uma comparação de métodos entre abordagens léxicas e bases de conhecimento com aspectos ontológicos, além de um novo *corpus* paralelo em português - inglês. Para tanto, efetuou-se, inicialmente, uma revisão de métodos com abordagens ontológicas aplicadas à AS com o objetivo primário de levantar discussões sobre possíveis estratégias de expansão para o campo de AS no Brasil. Espera-se, assim, despertar novas reflexões e um olhar mais profundo sobre humanidades digitais e suas abordagens híbridas, incluindo ontologias, no cenário luso-brasileiro.

Palavras-chave: Ontologias, Análise de Sentimento, Processamento de Linguagem Natural

Rio de Janeiro,

Agosto de 2024

ABSTRACT

Abordagens Baseadas em Ontologias para Análise de Sentimentos em Português do Brasil

Érica Carneiro Queiroz da Silva

Advisors:
Kele Teixeira Belloze, D.Sc.
Gustavo Paiva Guedes e Silva, D.Sc.

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.

With the advances of *Machine Learning* (ML), many statistical solutions were developed to solve *Sentiment Analysis* (SA) and *Natural Language Processing* (NLP) problems. However, until now, no statistical classification process would be able to resolve semantic and linguistic relationships in the same way as the human brain.

According to the literature, only extremely robust tools would be able to generate correlations between the meanings of words, bringing the real symbolic meaning to texts. Ontologies are capable of representing semantic structures in order to correlate concepts in a specific domain, just as human beings associate words and receive messages according to their cultural, historical and social reality.

Although recurrent in the area of Computer Science, ontologies are rarely applied to SA, with the exception of some examples in the English language, whose studies showed promising results for sentiment classification. Brazil currently does not have any robust *framework* that performs SA tasks with ontology-based knowledge base approaches and there are few studies using the methodology in Portuguese.

This work presents a comparison of methods between lexical approaches and knowledge bases with ontological aspects, in addition we present a sentiment Portuguese - English parallel corpus. To that end, a review of different methods with ontology-based approaches applied to SA was initially carried out. Its primary objective was raising more discussions about possible expansion strategies for the field of SA in Brazil. We hope, thus, to awaken new reflections and a deeper look at digital humanities and their hybrid approaches, including ontologies, in the Luso-Brazilian scenario.

Key-words: Ontologies, Sentiment Analysis, Natural Language Processing

Rio de Janeiro,

Agosto de 2024

Sumário

1 Introdução **17**

2 Fundamentação Teórica **23**

2.1 Processamento de Linguagem Natural: O significado além dos símbolos 23

2.2 Análise de sentimentos e suas abordagens 26

2.2.1 Léxicos 29

2.2.2 Ontologias 32

2.3 Ontologias, Mineração de Texto e Análise de Sentimentos 35

2.4 Processamento de Linguagem Natural e Análise de Sentimentos no Brasil 38

2.5 Abordagens para Análise de Sentimentos: Léxicos Afetivos e *Sentic Computing* 41

2.5.1 Léxicos Afetivos 41

2.5.2 *Sentic Computing* 44

2.6 Métricas de avaliação de modelos 51

3 Trabalhos Relacionados **53**

3.1 Abordagens ontológicas de análise de sentimentos de domínio fechado ou aberto em diferentes idiomas 54

3.2 *Benchmark* de métodos de análise de sentimentos 59

3.3 Trabalhos Relacionados a Métodos de Desenvolvimento de Corpora Paralelos 61

3.4 Considerações sobre os diferentes métodos e abordagens para análise de sentimentos 62

4 Metodologia **65**

4.1 Etapa 1 - Definição do Objetivo da Pesquisa 67

4.2 Etapa 2 - Contextualização e Revisão Bibliográfica 67

4.3 Etapa 3 - Identificação de Perguntas de Pesquisa 68

4.4 Etapa 4 - Escolha Criteriosa das Ferramentas 69

4.5 Etapa 5 - Seleção das Métricas de Avaliação e Conjuntos de Dados 70

4.6 Etapa 6 - Implementação dos Testes 70

4.7 Etapa 7 - Elaboração do Novo *Corpus* Paralelo 72

4.8 Etapa 8 - Análise dos Resultados dos Testes 73

4.9 Etapa 9 - Comparação do Impacto da Tradução no Desempenho de Tarefas de Análise de Sentimentos 73
4.10 Etapa 10 - Conclusão e Considerações 74

5 Avaliação Experimental 75
5.1 Conjuntos de Dados utilizados nos testes 75
5.1.1 MQD-1465 75
5.1.2 SADT 75
5.1.3 TOPIE 76
5.1.4 TweetSentBR 76
5.2 Experimento 1: SenticNet aplicada à análise de sentimentos de conjuntos de dados em português 76
5.2.1 Classificação em positivo, neutro e negativo 77
5.2.2 Classificação em positivo e negativo 80
5.3 Experimento 2: SenticNet aplicada à tradução para o inglês com o MQD Paralelo 83
5.3.1 Experimentos de Análise de Sentimentos das versões do MQD Paralelo 83
5.3.2 Classificação em positivo, neutro e negativo 84
5.3.3 Classificação em positivo e negativo 90
5.4 Experimento 3: Comparações com o GPT-3.5 94

6 Considerações Finais 98
6.1 PP1) Como as traduções automáticas e humanas se comparam em termos de impacto na classificação de Análise de Sentimentos? 98
6.2 PP2) Soluções mais robustas superariam os métodos tradicionais de classificação de sentimento baseados em léxicos? 99
6.3 PP3) O cenário de Processamento de Linguagem Natural em português do Brasil necessita de uma ontologia baseada em sentimentos para superar seus métodos tradicionais atuais de classificação de análise de sentimento? 100
Referências Bibliográficas 102

Lista de Figuras

2.1 Representação da estrutura de n-gramas 30

2.2 Reprodução de parte do grafo gerado pela ontologia de Emoções *Emotion Ontology* (MFOEM) na ferramenta Protégé [Musen, 2015] 36

2.3 Representação de Onto-Pares da SenticNet 6 na Protégé 47

2.4 Reprodução dos conceitos e atributos semânticos da OntoSenticNet em Português 50

4.1 Etapas da metodologia de pesquisa 66

4.2 Fluxograma simplificado de utilização da ferramenta SenticNet 71

4.3 Captura de tela contendo exemplo de execução do teste 72

5.3 Fluxo de processamento para análise de sentimentos com os léxicos LIWC e VADER. Ambos com etapas de pré-processamento de texto e preparação de *tokens* 84

5.16 Prompt e recorte da saída do ChatGPT para Classificação de Sentimentos na Língua Portuguesa. 94

Lista de Tabelas

2.1 Uma comparação entre abordagens lexicais e baseadas em ontologias para Computação Afetiva no Brasil, segundo o Idioma Original e Domínio (*EN = Inglês; PT = Português; PTB = Português Brasileiro). 40

3.1 Abordagens ontológicas levantadas em trabalhos de análise de sentimentos. 58

3.2 Bases de conhecimento e léxicos, ano da versão utilizada nos testes das versões 6 e 7 da SenticNet em Cambria et al. [2020] e Cambria et al. [2022] 61

5.1 Resultados das Médias de Precisão, Revocação e F1-Score e valor total da Acurácia da análise de sentimentos da SenticNet 6 para conjuntos de dados em português do Brasil. 77

5.2 Tabela de comparação entre léxicos afetivos e SenticNet, com a porcentagem de acertos para cada classe para o conjunto de dados MQD-1465. 80

5.3 Comparação de resultados das métricas de Precisão, Revocação, Acurácia e F1-Score de léxicos utilizados como comparação para o português e a SenticNet com tradução automática (BabelSenticNet). 81

5.4 Comparação de resultados das métricas de Precisão, Revocação, Acurácia e F1-Score para cada conjunto de dados em português (MQD-1465, SADT, TOPIE e TweetSentBR) excluindo o sentimento neutro (0) da classificação da AS com a SenticNet. 81

5.5 Tabela de comparação entre os *corpora* paralelos MQD00, MQD01, MQD02, MQD03, MQD04 e MQDGT, em inglês, com a Precisão, ou porcentagem de acertos para cada uma das bases de conhecimento/léxicos, SenticNet, LIWC2015 e VADER. 87

5.6 Comparação de resultados das métricas de Precisão, Revocação, Acurácia e F1-Score para cada um dos subconjuntos do *corpus* MQD Paralelo, em inglês, com o sentimento neutro (0) da classificação de AS para SenticNet, LIWC2015 e VADER. 89

5.7 Diferença Percentual de Desempenho da SenticNet entre o Experimento 1 e 2 89

5.8 Tabela de comparação entre os *corpora* paralelos MQD00, MQD01, MQD02, MQD03, MQD04 e MQDGT, em inglês, com porcentagem de acertos para cada uma das bases de conhecimento ou léxicos SenticNet, LIWC2015 e VADER, sem o sentimento neutro (0). 96

5.9 Comparação de resultados das métricas de Precisão, Revocação, Acurácia e F1-Score para cada o *corpus* paralelo MQD00, em inglês, sem o sentimento neutro (0) da classificação de AS para SenticNet, LIWC2015 e VADER. 97

Lista de Abreviações

AM Aprendizado De Máquina 19, 24, 25, 28, 37, 45, 46, 48, 51, 54, 63, 64, 83, 101

API *Application Programming Interface* 49, 70, 71, 73

AS Análise De Sentimentos 19, 20, 21, 22, 26, 27, 28, 29, 31, 35, 38, 39, 41, 44, 48, 49, 53, 54, 55, 56, 57, 59, 62, 63, 64, 65, 67, 68, 69, 70, 72, 73, 74, 75, 76, 77, 80, 81, 83, 84, 94, 95, 98, 99, 100, 101

BFO *Basic Formal Ontology* 35, 55, 58

BILSTM *Bidirectional LSTM* 56

BOW *Bag of Words* 38, 42, 45, 68, 69, 101

CEFET/RJ Centro Federal De Educação Tecnológica Celso Suckow Da Fonseca 72

CLG Curso De Linguística Geral 23

DO *Disease Ontology* 58, 63

DSN *DomainSenticNet* 49

EAV *Entity - Attribute - Value* 58

FCA *Formal Concept Analysis* 58, 64

FDSOT *Fuzzy Sentiment Ontology Tree* 63

GPT *Generative Pre-Trained Transformer* 94, 100

HL-SOT *High Hierarchical Learning - Sentiment Ontology Tree* 58

IA Inteligência Artificial 21, 37, 46, 48, 62, 71

IDF Inverso Da Frequência Nos Documentos 29, 30

IDO *Infectious Disease Ontology* 58, 63

KBRS *Knowledge-Based Recommendation System* 56, 57

LA Léxico Afetivo 41, 42, 43, 44

LIWC *Linguistic Inquiry and Word Count* 29, 42

LLM *Large Language Models* 69, 70, 77, 94, 95, 100, 101

LSTM *Long Short-term Memory* 46, 48

MC Matriz De Confusão 73, 76, 77, 78, 81, 84, 85, 87, 92

MFO *Mental Disease Ontology* 55

MFOEM *Emotion Ontology* 35, 36, 55, 58, 64, 69, 76, 100, 101

ML *Machine Learning* 39

MO Mineração De Opinião 19, 26, 28, 46, 54, 63

MPQA-PT *Multi-Perspective Question Answering* 42

MQD-1465 Meu Querido Diário Dataset 21, 70, 72, 73, 75, 77, 78, 81, 83, 84, 88, 93, 94, 99

MT Mineração De Texto 22, 25, 26, 35, 64, 75

NILC Núcleo Interinstitucional De Linguística Computacional 42

NLP *Natural Language Processing* 83

NTN *Recursive Neural Tensor Network* 28

OL *Ontology Learning* 64

OWL *Web Ontology Language* 33, 34, 37, 44, 46, 64, 101

PLN Processamento De Linguagem Natural 19, 20, 21, 24, 25, 27, 28, 38, 39, 41, 44, 45, 48, 57, 63, 64, 65, 67, 68, 69, 73, 74, 82, 93, 98, 99, 100, 101

RBC Raciocínio Baseado Em Casos 58

RDF *Resource Description Framework* 33, 34, 49

RKM *Regularized K-Means* 80

SA *Sentiment Analysis*

SADT *Sentiment Analysis Dataset Twitter-pt-br* 75, 77, 78, 81

SDK *Software Development Kit* 70

SMT *Statistic Machine Translation* 49

SOT *Sentiment Ontology Tree* 55, 58

SUMO *Suggested Upper Merged Ontology* 37, 44

SVM *Support Vector Machine* 55, 58

TF Frequência Do Termo 29, 30

UO *Upper Ontology* 37

VADER *Valence Aware Dictionary and sEntiment Reasoner* 44, 83, 84, 88, 89, 90, 93, 94, 98

VSO *Visual Sentiment Ontology* 55

XML *Extensible Markup Language* 33, 49

Capítulo 1 Introdução

"You do not get discoveries in the sciences by taking huge amounts of data, throwing them into a computer and doing statistical analysis of them: that's not the way you understand things, you have to have theoretical insights."
(Noam Chomsky)

Compreender pensamentos e opiniões das pessoas é um tema de interesse para indivíduos e indústrias há muito tempo. Desde o primeiro Curso de Linguística Geral de de Saussure [2008], em 1916, até o surgimento de algoritmos de linguística computacional, a exploração dos diversos significados que palavras, frases e expressões podem carregar e evocar nos indivíduos tem sido uma fonte constante de discussões. Apesar da notável evolução observada nos estudos de compreensão semântica, barreiras de comunicação e falhas interpretativas não são eventos raros.

As áreas de humanidades e tecnologia passaram por transformações significativas que trouxeram novas formas de comunicação e expressão na sociedade. O grande volume de conteúdo gerado por usuários levou ao desenvolvimento de ferramentas automatizadas capazes de extrair e analisar toda essa massa de dados [Freitas and Vieira, 2013], tornando a Análise de Sentimentos crucial para compreender opiniões em redes sociais e demais mídias online [Stavrianou et al., 2007].

Com a advento da internet e a imensa popularização das mídias sociais, o compartilhamento de opiniões e sentimentos tornou-se algo frequente e corriqueiro. A popularização de *smartphones* e *tablets* conectados à internet propicia um número crescente de usuários ativos nas mídias sociais, reforçando essa troca constante de informações.

Assim, as mídias sociais são fonte inestimável de expressão de opiniões e sentimentos. As mesmas refletem os hábitos e pontos de vista de boa parte da população, uma vez que a inclusão digital atinge proporções cada vez mais significativas. Esse empoderamento do usuário é um dos principais geradores de valor de aplicativos da *Web* 2.0 [Peñalver-Martinez et al., 2014]. Além disso, empresas de tecnologia criam constantemente mais recursos digitais, como *blogs* e novas redes sociais [Freitas and Vieira, 2019], que permitem sua aproximação da população.

Plataformas de interação para o compartilhamento de opiniões sobre produtos surgem a todo o

momento e, com elas, uma riquíssima fonte de dados capaz de ajudar a prever cenários políticos, contextos sociais, percepção socioeconômica, posicionamentos sócio-políticos, bem como opiniões sobre produtos e empresas [Borth et al., 2013]. Ou seja, tais plataformas são fontes essenciais para o monitoramento de satisfação do público e que podem ser de grande auxílio para que organizações compreendam melhor como os consumidores percebem seus produtos [Nguyen et al., 2021].

Além disso, o consumo de conteúdo e produtos *online* tornou-se ainda mais frequente. A união desses dois fatores, conteúdo e produto, permite que o consumidor leia comentários, especificações técnicas e classificações sobre uma mercadoria que queira adquirir. Esses conteúdos são avaliações geradas por outros usuários que já compraram o mesmo item, e são de grande auxílio nas tomadas de decisão de compra [Karthik and Ganapathy, 2021].

Tamanha é a relevância das postagens dos usuários, que diversas aplicações são rotineiramente desenvolvidas exclusivamente para extrair automaticamente a opinião e analisar os sentimentos do público [Reis et al., 2015]. Torna-se, então, crescente o interesse corporativo, industrial e acadêmico por ferramentas e métodos voltados para a detecção de sentimentos *online* [Freitas and Vieira, 2019].

Assim, nos últimos anos, a análise de conteúdos postados em mídias sociais tem sido utilizada para diversos fins relacionados a avaliação e monitoramento de hábitos e comportamentos de usuários, incluindo conceitos subjetivos como seus sentimentos e humor. Porém, esse amontoado de textos gerados na *web* excede a capacidade de processamento humano. Sendo assim, torna-se praticamente impossível que um indivíduo verifique manualmente todas as possíveis opiniões e postagens sobre um determinado assunto [Freitas and Vieira, 2013].

Tal problema motivou cientistas da área de Computação Afetiva a buscar métodos e ferramentas que automatizassem a recuperação de informações e a extração de conhecimento de repositórios da Web [Liu et al., 2012]. O termo Computação Afetiva foi empregado pela primeira vez em 1995 por Rosalind Picard [Picard, 2000]. Segundo a pesquisadora, a habilidade de uma máquina de "*compreender emoções humanas, bem como interagir e interpretá-las*", caracterizá-la-ia como um computador *"genuinamente inteligente"* [Picard, 2000; Cruz et al., 2017]. O principal objetivo dessa área é o desenvolvimento de pesquisas e ferramentas que reconheçam, interajam e interpretem sentimentos humanos [Picard, 2000].

A Computação Afetiva possui diversos métodos e subáreas com nomes aparentemente semelhantes, mas que executam tarefas diferentes. Análise de Sentimento; Mineração de Opinião e Classificação de Polaridades costumam lidar com tarefas de classificação de sentimentos, conferindo-lhes rótulos como positivo, negativo e neutro. Análise de Afeto, Análise de Emoção ou Reconhecimento de Emoção lidam com tarefas que visam identificar a emoção evocada por um determinado texto, geralmente com base nas seis emoções primitivas apresentadas por Parrott [2000] [Liu, 2012; Dragoni et al., 2018].

A AS, bem como suas técnicas e metodologias, tornaram-se fundamentais para avanços nos estudos das relações presentes entre diversos conteúdos e documentos [Stavrianou et al., 2007]. As abordagens e estratégias mais populares para AS utilizam métodos estatísticos de classificação das polaridades (positivo, negativo e neutro) [Cambria et al., 2017]. Alguns dos mais usados são: Aprendizado de Máquina (AM), léxicos e métodos de Processamento de Linguagem Natural (PLN), como *crowdsourcing*, escalas psicométricas ou uma combinação desses métodos [Reis et al., 2015].

Segundo a literatura, métodos mais abrangentes, que utilizam ferramentas mais robustas e aspectos ontológicos, poderiam trazer melhoras significativas às tarefas de PLN e AS [Cambria et al., 2017; Liu et al., 2012; Jiang and Tan, 2005]. Essas abordagens costumam ser de grande auxílio em problemas de desambiguação, já que a estrutura semântica e as relações presentes entre conceitos de entidades e atributos podem ajudar na segmentação de domínios de acordo com os *corpora*.

Nossas pesquisas apontam que, no Brasil, as contribuições científicas em AS utilizam, em sua maioria, léxicos ou tratam a análise de sentimentos como um problema de classificação. Embora existam algumas contribuições, como o Unilex, de de Souza et al. [2017], e o Reli Lex, de Freitas and Vieira [2013], os trabalhos com foco em PLN e AS para o idioma português ainda são escassos [Pereira, 2020; da Silva Conrado et al., 2014]. É importante observar que, apesar da variedade de recursos e técnicas para AS e Mineração de Opinião (MO), a maioria dessas aplicações é desenvolvida para a língua inglesa [Pereira, 2020; Steinberger et al., 2011]. Recentemente, novas contribuições vêm aparecendo em outros idiomas, como árabe, alemão e russo [Reis et al., 2015; Kontopoulos et al., 2013].

Além disso, a maioria das aplicações e ensaios empregados para PLN no país usam traduções automáticas do inglês para o português, tendo como base a estrutura linguística do idioma anglo-saxão, que diverge do português brasileiro [da Silva Conrado et al., 2014]. Segundo da Silva Conrado et al. [2014] existem diversas lacunas na área de PLN no Brasil. Para os autores, traduções automáticas trazem prejuízos às tarefas de PLN devido à diferença nas estruturas dos idiomas inglês e português. Os autores ainda acrescentam que, no Brasil, são poucos os trabalhos que utilizam abordagens ontológicas aplicadas à PLN [da Silva Conrado et al., 2014], havendo uma predileção pelo uso de léxicos, principalmente no campo de AS [Pereira, 2020].

As afirmações dos autores brasileiros parecem se refletir também no estado da arte no exterior. Foi observado que problemas de PLN e AS são comumente tratados como meras tarefas de classificação estatística [Cambria et al., 2017; Cambria and Hussain, 2012]. Tais abordagens puramente matemáticas desconsideram aspectos linguísticos e semióticos, intrínsecos às estruturas linguísticas, além de ignorarem características relacionadas à sócio e psicolinguística, bases fundamentais da comunicação e linguagem humanas [Cambria et al., 2017].

Em Cambria and Hussain [2012],Cambria and White [2014], Cambria and Hussain [2015], Cambria et al. [2017], Cambria et al. [2020] e Cambria et al. [2022] são discutidos argumentos que mostram que as tarefas de classificação tendem a depender exclusivamente de métodos estatísticos, negligenciando abordagens holísticas que envolvem análise semântica. Além disso, outros trabalhos também observaram que a maioria das aplicações de PLN no Brasil depende fortemente de traduções automáticas do inglês para o português [da Silva Conrado et al., 2014; de Souza et al., 2017]. Essa prática apresenta desafios para as tarefas de PLN devido às diferenças estruturais entre os dois idiomas. No contexto brasileiro, as contribuições científicas em AS predominantemente utilizam léxicos ou abordam a análise de sentimento como um problema de classificação [Freitas and Vieira, 2013; Rosa et al., 2019; Freitas and Vieira, 2019; Pereira, 2020; Carneiro et al., 2022].

Com uma população superior a 203 milhões de habitantes [IBGE, 2020], o Brasil possui uma das maiores comunidades online e ocupa o quarto lugar em atividade no X (anteriormente Twitter) [World Population Review, 2024]. Tais características resultam em um volume substancial de dados textuais, motivando uma exploração detalhada de seu cenário linguístico.

Desta forma, este trabalho foi dividido em diferentes etapas de pesquisa devido à escassez de material no estado da arte no Brasil, principalmente referente a ontologias e bases de conhecimento em português voltadas para tarefas de AS. Sendo assim, conforme proposto pelas premissas apresentadas em Cambria and Hussain [2012]; Cambria et al. [2017]; Jiang and Tan [2005]; e Liu et al. [2012]; e Rachid et al. [2018] de que as relações presentes nas ontologias podem ser de grande auxílio para tarefas de análise de sentimento, iniciamos um revisão bibliográfica sobre estudos que utilizam metodologias com abordagens ontológicas aplicadas à AS. Com isso, buscou-se avaliar a potencial necessidade de novas soluções e técnicas com abordagens ontológicas e semânticas para o português brasileiro.

Com base nas premissas levantadas e no contexto fornecido, esta dissertação tem como objetivo abordar três Perguntas de Pesquisa (PP):

- PP1) Como as traduções automáticas e humanas se comparam em termos de impacto na classificação de Análise de Sentimentos?
- PP2) Soluções mais robustas superariam os métodos tradicionais de classificação de sentimento baseados em léxicos?
- PP3) O cenário de PLN em português do Brasil necessita de uma ontologia baseada em sentimentos para superar seus métodos tradicionais atuais de classificação de análise de sentimento?

Assim, após a revisão bibliográfica, buscou-se por recursos que pudessem ser testados a fim de dar margem ao escopo inicial deste trabalho. Foi encontrado pouco material que focasse em domínio

aberto e análise de sentimentos. Ademais, apenas alguns poucos exemplos de estudos com domínios específicos, ou fechados, foram encontrados. Como exemplo temos a HOntology, de Freitas and Vieira [2013], que trata do domínio hoteleiro, e a KBRS [Rosa et al., 2019], cujo foco é a depressão e estresse.

Identificou-se a SenticNet [2022] (versão 6), um *framework* baseado em ontologias, como passível de testes. A ferramenta faz uso de técnicas de *sentic computing*, trata-se de uma metodologia que utiliza uma abordagem holística, que vai desde Inteligência Artificial (IA) simbólica e sub-simbólica, a ontopares de palavras e conceitos primitivos [Dragoni et al., 2018; Cambria et al., 2020; SenticNet, 2012a; Cambria et al., 2022; Cambria and Hussain, 2012; Cambria et al., 2017].

A opção pela SenticNet 6 ganhou mais força devido às suas características multilíngue, que inclui o português, com tradução automática, e por se tratar de uma ferramenta *open source*. Assim, a hipótese de da Silva Conrado et al. [2014] quanto ao prejuízo de tarefas de PLN realizadas com traduções automáticas também poderia ser avaliada.

Com base nas premissas de da Silva Conrado et al. [2014] quanto ao prejuízo que traduções automáticas trariam a problemas de PLN, executamos um trabalho de engenharia reversa, traduzindo a base de dados Meu Querido Diário Dataset (MQD-1465) [Azevedo et al., 2021a,b] manualmente para o inglês a fim de comparar o desempenho de diferentes métodos de análise de sentimentos nos *corpora* original e paralelo. O experimento tinha como objetivo investigar a hipótese de da Silva Conrado et al. [2014] e verificar se o novo *corpus* traduzido manualmente levaria a SenticNet a um melhor desempenho na análise de sentimentos do que uma tradução automática.

Os resultados da SenticNet 6, que utiliza a abordagem *sentic computing* desenvolvida em Cambria and Hussain [2012], Cambria and White [2014] e Cambria et al. [2017], foram comparados a métodos já consolidados de AS, incluindo alguns léxicos mais tradicionais, previamente analisados no trabalho de Nascimento [2021]. Com eles, foi possível verificar a necessidade de novas discussões sobre as consequências de modelos com tradução automática e seus impactos em tarefas de classificação, principalmente no que diz respeito às relações semânticas.

Como contribuição científica, este trabalho traz um novo *corpus* paralelo anotado para análise de sentimento que permite comparações de ferramentas tanto em português quanto em inglês. Além disso, analisamos se as traduções automáticas trazem prejuízos aos classificadores. Ademais, verificamos se soluções mais arrojadas e robustas, que utilizam aspectos semânticos, como as ontologias, têm melhor desempenho do que métodos lexicais mais tradicionais de classificação de sentimentos. Assim, buscou-se avaliar criticamente o estado do cenário de abordagens ontológicas dedicadas à análise de sentimentos no Brasil e, ainda expor a urgência no desenvolvimento de maiores estudos e esforços na elaboração de um *framework* mais robusto para a extração de opiniões específicas ao cenário brasileiro.

Este trabalho é organizado da seguinte maneira: a Seção 2 apresenta a fundamentação teórica, contendo referências necessárias à definição dos problemas encontrados e à compreensão teórica dos temas aqui abordados, a saber: (i) (Mineração de Texto (MT)); (ii) Análise de Sentimentos (AS) e suas abordagens; (iii) Léxicos, (iv) Ontologias; e, (v) Ontologias, Mineração de Texto (MT) e Análise de Sentimentos (AS). A Seção 3 dedica-se à exposição da revisão bibliográfica acerca das diferentes metodologias aplicadas à AS computacional. Na Seção 4, discute-se a metodologia aplicada para a definição do problema observado nas pesquisas. Ademais, apresenta-se a descrição dos experimentos computacionais realizados para avaliar a capacidade da SenticNet 6 na execução de tarefas de AS. A Seção 5 apresenta os resultados e discussões dos experimentos computacionais executados com a SenticNet 6. A seção 6, por fim, aborda as considerações advindas deste trabalho e propõe possíveis direções futuras para a pesquisa.

Capítulo 2 Fundamentação Teórica

"Tomada em seu todo, a linguagem é multiforme e heteróclita; um cavaleiro de diferentes domínios, ao mesmo tempo física, fisiológica e psíquica, ela pertence além disso ao domínio individual e ao domínio social; não se deixa classificar em nenhuma categoria de fatos humanos, pois não se sabe como inferir sua unidade."(SAUSSURE, 2004, p. 17).)

Neste capítulo, a fundamentação teórica é construída a partir das premissas estabelecidas nos trabalhos de da Silva Conrado et al. [2014] e Cambria and Hussain [2012], que serviram como alicerces para a concepção e desenvolvimento desta pesquisa. As observações e hipóteses de cada um desses autores são apresentadas e discutidas, além de conceitos essenciais da linguística e da linguística computacional que reforçam e ampliam os apontamentos realizados. Assim, busca-se proporcionar uma base sólida e abrangente para a compreensão dos fenômenos estudados, situando esta pesquisa no contexto mais amplo das investigações acadêmicas já realizadas na área.

2.1 Processamento de Linguagem Natural: O significado além dos símbolos

Em seu Curso de Linguística Geral (CLG), de Saussure [2008] apresenta diversos conceitos fundamentais ao desenvolvimento da linguística, que junto com o círculo de Praga fizeram dela a mais avançada das ciências humanas da era moderna [Broden, 2018]. Segundo o autor, a linguagem seria uma faculdade inerente ao ser humano, enquanto a língua seria um fenômeno social [de Saussure, 2008; Menezes and Silva, 2018].

Mais adiante em seu CLG, [de Saussure, 2008] observa que a língua é um conjunto de sinais sociais depositados no cérebro coletivamente, já a linguagem pode ter caráter individual, mas não necessariamente social. Porém a comunicação só ocorreria quando a linguagem é falada e transportada para o social [de Saussure, 2008; Menezes and Silva, 2018]. Desta forma, seria possível inferir que a língua trata de operações e processos que são traduzidos mentalmente para a substância lin-

guística [Broden, 2018].

As definições saussurianas são ponto fundamental para compreendermos as relações presentes entre os termos e como a concepção "mecânica" dos símbolos pode gerar conflitos na comunicação. Para o autor, a identidade dos símbolos "jamais pode ser fixada" [de Saussure, 2008], uma vez que recebe novo valor cada vez que apresentado à sociedade. Essas variações podem ocorrer no nível do nome, posição em relação aos outros símbolos, quanto ao caráter ou mesmo à função [de Saussure, 2008; Broden, 2018].

Ainda sem entrar nos aspectos da linguística computacional, como exemplo da própria flexibilidade dos símbolos, temos a própria tradução do termo "idioma" e "linguagem" para seu correspondente genérico em inglês "*language*". Este nome já gera ambiguidades quanto ao que são realmente aspectos de linguagem e aspectos de língua [Menezes and Silva, 2018]. Em seu livro *Natural Language Understanding*, Allen [1988], menciona diferentes aspectos linguísticos que transitam tanto pela morfologia quanto pela semiótica, uma vez que suas pesquisas propõem capacitar o leitor a criar modelos de PLN para executar diversas tarefas presentes na comunicação humana, como ler, escrever, ouvir e falar.

Porém, o autor logo ressalta que a continuidade dessas pesquisas poderia, um dia, levar a uma compreensão profunda do processamento da linguagem humana e que, para tanto, seria necessária a união dos esforços de linguistas, psicolinguistas, filósofos e cientistas da computação [Allen, 1988]. Desta forma, o autor em nenhum momento pretende utilizar a linguística computacional para tentar compreender as estruturas semânticas das palavras e frases ou qual significado real elas poderiam conter em diferentes contextos sociais e psicológicos.

Da mesma forma, Jurafsky and Martin [2008] mencionam a estrutura sintática das palavras. Os autores classificam termos ambíguos como aqueles que podem ter diferentes estruturas linguísticas. Porém, os autores tratam os problemas de desambiguação de acordo com a classe gramatical das palavras. Por exemplo, a palavra "casa", em português, pode ser tanto um verbo quanto um substantivo.

Além da morfossintaxe, qualquer língua é composta por figuras de linguagem, ou semânticas, que podem alterar o sentido lexical de uma palavra dentro de uma frase, mas que podem representar símbolos específicos dentro de um determinado grupo. Isto posto, Jurafsky and Martin [2008] afirmam que compreender o processo de linguagem humana, um dos objetivos da ciência cognitiva, seria útil para a construção de melhores modelos de máquina com foco em linguagem.

Os trabalhos de Allen [1988] e de Jurafsky and Martin [2008] representam bem a ascensão dos principais métodos utilizados em PLN nas décadas de 80 e 90 e na primeira década de 2000. O primeiro com foco em modelos probabilísticos baseados em dados, e o segundo com AM, abordagens estatísticas amplamente não supervisionadas, que culminaram na ascensão de modelos para tradução

automática e modelagem de tópicos [Jurafsky and Martin, 2008].

Desde 1940, cientistas tentam desvendar formas de representar matematicamente a linguagem humana [Jurafsky and Martin, 2008], seja por meio da teoria da linguagem formal de Chomsky e dos primeiros modelos de álgebra e teoria dos conjuntos para definir linguagens formais como sequências de símbolos na década de 40, seja pelos métodos estocásticos com viés bayesiano dos departamentos de estatística e engenharia elétrica das universidades [Jurafsky and Martin, 2008].

Entretanto, o advento da *internet* e da *web* 2.0 trouxe uma nova dinâmica de usabilidade, permitindo que o usuário, antes mero consumidor de textos, passasse a ter o poder de gerar conteúdo na *internet*, expressando sua opinião e avaliando personalidades e produtos. Esse novo cenário revelou a necessidade de novos estudos na área de Computação que fossem capazes de lidar de maneira rápida com a enorme quantidade de textos presentes na *web*.

Surgem, assim, as pesquisas em Mineração de Texto (MT), área com grande relevância para a descoberta de informações em bases de dados não estruturadas, presentes em conjuntos de dados textuais, que diferem das encontradas em sistema de base de dados estruturados [Stavrianou et al., 2007]. Simplificadamente, a MT é um subcampo da mineração de dados que tem fontes textuais como objeto de estudo e pesquisas que lidam constantemente com questões relacionadas à PLN. Com ela, é possível identificar informações importantes presentes no texto, ou *corpus*/*corpora* [Stavrianou et al., 2007].

Trata-se de um campo interdisciplinar que combina técnicas de mineração de dados com diversas outras áreas, tais como Extração de Informação, Recuperação de Informação, Linguística Computacional, Categorização e Busca por Tópicos e Relacionamento de Conceitos [Stavrianou et al., 2007]. Tais técnicas possibilitam a descoberta de informações incógnitas em coleções textuais.

As tarefas tradicionais de MT, como classificação e análise de agrupamento de dados (*clustering*), costumam ser desenvolvidas com viés taxonômico, ou seja, focam no conceito ou na morfossintaxe das palavras. Assim, ao longo da análise de extração de dados textuais, vários subprocessos podem ocorrer, como o reconhecimento de padrões, *clustering*, segmentação de palavras (tokenização), análises sintáticas e semânticas, além da aplicação de diferentes algoritmos [Stavrianou et al., 2007].

Algumas bibliotecas em *Python*, como a *NLTK*, *Scikit-Learn*, e métodos de AM como o *K-Means* e demais algoritmos de classificação, dentre outros, são de grande importância para o processamento e verificação em MT. Ressalta-se que diferentes métodos podem ser abordados e devem ser empregados de acordo com o objetivo final do processo de mineração textual. Alguns deles são: dicionários, léxicos e ontologias.

2.2 Análise de sentimentos e suas abordagens

A partir da MT passa a ser possível extrair e recuperar informações estruturadas de textos da internet, permitindo maior exploração das informações ali presentes, como classes, rótulos, similaridade entre os textos, dentre outros [ŽiŽka et al., 2019]. Uma das informações que podem ser extraídas com técnicas de MT são possíveis aspectos sentimentais e emotivos que eventualmente sejam evocados pelo texto.

Sentimentos, ou a faculdade de sentir, são uma das maiores manifestações de consciência mais intensas que um ser vivo pode ter, se não a maior. Com eles, pessoas, e mesmo alguns animais, são capazes de perceber o mundo ao seu redor e manifestar sua percepção sobre aquilo que os cerca. Por meio de emoções, amores, dores, prazeres, desejos e frustrações se manifestam.

Sobretudo, os sentimentos exprimem opiniões negativas, positivas - e mesmo neutras - que, devido à inerente faculdade humana de se comunicar, permitem o compartilhamento de convicções pessoais. São sentimentos, emoções e opiniões que tornam as decisões possíveis [Zehra et al., 2017].

Com o rápido crescimento e atual tamanho da internet, técnicas automatizadas se tornaram essenciais a fim de encontrar soluções práticas e minimamente eficientes que pudessem desempenhar análises de opinião emitidas por usuários na internet [Peñalver-Martinez et al., 2014]. Assim surge a Mineração de Opinião (MO), algumas vezes também denominada Análise de Sentimento (AS), ainda que Cambria et al. [2017] os considerem diferentes subcampos da MT. A AS conta com técnicas e métodos capazes de extrair, analisar e classificar os sentimentos de diferentes pessoas em relação a uma entidade-alvo e seus atributos relacionados [Liu et al., 2012].

É importante compreender que existem diferentes formas de avaliar as impressões passadas por um texto, classificando-o em classes que podem ser mais ou menos subjetivas. Desta forma, pode-se avaliar um *corpus* pelo sentimento ou emoção evocada.

A Análise de Sentimento tem como objetivo avaliar os sentimentos em classes negativas, positivas e neutras. Para avaliar a emoção evocada, como "felicidade", "compaixão" *etc*, entra-se na esfera da Análise de Afeto; Análise de Emoção ou Reconhecimento de Emoção. É comum que as classes de emoção sejam baseadas nas seis emoções primitivas apresentadas por Parrott [2000].

As postagens de usuários em blogs e redes sociais fornecem dados preciosos para empresas e sociedade [de Souza et al., 2017], tornando a análise de tais informações fundamental para as ciências do comportamento. O principal objetivo da AS é compreender e prever as tomadas de decisões da sociedade, permitindo o acompanhamento das percepções do público em relação a uma marca ou figura pública, realizar previsões sobre o mercado de ações, bem como de cenários políticos [Borth et al., 2013].

Por meio da AS é possível entender e recuperar potenciais sentimentos presentes em qualquer

conteúdo gerado pelo usuário, independentemente do formato de mídia. Alguns exemplos de formato são: texto, imagem, animação, vídeo *etc* [Narayanasamy et al., 2021]. Porém, a análise em formatos textuais é a mais comum, devido ao surgimento de plataformas de mídias sociais e outras fontes [Narayanasamy et al., 2021].

Segundo Narayanasamy et al. [2021], a extração de sentimentos ou emoções de um texto é uma tarefa complexa, pois é passível de grande gama de ambiguidades e anomalias presentes no conteúdo do *corpus*. Para Cambria et al. [2017] e Cambria and Hussain [2012], a AS depende de uma imensa quantidade de tarefas de PLN, ainda que o entendimento da maioria dos trabalhos a aborde como um simples problema de categorização ou classificação estatística.

Trata-se, porém, reforçando o posicionamento de Narayanasamy et al. [2021], bem como as primeiras impressões de Allen [1988] e Jurafsky and Martin [2008], de um desafio que lida com múltiplas tarefas de PLN, tais como detecção de subjetividade e sarcasmo, anáfora, resolução, desambiguação do sentido da palavra e muito mais [Cambria et al., 2020, 2017; Cambria and Hussain, 2012].

É importante ressaltar, que, apesar da evolução na área, muitas tarefas de PLN ainda são necessárias para resolvermos problemas de AS ou emoções. São processamentos que abrangem etapas nos níveis sintáticos, semânticos e pragmáticos, conforme observado nas pesquisas de Cambria et al. [2017] e Cambria and White [2014].

Todas essas camadas, possuem etapas que são fundamentais para a boa execução das tarefas de PLN. Problemas de desambiguação, por exemplo, estão intrinsecamente ligados a vários outros, uma vez que uma compreensão semântica inadequada pode levar a erros de interpretação. Palavras podem ter diferentes significados dependendo do seu contexto e, a compreensão errada de seu sentido pode afetar a análise de metáforas e outras figuras de linguagem.

Atualmente, idiomas evoluem muito rápido devido à *internet* e à globalização, criando termos novos a cada dia, o que dificulta bastante a criação de conjuntos de dados e modelos de treino de classificação [Jung et al., 2017]. Porém é importante ressaltar, que o uso inapropriado das tarefas de PLN podem prejudicar toda a avaliação que deveria ser extraída da informação. Por exemplo, reconhecer a entidade nomeada após o uso da lematização pode descaracterizar completamente o sentido do texto.

Um bom caso de problema de desambiguação na linguagem coloquial, muito comum a conjuntos de dados extraídos da internet, é o termo *"passada"*. Atualmente ele também é utilizado na linguagem cotidiana em frases que denotam surpresa, como em *"Estou passada"*. Neste caso, a não ser que o gênero textual seja um apólogo, onde personagens inanimados - como um ferro de passar roupa ou uma camiseta - ganham vida, dificilmente, veremos o adjetivo "passada" na primeira pessoa num texto formal. Sendo assim, *"passada"*, numa linguagem informal, passa a manifestar emoções,

entrando num domínio de sentimentos.

Levando em conta todas essas etapas da PLN e a própria característica psicossocial é compreensível o entendimento de Allen [1988], Jurafsky and Martin [2008] e Cambria and Hussain [2012]. Alguns desses atributos estão intrinsecamente ligados a fundamentos da linguística e da psicolinguística e é natural que os autores vislumbrem a necessidade de uma equipe multidisciplinar para solucionar problemas de PLN, principalmente aqueles ligados à AS.

Segundo Cambria et al. [2017] as abordagens de PLN e AS podem ser divididas em simbólicas e sub-simbólicas. A primeira utiliza métodos de lógica relacional, como ontologias e léxicos. Trata-se de uma antiga escola de pensamento sobre a construção de bases de conhecimento para a identificação de polaridade textual. Alguns exemplos de uso de abordagem simbólica são as bases de conhecimentos WordNet-Affect [Fondazione Bruno Kessler, 2009a], SentiWordNet [Esuli and Sebastiani, 2006; Baccianella et al., 2010; Esuli, 2019] e SenticNet [SenticNet, 2012a].

A segunda abordagem faz uso de métodos estatísticos de PLN para a mesma tarefa, bem como técnicas de AM supervisionado [Cambria et al., 2020]. Podemos exemplificar as abordagens sub-simbólicas com alguns algoritmos de uso comum em tarefas de classificação de polaridade, como os algoritmos Naïve Bayes [Cambria et al., 2020], que analisa a probabilidade de uma dada palavra ter sentimento positivo ou negativo; e *Clustering*, que verifica a proximidade de uma palavra em relação a outra dentro de um espaço vetorial e as agrupa em categorias. Tais categorias podem ser positivas, negativas ou neutras, ou tipos (o termo classe será evitado aqui a fim de afastar possíveis confusões futuras com algumas definições presentes no item Ontologias) de sentimentos, como alegria, tristeza, amor *etc.*

Pang et al. [2002] foram pioneiros no uso da abordagem sub-simbólica aplicada à MO. Os autores utilizaram diferentes algoritmos de AM em um conjunto de dados para a revisão de filmes obtendo 82% de precisão na detecção de polaridade de suas análises. Onze anos mais tarde, Socher et al. [2013] usariam o mesmo conjunto de dados de Pang et al. [2002], obtendo 85% de precisão com o método de AM *Recursive Neural Tensor Network* (NTN).

A abordagem baseada em AM tornou-se extremamente popular. Nela, um classificador de sentimento especializado é treinado para um dado *corpus* com o objetivo de prever a polaridade do texto. Normalmente, o classificador de sentimento é treinado com base em bigramas ou n-gramas, classificando o texto com as respectivas polaridades de sentimento. Sua principal desvantagem é a rotulação manual dos conjuntos de dados de treinamento [Cambria et al., 2020]. Tais conjuntos são bases de dados históricas utilizadas em modelos de algoritmos supervisionados e simulam previsões ou classificações reais como parâmetros para o modelo.

2.2.1 Léxicos

Léxicos são ferramentas da linguística computacional que analisam termos de um dado idioma, num determinado *corpus*, através da morfossintática [Stavrianou et al., 2007]. Há alguns anos, esses recursos são utilizados para AS. Segundo Hutto and Gilbert [2014], léxicos são listas com características lexicais rotuladas em positivas ou negativas, de acordo com sua orientação semântica.

Essas listas podem ser dicionários dedicados a palavras de opinião, que passam a ser denominados de dicionários de léxicos. É com base neles que o mapeamento das palavras de opinião seria realizado [Narayanasamy et al., 2021].

Para a construção de um léxico, listas de recursos de opinião são criadas manualmente, e, embora este seja um dos métodos mais confiáveis, é uma abordagem custosa e lenta. Sendo assim, a maioria das pesquisas aplicadas à AS utiliza léxicos pré-existentes e manualmente construídos [Hutto and Gilbert, 2014].

Os léxicos mais antigos e comumente utilizados são, respectivamente, o *General Inquirer* [1], criado em 1966; e o *Linguistic Inquiry and Word Count (LIWC)* [PortLEX, 2022], de 2005. O primeiro foi desenvolvido como uma ferramenta para análise de conteúdo, ajudando cientistas sociais, políticos e psicólogos a identificar características específicas em mensagens. O segundo, é fruto do trabalho conjunto de linguistas, psicólogos e sociólogos, que levaram uma década em suas pesquisas [Hutto and Gilbert, 2014].

Ao utilizar a abordagem lexical para análise de sentimentos, segmenta-se o texto em léxicos morfológicos e, em seguida, identificam-se as palavras de opinião em positivas ou negativas, *i.e.*, em polaridades. A polaridade dos sentimentos é rotulada com um algoritmo que utiliza uma medida estatística denominada Frequência do Termo (TF)/Inverso da Frequência nos Documentos (IDF). Segundo esse cálculo, uma palavra é mais ou menos relevante dependendo da quantidade de vezes que aparece em um dado documento [de Souza et al., 2017].

A Equação 2.1 apresenta a definição matemática de TF/IDF. Nela, Wt, d representa o peso do termo t em um documento d; a quantidade de vezes que aparece em uma amostra de texto é representada por $Freq_{t,d}$. Max é o número de ocorrências do termo mais frequente, enquanto N é o total de amostras de texto numa base de treino. Já nt representa o número de amostras na base de treino que possui um dado termo t [de Souza et al., 2017].

$$W_{t,d} = \frac{Freq_{t,d}}{Max} * log_2\left(\frac{N}{n_t}\right) \tag{2.1}$$

O principal problema dessa abordagem é a possibilidade de problemas de ambiguação. Sendo assim, apesar de ter resultados eficazes, não pode ser considerada como um paradigma da com-

[1] https://inquirer.sites.fas.harvard.edu/

putação afetiva pelos critérios de Picard [2000], segundo os quais, a habilidade de uma máquina de compreender emoções humanas, bem como interagir e interpretá-las, caracterizá-la-ia como um computador *"genuinamente inteligente"* [Picard, 2000; Cruz et al., 2017].

Segundo Jiang and Tan [2005] e da Silva Conrado et al. [2014], a relação TF/IDF pode gerar equívocos ou descaracterizar um determinado sentimento caso a análise esteja focada num domínio específico. Para Jiang and Tan [2005], em alguns casos, quanto menor for a relação TF/IDF, mais relevante e sutil será o sentimento apresentado, o que adentraria no campo semântico, ramo da linguística fundamental para a compreensão do significado e sentido das palavras. Desta forma, é possível diferenciar sentidos conotativos e denotativos.

No caso de ironias e figuras de linguagem, um dicionário léxico não seria capaz de identificar o real sentido de um termo num contexto semanticamente irônico, já que apenas atribuiria um valor oposto à intenção/sentimento capturados em relação ao sentimento comumente atribuído para o termo no contexto em questão. Outro exemplo que poderia ser utilizado para representar possíveis problemas de ambiguação na relação TF/IDF são termos como "ansioso" e "cismado". Ainda que ambos sejam sinônimos, o primeiro é comumente mais utilizado, consequentemente, tem uma relevância maior em relação ao segundo.

Outro problema do modelo léxico, seria a análise dos n-gramas. N-gramas são sequências de palavras num texto, cuja variável n equivale à quantidade de termos na cadeia de vocábulos de um dado documento [ŽiŽka et al., 2019].

Segundo da Silva Conrado et al. [2014], a maioria das pesquisas relacionadas à extração de termos de português brasileiro se restringe aos limites de unigramas, bigramas e trigramas. Assim, caso um documento seja um pouco maior e as relações entre os termos seja fundamental para sua plena compreensão - como nas figuras de linguagem -, a limitação dos n-gramas poderia trazer problemas de interpretação. A Figura 2.1 representa a estrutura de n-gramas, em que cada palavra ou sequência de palavras corresponde numericamente a n. Desta forma, temos uma palavra correspondendo a unigrama, uma sequência de dois termos a bigrama e assim sucessivamente.

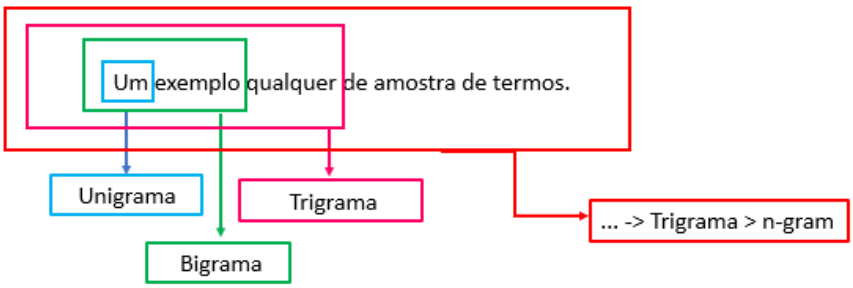


Figura 2.1: Representação da estrutura de n-gramas

Alguns dos dicionários léxicos desenvolvidos para o português falado no Brasil são baseados

em traduções de línguas estrangeiras, como PortLEX [2022], dentre outros [Cruz et al., 2017]. da Silva Conrado et al. [2014] menciona a priorização do uso de léxicos no estado da arte no Brasil, bem como os impactos de traduções automáticas nessas estruturas. Segundo os autores, léxicos que fazem traduções automáticas podem levar a problemas de classificação devido às estruturas dos idiomas inglês e português, que são completamente diferentes.

Um bom exemplo de conjunto de termos é apresentado por da Silva Conrado et al. [2014]. Os autores citam a expressão "*ambientes virtuais de aprendizagem*", cuja estrutura na língua portuguesa é [substantivo + adjetivo + preposição + advérbio], enquanto em inglês, em "*virtual environment for learning*", a estrutura é [adjetivo + substantivo + preposição + advérbio].

Palavras homônimas, mas com significados diferentes, também podem gerar equívocos semânticos. Na palavra "*esmalte*", por exemplo, dependendo de seu contexto/domínio na língua portuguesa, ela pode ser traduzida em inglês para "*polisher*", uma espécie de verniz que protege superfícies, ou como "*enamel*", um produto específico na fabricação de cerâmica [da Silva Conrado et al., 2014].

Talvez uma pré-classificação de domínio por contexto pudesse solucionar este impasse. Uma saída poderia ser o uso de uma estrutura inicial em árvore, cujos nós classificassem os termos em "indústria" e "cosmético" ou "produtos químicos para uso estético" e/ou "produtos químicos de uso industrial", ou seja, uma ontologia, ou árvore ontológica, de domínio químico. Tal árvore aplicada a técnicas de *web* semântica com consultas a outras bases de conhecimento poderia solucionar problemas de desambiguação como esse de maneira eficiente e automática.

Para Hutto and Gilbert [2014], um dos maiores problemas de léxicos consolidados é sua dificuldade em analisar palavras de uso diário, como gírias, jargões e, principalmente, *emoticons* e abreviações, cujo significado semântico é fundamental para a boa compreensão do contexto e AS. Ora, se tal problema ocorre mesmo no texto em seu idioma original, ao uni-lo com os obstáculos expostos por da Silva Conrado et al. [2014], *i.e*, a falta de ferramentas originais voltadas para a língua portuguesa brasileira, bem como o uso de traduções automáticas, podemos concluir que a maioria dos métodos e aplicações existentes não são suficientes para uma AS realmente eficaz de conjuntos de dados textuais do português do Brasil.

No contexto de redes sociais, em que a língua sofre rápida evolução, tais questões linguísticas podem ocorrer com mais frequência, uma vez que a evolução da linguagem, principalmente na *internet*, tem velocidade jamais vista, dificultando até mesmo a classificação com o uso de inteligências artificias [Jung et al., 2017]. Assim, os problemas levantados por da Silva Conrado et al. [2014] e também por Hutto and Gilbert [2014] poderiam ser representados pelo exemplo contido na frase: "*Estou passada!*", cuja interpretação pode variar de acordo com o contexto, cujo uso mais corriqueiro passa a ocorrer apenas num período recente da língua portuguesa, não pertencendo à norma culta vigente. Além disso, caso a frase fosse traduzida para o inglês denotativamente, ela perderia

seu sentido original em português.

2.2.2 Ontologias

É importante destacar que, ao longo dos séculos, o conceito de "ontologia" variou bastante com os avanços da humanidade, ciências e tecnologias. Porém independentemente do campo científico de sua aplicação, sua essência enquanto classificadora de conceitos hierárquicos permanece a mesma até os dias de hoje.

A fim de dar maior escopo a este trabalho e maior entendimento sobre o conceito de ontologias, algumas definições do termo são apresentadas a seguir. Para fins de formalismo conceitual das área de computação e filosofia, Ontologia, com letra maiúscula, aborda o campo filosófico, enquanto com letra minúscula, refere-se às ontologias presentes na Ciência da Computação [Guarino et al., 2009].

O termo "ontologia" (do grego *ontos*, "ente" e *"logoi"*) trata da "teoria do ser". Advindo da filosofia aristotélica, ocupa-se do estudo do ser ou da existência. Trata-se da teoria da natureza da existência, cujas categorias primitivas - como substância e qualidade - presumem explicar "Tudo o que É". Refere-se a um conceito sistemático da existência onde categorias são aplicadas a qualquer objeto possível de cognição [Thomasson, 2019], independentemente de sua existência no mundo [Guarino et al., 2009].

Após o advento da filosofia Kantiana, no século XIX, tornou-se comum denominar projetos de categorização de "descritivismo categorial". Segundo este conceito, seria possível descrever uma estrutura em classes - ou categorias - presentes no mundo, de acordo com o modo de pensar, experiência ou linguagem de um indivíduo, porém sem assumir compromissos se tais categorias estão ou não ocupadas ou são onticamente fundamentais [Thomasson, 2019].

O filósofo e matemático alemão Edmund Husserl, estabelece categorias de significados que podem ser usadas para traçar classes ônticas, *i.e.* categorias, de objetos com possíveis significados, porém sem se preocupar com a existência ou não desses objetos em várias categorias de uma Ontologia. Ainda no campo filosófico, é interessante observar que, nos últimos cem anos, o ceticismo sobre a possibilidade de um sistema unicamente verdadeiro e completo de categorias levou a discussões sobre mudanças na tentativa de oferecer sistemas completos de categorias para traçar distinções particulares, especialmente entre conceitos linguísticos [Thomasson, 2019].

Um sistema de categorias ônticos tem o potencial para muitos usos em filosofia, porém enfrentam uma variedade de dificuldades. Eles devem abordar quais são os métodos adequados de definição de categorias, quantas categorias existem e quais são e se existe - ou não - um único gênero *summum*, ou seja, paradigmático, que englobe todas as demais categorias. Deve-se ainda atentar quanto à distinção de um único sistema de categorias ou dimensões múltiplas de categorias - questões sobre as quais houve pouco acordo [Thomasson, 2019].

Apesar de sua longínqua origem histórica, as ontologias ganharam nova atribuição na área de Ciência da Computação, cujo termo técnico denota um *"artefato"* projetado para uma finalidade, que possibilita a modelagem de conhecimento sobre algum domínio, seja este real ou imaginário [Gruber, 2009]. Simplificadamente, uma ontologia consiste em conceitos, relações-conceito, axiomas e instâncias [Stavrianou et al., 2007].

Segundo [Gruber, 1993], ontologias são representações de um vocabulário específico sobre um domínio compartilhado (classes, relações, funções e objetos), cujas definições são padronizadas por um conjunto de ferramentas de análise e tradução com o objetivo de calcular predicados em linguagens relacionais.

Em 2009, Gruber [2009], atualizou sua definição de ontologia no campo da Ciência da Computação e Informação. De acordo com o novo parecer do autor, uma ontologia define um conjunto de premissas capazes de representar e modelar um domínio de conhecimento ou discurso. Tais premissas representacionais são, de modo geral, classes (ou conjuntos de classes), atributos (ou propriedades) e relacionamentos (ou relações entre os itens da classe). Ainda segundo os autores, as definições das primitivas - ou premissas - representacionais incluem informações sobre seu significado e restrições à sua aplicação para a construção de uma estrutura logicamente consistente.

A seleção de conceitos depende da tarefa e das informações de domínio que precisam ser capturadas [Stavrianou et al., 2007]. Porém, antes de defini-los, é importante saber para quais questões uma resposta será desejada, uma vez que a ontologia esteja construída [Stavrianou et al., 2007].

Ontologias diferem de bancos de dados ou esquemas de *Extensible Markup Language* (XML). Elas representam, principalmente, um domínio e os termos técnicos que o cercam podem ser usados a qualquer momento em que uma análise semântica seja necessária [Stavrianou et al., 2007]. Não há tipos de dados envolvidos nem restrições de integridade. As semânticas são definidas com base nos conceitos de domínio e não dependem de uma aplicação específica, como em um banco de dados [Stavrianou et al., 2007].

Em definições mais recentes, ontologias tratam das relações entre conceitos, axiomas e instâncias. Trata-se, ainda, de um método de conhecimento que facilita a interação entre humanos e computadores através de marcações como a *Resource Description Framework* (RDF) e *Web Ontology Language* (OWL) [Liu et al., 2012].

Segundo Gruber [2009], uma ontologia deve ser formulada em alguma linguagem de representação que especifique um nível semântico, ou seja, independente da estrutura e modelagem dos dados ou sua implementação. Ressalta-se que, ao contrário do *Extensible Markup Language* (XML), a semântica em ontologias não é definida pelo usuário, mas segue as regras impostas pelo domínio relevante [Stavrianou et al., 2007]. Ao contrário de um banco de dados convencional de modelo relacional, cujo principal identificador é uma chave primária que não denota um domínio, ontolo-

gias costumam ser desenvolvidas em linguagens mais próximas do poder expressivo de formalismos lógicos, tais como cálculos de predicados [Gruber, 2009].

Além disso, formalismos tipicamente ontológicos permitem a identificação de um indivíduo em uma classe ou atributo sem referências padrões de modelagens de banco de dados, como identificadores-chave primárias [Gruber, 2009]. Da mesma forma, uma ontologia pode representar restrições que se mantêm através das relações em uma simples declaração, como A é uma subclasse de B, por exemplo. Já em um modelo relacional, essa tarefa seria codificada por meio da definição de chaves estrangeiras. Tal etapa se torna desnecessária nas relações presentes nas ontologias [Gruber, 2009].

Ressaltamos que a Engenharia de ontologias se preocupa em fazer escolhas representacionais que capturem as distinções relevantes de um domínio ao mais alto nível de abstração, sendo o mais clara possível quanto aos significados dos termos [Gruber, 2009], ainda que ontologias genéricas não sejam comumente usadas [Stavrianou et al., 2007]. Desta forma, a estrutura de uma ontologia consiste em uma generalização hierárquica de conceitos, ou seja, uma *taxonomia* [Guarino et al., 2009].

Gruber [2009] destaca ainda como ontologias já codificadas em formalismos padronizados podem ser reutilizadas e projetadas através de relatos sistemáticos do conhecimento humano ou da linguagem. Desta forma, as ontologias incorporam resultados de pesquisas acadêmicas e oferecem um método operacional que possibilita a experiência empírica por meio de sistemas de bancos de dados [Gruber, 2009].

O autor também observa como ontologias são parte dos padrões da W3C [Gruber, 2009; W3C, 2024b,a] para *web* semântica. Tais padronizações são utilizadas para especificar vocabulários conceituais que permitem a troca de dados entre sistemas, realizam operações de consultas e respostas, publicam bases de conhecimento reutilizáveis, além de oferecer serviços que facilitam a interoperabilidade entre sistemas e bancos de dados múltiplos e heterogêneos [Gruber, 2009].

Guarino et al. [2009] definem ontologias computacionais como um artefato computacional que permite a modelagem formal da estrutura de um sistema. Ainda segundo o autor, ontologias apresentam todas as entidades e relações relevantes presentes na arquitetura da informação que podem ser úteis para a representação de um domínio fim. Em definições mais recentes, ontologias tratam das relações entre conceitos, axiomas e instâncias. Trata-se, ainda, de um método de conhecimento que facilita a interação entre humanos e computadores através de marcações como a RDF e OWL [Liu et al., 2012].

Desta forma, ontologias têm sido propostas para lidar com a heterogeneidade semântica ao extrair informações de várias fontes de texto, como a Internet [Stavrianou et al., 2007]. A partir dessa premissa, tanto ontologias quanto léxicos de domínio específicos poderiam apenas ser uma parte de todo o - ainda utópico [Hitzler, 2021] - sistema de *web* semântica. Assim, seria possível que

sistemas realizassem consultas e segmentassem suas buscas por informação e respostas conforme a tarefa ou problema a ser solucionado.

Compartilhar conhecimento, relacionar dados e extrair informações a partir dessas relações são princípios de engenharia ontológica. De acordo com Smith [2019], criador da Ontologia de Fundamentação *Basic Formal Ontology* (BFO), muitas ontologias são criadas equivocadamente, desconsiderando fundamentos teóricos relevantes para o estado da arte. A desatenção à teoria das ontologias tornaria esses sistemas em meras estruturas de dados em grafos, sem real compartilhamento de informação e/ou conexão com outras ontologias.

A Figura 2.2 representa um recorte da Emotion Ontology MFOEM, de Smith [2019], com foco nos sentimentos e suas valências positiva ou negativa, implementado na ferramenta Protégé [Musen, 2015]. Nela é possível observar a estrutura em grafos gerada pela ferramenta, com as relações entre as classes e atributos presentes no recorte.

2.3 Ontologias, Mineração de Texto e Análise de Sentimentos

A principal importância das ontologias é sua capacidade de representar um esquema para um determinado domínio, esclarecendo termos técnicos que aparecem num texto ou especificando a relação entre determinados conceitos de domínio. Durante o processo de MT, as ontologias podem ser usadas para fornecer conhecimento básico e especializado sobre um domínio [Stavrianou et al., 2007].

De acordo com Jiang and Tan [2005], uma estrutura ontológica poderia ter melhor desempenho ao tratar de tarefas de AS do que as contribuições pré-existentes com abordagem léxica. Assim, os autores apresentam sua ferramenta, a CRTOL, que, segundo os mesmos, é capaz de fazer relações semânticas entre os conceitos de maneira mais precisa do que as já descritas no estado da arte. Eles também apontam como a relação TF/IDF pode trazer ainda mais divergências nos resultados dos processamentos de AS. Para os autores, estruturas linguísticas mais elaboradas requerem ferramentas de MT mais robustas [Jiang and Tan, 2005].

Segundo os mesmos autores, ao analisar um *corpus* de domínio fechado, ou seja, uma base ou repositórios de textos e/ou documentos com um determinado assunto em comum, a relação TF/IDF traria divergência na análise de frases mais complexas. Os autores apresentaram os resultados de seus estudos apontando como os verbos auxiliares da língua inglesa, como *may do, have to do*, etc, podem ter seu significado prejudicado.

Saranya and Jayanthy [2017] também expuseram a deficiência da AS no que concerne à classificação textual. Para os autores, uma abordagem ontológica seria mais eficiente ao avaliar a emoção evocada em textos de usuários de mídias sociais. Suas pesquisas apontam que ontologias poderiam ajudar a classificar automaticamente textos, frases e documentos com base em suas relações semân-

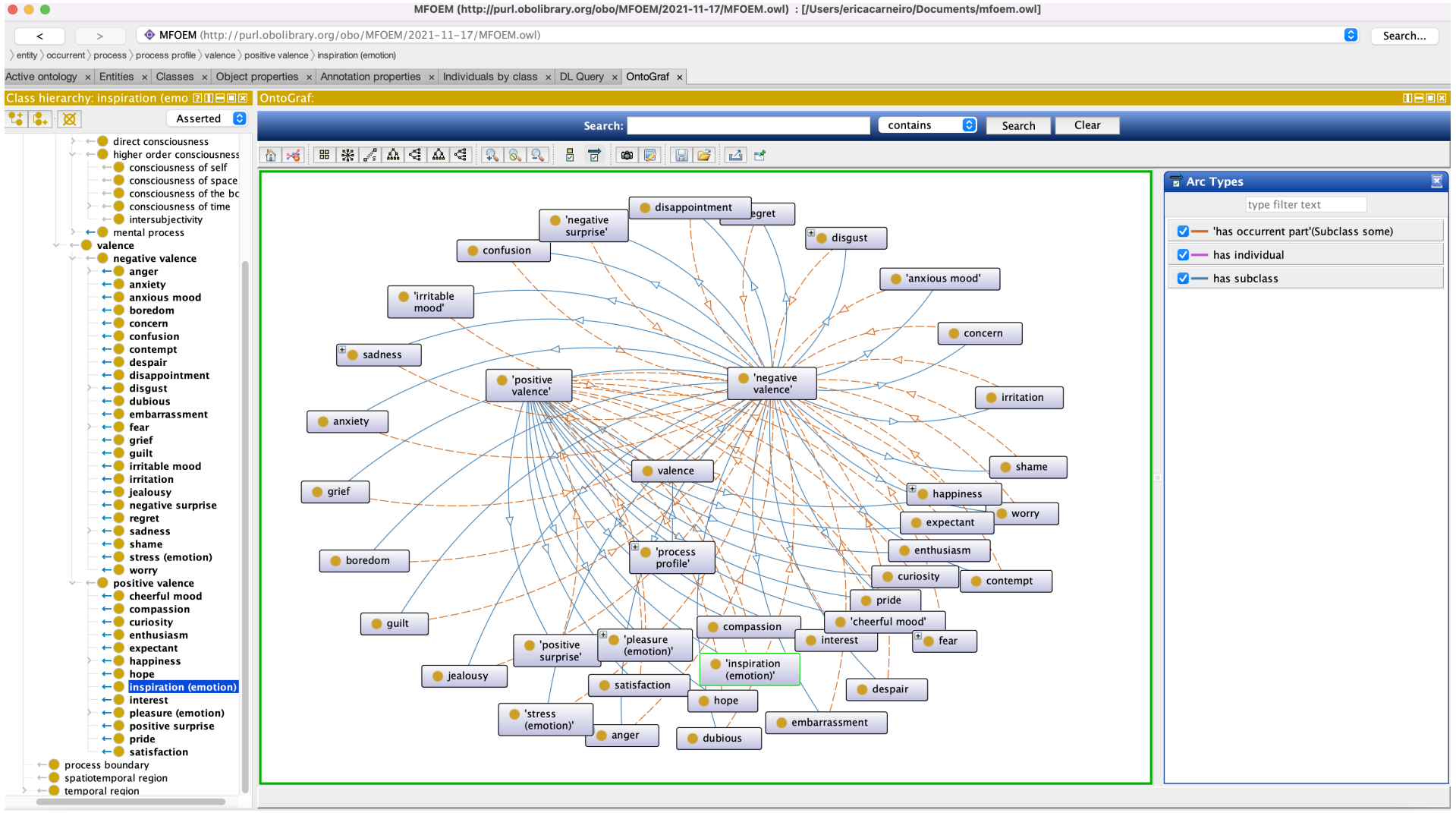


Figura 2.2: Reprodução de parte do grafo gerado pela ontologia de Emoções MFOEM na ferramenta Protégé [Musen, 2015]

ticas. Além disso, a combinação de uma ontologia com técnicas de AM traria melhores resultados do que uma análise morfossintática [Saranya and Jayanthy, 2017]. Para os autores, as ontologias mais eficientes utilizaram soluções que combinam análise de polaridades e AM.

Ainda segundo Saranya and Jayanthy [2017], uma ontologia poderia ser criada em um domínio fechado, trazendo significado semântico e relação entre os termos do *corpus*. Para os autores, ontologias *fuzzy* poderiam ser uma solução a ser considerada, devido a sua estrutura em árvore com nós e relações hierárquicas entre si [Saranya and Jayanthy, 2017]. Suas pesquisas mostram o funcionamento e eficiência da ferramenta WordNet, que classifica as emoções, hierarquizando-as por classe. Posteriormente, os sentimentos extraídos do *corpus* são assimilados a suas classes semânticas. É importante acrescentar que, apesar de ter surgido como um léxico, e até hoje ser referido como tal, o WordNet possui uma versão em OWL, mapeada na *Suggested Upper Merged Ontology* (SUMO), contando com relações de conhecimentos e estrutura em grafos [Niles and Pease, 2003]. Desta forma, está mais próxima de uma base de conhecimento da língua inglesa do que de um léxico computacional.

O uso de sistemas ontologicamente estruturados ganhou destaque no campo da IA, uma vez que ajudam no compartilhamento e reutilização de um conhecimento formal adquirido entre sistemas de IA [Gruber, 1993]. A existência de algumas ontologias genéricas, como uma Ontologia de Fundamentação ou *Upper Ontology (UO)*, cuja finalidade é ser reutilizável através de conceitos abstratos [Stavrianou et al., 2007], pode ajudar bastante nessas tarefas.

A *SUMO* é uma dessas UO. Trata-se de uma ontologia fundamental com conceitos gerais independentes de qualquer domínio, *i.e.* de domínio aberto, cuja estrutura pode ser utilizada como base para a criação de ontologias de domínio específico. Niles and Pease [2003] mapearam conceitos da SUMO para palavras semanticamente equivalentes, denominadas *synsets*, do WordNet. Segundo os autores, esses mapeamentos podem funcionar como um índice de linguagem natural para construção de conceitos de uma ontologia, atuando como um conector entre o texto livre e verificador de integridade sobre o conteúdo da ontologia [Niles and Pease, 2003].

Devido às características hierárquicas das ontologias, um método com abordagens ontológicas permitiria classificações que categorizem e agrupem informações em classes, além de permitir a reutilização desse conhecimento por outras ontologias. Teríamos, então, um fluxo constante de compartilhamento de conhecimento, corroborando as afirmações dos autores de que ontologias podem ser de grande auxílio em tarefas de conceituação, reutilização, compartilhamento dos recursos e co-referenciamento de termos.

Técnicas estatísticas ou de AM têm lidado com o problema de construir ontologias a partir de um texto, porém o maior desafio ainda reside na semântica do conteúdo e do resultado [Stavrianou et al., 2007]. Muitos pesquisadores se valem de relações simples entre os membros das ontologias

extraídas que dependem de técnicas avançadas de PLN e julgamento humano, porém para que uma ontologia de sentimentos desempenhe um papel relevante para a AS, muitos sinônimos e relações são exigidos [Jung et al., 2017].

Numa analogia com a semiótica, seres humanos sempre utilizaram a linguagem e seu conhecimento de mundo, ou de domínio, para compreender as mensagens enviadas por um emissor [Guarino et al., 2009]. Da mesma forma, ontologias também utilizam representações de linguagem formal para descrever os domínios de documentos [Narayanasamy et al., 2021].

Devido aos fatores supracitados, ontologias podem ser um meio de economizar o processamento ao lidar com análises de texto, que, em alguns casos, podem ser excessivamente custosas para diferentes modelos de informação. Além disso, ontologias foram amplamente utilizadas em problemas de desambiguação de conteúdos textuais, uma vez que sua estrutura permite uma descrição formal de representação de conceitos padronizados do mundo real [Narayanasamy et al., 2021].

Para Liu et al. [2012], ontologias de domínio criam relações entre um dado domínio e os conceitos presentes no mesmo. Sendo assim, abordagens ontológicas poderiam trazer a resposta para sanar possíveis falhas de classificação de AS computacional em contextos específicos. Assim, uma estrutura ontológica poderia ter desempenho superior a uma abordagem estatística quando aplicada à AS, conforme observado por Cambria et al. [2017]. Já para Vilares et al. [2018] o desenvolvimento de sistemas com acesso a grande quantidade de conhecimento, tanto sobre o mundo quanto sobre o domínio do discurso, seria capaz de ir além das abordagens de PLN já tradicionalmente utilizados, como *Bag of Words* (BOW), ou aplicativos com enorme dependência de domínio (como léxicos) e ferramentas de baixa cobertura, que, segundo estes autores tornam-se cada vez menos eficientes.

Como observado por Dragoni et al. [2018], apesar de esforços e pesquisas na área, a maioria das contribuições ontológicas para a AS são principalmente em inglês. Raras são as abordagens com ontologias de domínio de sentimentos voltadas para tarefas de classificações de texto. Na literatura, encontramos ontologias com capacidade de processar documentos multilíngues, uma desenvolvida para *corpora* em espanhol latino-americano [García-Díaz et al., 2020], duas para o português brasileiro [Freitas and Vieira, 2013; Rosa et al., 2019] e uma para o idioma sul-coreano [Jung et al., 2017]. Todas as demais ontologias trabalham com *corpus* em inglês. Entretanto, poucas ferramentas foram disponibilizadas ao público. Em contraste com outros idiomas, as abordagens baseadas em ontologia usando o *corpus* em inglês são predominantes.

2.4 Processamento de Linguagem Natural e Análise de Sentimentos no Brasil

da Silva Conrado et al. [2014] e Pereira [2020] apontam que a grande maioria dos métodos de PLN empregados para a língua portuguesa permite majoritariamente apenas a análise morfossintática das palavras e conceitos, sem que haja uma ferramenta que represente e contribua verdadeiramente

para o campo semântico da linguística computacional no Brasil. Segundo da Silva Conrado et al. [2014], são muitas as lacunas na área de PLN no Brasil. Para os autores, traduções automáticas não são tão eficientes para soluções de tarefas de PLN devido à diferença nas estruturas dos idiomas inglês e português.

A relevância do trabalho de da Silva Conrado et al. [2014] perdura na pesquisa contemporânea. Apesar da exploração detalhada dos métodos tradicionais, atualmente não há um resumo sistemático dos avanços nos sistemas supervisionados de extração de termos usando ML ou técnicas de aprendizagem profunda [Tran et al., 2023]. Segundo da Silva Conrado et al. [2014], identificar termos num domínio é subjetivo, com 60% dos especialistas admitindo dificuldades e desafios. Uma vez que a extração automática de termos permanece um campo impreciso e sem solução, passados 10 anos, a pesquisa de da Silva Conrado et al. [2014] ainda é relevante no campo extração de termos [Soliman, 2023].

A Tabela 2.1 é um recorte do cenário da língua portuguesa dos resultados de nossas pesquisas na rede Scopus sobre métodos com ontologias e léxicos para análise de sentimentos. A Tabela 2.1 lista alguns dos estudos mais importantes em português brasileiro e indica se o estudo aplicou uma tradução ou se foi desenvolvido originalmente em português. Vinte e dois estudos com foco em AS foram encontrados. De 10 trabalhos baseados em ontologias, apenas dois tiveram foco em AS. Nove trabalhos foram baseados em léxicos de AS. A Tabela 2.1 apresenta o título do estudo, sua abordagem e idioma original, o tratamento dado ao problema de linguagem, o domínio (sentimentos ou outros domínios) e o ano do desenvolvimento do trabalho. Como a SenticNet é uma ferramenta multilíngue sem pesquisa em português, não a incluímos na Tabela 2.1.

Tabela 2.1: Uma comparação entre abordagens lexicais e baseadas em ontologias para Computação Afetiva no Brasil, segundo o Idioma Original e Domínio (*EN = Inglês; PT = Português; PTB = Português Brasileiro).

Título do Estudo	**Abordagem**	**Idioma Original***	**Fonte Idiomática**	**Abordagem de AS**	**Domínio**	**Ano**
Wordnet AffectBR [Pasqualotti and Vieira, 2008]	Léxico	EN	Tradução Humana	X	N/A	2003
LIWC [PortLEX, 2022]	Léxico	EN	Tradução	X	N/A	2007
ANEW BR [Kristensen et al., 2011]	Léxico	EN	Tradução	X	N/A	2011
Opinion lexicon [Souza et al., 2011a]	Léxico	EN	Tradução via API	X	N/A	2011
Wordnet AffectBR_adapt [Longhi, 2011]	Léxico	EN	Tradução Humana	X	N/A	2011
Reli Lexicon [Freitas, 2013]	Léxico	PT	Brasil	X	N/A	2013
Polyglot [Chen and Skiena, 2014]	Léxico	EN	Tradução Automática	X	N/A	2014
Personalitem Lexicon [Machado et al., 2015]	Léxico	PTB	Brasil	X	N/A	2015
Sentilex-PT [Carvalho and Silva, 2015]	Léxico	PT	Portugal	X	N/A	2016
Unilex [de Souza et al., 2017]	Léxico	PTB	Brasil	X	N/A	2017
LIWC 2015pt [Carvalho et al., 2019]	Léxico	EN	Tradução	X	N/A	2018
AffectPT-br [Carvalho et al., 2018]	Léxico	EN	Brasil	X	N/A	2018
Linguateca [Fundação para a Computação Científica Nacional, 2022]	Ontologia	PT	Brasil e Europa		Linguística	1998
TEXTQUIM/lTEXTECC [Instituto de Letras da UFRGS, 2022]	Ontologia	PTB	Brasil		Química/Medicina	2003
NANOTERM [Kasama et al., 2011]	Ontologia	PTB	Brasil		Nanotecnologia	2006
OntoLP [OntoLP, 2008]	Ontologia	PTB	N/A		Contrutor de Ontologia Semi-Automático	2008
Bio-C [Pino, 2021]	Ontologia	PTB	Brasil		Bio-Combustível	2009
e-Termos [EMBRAPA, 2022]	Ontologia	PTB	Brasil		Eletrônica	2009
TermiNet [Di Felippo, 2010]	Ontologia	PTB	Tradução Humana		Linguística	2009
Hontology [Freitas and Vieira, 2013]	Ontologia	PTB	Brasil	X	Hotelaria e Acomodações	2012
TOPTAX [Moura et al., 2022]	Ontologia	PTB	Brasil		Química, Computação, Física, IFM	2013
KBRS [Rosa et al., 2019]	Ontologia	PTB	Brasil	X	Depressão e Estresse	2018

Além do problema com as traduções automáticas, a falta de diversidade em metodologia com forte apelo ao uso de léxicos, o cenário da PLN no Brasil sofre com a mesma escassez de recursos que países não falantes da língua inglesa. Segundo Steinberger et al. [2011], conjuntos de dados anotados para análise de sentimentos são raros em idiomas diferentes do inglês. Para os autores, a maioria das contribuições em idiomas alternativos, são pobres quando comparadas aos recursos em inglês.

2.5 Abordagens para Análise de Sentimentos: Léxicos Afetivos e *Sentic Computing*

A classificação de sentimentos em positivo, negativo ou neutro é mais comumente aplicada no nível de sentenças [Rachid et al., 2018]. Nesta seção, apresentaremos mais detalhes sobre os recursos utilizados para a avaliação de AS, discorrendo um pouco mais sobre a *sentic computing* e a SenticNet, bem como sobre os léxicos afetivos com os quais ela foi comparada.

2.5.1 Léxicos Afetivos

Os léxicos afetivos utilizados para este trabalho foram baseados nos estudos de Nascimento [2021] e Cambria et al. [2022]. A maioria dos léxicos, assim como a SenticNet, possuem traduções automáticas. Apenas o léxico VADER [Hutto and Gilbert, 2014] está exclusivamente em inglês.

AffectPT-br

O AffectPT-br [Carvalho et al., 2018] contém 1.139 palavras, das quais 479 são positivas, e 661 negativas. O Léxico Afetivo (LA) foi desenvolvido a partir da categoria de afeto do LIWC 2015 (*"affect"*), além de *"posemo"* e *"negemo"*, que correspondem respectivamente de às categorias de emoções positivas e negativas.

O AffectPT-br foi fruto de uma tradução automática via *Google Translate* [Google, 2024] e *Microsoft Translator* via *Bing Translator*[Microsoft, 2024]. Os autores também utilizaram o *Oxford Advanced Learner's Dictionary* para avaliar a ferramenta que utilizou termos que fizessem mais sentido ao traduzir as palavras para o português [Carvalho et al., 2018].

ANEW-Br

O ANEW-Br é um LA com 1.046 palavras, cujas medidas abordam duas dimensões emocionais do *Self-Assessment Manikin*, a saber: valência e excitação. A primeira visa obter uma avaliação entre $1,16$ a $8,80$ para o quão agradável ou desagradável um estímulo é percebido, sendo 1 para "muito agradável" e 9 para "muito desagradável". Na escala entre 4 e 6 encontram-se as palavras com valência neutra, já entre $2,22$ a $7,67$ se encontram os valores de excitação. Após uma tradução direta do inglês, as anotações de valência foram efetuadas por alunos 755 de graduação.

LIWC, LIWC2007-PT-br e LIWC2015-PT-br

O LIWC é um LA com mais de 6.626 palavras divididas entre 80 categorias. A partir de um método de BOW, o LIWC analisa a porcentagem de termos presentes em um texto que se enquadram a uma ou mais das 80 categorias. Segundo Newman et al. [2003], essas categorias são agrupamentos de palavras separadas em subdicionários de domínio.

O LA possui as categorias de afeto *affect*, *positive emotions* (posemo) e *negative emotions* (negemo) [Tausczik and Pennebaker, 2010; Newman et al., 2003], que correspondem aos sentimentos positivos e negativos. Desta forma, é possível avaliar o peso de uma sentença de acordo com a quantidade de palavras que pertencem a cada categoria.

O LIWC está disponível em diversos idiomas, incluindo o português. A seguir entraremos em mais detalhes dos LAs afetivos LIWC2007-PT-br [Filho et al., 2013] e LIWC2015-PT-br [Carvalho et al., 2019], que são as versões do LIWC disponíveis em português do Brasil.

O LIWC recebeu uma primeira versão em português, denominada LIWC2007-PT-br, desenvolvido pelo Núcleo Interinstitucional de Linguística Computacional (NILC). O LIWC2007-PT-br conta com 127.161 palavras que podem pertencer a uma ou mais de suas 64 categorias [Filho et al., 2013]. Com grande quantidade de palavras, porém, segundo Carvalho et al. [2018] não necessariamente traz benefícios ao LA, já que algumas palavras possuem erros de ortografia ou caracterização.

Posteriormente, uma nova versão do LIWC foi desenvolvida em português. Ela foi inspirada na versão mais recente do LA em inglês, o LIWC2015 [Pennebaker et al., 2015], e possui 14.459 palavras que podem ser divididas em 73 categorias [Carvalho et al., 2019].

É importante observar que ambas as versões do LIWC em português, possuem grande quantidade de palavras devido a possíveis flexões de gênero e plural de adjetivos e substantivos da língua portuguesa [Carvalho et al., 2019], característica que a difere mais uma vez da estrutura linguística do inglês.

MPQA-PT

O *Multi-Perspective Question Answering* (MPQA-PT) [Almeida et al., 2015] é tradução do inglês do *Multi-Perspective Question Answering* ou MPQA. A versão em português do LA possui 9.945 palavras e expressões, subdivididas em três categorias. A primeira delas representa as classes morfológicas das palavras, como Verbo (V) e Adjetivo (A). A segunda, trata da subjetividade ou intensidade dos termos em mais fortes ou mais fracos. A terceira aborda a valência dos sentimentos, que podem ser negativos, positivos ou neutros [Almeida et al., 2015].

O *corpus* do MPQA-PT foi anotado manualmente por três linguistas e conta com 80 tópicos de notícias extraídos de jornais, revistas e sites. Cada um desses tópicos possui 10 documentos [Almeida et al., 2015], sendo um bom exemplo de ferramenta para recuperação de informação em

português.

Opinion Lexicon

O Opinion Lexicon, ou OpLexicon, [Souza et al., 2011b] foi elaborado a partir de três diferentes abordagens. Elas podem ser baseadas em *corpus*, léxico e tradução.

A primeira possui 1.317 documentos e aproximadamente um milhão de palavras divididas em 970 textos jornalísticos sobre diferentes assuntos, além de 346 documentos sobre críticas de filme. A segunda abordagem, baseada no *thesaurus* TEP, contem 44.077 palavras anotadas. Já a última abordagem, baseou-se numa tradução automática do inglês para o português do Opinion Lexicon [Hu and Liu, 2004] com a ferramenta Google Translate.

Posteriormente, os autores juntaram as três abordagens e observaram possíveis interseções cujas valências pudessem corresponder a palavras de opinião, culminando num LA com 32.191 palavras.

Polyglot

O Polyglot [Chen and Skiena, 2014] é um LA e também foi obtido por meio de traduções automáticas via *Google Translate* para mais de 136 idiomas, dentre eles o português. A versão em português do Polyglot possui 3.953 palavras dentro de uma única categoria que atribui um valor de (-1) para sentimentos negativos e (1) para positivos.

Reli-Lex

Elaborado a partir de resenhas de livros publicadas presentes na internet, o ReLi-Lex [Freitas, 2013], foca em contextos informais do uso do idioma. O LA possui uma única categoria com 609 palavras, cujas valências podem ser negativa (-1) ou positiva (1).

Sentilex-PT

O LA Sentilex-PT [Carvalho and Silva, 2015] possui 7.014 palavras e expressões, divididas em quatro categorias que correspondem à classe gramatical, polaridade (1 para positiva, -1 para negativa e 0 para neutra), alvo da polaridade e anotação manual da polaridade. O principal objetivo do Sentilex-PT é analisar o sentimento evocado por determinadas entidades humanas, ou figuras públicas, em textos na língua portuguesa [Carvalho and Silva, 2015].

SentiStrength

Presente em 15 idiomas, incluindo o português, o SentiStrength [Thelwall, 2017] possui 2.212 palavras dentro de uma única categoria. Esta conta com uma escala de valências de -5 a 5, que

corresponde respectivamente aos atributos extremamente negativo e extremamente positivo. A valência 1 corresponderia ao sentimento neutro.

WordNetAffect, WordnetAffectBR e WordnetAffectBR(adapt)

Trata-se de uma extensão do *WordNet Domains*, que é um recurso lexical com o objetivo de aumentar os rótulos de domínio do *WorNet* de forma semi-automática [Fondazione Bruno Kessler, 2009b]. No caso do *WordNet Affect* o domínio trata de um subconjunto de *synsets* que representam sentimentos evocados em palavras afetivas [Fondazione Bruno Kessler, 2009a]. *Synsets* são palavras que expressam o mesmo significado em pelo menos um contexto, ou seja, um conjunto de sinônimos [Niles and Pease, 2003].

O WordNet surgiu como um léxico voltado para PLN, com categorias relacionadas a classes gramaticas, porém atualmente possui uma versão em OWL, mapeada na SUMO por Niles and Pease [2003] com relações de conhecimentos e estrutura em grafos.

Composto por 289 palavras e expressões e uma única categoria, que representa as valências (+) para sentimento positivo e (−) para o sentimento negativo, o WordnetAffectBR [Pasqualotti and Vieira, 2008] é a versão brasileira, em português, do *WordNet Affect*. Posteriormente, o WordnetAffectBR de Pasqualotti and Vieira [2008] recebeu uma extensão fruto da tese de doutorado de Longhi [2011], denominada WordnetAffectBR(adapt). A nova versão do LA em português conta com 1.023 palavras e expressões. Não há informações sobre o mapeamento em OWL das versões brasileiras do WordNet até a conclusão deste trabalho.

VADER

Com aproximadamente 7.500 palavras o *Valence Aware Dictionary and sEntiment Reasoner* (VADER) [Hutto and Gilbert, 2014] é um LA que indica a polaridade de um sentimento e um dado texto em positivo ou negativo. Possui um escala entre −4 (extremamente negativo) e 4 (extremamente positivo) que aponta a intensidade do sentimento [Hutto and Gilbert, 2014]. O VADER também possui uma pontuação para o sentimento neutro equivalente a 0.

2.5.2 *Sentic Computing*

Resumidamente, as abordagens de AS consolidadas na literatura podem ser divididas em três categorias: (a) localização de palavras-chave, que classifica um texto em categorias com base em palavras afetivas explícitas; (b) afinidade lexical, onde há atribuição arbitrária de uma palavra a uma emoção específica, devido à afinidade probabilística; (c) métodos estatísticos, que determinam o cálculo da valência de termos afetivos, bem como a frequência dessas palavras num conjunto de dados de treino. Em suas pesquisas, Cambria and Hussain [2012] visaram preencher a lacuna semântica

entre os dados presentes no nível das palavras e as opiniões em nível de conceito transmitidas por elas.

A *sentic computing* é um campo interdisciplinar que combina técnicas de PLN e Aprendizado de Máquina para analisar e compreender sentimentos e emoções humanas [Cambria and Hussain, 2015]. Além disso, pode ser aplicada tanto para a análise de sentimentos, quanto avaliação de opiniões, reconhecimento de emoções, previsão de tendências, detecção de notícias falsas, análise de discursos políticos, dentre outras. Deste modo, técnicas de PLN, AM e redes neurais são aplicadas para resolver as tarefas de análise de sentimentos [Cambria and Hussain, 2015].

Segundo Cambria and Hussain [2012], para o desenvolvimento da SenticNet, com os métodos de *sentic computing*, aplicou-se uma unificação das técnicas de mineração de grafos e de redução de multidimensionalidade, que reduz o número de variáveis aleatórias, em duas bases de conhecimento de senso comum. Por meio deste novo sistema de mineração de opinião de domínio aberto e análise de sentimento, a *sentic computing* possibilita a inferência dos aspectos conceitual e emocional das informações associadas a opiniões em linguagem natural. Através de uma análise semântica a nível de cláusulas textuais, torna-se viável a conversão de informações contidas em texto não estruturado em dados que podem ser processados por máquinas [Cambria and Hussain, 2012].

Essas abordagens são, ainda, baseadas em conceito e utilizam ontologias da *Web*, ou redes semânticas, para executar a análise semântica de textos. Tais técnicas permitem a detecção de sentimento sutilmente expressos, bem como a análise de expressões idiomáticas, que aparentemente não evocam nenhuma emoção, mas cujos conceitos aos quais estão relacionadas transmitem muita informação quanto ao sentimento do texto [Cambria and Hussain, 2015].

A concepção de BOW foi substituída por uma representação em *saco de conceitos*, ou *bag-of-concepts*, incorporando conceitos linguísticos que vão além da estatística. Essa abordagem pode enriquecer a representação textual por meio da relação de dependência entre as orações [Cambria and Hussain, 2015].

O principal objetivo da metodologia de *sentic computing* é desenvolver uma computação de senso comum com padrões linguísticos, que vão além da abordagem palavras-chave, de modo a preencher a lacuna cognitiva e afetiva entre os dados da linguagem natural no nível da palavra e as opiniões no nível do conceito transmitidas por eles. Além disso, a abordagem de análise de sentimento no nível de conceito possui uma visão multidisciplinar. As pesquisas de [Cambria and Hussain, 2012] e [Cambria and Hussain, 2015] foram extremamente extensas e culminaram num *framework* com abordagem de *sentic Computing*: a SenticNet.

SenticNet

A Senticnet 6 é um *framework* de MO público, cujo objetivo é gerar uma coleção de pares de palavras (onto-pares) de "conceito polar" comumente usadas e com polaridade positiva ou negativa relativamente forte. Na SenticNet, cada conceito c está associado a um valor de Polaridade P de apenas um Pc, entre (-1, 1), representando sua polaridade de modo a evitar redundâncias e facilitar a representação dos onto-pares da SenticNet como uma rede semântica. Como exemplo, podemos citar a expressão *"make good impression"*, cuja pontuação do par ontológico é 1.1, por exemplo, pois possui dois termos que combinados reforçam a positividade da expressão idiomática [Cambria and Hussain, 2012]. A estrutura da base de conhecimento da SenticNet é composta de pares de palavras e se encontra no formato *Web Ontology Language* (OWL), representando muito bem a construção de sua rede semântica.

O *framework* emprega IA sub-simbólica para reconhecer padrões significativos em textos, representando-os em uma base de conhecimento. O aprendizado profundo, é capaz de generalizar palavras e expressões multipalavras em primitivas, posteriormente definidas em termos de super primitivos. Assim, temos o exemplo: "*shop_for_iphone11*" ou "*buy_huawei*" generalizados para "*BUY(PHONE)*", que pode, ainda, ser reduzido para "*BUY(X)*".

A *sentic algebra* cria a lógica por trás dos onto-pares de palavras, diferenciando-os de uma análise meramente estatística ou apenas com base na classificação dos sentimentos dos vocábulos dentro do léxico. Por exemplo, $grow_profit$, $enhance_reward$ ou $intensify_benefit$ são todos generalizados como AUMENTO ($GAIN$). Logo, classificados como positivos. Numa abordagem puramente estatística, expressões como $lessen_agony$, $lessen_affliction$ ou $diminuir_sofrimento$ seriam classificadas como negativas, devido ao resultado estatisticamente negativo composto por tais palavras. Desta forma, como ambas são palavras negativas e $(- * - = +)$, o retorno é um resultado positivo. Logo, a classificação desses pares de palavras também serão sentimentos positivos.

A Figura 2.3 apresenta um exemplo do onto-par de palavras "diminuir_sofrimento", e mostra como essa combinação de termos negativos retorna, na verdade, um sentimento positivo, cujo valor da polaridade é 0,788.

Além da *sentic algebra*, Dragoni et al. [2018] e também Cambria et al. [2020] desenvolveram a SenticNet 6 como uma ferramenta robusta, capaz de extrair Primitivos utilizando algoritmos de AM, como o *Long Short-term Memory* (LSTM). O modelo, de aprendizado profundo, permite a descoberta automática de agrupamentos de conceitos semanticamente relacionados que compartilham uma função lexical semelhante [Cambria et al., 2020].

A SenticNet atribui valores de polaridade contextual em positivo, negativo e neutro como uma de suas tarefas mais básicas. Tais conceitos são rotulados por meio de recursos de IA neuro simbólica. Trata-se de um conceito que tem como objetivo minimizar comportamentos inesperados ou mesmo

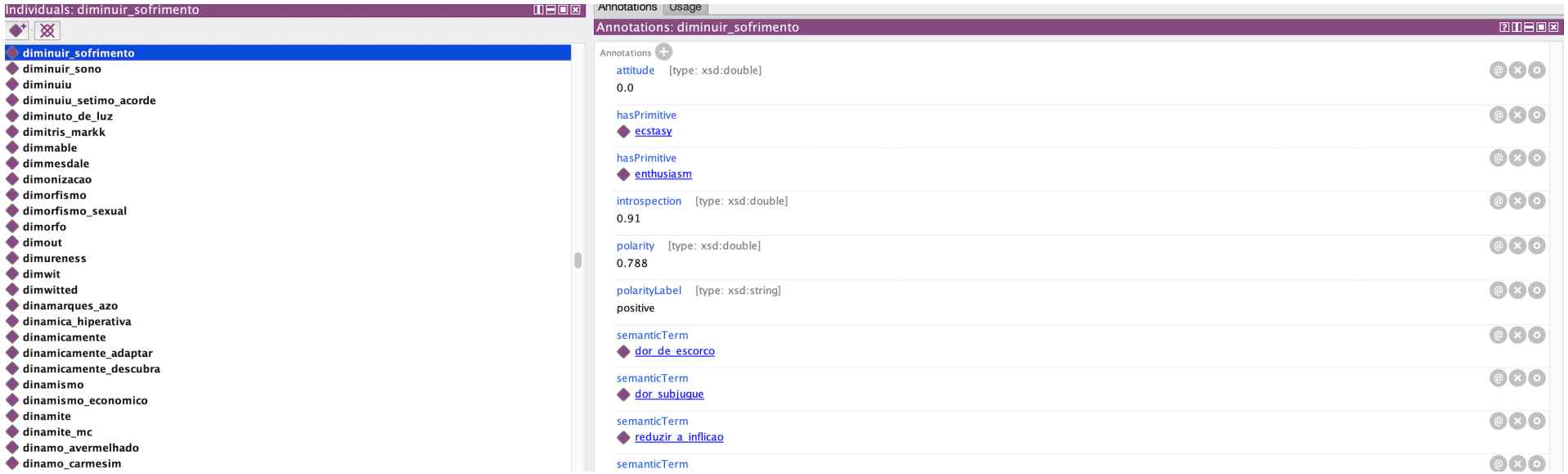


Figura 2.3: Representação de Onto-Pares da SenticNet 6 na Protégé

incluir interpretabilidade, ao juntar os *"parâmetros, i.e, neuro, com modelos de conhecimento pré-definidos (simbólicos)"* [Cambria et al., 2022].

Assim, essa polaridade flui entre arcos de dependência que concedem um rótulo de polaridade final a cada frase por meio de padrões sênticos [Poria et al., 2014]. Os padrões sênticos são parte da *sentic computing* [SenticNet, 2012b], método utilizado pela SenticNet, com o objetivo de sanar a lacuna entre a PLN estatística e disciplinas como Linguística, Raciocínio de Senso Comum, Semiótica e Computação Afetiva, que são necessárias para a compreensão da linguagem humana [Poria et al., 2014].

A SenticNet permitiria, então, uma melhor compreensão do papel contextual de cada conceito no texto e um melhor entendimento do papel contextual de cada conceito no *corpus*. Desta forma, seria possível obter uma inferência de polaridade dinâmica capaz de superar a precisão e tempo de treinamento de métodos estatísticos de última geração [SenticNet, 2024].

OntoSenticNet

A OntoSenticNet desenvolvida por Dragoni et al. [2018], é uma aplicação com abordagem ontológica para AS baseada em métodos de *fuzzy ontology* e verificação de polaridade dos conceitos primitivos presentes no *corpus*. A ferramenta é parte da SenticNet, um repositório semântico com 100.000 termos e seus significados em diversos idiomas.

Uma das principais diferenças e características da Senticnet e da OntoSenticNet é sua relação hierárquica conceitual, que conta com propriedades que relacionam conceitos e valores de sentimento. Além disso, a ferramenta também utiliza métodos linguísticos e estatísticos em suas análises, AM e IA simbólica e sub-simbólica [Dragoni et al., 2018].

Na OntoSenticNet, o conhecimento de senso comum é organizado em entidades, que, por sua vez, ligam-se a conceitos conectados a primitivos conceituais. Neles, o significado é codificado em termos de polaridade e emoções. Assim, SenticNet pode extrair primitivos usando algoritmos de AM, como *LSTM*. Para inferir sentimentos dependentes, Dragoni et al. [2018] desenvolveram um modelo de *LSTM* com mecanismo de atenção empilhada para níveis de alvo e de sentença. Além disso, os autores propuseram um Sentic *LSTM*, que integra conhecimento explícito e implícito através de uma célula *LSTM* com um portão de saída adicional. Um híbrido do Sentic *LSTM* com uma rede aditiva recorrente simula padrões sênticos. Desta forma, o modelo de aprendizado profundo permite a descoberta automática de *clusters* de conceitos semanticamente relacionados que compartilham uma função lexical semelhante [Cambria et al., 2020]. Seu idioma principal é o inglês, mas a ferramenta é multilíngue e inclui o português. A Figura 2.4 apresenta um recorte das relações semânticas presentes no conceito diminuir_sofrimento.

Assim como a SenticNet, a Ontosenticnet foi desenvolvida em inglês e é capaz de analisar con-

juntos de dados multilíngue. Em fevereiro de 2021, a OntoSenticnet recebeu uma extensão, a *DomainSenticNet* (DSN). Ao contrário de sua precursora, a DSN está em análise para ser disponibilizada ao público via *Kickstarter* [Distante et al., 2021].

BabelSenticNet

Desenvolvida na versão 3 da SenticNet, a BabelSenticNet foi a primeira base de conhecimento de conceito multilíngue para AS. As bases foram disponibilizadas à comunidade científica em 40 idiomas em RDF/XML e via *Application Programming Interface* (API) [Vilares et al., 2018].

O objetivo da BabelSenticNet é gerar a SenticNet para diversos idiomas. Os autores utilizaram ferramentas estatísticas de tradução automática *Statistic Machine Translation (SMT)* de alta cobertura, bem como um algoritmo com técnicas de mapeamento baseado no WordNet e suas versões multilíngues, a fim de aumentar a robustez dos recursos traduzidos. Assim, a combinação dos conceitos da SenticNet com *synsets* do WordNet permitem que a tradução sejam encontrada no WordNet de destino [Vilares et al., 2018].

Segundo Vilares et al. [2018] a BabelSenticNet apenas envolve os conjuntos e conceitos que pertence a um grafo de associação de termos em inglês, já que alguns conceitos podem ter seu atributo afetivo modificado devido a diferenças culturais. Os autores dão o exemplo do conjunto de termos *buy_christmas_present* (comprar_presente_de_natal) que para países ocidentais no hemisférios norte podem também estar relacionados a conceitos como *annual_celebration* (celebração_anual) e *winter_time*. No Brasil, o Natal acontece no verão, o que lhe concede características totalmente diferentes, e evocar sentimentos e percepções diferentes de um habitante do norte do planeta.

Desta forma, em um Grafo (G) em inglês (in) Gin e um conjunto de línguas-alvo T, o objetivo é obter um conjunto de grafos traduzidos (T). Dicionários presentes na internet foram utilizados para encontrar a tradução referente a um conceito em inglês, para então chegar a um conjunto semântico de termos. Esta abordagem foi inicialmente aplicada para o chinês, porém pode ser utilizada em outros idiomas, ainda que necessite de uma grande quantidade de recursos em para obter uma versão não inglesa do SenticNet.

Vilares et al. [2018] executaram um experimento em que solicitaram a anotadores humanos que avaliassem a qualidade da tradução da abordagem utilizada nos idiomas espanhol, português, italiano, hindi e chinês. Os entrevistados deveriam classificar se uma palavra ou conjunto de termos é, de fato, semanticamente relacionada ao seu conceito. A palavra cachorro, por exemplo, além de evocar, termos como animal_de_estimação e mamífero, poderia trazer à tona conceitos como "passear", "adestramento", dentre outros. [Vilares et al., 2018].

É importante observar que relações semânticas, dentro do conceito linguístico, depende primariamente da experiência de um indivíduo. Vilares et al. [2018] também apontam para esse fenômeno

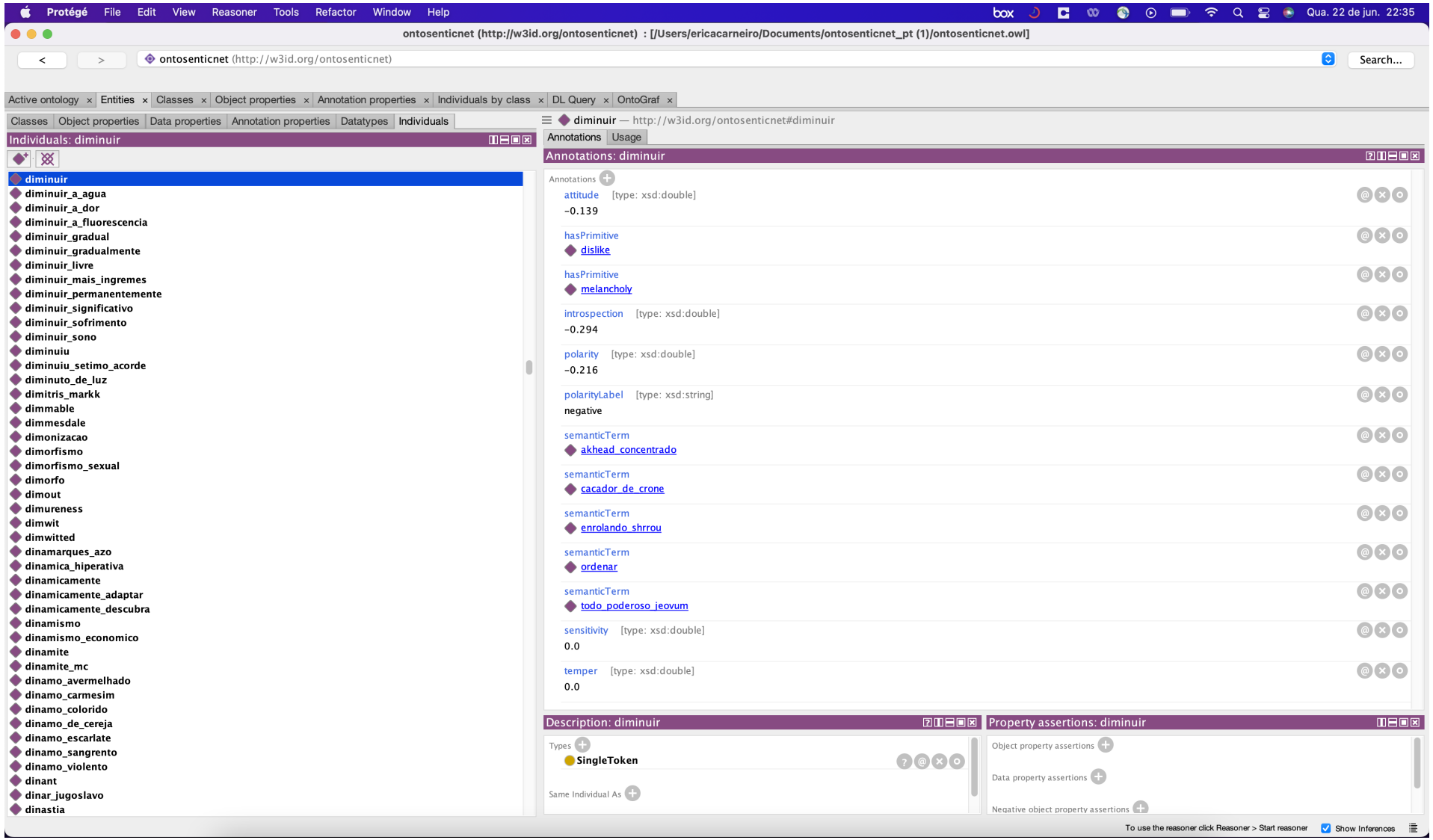


Figura 2.4: Reprodução dos conceitos e atributos semânticos da OntoSenticNet em Português

que poderia causar ruído nas anotações humanas, já que a SenticNet também conta com relações e conceitos sobre áreas técnicas muito específicas que poderiam não ser do domínio dos anotadores. No que concerne à acurácia das polaridades da análise de sentimento, os anotadores humanos apontaram que 64,4% das polaridades atribuídas estavam corretas [Vilares et al., 2018].

2.6 Métricas de avaliação de modelos

Algumas métricas de avaliação são essenciais para verificação do desempenho de modelos e ferramentas de AM e IA [Arendt et al., 2021]. Além disso, essas métricas permitem inferências que auxiliam na tomada de decisão do melhor modelo a ser utilizado em cada caso [Han et al., 2011]. A seguir, apresentamos quatro termos fundamentais para a compreensão da avaliação de modelos, incluindo os apresentados neste trabalho. As definições abaixo são inspiradas nos estudos de Han et al. [2011].

Ressaltamos que as métricas aqui apresentadas são utilizadas em modelos de classificação. Modelos de Regressão possuem outras métricas que não serão abordadas neste trabalho.

- Verdadeiros Positivos (VP): são valores previstos pelo modelo e rotulados no conjunto de dados de treino como verdadeiros. *E.g*: no caso de uma tarefa de análise de sentimentos, são valores anotados como 1 (positivo) no conjunto original e classificados como 1 pelo modelo;

- Verdadeiros Negativos (VN): assim como os Verdadeiros Positivos referem-se aos valores corretamente classificados pelo modelo. Os Verdadeiros Negativos representam os valores que o modelo acertou anotados como Negativos. No caso de uma análise de sentimento são as classificações -1 (negativas), rotuladas como negativas tanto no conjunto de dados anotado, quanto no classificado pelo modelo;

- Falsos Positivos (FP): são as classificações positivas (1) geradas pelo modelo incorretamente. Ou seja, foram anotadas originalmente como negativas (-1), mas classificadas como positivas (1) pelo modelo;

- Falsos Negativos (FN): trata-se dos valores negativos erroneamente classificados pelo modelo. Originalmente, foram anotados como positivos (1), porém classificados como negativos (-1) pelo modelo.

Além das categorias relativas às classes positivas e negativas supracitadas, modelos de predição também possuem métricas de avaliação [Olson and Delen, 2008]. Tais valores permitem observar se um modelo tem um desempenho próximo da excelência. Com eles também é possível avaliar se o modelo de AM sofre com viés ou sobre-ajuste, acertando bem os dados de treino, mas com péssimo desempenho para os conjuntos de teste; ou sub-ajuste, *i.e.*, não aprendeu suficientemente com os

dados de treino. Assim, as principais métricas para avaliação de modelos de predição são: Precisão, Revocação, *F1-Score* e Acurácia, descritas a seguir.

A Precisão é uma medida utilizada para avaliar o desempenho do modelo em relação aos Verdadeiros Positivos (VP). A medida leva em conta os acertos como verdadeiros positivos e os erros como falsos positivos [Han et al., 2011]. O cálculo matemático para Precisão (P) é representada pela equação 2.2:

$$P = \frac{VP}{VP + FP} \tag{2.2}$$

A Revocação reflete a taxa de Verdadeiros Positivos (VP) ou valor preditivo positivo. Logo, trata-se da medida que reflete a probabilidade de acerto de classes positivas por um modelo [Han et al., 2011]. A Revocação (R) pode é representada pela seguinte equação 2.3.

$$R = \frac{VP}{VP + FN} \tag{2.3}$$

A média uniformemente ponderada entre Precisão (P) e Revocação (R) é obtida pela medida de *F1-Score*, também denominada F1 [Han et al., 2011]. Trata-se da média aproximada de P e R, ou ainda, pode ser considerada como a divisão da média geométrica pela média aritmética. O cálculo matemático para *F1-Score* pode ser obtido pela equação 2.4:

$$F1 = \frac{2 * P * R}{P + R} = \frac{2 * VP}{2 * VP + FP + FN} \tag{2.4}$$

A Acurácia (ACC) é a medida que permite a avaliação de quanto o modelo acertou das previsões possíveis entre erros e acertos [Han et al., 2011], ou seja, dentre os possíveis resultados de verdadeiros e falsos Positivos qual é a taxa de acerto do modelo. A representação matemática para Acurácia (ACC) é representada pela equação 2.5.

$$ACC = \frac{VP + VN}{VP + VN + FP + FN} \tag{2.5}$$

Capítulo 3 Trabalhos Relacionados

Conforme observado nas pesquisas realizadas para este texto, a escassez de estudos no campo semântico da língua portuguesa brasileira é uma lacuna presente na área de computação afetiva, uma vez que a maioria das pesquisas tem como base análises de viés morfossintático [da Silva Conrado et al., 2014; Pereira, 2020]. Ressalta-se que já existem ontologias eficientes em outros idiomas e, sendo o Brasil o sexto país mais populoso do mundo [IBGE, 2020], seria importante avaliar a necessidade do desenvolvimento de pesquisas mais profundas que prestigiem o idioma nacional.

Neste capítulo apresentamos estudos na área de Ciência da Computação com foco em métodos que utili zaram ontologias para AS. A investigação teve seu início com buscas nos portais de periódicos Scopus e Scielo, limitadas ao período 2016-2022, a fim de encontrar metodologias e técnicas mais recentes. Foram encontrados apenas três artigos utilizando a seguinte *string* de busca: *((“sentiment analysis”) AND (ontology) AND (Portuguese))*. Posteriormente, expandimos nossa pesquisa para *((“sentiment analysis”) AND (ontology))*, levando-nos a diferentes soluções que poderiam ser aplicadas no cenário português-brasileiro.

Nossas pesquisas foram divididas em três etapas: i) prospecção de abordagens ontológicas de AS; ii)] comparação dos principais métodos de AS; iii) técnicas de desenvolvimento de *corpus*.

A primeira etapa da pesquisa bibliográfica visou fundamentar a importância de estudos com base em ontologias. Ademais, também foi responsável por direcionar o desenvolvimento de diferentes aspectos deste trabalho como, por exemplo, a escolha do *framework* SenticNet e a construção de *corpora* paralelos. Uma segunda etapa de revisão da literatura se dedicou a uma comparação de diferentes estruturas de dados aplicáveis à AS, em especial, os métodos baseados em léxicos. Por fim, durante o desenvolvimento dos experimentos computacionais, as pesquisas bibliográficas focaram em artigos que tratassem sobre traduções automáticas e construções de *corpora* paralelos.

Uma revisão detalhada dos resultados das execuções das *strings* de busca foi feita de forma a selecionar os artigos mais relacionados ao tema dessa dissertação. Tais referências serviram de base para as metodologias de *backward* e *forward snowballing*. A etapa de descoberta inicial nos levou a diferentes métodos com abordagens ontológicas, cuja maioria não era de domínio aberto, ou seja, lidavam com algum tema específico. A seleção de estudos nesta primeira etapa de nossa pesquisa teve como a base os seguintes critérios: (a) abordagens baseadas em ontologia para AS; (b) idioma dos *corpora*; (c) a possibilidade de adaptação de domínio, especialmente em trabalhos

que utilizaram bases de dados em inglês.

Neste trabalho, maior foco foi dado a estudos que utilizaram ontologias em seus métodos de AS. Espera-se que este levantamento possa auxiliar no desenvolvimento potencial de novas estratégias e abordagens no cenário da língua portuguesa brasileira. Não foram abordadas todas as possíveis ontologias de domínio devido à grande variedade de assuntos. Todavia, é possível encontrar alguns métodos de classificação de sentimentos e MO baseados em ontologias de domínios diferentes e específicos.

É importante ressaltar que, apesar de muitos dos métodos apresentados utilizarem AM e mesmo Aprendizado Profundo, este não é o foco deste trabalho. O objetivo foi analisar, por meio da bibliografia, as impressões de Cambria et al. [2017], que afirmam não ser a AS apenas um problema de classificação estatística. Também é importante observar as impressões de da Silva Conrado et al. [2014], que notaram a falta de ontologias no estado da arte em Ciência da Computação para a língua portuguesa.

Em contraste com outros idiomas, as abordagens baseadas em ontologia para AS usando o *corpus* em inglês são predominantes. Porém alguns deles podem influenciar novos trabalhos em diferentes idiomas, desta forma, também foram incluídas algumas metodologias para AS que trabalham com conjuntos de dados em inglês.

3.1 Abordagens ontológicas de análise de sentimentos de domínio fechado ou aberto em diferentes idiomas

Como mencionado anteriormente, a etapa inicial da revisão bibliográfica objetivou a busca de abordagens de AS com aspectos ontológicos. Notadamente, foi encontrada apenas uma abordagem que permitisse, também, testes com conjuntos de dados na língua portuguesa. Tal escassez de métodos para AS baseados em ontologia em línguas que não a inglesa já foi discutida anteriormente na literatura por Dragoni et al. [2018].

A seguir são destacados alguns desses estudos de modo a apresentar ao leitor algumas possíveis abordagens ontológicas para análise de sentimentos. Espera-se que este levantamento possa auxiliar no desenvolvimento potencial de novas estratégias e abordagens no cenário da língua portuguesa brasileira. Posteriormente, são discutidos os trabalhos relacionados que inspiraram o desenvolvimento das análises presentes nesta pesquisa

Kontopoulos et al. [2013] desenvolveram um protótipo de uma ontologia de AS com foco na descoberta de conceitos. O *corpus* era composto por *tweets* em inglês, abrangendo o domínio fechado de *smartphones* e seus principais recursos. A ontologia fez uso de métodos e técnicas de aprendizado de máquina. O experimento dos autores avaliou as métricas de revocação das duas arquiteturas desenvolvidas em seus estudos: a) um sistema semanticamente habilitado baseado em ontologia

completa; e, b) o mesmo sistema sem o aumento de sinônimo/hipônimo. Segundo os autores, que utilizaram a Revocação como métrica de avaliação, o método (a) atingiu $0,94$ enquanto o método (b) apenas $0,26$ em uma de suas amostras. Em todas as demais amostragens de seu experimento, o método baseado em ontologia superou o sem.

A *Safe Traveling Ontology* de Ali et al. [2017] é uma ontologia de domínio específico sobre aspectos do trânsito. Os autores desenvolveram uma abordagem de ontologia *fuzzy* que inferisse a polaridade da qualidade de transporte e popularidade de locais em uma cidade a partir de *tweets* e *reviews*. de Kok et al. [2018] desenvolveram o protótipo de uma ontologia cujo domínio aborda *reviews* de restaurantes. A polaridade dos sentimentos era classificada por meio de um léxico afetivo. Os autores demonstraram que as métricas de precisão, revocação e acurácia atingiram 98%, 76%, 86% respectivamente para métodos com *fuzzy ontology* e 66%, 77%, 66% para ontologia clássica.

Ceci et al. [2016], por sua vez, desenvolveram uma ontologia de domínio sobre câmeras e equipamentos de fotografia com base em revisões de textos no *site* da Amazon. Os autores utilizaram a *Movie Ontology* e métodos de AS em *reviews* de produtos com *Sentiment Ontology Tree (SOT)*. Segundo os autores, os resultados do modelo com base em ontologia superou tanto as classificações com método Naïve Bayes quanto com *Support Vector Machine* (SVM). Os resultados obtidos para Acurácia foram respectivamente $0,85$, $0,787$ e $0,786$ no domínio de filmes; e $0,91$, $0,845$ e $0,845$ para o domínio de câmeras.

A *Visual Sentiment Ontology (VSO)* [Borth et al., 2014, 2013] tem como base a roda de emoções de Plutchik [Plutchik, 1980]. Ao contrário das demais ontologias presentes nesta pesquisa, a ferramenta tem como domínios imagens e vídeos das redes sociais Flickr e YouTube [Borth et al., 2013].

A MFOEM [The National Center for Biomedical Ontology, 2022; Hastings et al., 2011] de Smith [2018] é baseada na BFO [Smith, 2019] e suas terminologias foram definidas conforme os termos presentes na *Mental Disease Ontology* (MFO) (também denominada *Ontolology of Mental Disease*). Juntas, elas possuem informações e relações entre diferentes tipos de dados que incluem: neuroimagem, estudos farmacêuticos, monitoramento de comportamento, questionários, *etc.* Apesar de sua imensa contribuição científica, a MFOEM não classifica sentimentos automaticamente a partir de um texto no nível das frases.

Os *frameworks* baseados em ontologia em inglês também podem ser testados ou até treinados para outros idiomas. O Senpy é um *framework* em Python que analisa sentimentos com uma arquitetura de *plugins*. Ele oferece uma abordagem de *web* semântica usando vocabulários semânticos nos formatos NIF, Marl, Onyx, Turtle, JSON-LD e XML-RDF [Sánchez-Rada et al., 2016]. O Senpy contém bases de conhecimento e léxicos, incluindo VADER, SenticNet e Wordnet Affect.

Além da língua inglesa, merecem destaque os estudos em países asiáticos de Liu et al. [2012]

e também de Jung et al. [2017]. Liu et al. [2012] propuseram um novo mecanismo de AS baseado em um algoritmo de árvore ontológica com domínio *fuzzy*. Com ele, seria possível construir automaticamente uma árvore ontológica de domínio pautado nas revisões de produtos. O sistema incluiria a extração de palavras de sentimento, características do produto e as relações entre tais características. O *corpus* utilizado no teste consistiu de 2.400 análises em chinês de produtos para *laptops* com um comentário positivo e um negativo para cada revisão. Segundo o autor o método com ontologia *fuzzy* obteve $0,673$, $0,724$ e $0,629$ para F1-Score, precisão e revocação. Já a ontologia sem método *fuzzy* chegou a $0,534$, $0,584$ e $0,492$.

Com a crescente onda de depressão e suicídio entre adolescentes sul-coreanos, Jung et al. [2017] desenvolveram uma ontologia com terminologias comumente presentes em textos sobre depressão. Tais terminologias são sinônimos de conceitos que podem ajudar a integrar várias formas de linguagem natural. O grande diferencial da ontologia de Jung et al. [2017] é a construção de seu domínio e escopo, que foram desenvolvidos utilizando questões de competência. Ao todo, a ontologia de depressão em adolescentes dos autores contém 443 classes, com 60 relações entre si, além de uma terminologia com 1.682 sinônimos presentes nas 443 classes. Segundo os autores, acompanhar a rápida evolução linguística dos adolescentes em redes sociais é uma tarefa altamente complexa [Jung et al., 2017]. As classes presentes na ontologia estão em inglês, porém seu *corpus* é uma mescla do idioma sul-coreano com inglês. Com base nas questões de competência, a taxa geral de mapeamento na ontologia foi de $88,7\%$.

Na América Latina, encontramos alguns estudos desenvolvidos para o espanhol e o português, dos quais merecem destaque as pesquisas de García-Díaz et al. [2020], que desenvolveram um método de AS baseado em aspectos ontológicos com objetivo de classificar a polaridade dos sentimentos evocados pelas epidemias de Dengue, Zika e Chikungunya na América Latina. O *corpus*, em espanhol latino-americano, foi compilado do Twitter e posteriormente rotulado por voluntários como positivo, negativo e neutro. Os autores utilizaram diversos modelos de aprendizado profundo para a classificação de sentimentos. A melhor precisão média foi de $55,3\%$, obtida pela combinação de recursos linguísticos com *Perceptron Multilayer*, seguida pelo uso do algoritmo *Bidirectional LSTM* (BiLSTM) combinado ao modelo de recursos linguísticos, cuja precisão média foi de $54,2\%$.

Freitas and Vieira [2013] desenvolveram uma ferramenta para identificação de *features* no domínio hoteleiro. Os autores também observaram a falta de recursos com suporte para a língua portuguesa. A HOntology possui 282 conceitos categorizados em 16 conceitos de nível superior. Já Rosa et al. [2019] criaram a *Knowledge-Based Recommendation System (KBRS)*, um sistema de recomendação baseado em conhecimento em um aplicativo de monitoramento de saúde emocional para detectar depressão e estresse em usuários. Dependendo dos resultados do monitoramento, o KBRS, baseado em ontologias e AS, é ativado para enviar mensagens motivacionais [Rosa et al.,

2019]. Redes neurais recorrentes são responsáveis por identificar frases com contextos depressivos e estressantes. De acordo com Rosa et al. [2019], resultados experimentais sugeriram que o KBRS atinge uma classificação de 94% de usuários satisfeitos em comparação com 69% sem o uso de métrica de sentimento e ontologias. Além disso, o método também alcançou uma precisão de $0,89$ ao detectar usuários deprimidos e $0,90$ para estressados.

Apesar de alguns resultados promissores, conforme observado nas métricas de avaliação, a revisão aqui apresentada indica que as abordagens baseadas em ontologia não são a principal opção para classificação de sentimentos. Conforme declarado no Capítulo 2, da Silva Conrado et al. [2014] já observaram que a maioria das contribuições científicas para PLN no Brasil são léxicos, que tratam a AS como um problema de classificação de sentimentos positivos, neutros e negativos [Cambria et al., 2017], independentemente das relações semânticas presentes no *corpus*.

Na Tabela 3.1, encontra-se um resumo das abordagens ontológicas para AS levantadas neste trabalho. Nela, estão dispostas algumas informações relevantes de cada ontologia como, por exemplo: a relação com outra ontologia de domínio; quantidade de classes, conceitos e sinônimos; idioma do *corpus*; e abordagem aplicada. Destaca-se, aqui, a SenticNet devido à grande quantidade de sinônimos presentes e aspecto multilíngue.

Tabela 3.1: Abordagens ontológicas levantadas em trabalhos de análise de sentimentos.

Ontologia	**Ontologia Relacionada**	**Tradução Automática**	**Domínio**	**Quantidade de Classes e Conceitos**	**Quantidade de Sinônimos**	**Idioma da Ontologia**	**Idioma do Corpus**	**Abordagem**
Ontologia Infodemiológica Latino Americana	*Disease Ontology* (DO) e *Infectious Disease Ontology* (IDO)	Não	Doenças e Infectologia	12	Múltiplos	Inglês	Espanhol da América Latina	Estatística
SenticNet e Ontosenticnet	N/A	Sim	Sentimentos	32	400000	Inglês	Inglês e Multilíngue	Sentic Computing, Ampulheta de Sentimentos e Onto-pares
Geração Automática de Ontologia de Domínio	N/A	Não	Smartphones. (Escolhido para extração. O domínio é gradualmente extraído.)	3	Não informado	Inglês	Inglês	Formal Concept Analysis
Visual Sentiment Ontology (VSO)	N/A	Não	Imagens e Vídeos	1200	1200	Inglês	N/A	LibSVM
Safe Traveling Ontology	N/A	Não	Trânsito	29	N/A	Inglês	Inglês	Comparação entre ontologia clássica e fuzzy
Ontologia de Equipamentos Fotográficos (Rating)	*Movie Ontology*	N/A	Equipamentos fotográficos	N/A	N/A	Inglês	Inglês	Raciocínio Baseado em Casos (RBC) e SOT
Restaurant Ontology	N/A	Não	Restaurantes	Múltiplas	N/A	Inglês	Inglês	*High Hierarchical Learning - Sentiment Ontology Tree* (HL-SOT)
Ontologia para Detecção de Depressão Adolescente	BFO	Não	Depressão Adolescente	443	1682	Inglês	Inglês e Sul Coreano	Modelagem de dados em trios *Entity - Attribute - Value* (EAV)
HOntology	N/A	Não	Hotelaria	274	Não informado	Inglês	Multilíngue (Português, Espanhol e Francês)	Web Semântica Multilíngue
EOSentiMiner	HowNet	Não	Sentimentos	8	Não informado	N/A	Chinês	Fuzzy
Ontologia para Detecção de Emoções no COVID-19	MFOEM	Não	Covid	N/A	N/A	Inglês	Inglês	*Formal Concept Analysis* (FCA) e Aprendizado de Ontologias
ALDONAr - Ontologia de Domínio Lexicalizada	N/A	Não	Sentimentos	2	N/A	Inglês	Inglês	Neural Attention model e SVM
LexiPers	*FarsNet ontology*	Não	Sentimentos	N/A	N/A	N/A	Persian	ExpandSimple e K-nearest
KBRS	Nuadu	Não	Depressão e Estresse	6	Não informado	Inglês	Português	Sistema de Recomendação de Ontologias, Deep Learning e Rede Neural

3.2 *Benchmark* de métodos de análise de sentimentos

Cruz et al. [2017] relacionaram os léxicos afetivos com foco na língua portuguesa brasileira. Segundo os autores, os mesmos são recursos úteis para a melhoria nos processos de ensino e aprendizagem. Seu trabalho conta, ainda, com uma revisão sistemática dos recursos e suas aplicações com foco em AS. É importante destacar que os estudos apresentados pelos autores objetivaram relacionar os léxicos afetivos já criados. O artigo traz escopo ao desenvolvimento deste trabalho, uma vez que em comparação ao estado da arte internacional, as pesquisas em computação afetiva ainda são escassas no Brasil, evidenciando a necessidade de atenção à elaboração de maiores estudos no campo semântico.

Nascimento [2021] realizou um *benchmarking* de 13 léxicos desenvolvidos para a língua portuguesa. O trabalho comparou o desempenho de diversos léxicos com suporte para o português. Foram analisados quatro conjuntos de dados na língua portuguesa. Seu trabalho foi base para as comparações, elaboradas nesta dissertação, entre métodos já consolidados na literatura e a SenticNet, com sua abordagem ontológica e *sentic computing*.

O autor testou as métricas de F1-*Score*, acurácia, precisão e revocação em todos os conjuntos de dados. Seus testes demonstraram que os léxicos AffectPT-br, LIWC2015-PT-br e SentiStrength obtiveram os melhores resultados. Os léxicos testados com suporte para o português foram:

- AffectPT-br [Carvalho et al., 2018]: desenvolvido a partir da categoria de afeto (*"affect"*) da versão original de 2015 (em inglês) do LIWC, contendo 1.139 palavras, sendo 479 palavras de emoções positivas (*"positive emotion"*) e 661 de emoções negativas *"negative emotion"*;

- ANEW-Br [Kristensen et al., 2011]: traduzido diretamente do original em inglês e posteriormente anotado por juízes humanos, é composto por 1.046 palavras;

- LIWC2007-PT-br [Filho et al., 2013]: trata-se de um léxico afetivo na língua portuguesa traduzido da versão de 2007 do LIWC. Possui 127.161 palavras e 64 categorias;

- LIWC2015-PT-br [Carvalho et al., 2019, 2023]: trata-se da versão mais atual do LIWC em português brasileiro, desenvolvida a partir da versão original de 2015 (em inglês) do LIWC. Contém 14.459 palavras e 73 categorias e subcategorias;

- MPQA-PT [Almeida et al., 2015]: tradução do léxico afetivo em inglês *Multi-Perspective Question Answering* ou MPQA. É composto por um *corpus* manualmente anotado por três linguistas e possui 80 tópicos de notícias extraídos de jornais, revistas e sites. Cada um desses tópicos possui 10 documentos;

- OpLexicon [Souza et al., 2011b]: tradução para o português via Google Translate do Opinion Lexicon [Hu and Liu, 2004]. Possui 44.077 palavras anotadas;
- Polyglot [Chen and Skiena, 2014]: composto por 3.953 palavras, trata-se de um léxico afetivo com tradução automática para mais de 136 línguas, incluindo o português. As traduções foram efetuadas pelo Google Tradutor;
- Reli-Lex [Freitas, 2013]: possui 609 palavras e expressões em uma única categoria (valência), na língua portuguesa. O *corpus* do léxico afetivo foi criado com resenhas de livros publicados na Internet;
- Sentilex-PT [Carvalho and Silva, 2015]: com 7.014 palavras e expressões, tem como objetivo analisar o sentimento evocado por uma determinada figura pública, ou seja, entidades humanas;
- SentiStrength [Thelwall, 2017]: criado em 2010 com 2.212 palavras, está disponível em 15 idiomas, incluindo o português;
- WordnetAffectBR [Pasqualotti and Vieira, 2008]: trata-se da versão brasileira em português do WordNetAffect [Miller, 1995]. O léxico afetivo é composto por 289 palavras e expressões em uma única categoria (valência).
- WordnetAffectBR(adapt) [Longhi, 2011]: trata-se da atualização e aprimoração do WordnetAffectBR [Pasqualotti and Vieira, 2008]. Possui 1.023 palavras e expressões.

Musto et al. [2014] desenvolveram uma abordagem explorando léxicos e bases de conhecimento de domínio aberto, tais como SentiWordNet, WordNet-Affect, MPQA e SenticNet. Os autores utilizaram dois conjuntos de dados em inglês, o *SemEval-Test 2013* e o *STS-Test*, a fim de avaliar a métrica de acurácia de seu modelo. Suas pesquisas trataram os conjuntos com quatro estruturas diferentes: básico, normalizado, enfatizado e normal-enfatizado.

Segundo as pesquisas dos autores, o SentiWordNet possuía o maior vocabulário, seguido pela SenticNet. Após as comparações, MPQA e SentiWordNet obtiveram os melhores resultados de acurácia no conjunto de dados *SemEval-Test 2013*, ambas as ferramentas obtiveram em torno de 58% de acurácia.

O conjunto de dados *STS-Test* teve melhor desempenho de acurácia com a utilização da SenticNet, que chegou a $74,65\%$ de acurácia, seguida pelo SentiWordNet, com $71,59\%$, ambas com os dados enfatizados. Suas pesquisas não incluíram as demais métricas de precisão, revocação e F1-*Score*.

Desde 2012, Cambria and Hussain [2012] vêm aprimorando sua metodologia, a *sentic computing* e o *framework* SenticNet. Diversas atualizações da ferramenta foram desenvolvidas após sua versão

inicial. O experimento realizado neste trabalho focou principalmente nas versões 6 e 7. Tais versões utilizaram uma série de conjuntos de dados em inglês amplamente populares, aplicando-os em uma série de testes [Cambria et al., 2020, 2022]. Os resultados foram comparados com outros léxicos de sentimentos, conforme apresentado na 3.2.

Tabela 3.2: Bases de conhecimento e léxicos, ano da versão utilizada nos testes das versões 6 e 7 da SenticNet em Cambria et al. [2020] e Cambria et al. [2022]

Léxico/Base de Conhecimento	Ano	Termos	SenticNet 6	SenticNet 7
LIWC	1993	2290		x
ANEW	1999	1030	x	x
WordNet-Affect	2004	4787	x	x
Opinion Lexicon	2004	6789	x	x
Opinion Finder	2005	8221	x	x
Micro WNOp	2005	5636	x	x
Sentiment140	2009	62466	x	x
SentiStrength	2010	2546	x	x
SentiWordNet	2010	23089	x	x
General Inquirer	2011 e 1966	8639	x	x
AFINN	2011	2477	x	x
EmoLex	2013	5636	x	x
NRC HS Lexicon	2014	54128	x	x
VADER	2014	7503	x	x
MPQA	2015	8222	x	x
SenticNet 5	2018	100000	x	
SenticNet 6	2020	100000	x	
SO-CAL	2011	N/A		x
Novad	2013	13915		x
Senti Words	2016	155000		x
HSSWE	2017	N/A		x
Lingmotif-Lex	2018	67400		

3.3 Trabalhos Relacionados a Métodos de Desenvolvimento de Corpora Paralelos

da Silva Conrado et al. [2014] apresentaram as lacunas presentes em trabalhos relacionados à extração automática de termos na língua portuguesa no Brasil. Segundo os autores, são poucas as pesquisas na área que contemplam o idioma. Além disso, a grande maioria desses estudos permite apenas a análise morfossintática das palavras e conceitos, porém não há uma ferramenta capaz de contribuir verdadeiramente no âmbito semântico nacional. Para os autores, existe uma grande discrepância entre o estado da arte do português brasileiro e outros idiomas, como o inglês, o espanhol e o francês. Ainda segundo os autores, essa lacuna deveria ser preenchida e esforços seriam necessários a fim de encontrar novas propostas e soluções.

Os autores mostram como as diferentes abordagens, estatística, linguística e híbrida com aprendizado de máquina podem contribuir para a melhoria do cenário dessas pesquisas no Brasil e para a solução de problemas relacionados ao uso de cada uma delas. da Silva Conrado et al. [2014] expli-

cam os riscos do uso de *corpus* que se baseiam na tradução automática de documentos, preterindo a estrutura idiomática original das línguas. Esta observação apontaria a necessidade do desenvolvimento de pesquisas na área de Computação que englobem o campo semântico da língua portuguesa no Brasil.

Em Rossetti and Gaspari [2017], os autores descrevem um método de tradução realizado com seis estagiários tradutores. O experimento tem como objetivo propor um modelo de análise do impacto do uso da memória de tradução e pós-edição completa em traduções automáticas brutas tanto no nível de dificuldade percebida quanto no tempo gasto pelos tradutores estagiários. Desta forma, os seis estudantes italianos do curso de mestrado em tradução deveriam produzir textos-alvo de alta qualidade ao traduzir material semi-técnico do inglês para o italiano nativo.

Os estudos de Jiang et al. [2022] descrevem a criação do BWB, um *corpus* bilíngue do chinês para o inglês a partir de romances. Em seu trabalho, os autores consultaram tradutores especialistas em inglês que auxiliaram na criação de um conjunto de teste anotado. O referido *dataset* foi, então, projetado para sondar a capacidade de modelagem de fenômenos discursivos em sistemas de tradução automática.

O processo reverso de anotação de *corpus* paralelo pode ser encontrado em Han et al. [2020]. Em seu estudo, os autores primeiro realizam a tradução automática do *corpus*, seguido pela de pós-edição humana e anotações manuais de expressões com várias palavras (em inglês, *multiword expressions* ou MWEs para os autores) de destino. Segundo os pesquisadores, a tradução precisa dessas expressões representa uma difícil tarefa para ferramentas de tradução ou *machine translations*. Portanto, consiste em um desafio para os sistemas de MT. Assim, além das anotações humanas, os autores utilizaram quatro modelos populares de tradução automática: Microsoft Bing Translator, GoogleMT, Baidu Fanyi e DeepLMT. Seus estudos criaram o AlphaMWE, que são *corpora* paralelos com foco em expressões com várias palavras nos idiomas inglês, chinês, polonês e alemão.

3.4 Considerações sobre os diferentes métodos e abordagens para análise de sentimentos

Nesta seção, são levantadas discussões sobre os diferentes métodos e abordagens apresentados anteriormente. Conforme observado por Gruber [1993], muitas são as contribuições da aplicação de ontologias na área de IA e Ciências da Informação. Desta forma, a revisão de abordagens baseadas em ontologia para AS aqui dispostos, poderia trazer sugestões melhorias, bem como provocar novas discussões e pesquisas no cenário brasileiro de Computação Afetiva.

No caso de García-Díaz et al. [2020], talvez uma solução com aprendizado profundo pudesse realmente contribuir para o trabalho dos autores, principalmente ao executar tarefas complexas de *web* semântica, como a adição de relações em grafos de conceitos e eliminação de ruídos [Hitzler,

2021]. Além disso, os autores realizaram um incrível e árduo trabalho traduzindo todos os termos do domínio, contribuindo profundamente tanto para o campo de PLN, quanto de bio-informática e ontologias. Essa contribuição poderia, no futuro, se tornar uma nova DO ou IDO, com foco no espanhol da América Latina.

Os autores efetuaram diversos testes com diferentes algoritmos de aprendizado profundo, porém, segundo os mesmos, um teste com engenharia reversa aplicado à linguística computacional poderia ser executado. Desta forma, *i.e.*, ao invés de traduzir os conceitos e classes da ontologia, traduzir o *corpus* para o inglês, ainda que utilizando o *Google Translate*, e testá-los nas ontologias originais com os algoritmos propostos de classificação, a fim de avaliar se ocorreria uma melhora no desempenho dos modelos [García-Díaz et al., 2020].

Os testes realizados por Liu et al. [2012] mostraram que uma MO com base em uma árvore ontológica de sentimento sem conjuntos *fuzzy* usados como mapa conceitual, tiveram menor desempenho nas métricas de *precision*, *recall* e *F-measure*. Isso significa que o desempenho da AS não é altamente eficaz quando depende apenas da árvore de ontologia de sentimentos. Os pesquisadores observaram, ainda, que consumidores sempre comentam o produto e, em seguida, apontam vantagens e desvantagens de seus atributos. Adjetivos padrão dos recursos do produto, como "custo-benefício", são frequentemente usados em opiniões de consumidores. Sendo assim, a identificação de padrões de relações na estrutura da linguagem pode auxiliar na construção de relações semânticas que poderiam ser aplicadas a modelos de treino de AM.

Segundo Liu et al. [2012] também foi possível identificar atributos não tão comuns presentes na extração da *Fuzzy Sentiment Ontology Tree* (FDSOT). Os autores argumentam que as funcionalidades presentes no topo da árvore são palavras de sentimento correspondentes, amplas e diversas. Entretanto, a parte inferior da ontologia possui termos de sentimento correspondentes simples e específicos. Os autores também notaram que, as palavras com polaridade de sentimento, são, principalmente, substantivos, verbos, adjetivos e advérbios. Assim, a remoção de *stopwords*, como artigos e preposições pode auxiliar bastante tanto para análise quanto para o processamento, já que elimina ruídos desnecessários.

Apesar de eficaz, o método de Liu et al. [2012] aparenta ser excessivamente custoso e uma extração automática de entidades poderia facilitar o processo de construção da ontologia. Em um cenário de mercado, a informação de que uma determinada funcionalidade de um produto possui atributo negativo permite análises rápidas e imediatas que poderiam culminar em tomadas de decisão cruciais para uma marca [Liu et al., 2012].

O sistema criado por Kontopoulos et al. [2013] coleta automaticamente *tweets* e os envia para a ferramenta Open Dover [OpenDover, 2022], que, automaticamente, executa a tarefa de classificação de sentimentos. Os valores de sentimento resultantes para cada par de atributo-objeto são, então,

armazenados.

Segundo Kontopoulos et al. [2013], tanto a ontologia criada via *FCA* quanto o método com *Ontology Learning (OL)* seriam em essência uma taxonomia simples de conceitos e atributos. Porém, para aumentar a semântica, os autores precisaram enriquecer a ontologia com sinônimos e hipônimos dos atributos identificados. Jung et al. [2017] também mencionam a importância da criação de sinônimos para o sucesso de um projeto ontológico voltado para a AS.

É interessante observar como o método proposto por Kontopoulos et al. [2013] é capaz de identificar e criar domínios e classes automaticamente. Os autores se preocuparam com as diferentes abordagens possíveis para a criação de ontologias, delegando o problema de classificação de sentimentos a uma ferramenta paga, já existente, simplificando essa etapa. Desta forma, os pesquisadores focaram nas ontologias como arquitetura de recuperação de informação devido à escassez de abordagens que utilizam esse modelo computacional.

Ao analisar problemas de léxicos consolidados, Hutto and Gilbert [2014], apontaram sua dificuldade em analisar palavras de uso diário, como gírias. O mesmo problema foi levantado por Jung et al. [2017] com sua Ontologia Coreana de Depressão para adolescentes. Ou seja, não há máquina capaz de acompanhar rápido o suficiente as céleres evoluções linguísticas do cotidiano na era da *Web* 2.0.

Apesar de sua significativa contribuição científica, a MFOEM de Smith [2019] não classifica sentimentos no nível da frase. Entretanto, ela oferece um rico vocabulário emocional aplicável a abordagens de MT para detectar menções de emoções específicas. Segundo Dragoni et al. [2018], combinar a MFOEM com a SenticNet seria altamente vantajoso. A OntoSenticNet, focada no domínio emocional, poderia usar a MFOEM como uma *Upper Domain Ontology*, alinhando conceitos com a SenticNet. Essa integração criaria uma ontologia abrangente, unindo as categorizações da SenticNet com as propriedades emocionais do mundo real, como documentos, vídeos e imagens, fornecidas pela MFOEM.

Podemos observar que o aprendizado de ontologias combinado a técnicas de AM para aquisição de conhecimento são fundamentais no desenvolvimento de pesquisas em *web* semântica [Liu et al., 2012]. Em complemento, o uso de *kits* de ferramentas de PLN e uma base de dados léxica para análise das *synsets*, *i.e.*, uma base de sinônimos pré-existente, como o Wordnet, são metodologias que deveriam ser consideradas no desenvolvimento de uma ferramenta de AS com estrutura ontológica [Saranya and Jayanthy, 2017]. O trabalho de Niles and Pease [2003] e seu mapeamento da WordNet também poderiam servir como guias para a construção e mapeamento da WordNetAffectBR em OWL. Assim, as relações presentes no *corpus* também poderiam ser utilizadas em análises com abordagens semânticas.

Capítulo 4 Metodologia

Neste capítulo, detalhamos a metodologia aplicada em nossa pesquisa, cujo objetivo é desenvolver e validar métodos robustos para AS e PLN no contexto brasileiro, com foco na língua portuguesa. A pesquisa é fundamentada nas premissas estabelecidas por Cambria and White [2014] e também por da Silva Conrado et al. [2014], que destacam a necessidade de maior robustez nas técnicas de AS e PLN. Para alcançar nossos objetivos, utilizamos uma abordagem multifacetada que inclui a revisão bibliográfica, seleção criteriosa de ferramentas e métodos, implementação de testes em Python, e avaliação dos impactos de diferentes métodos de tradução automática. Em particular, investigamos o desempenho da SenticNet 6 e a influência das estruturas linguísticas do português e do inglês em tarefas de AS. Este capítulo descreve, de forma detalhada, cada etapa do processo, desde a concepção do *corpus* paralelo até a análise dos resultados, proporcionando uma visão abrangente das técnicas e procedimentos empregados para validar a eficácia das metodologias propostas.

O objetivo deste estudo éra encontrar um método aplicável a testes de AS que validassem as premissas de Cambria and White [2014] e da Silva Conrado et al. [2014], como a necessidade de maior robustez no desenvolvimento de métodos para tarefas de PLN [Cambria and White, 2014]. A pesquisa também visa comprovar a necessidade de abordagens mais robustas no contexto brasileiro, com foco na língua portuguesa, além de avaliar a influência de métodos de tradução automática para a AS [da Silva Conrado et al., 2014]. Para alcançar esses objetivos, foi realizada uma seleção criteriosa das ferramentas a serem analisadas, com destaque para a escolha da SenticNet e o desenvolvimento dos testes subsequentes. Este capítulo detalha os critérios e justificativas para a seleção de métodos e conjuntos de dados, bem como a implementação dos testes em Python. Também são descritos os métodos de análise utilizados nos experimentos. A Figura 4.1 apresenta uma breve descrição do passo a passo para o desenvolvimento deste trabalho.

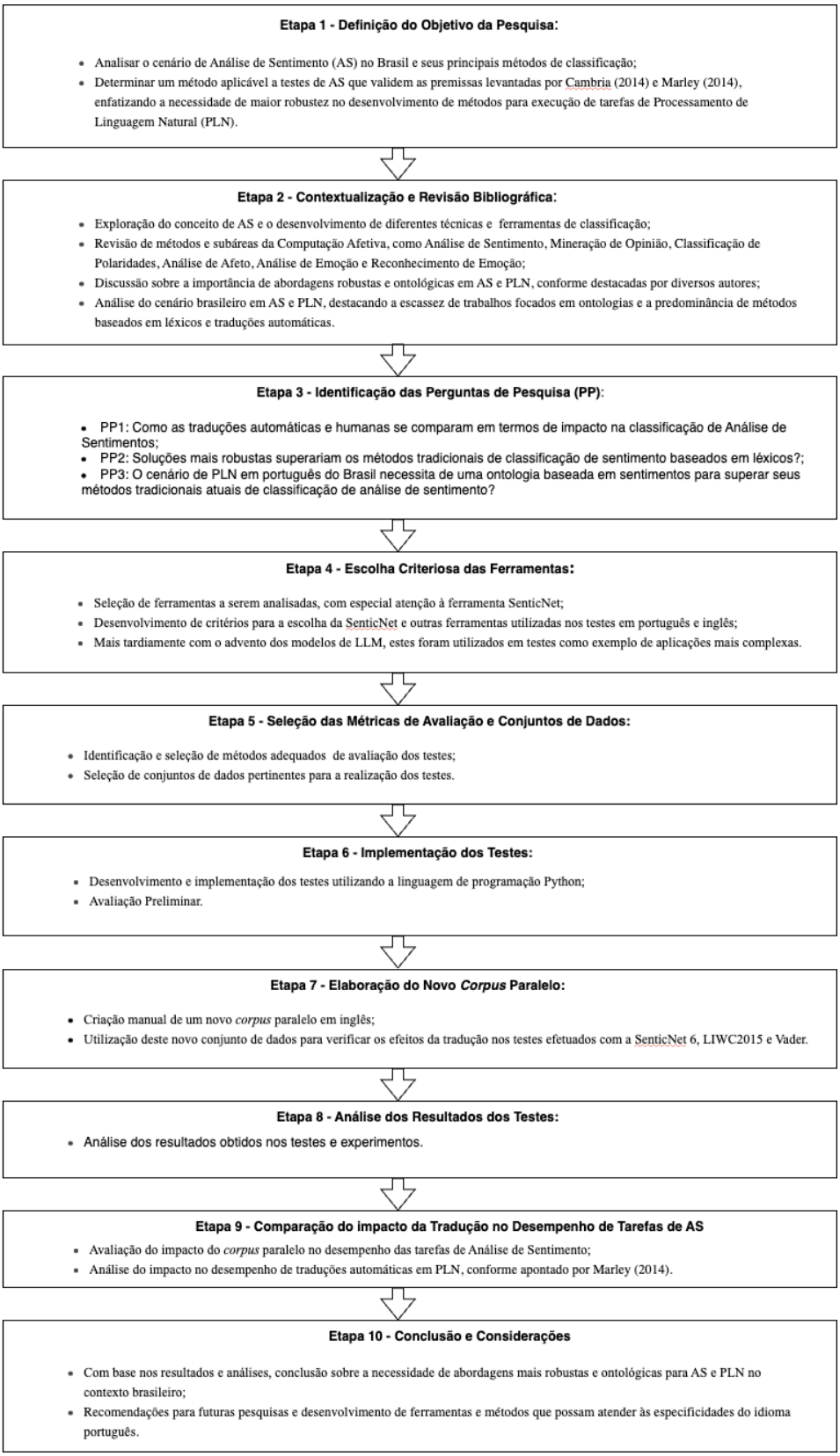


Figura 4.1: Etapas da metodologia de pesquisa

4.1 Etapa 1 - Definição do Objetivo da Pesquisa

Inicialmente, realizou-se uma pesquisa na base de literatura indexada Scopus sobre o estado da arte no campo de AS no Brasil. Identificou-se, com base nas premissas de da Silva Conrado et al. [2014], a necessidade de novos estudos que investigassem a possibilidade de novos recursos que pudessem superar as limitações dos métodos atuais do cenário de PLN e AS no Brasil, predominantemente baseados em métodos estatísticos e traduções automáticas do inglês para o português. Além disso, buscamos explorar a aplicabilidade de abordagens ontológicas e semânticas, que têm demonstrado potencial significativo em melhorar a desambiguação e a precisão em tarefas de AS.

A definição das premissas observadas nos trabalhos de da Silva Conrado et al. [2014] norteiam todas as etapas subsequentes da pesquisa, desde a seleção criteriosa de ferramentas até a implementação e análise de testes, garantindo que cada ação esteja alinhada com a necessidade de inovação e aprimoramento das técnicas de AS e PLN no cenário brasileiro. Além disso, buscou-se comparar o impacto que a forma de tradução dos conjuntos de dados exerce no desempenho de tarefas de AS. Em outras palavras, foi necessário avaliar se diferenças nas estruturas linguísticas dos idiomas português e inglês implicam num pior desempenho devido a traduções automáticas em PLN, conforme apontado por da Silva Conrado et al. [2014]. Portanto, elaborou-se manualmente um novo *corpus* paralelo em inglês. Neste trabalho, tal conjunto de dados foi utilizado para se verificar os efeitos da forma de tradução por meio dos testes efetuados com a SenticNet 6.

4.2 Etapa 2 - Contextualização e Revisão Bibliográfica

Nesta etapa, contextualizamos a fundamentação teórica e metodológica que embasa nossa pesquisa, fornecendo uma revisão sistemática abrangente da literatura pertinente do período de 2018 a 2021. Além do levantamento de métodos mais comumente utilizados no Brasil para AS, ampliamos esta pesquisa na base Scopus também para outros idiomas além do português. A partir desta busca, encontramos os trabalhos Cambria et al. [2020] e por meio de técnicas de *backward snowballing*, identificamos os demais trabalhos da SenticNet e da metodologia de *sentic computing*, bem como o trabalho de da Silva Conrado et al. [2014].

A literatura apontou que a maioria das abordagens de AS utiliza métodos estatísticos e léxicos para classificação de polaridades, mas destaca a importância de métodos mais abrangentes que incorporam aspectos ontológicos e semânticos para uma análise mais precisa [Cambria et al., 2017; Jiang and Tan, 2005]. No contexto brasileiro, apesar de contribuições importantes como o Unilex de de Souza et al. [2017] e o Reli Lex [Freitas and Vieira, 2013], a aplicação de técnicas robustas e ontológicas em PLN ainda é limitada [da Silva Conrado et al., 2014; Freitas and Vieira, 2013, 2019; Rosa et al., 2019].

Após essa análise inicial, levantamos trabalhos que tivessem *benchmarkings* com léxicos em português no Brasil no âmbito da análise de sentimentos, a fim de analisar o desempenho dessas ferramentas e técnicas de BOW. Como não encontramos estudos semelhantes utilizando ontologias, levantamos diferentes métodos que possibilitassem testes com uma abordagem para AS com textos em português.

4.3 Etapa 3 - Identificação de Perguntas de Pesquisa

A identificação das Perguntas de Pesquisa (PP) foi uma etapa crucial para orientar e delimitar nossa investigação científica, cujo escopo poderia tornar-se extremamente abragente após as investigações de métodos de AS com ontologia em diferentes idiomas. Desta forma, foi possível garantir foco e relevância nas análises voltadas para o contexto brasileiro. Com base na revisão bibliográfica e nas lacunas identificadas na literatura, formulamos três perguntas principais que guiaram nossa pesquisa:

PP1: Como as traduções automáticas e humanas se comparam em termos de impacto na classificação de Análise de Sentimentos?

Esta pergunta visa investigar a influência das diferentes técnicas de tradução na precisão e eficácia das tarefas de AS. Dados os desafios inerentes às diferenças estruturais e semânticas entre o inglês e o português, conforme discutido por da Silva Conrado et al. [2014], é essencial avaliar se traduções automáticas prejudicam as tarefas e métricas de AS em comparação com traduções realizadas por humanos.

PP2: Soluções mais robustas superariam os métodos tradicionais de classificação de sentimento baseados em léxicos?

Aqui, exploramos se abordagens mais sofisticadas, incluindo métodos ontológicos e semânticos, podem oferecer desempenho superior aos métodos tradicionais baseados em léxicos para a classificação de sentimentos. A literatura sugere que métodos estatísticos e léxicos predominantes podem ser insuficientes para capturar a complexidade das emoções humanas [Cambria and Hussain, 2012; Cambria and White, 2014; Cambria and Hussain, 2015; Cambria et al., 2017, 2022; Liu et al., 2012], justificando a necessidade de explorar soluções mais robustas.

PP3: O cenário de PLN em português do Brasil necessita de uma ontologia baseada em sentimentos para superar seus métodos tradicionais atuais de classificação de análise de sentimento?

Esta pergunta investiga a viabilidade e a necessidade de desenvolver uma ontologia específica para sentimentos no contexto da língua portuguesa, com o objetivo de melhorar as tarefas de AS. Assim, abordagens ontológicas poderiam fornecer uma estrutura mais rica e contextualizada para a análise de sentimentos.

A formulação dessas perguntas de pesquisa não apenas orienta a direção metodológica do estudo,

mas também destaca áreas críticas que necessitam de investigação aprofundada. Ao responder a essas perguntas, esperamos contribuir significativamente para o avanço do estado da arte em AS e PLN, especialmente no contexto brasileiro.

4.4 Etapa 4 - Escolha Criteriosa das Ferramentas

Para viabilizar a execução dos testes pretendidos, identificamos duas abordagens principais com domínio aberto e foco em AS: a MFOEM [Smith, 2018] e a SenticNet [Cambria and Hussain, 2012; Cambria and White, 2014; Cambria and Hussain, 2015; Cambria et al., 2017, 2020, 2022]. Embora a MFOEM apresente uma estrutura formal adequada e alinhada aos paradigmas de troca de conhecimento, ela não oferece suporte para a língua portuguesa. Portanto, optamos por realizar os testes com a SenticNet, que é um *framework open source* com resultados positivos em inglês, estrutura baseada em onto-pares de palavras e conceitos primitivos, usando uma lógica denominada *sentic algebra*.

Dada a necessidade de suporte para a língua portuguesa e a disponibilidade como software livre, a SenticNet 6 foi selecionada para avaliar a eficácia das traduções automáticas e a influência de sua estrutura ontológica em tarefas de AS. A ferramenta proporciona um cenário ideal para testar e validar o desenvolvimento de uma ontologia de sentimentos específica para a língua portuguesa, fundamentada na metodologia de *sentic computing* desenvolvida Cambria and Hussain [2012]. A SenticNet, com sua OntoSenticNet e a *sentic algebra*, oferece um ambiente multidisciplinar baseado em ontologias e raciocínio de senso comum, que permitem avaliar as hipóteses da literatura sobre o desempenho de ontologias em classificação de sentimentos.

Além da SenticNet, consideramos incluir o LIWC2015 (em inglês) e o VADER nos testes em inglês por terem o melhor desempenho nos estudos comparativos de Cambria et al. [2020] e Cambria et al. [2022] com a SenticNet em inglês. Adicionalmente, testamos modelos de *Large Language Models* (LLM), com o Chat-GPT, assim que este fora lançado. O objetivo desses testes adicionais era trazer um escopo maior e comparar os métodos de BOW e *sentic computing* com um sistema de aprendizado profundo tão robusto quanto a SenticNet, proporcionando uma avaliação abrangente das técnicas modernas de AS. Os léxicos utilizados nesses testes foram o LIWC2015-PT-br e o Affect, que tiveram maior destaque no *benchmarking* de Nascimento et al. [2019].

Dessa forma, nossa escolha criteriosa das ferramentas incluiu tanto abordagens tradicionais baseadas em léxicos quanto métodos modernos de PLN e uma abordagem ontológica, garantindo um escopo amplo e consistente para a análise comparativa. Os experimentos realizados com a SenticNet 6 e outros recursos visaram validar a hipótese de que ferramentas baseadas em ontologias podem oferecer resultados significativos e justificar o desenvolvimento de novas ontologias ou sitemas com abordagens ontológicas para AS em português brasileiro.

4.5 Etapa 5 - Seleção das Métricas de Avaliação e Conjuntos de Dados

A seleção de métricas de avaliação é uma etapa crítica na metodologia de pesquisa, pois garante a precisão e a confiabilidade dos resultados obtidos. Para esta pesquisa, utilizamos um conjunto abrangente de métricas comumente empregadas na avaliação de tarefas de AS, a fim de capturar diferentes aspectos do desempenho dos métodos e ferramentas testadas. Desta forma, as métricas levantadas para o canário dos testes foram: Acurácia, utilizada como referência nos testes de Cambria et al. [2020] e Cambria et al. [2022], Precisão, Revocação e F1-Score, devido a sua importância para avaliações por matriz de confusão.

Para garantir a relevância e a robustez dos testes, selecionamos conjuntos de dados em português presentes no *benchmarking* de Nascimento et al. [2019]. Esses conjuntos de dados foram escolhidos devido à sua abrangência e à representatividade no contexto da AS em português, refletindo a diversidade linguística e estilística do português brasileiro. O foco adicional no MQD-1465 (explicado mais detalhadamente no Capítulo 5, seção 5.1.1) permite uma análise mais profunda das capacidades dos modelos em lidar com emotivos.

As métricas e os conjuntos de dados selecionados foram integrados no processo de avaliação dos métodos de AS, incluindo tanto léxicos tradicionais (LIWC2015-PT-br, Affect para português e Liwc2015 e Vader para o inglês) quanto abordagens modernas baseadas em LLMs e quanto abordagens ontológicas com a SenticNet. Esta integração permitiu uma comparação detalhada e rigorosa, alinhada com os objetivos de explorar e validar metodologias robustas para a AS em português brasileiro.

4.6 Etapa 6 - Implementação dos Testes

Para a implementação dos testes, seguimos as instruções fornecidas pelo *framework* da SenticNet [SenticNet, 2012a], que indicam a necessidade de solicitar o *Software Development Kit* (SDK) diretamente com a equipe da SenticNet. Após a obtenção das chaves únicas de API, iniciamos a implementação dos testes utilizando a linguagem Python. É recomendado pela SenticNet que o *corpus* não seja pré-processado antes da execução dos testes, garantindo que a análise seja realizada no texto bruto. A figura 4.2 representa um fluxograma das etapas do experimento com a SenticNet.

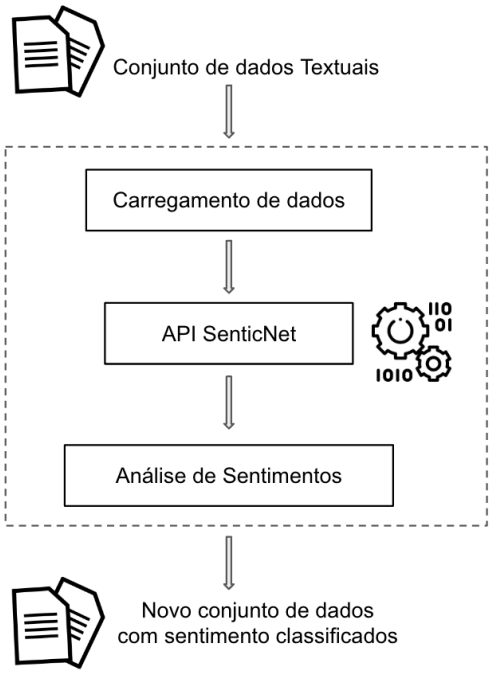


Figura 4.2: Fluxograma simplificado de utilização da ferramenta SenticNet

A tarefa de classificação de sentimentos foi conduzida utilizando a API da SenticNet, que atribui polaridade contextual a conceitos por meio de IA neuro simbólica. A API espera como entrada um trecho de texto (uma frase ou um parágrafo) e retorna um dos seguintes rótulos: positivo, negativo ou neutro. Devido à validade limitada das chaves de API (aproximadamente 30 dias), foram necessárias várias renovações para executar os diferentes testes ao longo do trabalho.

O código para chamar as chaves da API de classificação da SenticNet foi disponibilizado em um repositório no GitHub [Carneiro, 2022][1]. Esse código mostra a saída dos testes de classificação de cada conjunto de dados testado, onde as classificações positivas são rotuladas como *Positive*, as negativas como *Negative* (em vermelho) e as neutras como *Neutral* (em preto). Um exemplo dos resultados obtidos após a execução dos testes está ilustrado na Figura 4.3.

[1]https://github.com/EricaCQ/senticnet-6-test-2022.git

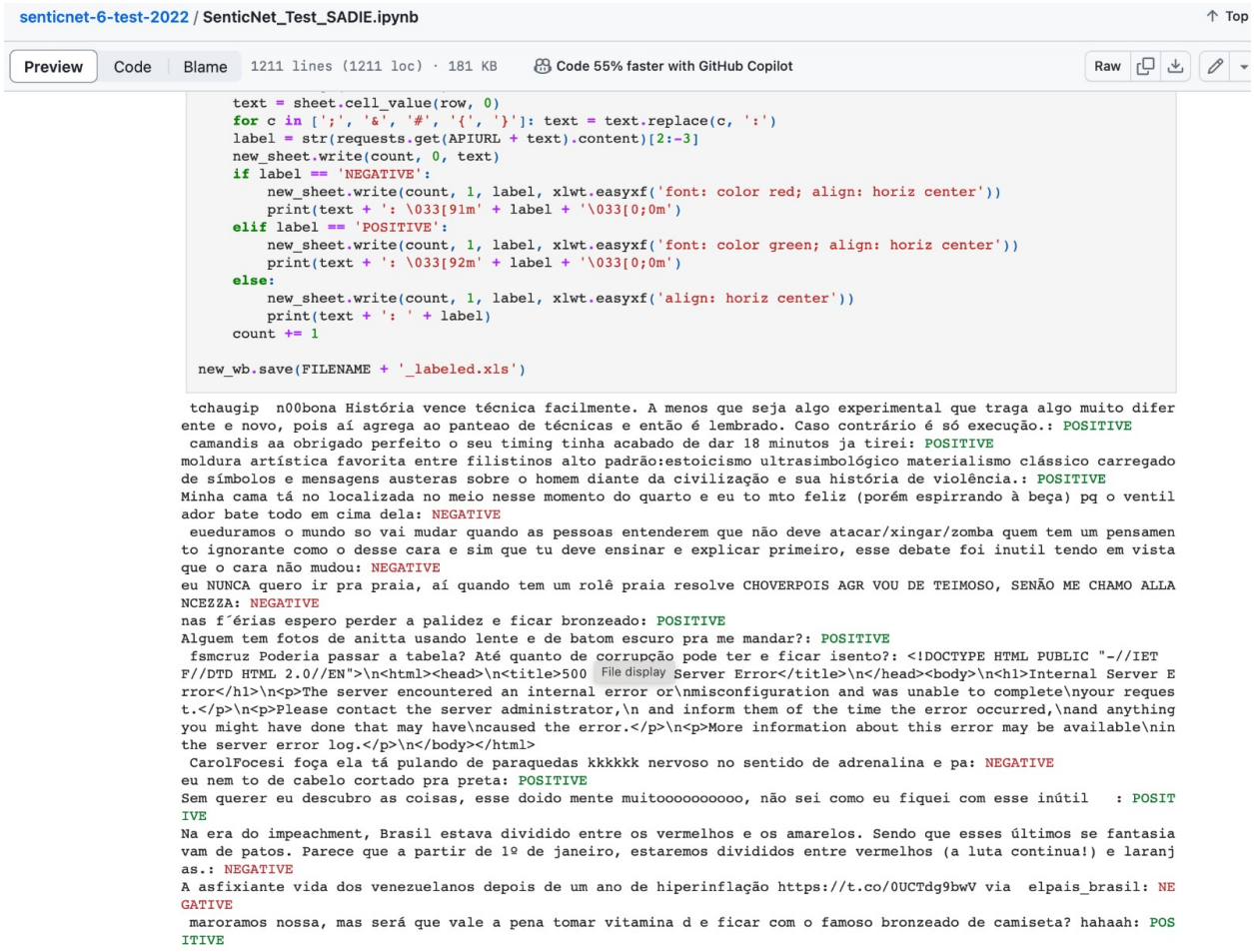


Figura 4.3: Captura de tela contendo exemplo de execução do teste

Para a avaliação dos resultados, utilizamos a biblioteca Scikit-Learn do Python. As métricas de Precisão, Revocação, F1-Score e Acurácia foram calculadas e comparadas por meio de matrizes de confusão, permitindo uma análise detalhada do desempenho da SenticNet em diferentes conjuntos de dados. Essas métricas fornecem uma visão abrangente da eficácia do *framework* em tarefas de AS, fundamentando a validação de sua aplicabilidade no contexto do português brasileiro.

4.7 Etapa 7 - Elaboração do Novo *Corpus* Paralelo

A criação do *corpus* paralelo do MQD-1465, doravante denominado MQD Paralelo foi realizada manualmente, um processo que demandou consideráveis recursos humanos e tempo, evidenciando seu caráter oneroso, conforme descrito no experimento de García-Díaz et al. [2020]. Para viabilizar a conclusão do presente trabalho dentro do prazo estipulado, optou-se pela tradução manual de apenas um dos quatro conjuntos de dados em português testados na primeira etapa do experimento.

Para evitar vieses sócio e psicolinguísticos, o MQD-1465 foi traduzido manualmente por cinco pesquisadores do Centro Federal de Educação Tecnológica Celso Suckow da Fonseca (CEFET/RJ), incluindo um mestrando e quatro graduandos em Ciência da Computação, todos com proficiência

em inglês. Além das traduções manuais, uma sexta versão do MQD Paralelo foi criada utilizando a API do *Google Translate* [Google, 2024].

O MQD Paralelo é composto pelos subconjuntos de dados MQD00, MQD01, MQD02, MQD03, MQD04, que representam cada uma das traduções dos cinco estudantes, e MQDGT, que representa a tradução utilizando *Google Translate* [Google, 2024]. A inclusão do MQDGT permitiu verificar avaliar se as traduções automáticas realmente prejudicam as tarefas de PLN, bem como verificar se os baixos resultados da classificação da SenticNet no MQD-1465 original em português eram decorrentes da tradução automática do português para o inglês, realizada pela BabelSenticNet antes da etapa de análise de sentimentos.

4.8 Etapa 8 - Análise dos Resultados dos Testes

A análise dos resultados dos testes foi realizada utilizando a biblioteca Scikit-Learn em Python. Foram calculadas as principais métricas de avaliação: Precisão, Revocação, *F1-Score* e Acurácia. Para facilitar a comparação dos resultados, utilizamos Matriz de Confusão (MC)s, que são tabelas que mostram a frequência de acertos e erros previstos pelos modelos. Esse método permite uma avaliação detalhada do desempenho de modelos e métodos de classificação, incluindo AS, destacando onde ocorriam erros de classificação e facilitando a identificação de padrões de desempenho em diferentes conjuntos de dados. A análise dessas métricas proporcionou uma visão abrangente da eficácia das abordagens testadas.

4.9 Etapa 9 - Comparação do Impacto da Tradução no Desempenho de Tarefas de Análise de Sentimentos

A comparação do impacto da tradução no desempenho das tarefas de AS foi uma etapa crucial deste estudo. Para isso, analisamos as diferenças entre as classificações obtidas a partir do MQD-1465 original em português, suas traduções manuais e a tradução automática (MQDGT) realizada pelo *Google Translate*.

Utilizamos as mesmas métricas de avaliação (Precisão, Revocação, *F1-Score* e Acurácia) para comparar os resultados das diferentes versões do *corpus* e o desempenho dos diferentes métodos de classificação de sentimentos abordados em cada um dos conjuntos de dados. Através das matrizes de confusão, foi possível identificar como a qualidade da tradução afetava o desempenho da SenticNet, do Liwc2015 e do Vader para o MQD Paralelo (inglês).

4.10 Etapa 10 - Conclusão e Considerações

A conclusão deste trabalho foi embasada na fundamentação teórica, trabalhos relacionados e experimentos deste trabalho. A contextualização apresentada não só fundamenta a relevância e originalidade da nossa pesquisa, mas também estabelece um marco teórico robusto para a análise e desenvolvimento de diferentes métodos de AS, garantindo que nossa investigação esteja bem ancorada nas discussões mais avançadas do campo de AS e PLN.

Capítulo 5 Avaliação Experimental

Os experimentos tiveram como objetivo aplicar a SenticNet na AS de conjuntos de dados em português, um idioma que frequentemente enfrenta desafios devido à menor quantidade de recursos linguísticos em comparação ao inglês. Para alcançar esse objetivo, desenvolvemos um *corpus* paralelo que serve como base para os experimentos subsequentes. Os resultados obtidos são avaliados e comparados com as versões do MQD Paralelo, proporcionando uma análise da eficácia da SenticNet na interpretação de sentimentos em textos em português.

5.1 Conjuntos de Dados utilizados nos testes

Os testes realizados para este trabalho foram conduzidos utilizando quatro conjuntos de dados em língua portuguesa: MQD [Azevedo et al., 2021b], *Sentiment Analysis Dataset Twitter-pt-br* (SADT) [Silva et al., 2019], TOPIE [Souza et al., 2016] e TweetSentBR [Brum and Volpe Nunes, 2018]. A escolha deu-se a fim de efetuar uma comparação com o *benchmarking* de Nascimento [2021].

5.1.1 MQD-1465

Desenvolvido por alunos do Programa de Pós-graduação em Ciência da Computação (PPCIC), do CEFET/RJ, como parte das atividades da disciplina Mineração de Textos, o MQD-1465 teve como fonte do *corpus* o site Meu Querido Diário. Nele, os usuários postam seus pensamentos sobre diversos assuntos. O conjunto de dados é composto por 1.458 frases analisadas por alunos do curso de MT e classificadas em positivas (507), negativas (542) ou neutras (409). Além disso, o MQD-1465 em seu formato original possui 26.692 palavras e 144.879 caracteres.

5.1.2 SADT

O conjunto de dados SADT [Silva et al., 2019] é composto por 1.867 frases classificadas em positivas (559), negativas (597) e neutras (711). Ao todo são 38.930 palavras e 197.261 caracteres, cujas médias por frases correspondem a $20,84$ e $105,60$ respectivamente. O conjunto de dados SADT é composto por comentários, em português, extraídos do Twitter a partir da busca por adjetivos presentes no *Thesaurus* TeP 2.0., entre 24/09/2018 e 06/12/2018. A rotulação das frases foi parte

de atividade desenvolvida pelos alunos do curso de Bacharelado em Sistemas de Informação da Universidade Federal da Paraíba.

5.1.3 TOPIE

Trata-se de um conjunto de dados desenvolvido por pesquisadores do grupo de estudos *MiningBR Research Group*, da Universidade Federal Rural de Pernambuco (UFRPE). O TOPIE contém 30.0000 *tweets*, extraídos por meio de um *webcrawler*, referentes a empresas que mais receberam reclamações, em 2013, de acordo com a Agência do Programa de Defesa e Proteção ao Consumidor (PROCON) [Souza et al., 2011a]. Foram selecionadas 498 frases anotadas como positivas (166), negativas (166) e neutras (166). Ao todo, o conjunto de dados utilizado nos testes consiste em 10.244 palavras e 52.116 caracteres, cujas médias por frase são respectivamente $20,52$ e $104,44$.

5.1.4 TweetSentBR

Composto por comentários extraídos do Twitter entre janeiro e julho de 2017, o TweetSentBR tem como tema postagens sobre programas de televisão em tempo real, ou seja, que foram escritos simultaneamente à exibição das atrações. Foi observado que muitos telespectadores postavam enquanto assistiam aos programas, utilizando *hashtags* (e.g. #MasterChef) [Brum and Volpe Nunes, 2018].

Desta forma, os próprios telespectadores segmentavam os comentários (frases) em um domínio específico, *i.e.* o programa ao qual assistiam. O conjunto de dados contém 10.156 comentários distintos, classificados em positivos (4.589), negativo (2.927) e neutro (2.640). A amostra possui um total de 117.797 palavras e 643.800 caracteres, cujas médias por frase são, respectivamente, $11,60$ e $63,38$

5.2 Experimento 1: SenticNet aplicada à análise de sentimentos de conjuntos de dados em português

O objetivo deste experimento foi validar se uma abordagem ontológica teria melhor desempenho do que apenas o uso de léxicos ou métodos de classificação estatística. Assim, verificaríamos a necessidade e possibilidade de criação de uma ontologia de AS, chamando à reflexão a comunidade científica no Brasil para o problema já observado por da Silva Conrado et al. [2014]. Devido à imensa quantidade de sinônimos e alta complexidade do mapeamento do domínio que uma ontologia de AS - como a MFOEM, a SenticNet e mesmo a ontologia de depressão de Jung et al. [2017] - exigiria, optou-se por não desenvolver um protótipo neste momento.

Após os testes iniciais com a ferramenta SenticNet 6, analisou-se possíveis onto-pares que influenciassem na classificação da ferramenta. Posteriormente, efetuou-se uma avaliação por MC dos

resultados da SenticNet 6. Da mesma forma, Cambria et al. [2020] realizara os testes similares e comparações para léxicos da língua inglesa e seu desempenho em relação à SenticNet.

O desempenho da análise de sentimentos efetuada por meio da SenticNet 6 foi comparado aos resultados obtidos pelo uso de léxicos extraídos do trabalho de Nascimento [2021]. Além disso, os resultados obtidos nos testes realizados também foram comparados com a capacidade do popular modelo de LLM ChatGPT [OpenAI, 2022] em efetuar AS. Para tanto, uma abordagem simplificada de *Chain of Thoughts* [Wei et al., 2023] foi utilizada, onde a plataforma foi inquirida quanto à classificação das frases contidas nos conjuntos de dados MQD-1465.

5.2.1 Classificação em positivo, neutro e negativo

Os resultados do Experimento 1 foram inconclusivos para esta avaliação, conforme apresentado a seguir. Os resultados obtidos pela execução de AS utilizando a API da SenticNet 6 foram comparados com as anotações manuais em seus respectivos conjuntos de dados: MQD-1465, Topie, SADT e TweetSentBR. Posteriormente, quatro Matrizes de Confusão foram geradas correspondendo a cada um dos conjuntos de dados. Após essa etapa, efetuou-se a avaliação cruzada dos resultados de modo a obter as métricas de Acurácia e médias de F1-Score, Precisão e Revocação.

Os valores de Acurácia obtidos pelos resultados de classificação estão dispostos na Tabela 5.1. Os valores obtidos foram de aproximadamente 30%, a saber: MQD-1465 (37%), SADT (32%) , TOPIE (31%) e TweetSentBR (36%). Pode-se concluir, então, que embora os resultados de classificação da Senticnet 6 para conjuntos de dados em inglês estivessem acima de 80% [Cambria et al., 2020], os resultados dos testes para *datasets* em português brasileiro foram inferiores.

Tabela 5.1: Resultados das Médias de Precisão, Revocação e F1-Score e valor total da Acurácia da análise de sentimentos da SenticNet 6 para conjuntos de dados em português do Brasil.

Conjunto de Dados	Idioma	Precisão	Revocação	F1-Score	Acurácia
MQD-1465	Português	0,39	0,34	0,30	0,37
SADT	Português	0,29	0,34	0,26	0,32
TOPIE	Português	0,35	0,31	0,26	0,31
TweetSentBR	Português	0,33	0,33	0,31	0,36

Quando comparados aos resultados médios de AS realizadas em inglês por [Cambria et al., 2020], os testes revelaram um desempenho relativamente baixo para a classificação dos sentimentos quando traduções automáticas feitas pela BabelSenticnet foram utilizadas. Sendo assim, ainda que a Senticnet 6 use uma abordagem ontológica e relações semânticas entre conceitos, são necessárias mais discussões e avaliações quanto aos possíveis impactos causados pelas diferenças entre as estruturas idiomáticas em modelos de classificação, principalmente quando se trata de relações semânticas.

Os resultados também podem ser analisados por meio de matrizes de confusão geradas para cada experimento com os quatro conjuntos de dados em português. A Figura 5.1 apresenta as quatro MCs

que confrontam os resultados da SenticNet versus juízes humanos. Nas MCs, as colunas rotuladas como 1 correspondem a sentimentos positivos, as como 0 aos neutros e as de -1 se referem aos sentimentos negativos. O eixo y, *Predito*, representa as classificações realizadas pela SenticNet 6, enquanto o eixo x, *Atual*, corresponde aos valores reais anotados pelos juízes humanos. Cada um dos quadros das classes 1 (positivo), 0 (neutro), e -1 (negativo) apresenta os valores obtidos de uma porcentagem de acertos em relação ao total de frases classificadas para cada um dos conjuntos de dados.

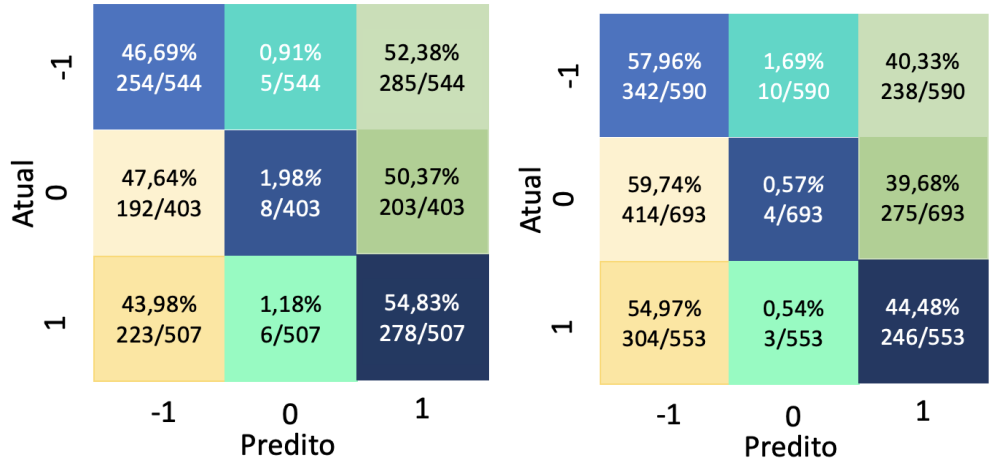
Cada um dos quadros na diagonal principal das MCs apresentam os valores de acertos da SenticNet divididos pelas anotações dos juízes, em cada uma das frases, para cada uma de suas respectivas classes (-1, 0, 1), bem como suas porcentagens. Os demais quadrantes mostram quantos valores foram sugeridos pela SenticNet para a classe errada, divididas pelas anotações dos juízes. A Acurácia (ACC) também pode ser obtida pela soma dos acertos de cada classe (-1, 0 e 1), dividida pelo Total (T) de frases classificadas, multiplicadas por 100.

Para o conjunto de dados MQD-1465, a ferramenta acertou 254 de 544 ($46,69\%$) de sentimentos negativos, 278 de 510 positivos ($54,5\%$) e apenas 8 de 410 neutros ($1,95\%$), como pode ser observado na MC (a), na Figura 5.1. Já para o SADT, a SenticNet classificou corretamente 342 de 590 ($57,97\%$) de sentimentos positivos, 246 de 553 ($44,48\%$) de negativos e somente 4 de 693 sentimentos neutros ($0,58\%$), presente MC (b) na Figura 5.1.

Os resultados dos *datasets* TOPIE e TweetSentBr encontram-se na MC (c) e (d) na Figura 5.1. A SenticNet 6 obteve um desempenho de 66 de 161 ($40,99\%$) ao classificar acertadamente sentimentos positivos do conjunto de dados TOPIE, 79 de 166 ($47,59\%$) para os negativos e 7 de 166 ($4,22\%$) para os neutros. Já no TweetSentBR, os resultados corretos foram de 342 de 590 ($57,97\%$) para os sentimentos positivos, 246 de 553 ($44,48\%$) para os negativos e 4 de 693 ($0,58\%$) para os neutros.

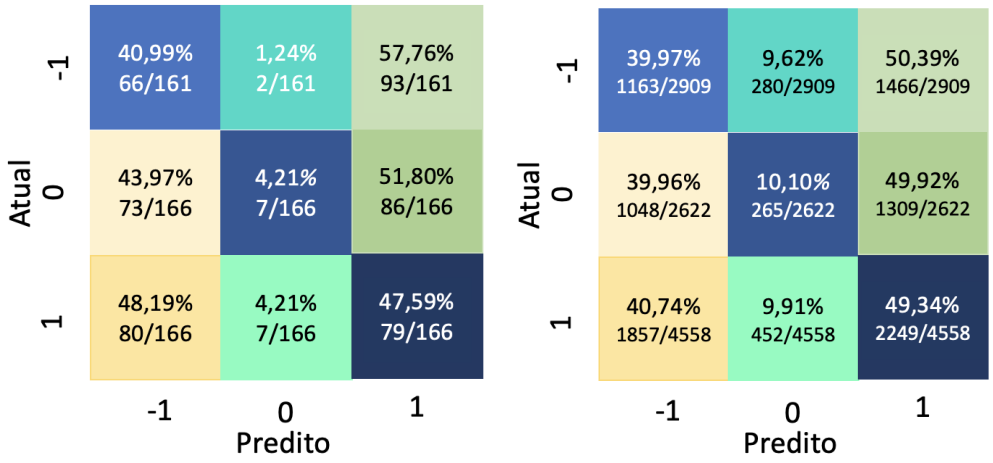
Os resultados das MCs também foram confrontados com aqueles presentes no trabalho de Nascimento [2021]. Os autores compararam as métricas de classificação de 13 léxicos afetivos para a língua portuguesa, dando maior ênfase aos com melhor resultado, a saber: AffectPT-br, LIWC2015-PT e SentiStrength.

Segundo Nascimento [2021], ao testar o conjunto de dados MQD-1465, o AffectPT-br detectou corretamente 310 de 510 exemplos positivos (61%). Porém, teve baixo desempenho ao lidar com sentimentos negativos, acertando apenas 251 de 544 exemplos (46%). O LIWC2015-PT-br obteve resultado semelhante para sentimentos positivos, detectando corretamente 314 de 510 exemplos ($61,6\%$), mas com baixo desempenho para sentimentos negativos, detectando 241 de 544 exemplos ($44,3\%$). Já o SentiStrength também apresentou desempenho similar ao lidar com sentimentos negativos, acertando 334 de 544 exemplos ($61,4\%$), mas não teve um bom desempenho com senti-



(a) Matriz de confusão do MQD-1465 com Neutros.

(b) Matriz de confusão do SADT com Neutros.



(c) Matriz de confusão do TOPIE com Neutros.

(d) Matriz de confusão do TweetSentBR com Neutros.

Figura 5.1: Matrizes de confusão incluindo sentimentos positivos, negativos e neutros do: (a) MQD-1465, (b) SADT, (c) TOPIE e (d) TweetSentBR. Os rótulos positivo, negativo e neutro são representados, respectivamente, por '1', '-1' e '0'. O eixo X representa as classes preditas e o eixo Y as classes atuais.

mentos neutros, acertando apenas 197 de 410 exemplos (48%). Esse desempenho na classificação de sentimentos neutros foi bastante inferior aos obtido pelo os outros dois léxicos (ambos superiores à 60%) [Nascimento, 2021].

Em complemento ao trabalho de Nascimento [2021], esta pesquisa realizou o mesmo experimento com a SenticNet, a fim detectar a acurácia, precisão, revocação e *F1-Score* de uma base de conhecimento com estruturas ontológicas, ainda que executando traduções automáticas com a BabelSenticnet. A ferramenta obteve resultados próximos quando comparados aos léxicos levantados por Nascimento et al. [2019], principalmente para sentimentos positivos e negativos, porém teve baixo desempenho ao classificar sentimentos neutros, conforme apresentado na Tabela 5.2.

A Tabela 5.2 dispõe a porcentagem de acertos obtidos pelas classificações com os léxicos afetivos

Tabela 5.2: Tabela de comparação entre léxicos afetivos e SenticNet, com a porcentagem de acertos para cada classe para o conjunto de dados MQD-1465.

Léxico/ Base de Conhecimento	Positivo	Negativo	Neutro
AffectPT-br	61,00%	46,60%	64,10%
LIWC2015-PT	61,80%	**44,70%**	61,00%
SentiStrengh	50,60%	62,00%	48,00%
LIWC2007-PT-br	**77,00%**	36,50%	39,50%
Polyglot	48,80%	50,50%	38,80%
MPQA-PT	52,80%	50,50%	33,90%
Sentilex-PT	34,80%	33,60%	71,70%
OpLexicon	54,30%	32,30%	42,20%
ANEW-Br	53,30%	31,40%	40,70%
WordNetAffectBR(adapt)	22,00%	30,80%	85,40%
ReLi-Lex	37,80%	17,30%	76,30%
WordNetAffectBR	20,30%	17,30%	**92,90%**
SenticNet.	*54,83%*	*44,69%*	*1,99%*

e da SenticNet. Ao comparar os resultados obtidos pela SenticNet 6 e pela utilização dos léxicos com tradução automática (ANEW-BR, LIWC, LIWC 2015 e Opinion Léxico), é possível verificar a influência da tradução automática ou humana nas ferramentas de classificação de AS. A SenticNet 6 teve desempenho pouco abaixo em relação aos sentimentos positivos e negativos. Quanto aos sentimentos neutros, a própria natureza da SenticNet, de busca por um dos extremos dentro do espaço vetorial entre -1 e 1 [Dragoni et al., 2018], faz com que ela não execute bem a classificação de sentimentos neutros. Segundo Cambria et al. [2017], os testes originais com a SenticNet não utilizaram sentimentos neutros.

Notadamente, léxicos que utilizam tradução humana ou foram originalmente criados para o português, como WordNetAffectBR, WordNetAffectBR(adapt) e ReLi-Lex apresentaram desempenho superior a qualquer outro para sentimentos neutros. Tais léxicos afetivos apresentaram uma porcentagem de acerto superior à 75%.

A Tabela 5.3 apresenta os resultados de Precisão, Revocação, Acurácia e *F1-Score* para todos os léxicos, que utilizam a língua portuguesa, confrontados com a Senticnet, utilizando a BabelSenticNet para tradução automática do português. Ao classificar sentimentos negativos e positivos, a SenticNet 6 ficou na quarta posição dentre os métodos de AS da Tabela 5.2. Porém, seu desempenho ao classificar sentimentos neutros ficou bem abaixo dos demais. Além disso, apesar de resultados dentro da média de acertos para as classes positiva (1), negativa (-1) e neutra (0), a SenticNet teve o pior desempenho nas métricas de avaliação de modelo, como apresentado na 5.3.

5.2.2 Classificação em positivo e negativo

Devido ao uso de um algoritmo de chaves polares e outro de aprendizado profundo denominado *Regularized K-Means* (RKM) [Cambria et al., 2020], a SenticNet tende a buscar os extremos das

Tabela 5.3: Comparação de resultados das métricas de Precisão, Revocação, Acurácia e F1-Score de léxicos utilizados como comparação para o português e a SenticNet com tradução automática (BabelSenticNet).

Léxico/ Base de Conhecimento	Precisão	Revocação	Acurácia	F1-Score
AffectPT-br	0,61	0,56	0,56	0,57
LIWC2015-PT	0,59	0,55	0,55	0,55
SentiStrengh	0,55	0,54	0,54	0,54
LIWC2007-PT-br	0,54	0,51	0,51	0,50
Polyglot	0,47	0,46	0,46	0,47
MPQA-PT	0,46	0,46	0,46	0,46
Sentilex-PT	0,53	0,44	0,44	0,44
OpLexicon	0,44	0,43	0,42	0,42
ANEW-Br	0,44	0,41	0,41	0,41
WordNetAffectBR(adapt)	0,58	0,43	0,43	0,40
ReLi-Lex	0,52	0,41	0,41	0,38
WordNetAffectBR	0,62	0,39	0,39	0,34
SenticNet.	*0,39*	*0,34*	*0,37*	*0,30*

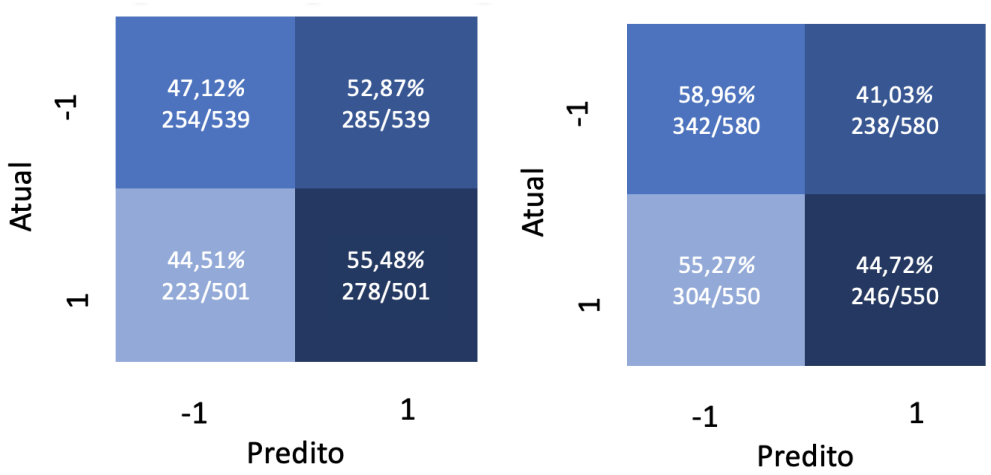
polaridades. Portanto, reproduzimos os testes com a SenticNet excluindo as frases anotadas como neutras (0) pelos juízes e então comparamos apenas frases classificadas como positivas (1) e negativas (-1).

Após análise dos resultados sem o sentimento neutro (0), as taxas de verdadeiros positivos foi de $55,48\%$ para o conjunto de dados MQD-1465, $44,72\%$ para o SADT, $49,68\%$ no TOPIE e $60,64$ para o TweetSentBR. Já os valores de verdadeiros negativos foi de $47,12\%$ para o MQD-1465, $58,96\%$ no SADT, 41.50% para o conjunto de dados TOPIE e $38,51\%$ no TweetSentBR. As taxas de falsos positivos e falsos negativos superaram às de verdadeiros positivos e negativos em todos os conjuntos de dados em português. Os resultados das MCs, sem o sentimento neutro (0), são mostrados na Figura 5.2.

No caso específico do conjunto de dados MQD-1465, esse aumento foi de 12% para Precisão, 17% para Revocação, 14% para Acurácia e 21% para F1-*Score*. Os resultados podem ser observados na Tabela 5.4. Os novos valores para as métricas de classificação da SenticNet se aproximam, então, do léxico LIWC2007-PT-br, lembrando que, no caso do léxico, a AS abrangia todas as classes, incluindo a neutra (0). Porém, tais resultados se encontram longe dos obtidos em testes em inglês por Cambria et al. [2020] e também Cambria et al. [2022], que chegaram a 85% de Acurácia.

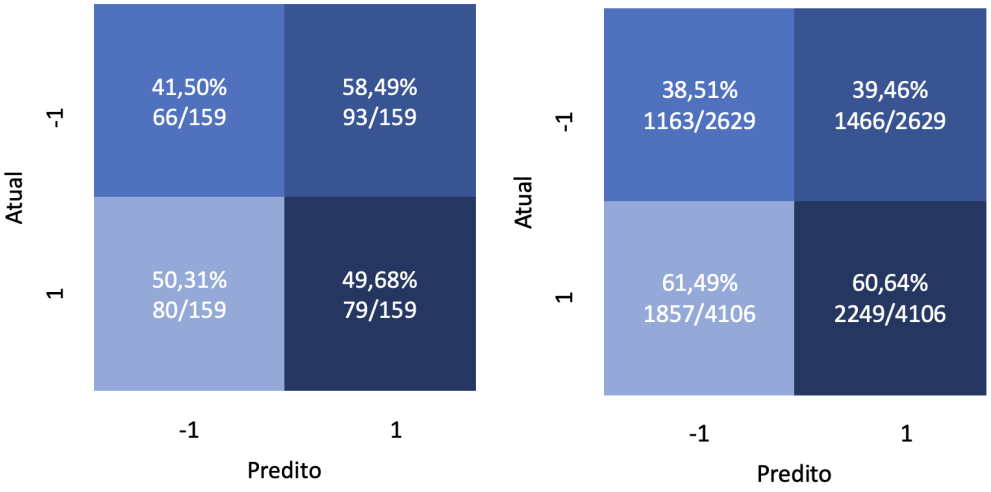
Tabela 5.4: Comparação de resultados das métricas de Precisão, Revocação, Acurácia e F1-Score para cada conjunto de dados em português (MQD-1465, SADT, TOPIE e TweetSentBR) excluindo o sentimento neutro (0) da classificação da AS com a SenticNet.

Conjunto de Dados	Precisão	Revocação	Acurácia	F1-Score
MQD-1465	0,51	0,51	0,51	0,51
SADT	0,52	0,52	0,52	0,52
TOPIE	0,46	0,46	0,46	0,46
TweetSentBR	0,50	0,50	0,51	0,50



(a) Matriz de confusão do MQD-1465 sem Neutros.

(b) Matriz de confusão do SADT sem Neutros.



(c) Matriz de confusão do TOPIE sem Neutros.

(d) Matriz de confusão do TweetSentBR sem Neutros.

Figura 5.2: Matrizes de confusão sem sentimentos neutros do: (a) MQD-1465, (b) SADT, (c) TOPIE e (d) TweetSentBR. Os rótulos positivo e negativo são representados, respectivamente, por '1' e '-1'. O eixo X representa as classes preditas e o eixo Y as classes atuais.

Como os resultados do Experimento 1 não foram significativos e não foram obtidos utilizando-se o idioma original da SenticNet, o experimento se mostrou inconclusivo para responder duas das três perguntas propostas na pesquisa:

a) Tarefas de PLN e análises de sentimento exigem abordagens de classificação mais robustas do que métodos estatísticos ou aprendizado de máquina? [Cambria and White, 2014; Cambria and Hussain, 2012]

b) Traduções automáticas prejudicam a execução de tarefas de PLN em português [da Silva Conrado et al., 2014]?

5.3 Experimento 2: SenticNet aplicada à tradução para o inglês com o MQD Paralelo

O segundo experimento computacional consistiu em utilizar a SenticNet 6 para efetuar a AS sobre a tradução do conjunto de dados MQD-1465 para o inglês. Para tanto, desenvolveu-se um *corpus* paralelo em inglês a partir do MQD-1465. Esse experimento de engenharia reversa tinha como objetivo fundamental verificar as premissas levantadas quanto à ineficiência de traduções automáticas para soluções de tarefas de NLP e aos impactos da diferença estrutural dos idiomas inglês e português [da Silva Conrado et al., 2014].

Com isso, poderíamos avaliar empiricamente as premissas de Cambria et al. [2017] quanto à necessidade de metodologias mais robustas do que métodos estatísticos ou aprendizado de máquina para a análise de sentimentos. Ao testar a Senticnet6 com o conjunto de dados MQD-1465 manualmente traduzido para o inglês, podemos verificar que tarefas de classificação de polaridade que visem alcançar o mais próximo da excelência (>80%) devem seguir abordagens mais complexas do que traduções manuais e abordagens com base em estatística ou AM.

5.3.1 Experimentos de Análise de Sentimentos das versões do MQD Paralelo

As cinco diferentes traduções, para o inglês, do *corpus* MQD Paralelo foram, então, testadas tanto na SenticNet quanto no léxico VADER. Tal léxico foi selecionado por constar duas vezes seguidas nos experimentos originais e presente no pódio de melhores resultados de acurácia apresentados em Cambria et al. [2022] e em Cambria et al. [2020].

Os resultados também foram comparados ao léxico em inglês LIWC2015. O LIWC2015 também foi selecionado por constar nos experimentos de Nascimento [2021] em sua versão em português, trazendo mais embasamento para os experimentos relacionados à tradução automática com o *Google Translate*. Além disso, trata-se de um dos léxicos em português com melhores resultados, assim, uma comparação com sua versão em inglês e com os resultados da SenticNet trariam mais embasamento para a análise do desempenho do método de *sentic computing*.

Em termos práticos e documentais, a principal diferença entre o uso da ferramenta SenticNet e todos os demais métodos já consolidados, como o LIWC2015 e o VADER reside na exclusão da etapa do pré-processamento dos dados textuais, que incluem o *lowering case i.e*, alteração de todas as letras em minúsculas; e a remoção de *stopwords*, que são palavras comuns e sem valor, como artigos indefinidos e preposições. Sendo assim, no Experimento 2, tivemos dois processos distintos para cada método, um sem pré-processamento (SenticNet) e outro com (LIWC2015 e VADER), conforme pode ser observado na Figura 5.3. No caso dos léxicos, além do pré-processamento, a etapa de preparação de *token* também foi necessária.

Os testes foram executados em Python 3.10 com as três diferentes bases de conhecimento com suas respectivas metodologias, conforme Figuras 4.2 e 5.3. Posteriormente, os *corpora* paralelos foram comparados com as anotações dos juízes em português e os resultados foram avaliados por meio de matrizes de confusão. O Experimento 2 levou em consideração as mesmas avaliações do Experimento 1 e uma etapa de pré-processamento dos textos.

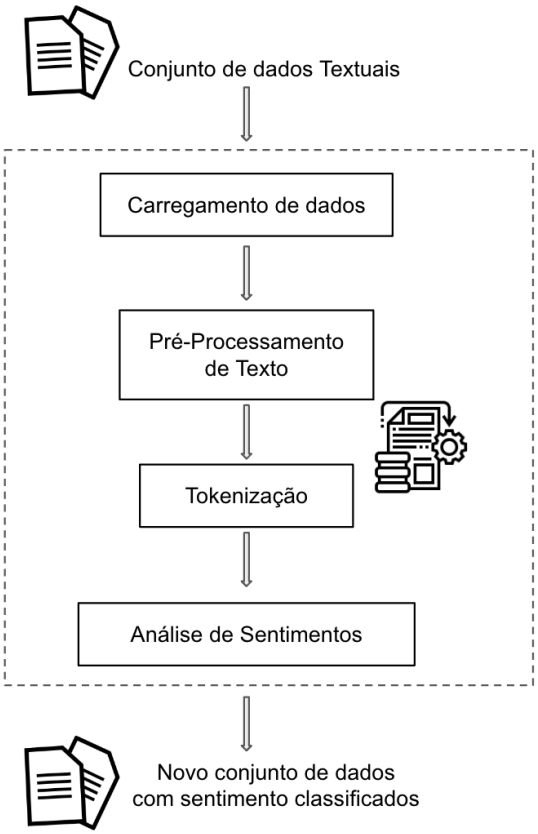


Figura 5.3: Fluxo de processamento para análise de sentimentos com os léxicos LIWC e VADER. Ambos com etapas de pré-processamento de texto e preparação de *tokens*

5.3.2 Classificação em positivo, neutro e negativo

Após a o Experimento 1, foi possível concluir que a SenticNet não chegou próximo de valores de acurácia de excelência para a língua portuguesa, ao contrário dos teste em inglês que chegaram a mais de 80% nos experimentos dispostos em Cambria et al. [2020] e em Cambria et al. [2022]. Com isso, optou-se por realizar o Experimento 2. Conforme explicado no capítulo 4, buscou-se avaliar o desempenho da SenticNet quanto à AS de traduções para o inglês do conjunto de dados MQD-1465.

As Figuras 5.4-5.9 apresentam as MCs geradas pelos *corpora* paralelas oriundas do MQD-1465 para cada um dos métodos utilizados pela SenticNet, LIWC2015 e VADER. Seguindo o padrão do primeiro experimento, a diagonal principal das MCs é a divisão dos valores de acertos pelas

anotações dos juízes para cada uma das classes (−1, 0, 1). O restante dos quadros das MCs apresentam quantos valores foram para uma classe divergente divididas pelas anotações dos juízes.

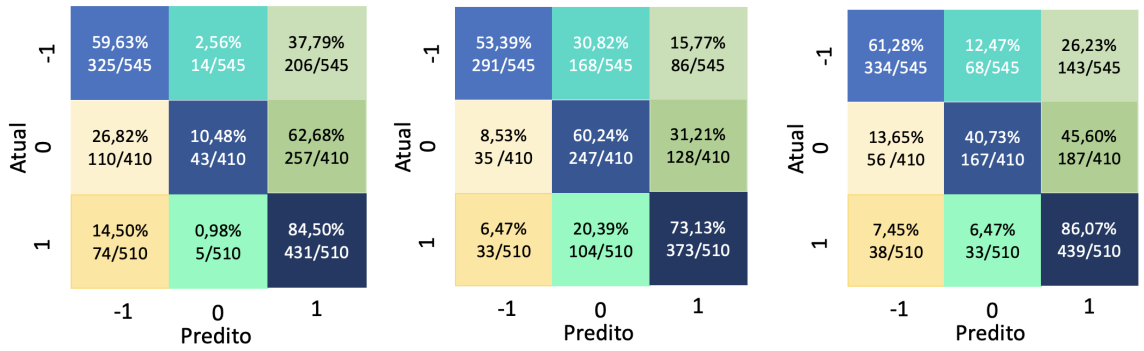


Figura 5.4: Matrizes de confusão do conjunto de dados MQD00 com: (a) SenticNet, (b) LIWC2015 e (c) VADER. Os rótulos positivo, negativo e neutro são representados, respectivamente, por '1', '-1' e '0'. O eixo X representa as classes preditas e o eixo Y as classes atuais.

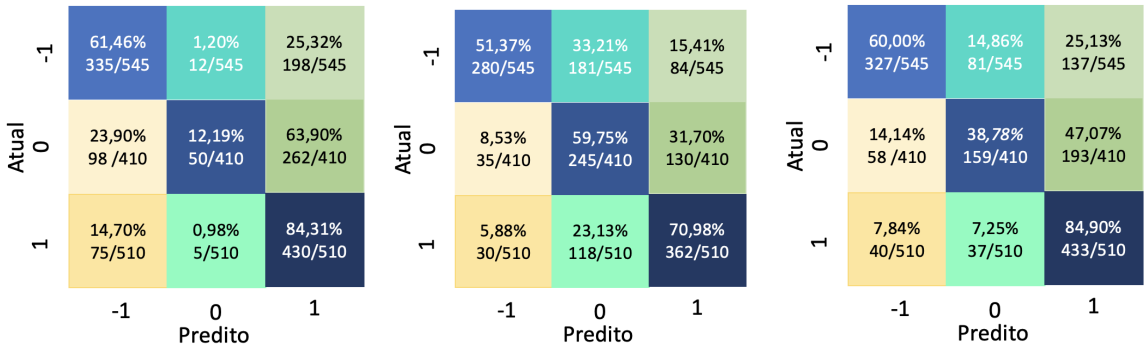


Figura 5.5: Matrizes de confusão do conjunto de dados MQD01 com: (a) SenticNet, (b) LIWC2015 e (c) VADER. Os rótulos positivo, negativo e neutro são representados, respectivamente, por '1', '-1' e '0'. O eixo X representa as classes preditas e o eixo Y as classes atuais.

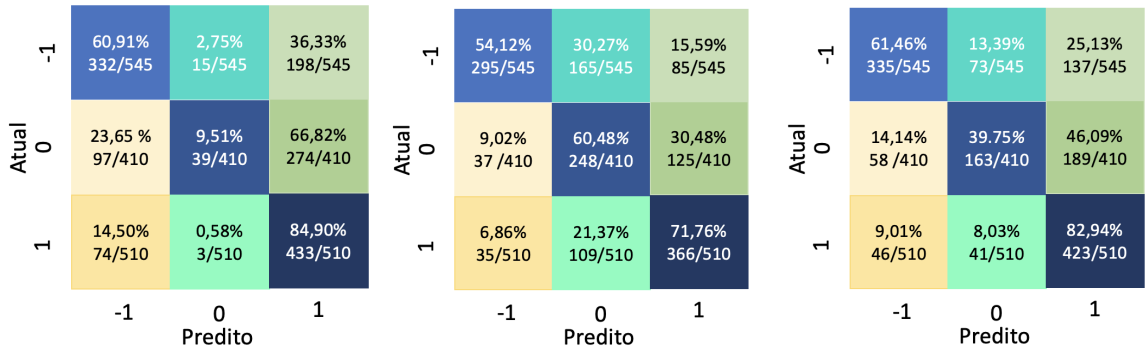


Figura 5.6: Matrizes de confusão do conjunto de dados MQD02 com: (a) SenticNet, (b) LIWC2015 e (c) VADER. Os rótulos positivo, negativo e neutro são representados, respectivamente, por '1', '-1' e '0'. O eixo X representa as classes preditas e o eixo Y as classes atuais.

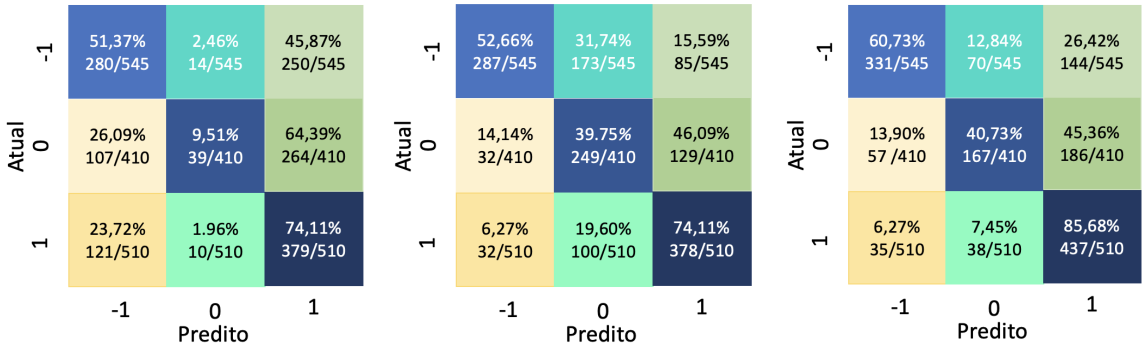


Figura 5.7: Matrizes de confusão do conjunto de dados MQD03 com: (a) SenticNet, (b) LIWC2015 e (c) VADER. Os rótulos positivo, negativo e neutro são representados, respectivamente, por '1', '-1' e '0'. O eixo X representa as classes preditas e o eixo Y as classes atuais.

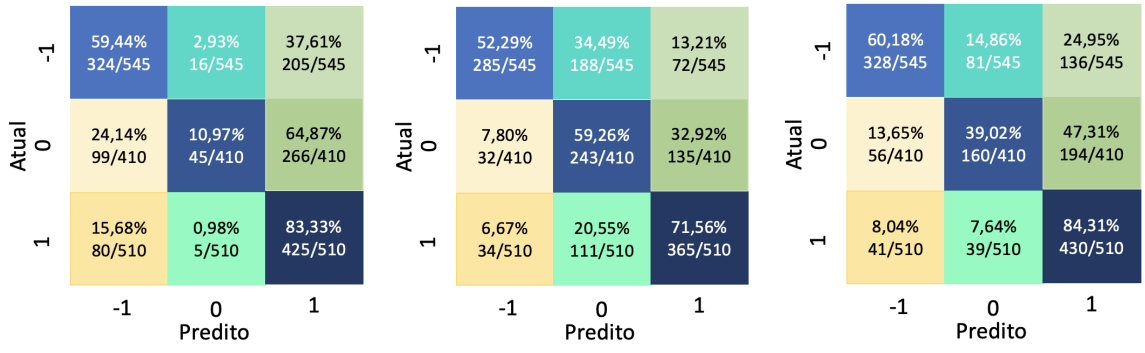


Figura 5.8: Matrizes de confusão do conjunto de dados MQD04 com: (a) SenticNet, (b) LIWC2015 e (c) VADER. Os rótulos positivo, negativo e neutro são representados, respectivamente, por '1', '-1' e '0'. O eixo X representa as classes preditas e o eixo Y as classes atuais.

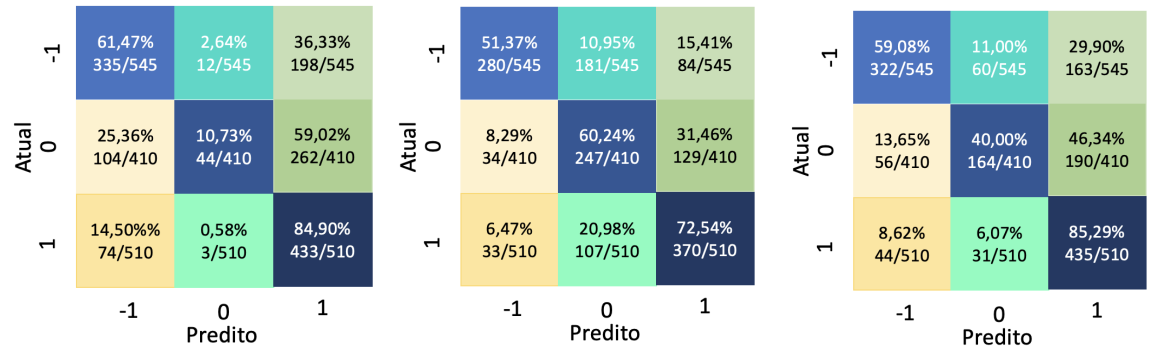


Figura 5.9: Matrizes de confusão do conjunto de dados MQDGT com: (a) SenticNet, (b) LIWC2015 e (c) VADER. Os rótulos positivo, negativo e neutro são representados, respectivamente, por '1', '-1' e '0'. O eixo X representa as classes preditas e o eixo Y as classes atuais.

A Tabela 5.5 apresenta um resumo de todos os valores classificados corretamente e presentes

nas MCs Figuras 5.4, 5.5, 5.6, 5.7, 5.8 e 5.9. Nela, é possível observar que o desempenho de acerto da SenticNet subiu consideravelmente em todos os *corpora* paralelos (em inglês) em relação ao Experimento 1, com o conjunto de dados em português, sendo traduzido automaticamente pela BabelSenticNet. Ressaltamos que os valores encontrados no primeiro experimento para as classes positivo (1), negativo (−1) e neutro (0) foram respectivamente: $44,69\%$, $54,83\%$ e $1,99\%$. Especificamente, os valores para o conjunto de dados MQD00 foram, naquela mesma ordem, de $84,50\%$, $59,69\%$, $10,48\%$. Destaca-se, aqui, a diferença de $39,81\%$ superior para acertos no sentimento positivo (1) e $8,49\%$ no neutro. A classe de sentimento negativo (−1), não apresentou uma variação considerável, sendo esta de apenas $4,86\%$ em relação ao Experimento 1.

Tabela 5.5: Tabela de comparação entre os *corpora* paralelos MQD00, MQD01, MQD02, MQD03, MQD04 e MQDGT, em inglês, com a Precisão, ou porcentagem de acertos para cada uma das bases de conhecimento/léxicos, SenticNet, LIWC2015 e VADER.

Conjunto de dados: MQD00				
	Léxico/Base de Conhecimento	Positivo	Negativo	Neutro
	SenticNet	84,50%	59,69%	10,48%
	LIWC2015	73,13%	53,39%	60,24%
	VADER	86,07%	61,28%	40,73%
Conjunto de dados: MQD01				
	Léxico/ Base de Conhecimento	Positivo	Negativo	Neutro
	SenticNet	84,31%	61,46%	12,19%
	LIWC2015	70,98%	51,37%	59,75%
	VADER	84,90%	60,00%	38,78%
Conjunto de dados: MQD02				
	Léxico/ Base de Conhecimento	Positivo	Negativo	Neutro
	SenticNet	84,90%	60,91%	9,51%
	LIWC2015	71,76%	54,12%	60,48%
	VADER	82,94%	61,46%	39,75%
Conjunto de dados: MQD03				
	Léxico/ Base de Conhecimento	Positivo	Negativo	Neutro
	SenticNet	74,11%	51,37%	9,51%
	LIWC2015	74,11%	52,66%	39,75%
	VADER	85,66%	60,73%	40,73%
Conjunto de dados: MQD04				
	Léxico/ Base de Conhecimento	Positivo	Negativo	Neutro
	SenticNet	83,33%	59,44%	10,97%
	LIWC2015	71,56%	52,29%	59,26%
	VADER	84,31%	60,18%	39,02%
Conjunto de dados: MQDGT				
	Léxico/ Base de Conhecimento	Positivo	Negativo	Neutro
	SenticNet	84,90%	61,47%	10,73%
	LIWC2015	72,54%	51,37%	60,24%
	VADER	85,29%	59,08%	40,00%

Os demais *corpora* paralelos mantiveram valores bem próximos desses, porém a versão do MQD03, teve uma queda de aproximadamente 10 pontos em relação às demais versões para o

sentimento neutro, demonstrando como a tradução humana pode influenciar nas classificações. O conjunto de dados com melhor desempenho, *i.e*, que melhor se aproximou dos juízes humanos nas classificações da SenticNet, foi a versão traduzida pelo *Google Translate*. Esta etapa do experimento indicou que a SenticNet pode chegar a valores de acerto bem superiores em seu idioma original, o inglês, porém, seu recurso de tradução, a BabelSenticnet não é suficiente para fornecer ao seu modelo de análise de sentimentos informações ou traduções adequadas que permitam um bom desempenho do mesmo.

Quando comparada aos demais léxicos, LIWC2015 (em inglês) e VADER, a SenticNet não teve resultados unanimemente superiores em nenhum dos *corpora* paralelos. Com exceção do MQD02, cujo vencedor foi a SenticNet, o VADER obteve pontuação superior para o sentimento positivo em todos os conjuntos de dados. Já para a classe de sentimentos negativo (-1), o VADER teve melhor desempenho em praticamente todas as versões em inglês do MQD, exceto para MQD01 e MQDGT, cujo melhor desempenho ficou com a SenticNet. O LIWC2015 foi o léxico com melhor resultado ao classificar sentimentos neutros (0). A única ocasião em que o melhor resultado para o sentimento neutro (0) foi obtido pelo VADER ocorreu no conjunto de dados MQD003. A SenticNet, mais uma vez, devido à busca pelas extremidades do eixo de sentimentos $(+)$ e $(-)$, teve desempenho muito baixo ao classificar sentimentos neutros.

É importante observar que a versão inglesa do LIWC2015 também teve resultados superiores em inglês, quando comparado aos testes de Nascimento [2021]. O léxico obteve melhor resultado na versão MQD00, cuja variação foi de $11,33$ pontos superiores no sentimento positivo (1) e $8,69$ no negativo (-1). Já a classe de sentimento neutra obteve 1 percentual a menos no MQD00 em relação ao original. O resultado levemente inferior se reproduziu em todas as versões em inglês, demonstrando que o LIWC2015 é o léxico mais efetivo ao classificar sentimentos neutros seja em inglês ou português. Embora o LIWC2015 tenha obtido resultado pouco inferior à SenticNet para as classes positivas e negativas, o léxico apresenta certa estabilidade nos resultados, tornando seu uso mais confiável para tarefas de classificação, principalmente quando envolverem traduções. Como o VADER não possui uma versão em português, a comparação deste léxico com diferentes idiomas não pôde ser realizada.

A Tabela 5.6 mostra os resultados da métricas de Precisão, Revocação, Acurácia e F1-Score para todos os *corpora* paralelos MQD00, MQD01, MQD02, MQD03, MQD04 e MQDGT classificados pelos léxicos LIWC2015, VADER e base de conhecimento SenticNet. No Experimento 1, com o MQD-1465 original, os resultados foram respectivamente $0,39$, $0,34$ e $0,37$ e $0,30$ de Acurácia.

O Experimento 2, com os *corpora* paralelos mostrou melhora em todas as métricas da SenticNet em todas as versões do MQD. Os percentuais de diferença podem ser observados na Tabela 5.7 e mostram que os léxicos MQD01 e MQDGT tendem a ter os melhores incrementos em relação ao

Tabela 5.6: Comparação de resultados das métricas de Precisão, Revocação, Acurácia e F1-Score para cada um dos subconjuntos do *corpus* MQD Paralelo, em inglês, com o sentimento neutro (0) da classificação de AS para SenticNet, LIWC2015 e VADER.

Conjunto de dados: MQD00					
	Léxico/ Base de Conhecimento	Precisão	Revocação	Acurácia	F1-Score
	SenticNet	60%	52%	55%	47%
	LIWC2015	64%	62%	62%	62%
	VADER	65%	63%	64%	62%
Conjunto de dados: MQD01					
	Léxico/ Base de Conhecimento	Precisão	Revocação	Acurácia	F1-Score
	SenticNet	63%	53%	56%	49%
	LIWC2015	63%	61%	61%	60%
	VADER	75%	73%	73%	73%
Conjunto de dados: MQD02					
	Léxico/ Base de Conhecimento	Precisão	Revocação	Acurácia	F1-Score
	SenticNet	61%	52%	55%	47%
	LIWC2015	64%	62%	62%	62%
	VADER	64%	61%	63%	61%
Conjunto de dados: MQD03					
	Léxico/ Base de Conhecimento	Precisão	Revocação	Acurácia	F1-Score
	SenticNet	53%	45%	48%	41%
	LIWC2015	64%	63%	62%	62%
	VADER	65%	62%	64%	62%
Conjunto de dados: MQD04					
	Léxico/ Base de Conhecimento	Precisão	Revocação	Acurácia	F1-Score
	SenticNet	60%	51%	54%	47%
	LIWC2015	63%	61%	61%	61%
	VADER	64%	61%	63%	61%
Conjunto de dados: MQDGT					
	Léxico/ Base de Conhecimento	Precisão	Revocação	Acurácia	F1-Score
	SenticNet	63%	52%	55%	48%
	LIWC2015	63%	61%	61%	61%
	VADER	65%	61%	63%	61%

MQDPT em todas as métricas.

Tais resultados estão a longe dos mais de 80% de Acurácia apresentados em Cambria et al. [2022] e em Cambria et al. [2020]. Além disso, o método de *sentic computing* obteve os piores resultados quando comparados ao demais léxicos testados; LIWC2015, em sua versão em inglês; e VADER.

Tabela 5.7: Diferença Percentual de Desempenho da SenticNet entre o Experimento 1 e 2

Métrica	**MQD00**	**MQD01**	**MQD02**	**MQD03**	**MQD04**	**MQDGT**
Precisão	53,85%	61,54%	56,41%	35,90%	53,85%	61,54%
Revocação	52,94%	55,88%	52,94%	32,35%	50,00%	52,94%
Acurácia	48,65%	51,35%	48,65%	29,73%	45,95%	48,65%
F1-Score	56,67%	63,33%	56,67%	36,67%	56,67%	60,00%
Média	53,53%	58,03%	53,67%	33,66%	51,87%	55,78%

O melhor resultado de Acurácia do VADER no experimento de Cambria et al. [2020] foi com o STS Dataset. O valor obtido foi de $64,90\%$, ou $0,65$, em uma classificação binária, ou seja, sem os sentimentos neutros. Esse resultado é bem próximo dos resultados encontrados nas métricas de Acurácia dos *corpora* paralelos MQD00, MQD02, MQD03 3 MQDGT. O MQD01 obteve $0,08$

pontos a mais do que Cambria et al. [2020], conforme pode ser observado na Tabela 5.6, ressaltando que esta etapa do experimento contemplou a classe neutra, ao contrário de [Cambria et al., 2020]. Isso nos leva a inferir que VADER tem um desempenho superior para analisar sentimentos em inglês quando comparado a LIWC2015 e SenticNet.

Em Cambria et al. [2022], os autores repetiram os testes com o VADER e SenticNet (versão 7) com conjuntos de dados diferentes, novamente, excluindo o sentimento neutro de sua classificação e utilizando a acurácia como métrica principal. Os resultados apontaram melhor desempenho da SenticNet em relação ao léxico VADER. Os resultados obtidos foram, respectivamente 83.60% e 75.18%. Em seu mais recente experimento [Cambria et al., 2022], o autor também acrescentou o LIWC em seu *benchmarking*, porém, sem mencionar a versão. De modo a tornar nosso experimento justo, utilizamos a versão em inglês do LIWC2015, por ser a mais próxima das versões da SenticNet do que o LIWC2007. Segundo os autores, o melhor resultado de acurácia para o LIWC foi com o conjunto de dados STS [Cambria et al., 2022], cujo resultado foi de $67,68\%$.

5.3.3 Classificação em positivo e negativo

Da mesma forma como no Experimento 1, reproduzimos os testes com todos os *corpora* paralelos sem o sentimento neutro, de modo a tornar o experimento mais justo e semelhante ao de Cambria et al. [2020] e Cambria et al. [2022]. As comparações entre os resultados das classes positiva (1) e negativa (-1), para SenticNet, LIWC2015 e VADER, se encontram nas Figuras 5.10-5.15.

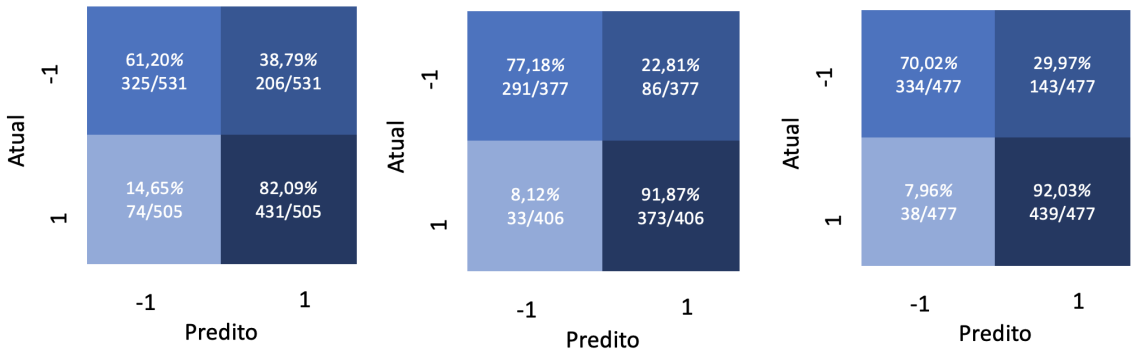
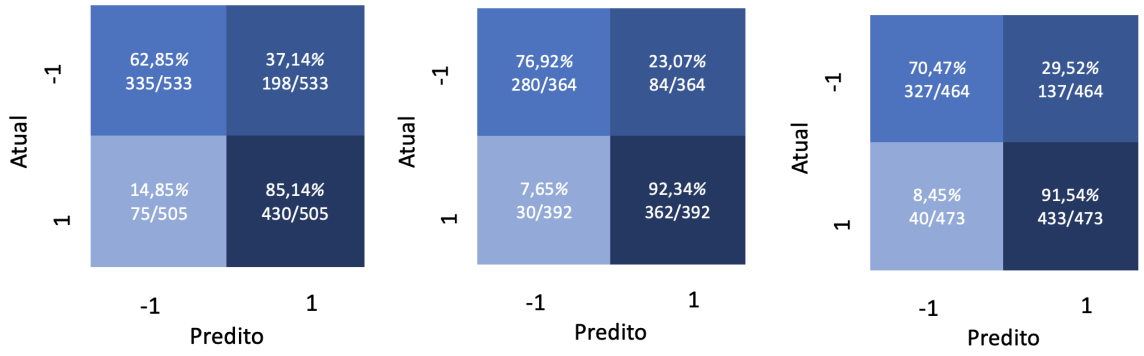
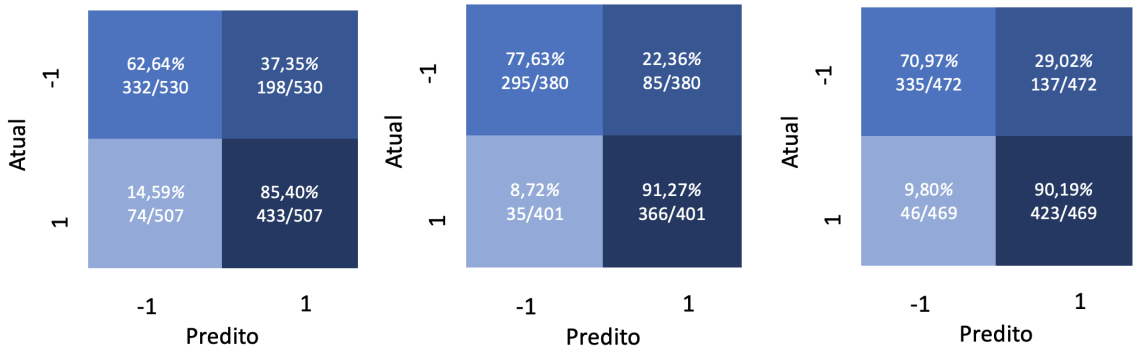


Figura 5.10: Matrizes de confusão do conjunto de dados MQD00: (a) SenticNet, (b) LIWC2015 e (c) VADER. Os rótulos positivo e negativo são representados, respectivamente, por '1' e '-1'. O eixo X representa as classes preditas e o eixo Y as classes atuais.



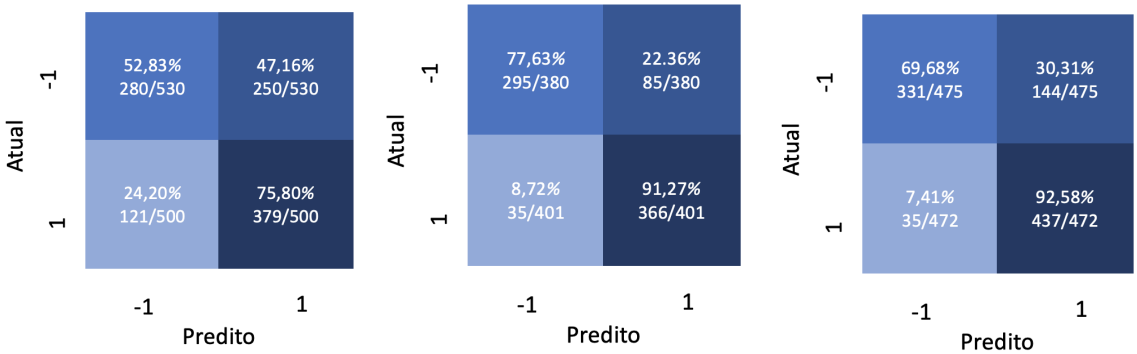
(a) (b) (c)

Figura 5.11: Matrizes de confusão do conjunto de dados MQD01: (a) SenticNet, (b) LIWC2015 e (c) VADER. Os rótulos positivo e negativo são representados, respectivamente, por '1' e '-1'. O eixo X representa as classes preditas e o eixo Y as classes atuais.



(a) (b) (c)

Figura 5.12: Matrizes de confusão do conjunto de dados MQD02: (a) SenticNet, (b) LIWC2015 e (c) VADER. Os rótulos positivo e negativo são representados, respectivamente, por '1' e '-1'. O eixo X representa as classes preditas e o eixo Y as classes atuais.



(a) (b) (c)

Figura 5.13: Matrizes de confusão do conjunto de dados MQD03: (a) SenticNet, (b) LIWC2015 e (c) VADER. Os rótulos positivo e negativo são representados, respectivamente, por '1' e '-1'. O eixo X representa as classes preditas e o eixo Y as classes atuais.

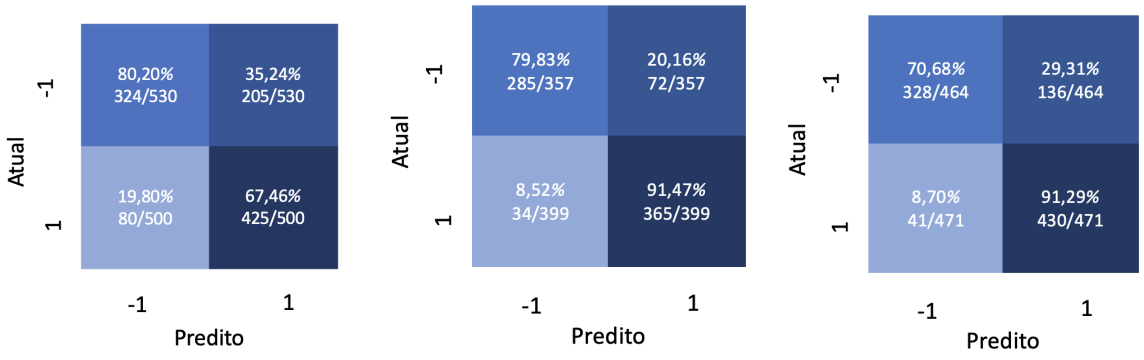


Figura 5.14: Matrizes de confusão do conjunto de dados MQD04: (a) SenticNet, (b) LIWC2015 e (c) VADER. Os rótulos positivo e negativo são representados, respectivamente, por '1' e '-1'. O eixo X representa as classes preditas e o eixo Y as classes atuais.

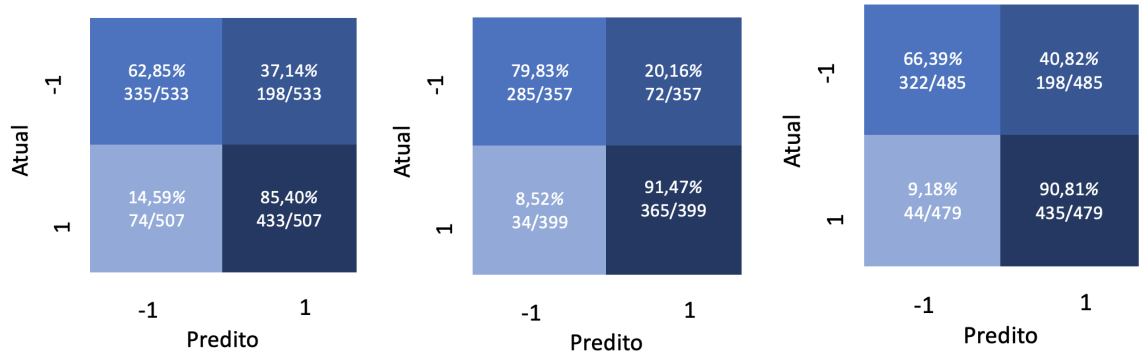


Figura 5.15: Matrizes de confusão do conjunto de dados MQDGT: (a) SenticNet, (b) LIWC2015 e (c) VADER. Os rótulos positivo e negativo são representados, respectivamente, por '1' e '-1'. O eixo X representa as classes preditas e o eixo Y as classes atuais.

Os resultados das MCs foram consolidados na Tabela 5.8. No experimento com o MQD original, em português, a classe de sentimento positivo (1) teve $55,48\%$ de acertos, já a classe o negativo (-1) obteve $44,12\%$ de classificações corretas. Ao comparar com o segundo experimento com os *corpora* paralelos sem o sentimento neutro, este índice aumentou significativamente, chegando a mais de 80% para sentimentos positivos e 77% para os negativos.

Os conjuntos de dados traduzidos para o inglês com melhor desempenho para a SenticNet foram o MQD01, MQD02 e MQDGT, que atingiram, respectivamente, $85,14\%$, $85,40\%$ e $85,40\%$ na classe de sentimentos positivos (1). Já para a classe de sentimentos negativos (-1), os valores foram $62,85\%$, $62,84\%$ e $62,85\%$. Desta forma, o MQDGT teve desempenho minimamente maior em relação aos demais conjuntos de dados traduzidos por humanos, com exceção do MQD03, que obteve apenas $75,80\%$ de acertos na classe positiva e $52,83\%$ na negativa. No conjunto de dados MQD04, a SenticNet chegou a $80,20\%$ de acertos na classe Negativa (-1), gerando um desvio no

padrão observado nos demais conjuntos de dados. Isso nos leva a concluir que a tradução pode, de fato, influenciar e enviesar linguisticamente as análises de sentimento.

Ainda que o desempenho da SenticNet tenho melhorado bastante em relação ao Experimento 1, não houve significativa mudança em relação à classificação dos sentimentos neutros presentes nos *corpora* paralelos. Além disso, a abordagem de sentic computing não obteve resultados superiores quando comparada a métodos tradicionais baseados em léxicos. Na Tabela 5.8, observa-se que a versão em inglês do LIWC2015 obteve a maior precisão na maioria das vezes para ambas as classes. No entanto, as métricas de melhor desempenho ficaram com o VADER nos conjuntos de dados MQD00 e MQD03, com acertos nos sentimentos positivos de 92,03% e 92,58%, respectivamente.

Embora os resultados utilizando a abordagem de sentic computing tenham mostrado alguma melhora, eles ainda ficaram aquém dos métodos tradicionais, como o LIWC2015 e o VADER. As métricas de Precisão, Revocação, Acurácia e F1-Score indicaram que o LIWC2015 teve um desempenho superior em comparação com a SenticNet e o VADER, mesmo sem incluir a classificação de sentimento neutro (0). Os conjuntos de dados com melhores resultados foram MQD04 e MQDGT, apresentando valores de 0,86 em todas as métricas.

Apesar de o conjunto de dados MQDGT ter sido a versão com os valores mais altos de acertos nos métodos de classificação, esses valores não são significativos o suficiente para se comparar os efeitos da utilização de traduções automáticas quando confrontados com traduções realizadas por um indivíduo com bons conhecimentos de inglês. Sendo assim, a premissa de que as traduções automáticas trariam prejuízo a tarefas de PLN [da Silva Conrado et al., 2014], não podem ser sustentadas no caso de análises de sentimento. Desta forma, os experimentos foram inconclusivos para averiguar tal questão, carecendo de novos experimentos que abordassem todos os possíveis problemas de PLN e não apenas análise de sentimento. Porém, foi possível demonstrar as influências sócio e psico-linguísticas dos tradutores na elaboração de novos conjuntos de dados originários de *corpora* paralelos.

Os resultados das métricas de Precisão, Revocação, Acurácia e F1-Score se encontram na Tabela 5.9 para todos os *corpora* paralelos MQD00, MQD01, MQD02, MQD03, MQD04 e MQDGT (sem o sentimento Neutro) e seus respectivos métodos de classificação: SenticNet, LIWC2015 e VADER. O primeiro teste, com o MQD-1465, original em português, obteve os seguintes resultados para Precisão, Revocação, Acurácia e F1-Score: $0,51$, $0,51$ e $0,51$ e $0,51$, respectivamente.

Novamente, o experimento com os conjuntos de dados em inglês obteve melhora quando comparado ao Experimento 1, ainda que sem a classe de sentimentos neutros, cujo melhor desempenho foi observado na utilização do conjunto de dados MQD02. Os resultados obtidos nesse caso foram de $0,75$ para Precisão, $0,74$ para Revocação, $0,74$ para Acurácia e $0,74$ para F1-Score. Os demais *corpora* paralelos obtiveram resultados muito próximos, conforme pode ser observado na Tabela 5.8.

Ao compararmos os resultados obtidos neste trabalho reproduzindo os mesmos testes elaborados em Cambria et al. [2020] e em Cambria et al. [2022], é possível observar que a SenticNet ainda não atingiu resultados esperados de excelência, com valores superiores a 80% de Acurácia. Além disso, tanto o LIWC2015 quanto o VADER obtiveram valores bem superiores aos apresentados nos testes originais mais recentes com o método de *sentic computing* [Cambria et al., 2020, 2022].

5.4 Experimento 3: Comparações com o GPT-3.5

Com as recentes inovações advindas do uso de modelos *trasnformers* e LLMs [Qihuang et al., 2023], executou-se um teste com a nova ferramenta da OpenAI, o ChatGPT [OpenAI, 2022]. O modelo foi lançado em Novembro de 2022 e é a plataforma com maior crescimento de usuários da atualidade [Leiter et al., 2024].

A realização deste experimento teve como objetivo testar o desempenho do novo método que utiliza modelos de *Generative Pre-Trained Transformer* (GPT) com LLM em tarefas de classificação e AS na língua portuguesa. Ademais, por tratar-se de um modelo de aprendizado profundo, o experimento com o ChatGPT [OpenAI, 2022], ainda que embrionário, traz escopo comparativo a este experimento e modelos de aprendizado profundo executados dentro da SenticNet 6 e 7.

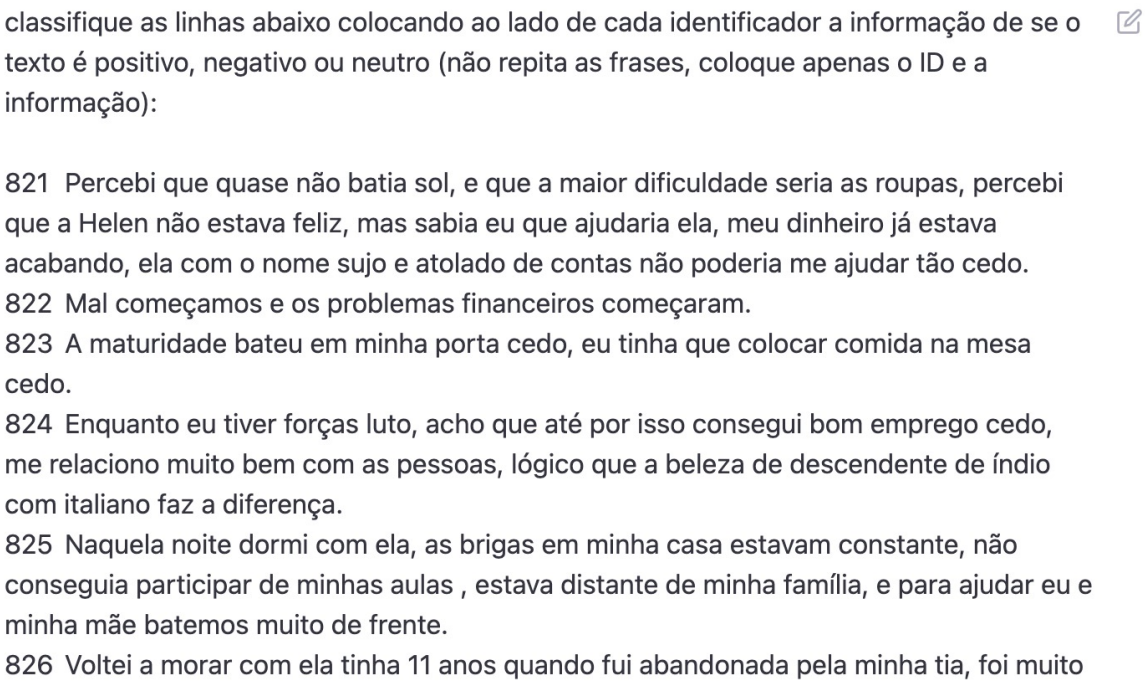
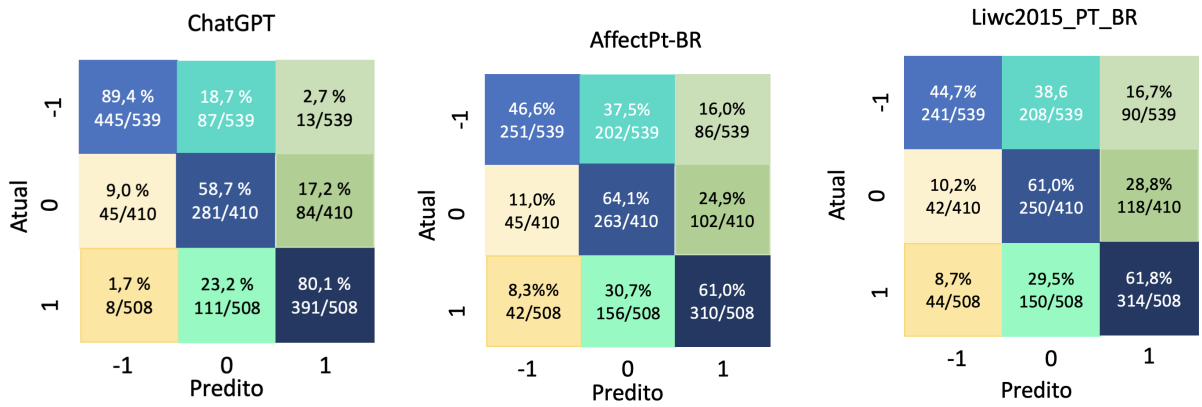


Figura 5.16: Prompt e recorte da saída do ChatGPT para Classificação de Sentimentos na Língua Portuguesa.

Utilizou-se o conjunto de dados MQD-1465 original, em português, e comparou-se o ChatGPT com o método de *sentic computing*, por meio da SenticNet, e com léxicos de melhor desempenho no *benchmarking* de Nascimento [2021]. Os resultados podem ser observados na Figura 5.17.



(a) Matriz de confusão do ChatGPT. (b) Matriz de confusão do AffectPT-BR. (c) Matriz de confusão do LIWC2015-PT-br.

Figura 5.17: Matrizes de confusão do: (a) ChatGPT, (b) AffectPT-BR e (c) LIWC2015-PT-br. Os rótulos positivo, negativo e neutro são representados, respectivamente, por '1', '-1' e '0'. O eixo X representa as classes preditas e o eixo Y as classes atuais.

A análise indicou que o ChatGPT teve o melhor desempenho em tarefas de AS em português, ainda que não haja um consenso sobre sua superioridade no que concerne a tarefas de classificação no idioma inglês [Qihuang et al., 2023; Amin et al., 2023].

Em comparação com a SenticNet [Cambria et al., 2022], que obteve $44,69\%$ de acertos na classe positiva (1), $54,83\%$ na negativa (-1) e $1,99\%$ na neutra (0), o ChatGPT obteve $44,71\%$ pontos percentuais a mais para sentimentos positivos, *i.e* $89,4\%$; $25,27\%$ para os negativos e $80,0\%$ para os neutros. O léxico AffectPT-BR [Carvalho et al., 2018] foi a ferramenta com melhor desempenho nos experimentos de Nascimento [2021], seguido pelo LIWC2015-PT-br [Carvalho et al., 2019, 2023].

Já quanto ao desempenho de métricas de avaliação de modelos de aprendizado de máquina, o ChatGPT obteve 76% de Precisão, 76% de Revocação, 76% de Acurácia e 75% para *F1-Score*. O AffectPT-br, obteve 61%, 57%, 57%, 57%, para cada métrica respectivamente. Quanto ao LIWC2015-PT-br, seus resultados foram de 60%, 55%, 55% e 56%. A SenticNet ficou em último lugar, com $0,39$ de precisão, 34% em revocação, 37% de Acurácia e 30% de *F1-Score*. Levando em conta o padrão de avaliação dos experimentos de Cambria et al. [2020] e Cambria et al. [2022] com a Acurácia, e de [Nascimento, 2021] que utilizou como padrão ouro de análise o *F1-Score*, a SenticNet perde em 39% para Acurácia obtida pelo ChatGPT e 46% no *F1-Score*.

Este experimento indica que os modelos de LLM possuem grande potencial para tarefas de classificação e AS no idioma português, ainda que seus resultados em inglês não tenham sido significativos para alguns autores [Qihuang et al., 2023; Amin et al., 2023]. Além disso, o teste pôde trazer mais um método de comparação diverso, com uso de um *framework* tão robusto quanto a SenticNet, além da avaliação em relação aos léxicos.

Tabela 5.8: Tabela de comparação entre os *corpora* paralelos MQD00, MQD01, MQD02, MQD03, MQD04 e MQDGT, em inglês, com porcentagem de acertos para cada uma das bases de conhecimento ou léxicos SenticNet, LIWC2015 e VADER, sem o sentimento neutro (0).

Conjunto de dados: MQD00		
Léxico/ Base de Conhecimento	Positivo	Negativo
SenticNet	82,09%	61,20%
LIWC2015	91,87%	77,18%
VADER	92,03%	70,02%
Conjunto de dados: MQD01		
Léxico/ Base de Conhecimento	Positivo	Negativo
SenticNet	85,14%	62,85%
LIWC2015	92,34%	76,92%
VADER	91,54%	70,47%
Conjunto de dados: MQD02		
Léxico/ Base de Conhecimento	Positivo	Negativo
SenticNet	85,40%	62,64%
LIWC2015	91,27%	77,63%
VADER	90,19%	70,97%
Conjunto de dados: MQD03		
Léxico/ Base de Conhecimento	Positivo	Negativo
SenticNet	75,80%	52,83%
LIWC2015	91,27%	77,63%
VADER	92,58%	69,68%
Conjunto de dados: MQD04		
Léxico/ Base de Conhecimento	Positivo	Negativo
SenticNet	67,46%	80,20%
LIWC2015	91,47%	79,83%
VADER	91,29%	70,68%
Conjunto de dados: MQDGT		
Léxico/ Base de Conhecimento	Positivo	Negativo
SenticNet	85,40%	62,85%
LIWC2015	91,47%	79,83%
VADER	90,81%	66,39%

Tabela 5.9: Comparação de resultados das métricas de Precisão, Revocação, Acurácia e F1-Score para cada o *corpus* paralelo MQD00, em inglês, sem o sentimento neutro (0) da classificação de AS para SenticNet, LIWC2015 e VADER.

Conjunto de dados: MQD00					
	Léxico/ Base de Conhecimento	Precisão	Revocação	Acurácia	F1-Score
	SenticNet	75%	73%	73%	73%
	LIWC2015	86%	85%	85%	85%
	VADER	83%	81%	81%	81%
Conjunto de dados: MQD01					
	Léxico/ Base de Conhecimento	Precisão	Revocação	Acurácia	F1-Score
	SenticNet	75%	74%	74%	73%
	LIWC2015	86%	85%	85%	85%
	VADER	83%	81%	81%	81%
Conjunto de dados: MQD02					
	Léxico/ Base de Conhecimento	Precisão	Revocação	Acurácia	F1-Score
	SenticNet	75%	74%	74%	74%
	LIWC2015	85%	84%	85%	85%
	VADER	82%	81%	81%	80%
Conjunto de dados: MQD03					
	Léxico/ Base de Conhecimento	Precisão	Revocação	Acurácia	F1-Score
	SenticNet	65%	64%	64%	64%
	LIWC2015	86%	85%	85%	85%
	VADER	83%	81%	81%	81%
Conjunto de dados: MQD04					
	Léxico/ Base de Conhecimento	Precisão	Revocação	Acurácia	F1-Score
	SenticNet	74%	73%	72%	72%
	LIWC2015	86%	86%	86%	86%
	VADER	82%	81%	81%	81%
Conjunto de dados: MQDGT					
	Léxico/ Base de Conhecimento	Precisão	Revocação	Acurácia	F1-Score
	SenticNet	75%	73%	74%	74%
	LIWC2015	86%	86%	86%	86%
	VADER	80%	79%	79%	78%

Capítulo 6 Considerações Finais

Este trabalho apresentou uma revisão de abordagens ontológicas de AS, bem como o levantamento de algumas reflexões sobre possíveis estratégias de expansão para o campo de AS no cenário do português brasileiro. O principal objetivo desta pesquisa foi tentar responder a três questões de pesquisa centrais:

- PP1) Como as traduções automáticas e humanas se comparam em termos de impacto na classificação de Análise de Sentimentos?
- PP2) Soluções mais robustas superariam os métodos tradicionais de classificação de sentimento baseados em léxicos?
- PP3) O cenário de PLN em português do Brasil necessita de uma ontologia baseada em sentimentos para superar seus métodos tradicionais atuais de classificação de análise de sentimento?

6.1 PP1) Como as traduções automáticas e humanas se comparam em termos de impacto na classificação de Análise de Sentimentos?

O principal objetivo com os experimentos deste trabalho foi avaliar as premissas levantadas por da Silva Conrado et al. [2014], que mencionam o prejuízo trazido por traduções automáticas a tarefas de PLN para a língua portuguesa, devido à diferença estrutural idiomática em relação ao inglês. Além disso, os testes visaram avaliar as proposições de Cambria and Hussain [2012]; Cambria et al. [2017]; Jiang and Tan [2005]; Liu et al. [2012] e Rachid et al. [2018] quanto à possibilidade de um melhor desempenho em tarefas de AS com o uso de ontologias, bem como avaliar a premissa de Cambria and White [2014] quanto à necessidade de modelos mais robustos para tarefas de PLN e AS.

A comparação da SenticNet 6 com o *benchmarking* de léxicos, levantado por Nascimento [2021], indicou que léxicos e demais ferramentas traduzidas diretamente por humanos têm melhor desempenho ao analisar sentimentos neutros. Possivelmente por terem maior capacidade ao analisar as sutilezas linguísticas que tornam um comentário neutro. Porém, numa avaliação geral, o conjunto de dados MQDGT teve desempenho semelhante ao idioma inglês para os léxicos LIWC2015 e VADER. A ferramenta com maior discrepância para sentimentos neutros foi a SenticNet, tanto em portu-

guês quanto em inglês devido às características de análise percorrendo os extremos das polaridades. Desta forma, é possível inferir que o relativo prejuízo na classificação de sentimentos neutros é uma característica das próprias ferramentas.

Nossa investigação revelou que a SenticNet, na sua forma original, não apresentou bons resultados em português. No entanto, ao traduzir os textos, criando um *corpus* paralelo, observamos uma melhora significativa nas métricas de desempenho. Esta descoberta é de grande importância, pois mostra que a simples avaliação da ferramenta original teria levado a resultados insatisfatórios. Ao adotar uma abordagem mais profunda e explorar soluções alternativas, conseguimos aprimorar os resultados, evidenciando a relevância e o valor da nossa análise para futuras pesquisas no campo e revelando a importância do uso de traduções para avaliações e testes com *frameworks* com tradução automática.

Seria importante acrescentar aos conjuntos de *corpora* paralelo uma tradução formal, executada por um nativo da língua inglesa, porém os recursos necessários para tal seriam muito custosos, tanto temporal quanto monetário. É importante ressaltar que traduções manuais, sejam eles com nativos ou não, demandam bastante tempo, conforme já fora observado por García-Díaz et al. [2020], desta forma, parte do tempo para a execução de teste é gasto com traduções. Não obstante, o *corpus* paralelo desenvolvido nessa pesquisa permitirá novos testes de engenharia reversa, possibilitando a avaliação de novas ferramentas ou soluções de AS que considerem as particularidades de estruturas linguísticas.

Os testes demonstraram que o suporte da SenticNet para a língua portuguesa é limitado, porém, apesar de não ter tido o melhor desempenho em todos os casos, a ferramenta *Google Translate* se mostrou bastante eficiente na tradução do conjunto de dados MQD-1465, ficando pouco atrás dos conjuntos de dados paralelos traduzidos por humanos. O uso da API, com chamada para a ferramenta, acelera bastante o processo de tradução e o experimento demonstrou que a mesma não trouxe prejuízos para as tarefas de AS em nenhum dos léxicos ou base de conhecimento utilizados nos testes.

6.2 PP2) Soluções mais robustas superariam os métodos tradicionais de classificação de sentimento baseados em léxicos?

Embora existam novas abordagens de ontologia, com destaque para os trabalhos de Rosa et al. [2019] e Freitas and Vieira [2013], a maioria das pesquisas para o português brasileiro se concentra em abordagens léxicas que reduzem a PLN e a AS a um simples problema de classificação estatística. Vários estudos no exterior, principalmente em inglês, já comprovam que estruturas linguísticas mais elaboradas requerem ferramentas de mineração de texto mais robustas [Liu et al., 2012].

Estudos e abordagens ontológicas poderiam ser de grande valor quando aplicados às áreas de

marketing ao analisar avaliações de usuários para os diferentes setores da sociedade, uma vez que, no caso de *reviews* de produtos, pessoas ou governos, as opiniões poderiam guardar relações entre si [Liu et al., 2012]. Porém, até a elaboração deste trabalho, há poucos estudos e aplicações ontológicas com foco em AS. Parte da pesquisa apresentada neste trabalho foi publicada no ONTOBRAS 2022, artigo intitulado *A Review of Ontology-Based Approaches for Sentiment Analysis: Possible Improvements on the Brazilian Affective Computing Scenario* [Carneiro et al., 2022].

Assim, ainda que os estudos levantados tenham sugerido que estruturas de dados semanticamente mais robustas possam ter melhor desempenho do que léxicos mais tradicionais, este não foi o caso da SenticNet. É importante destacar que métodos de aprendizado profundo, como LLM tiveram melhor desempenho do que léxicos e a SenticNet.

6.3 PP3) O cenário de Processamento de Linguagem Natural em português do Brasil necessita de uma ontologia baseada em sentimentos para superar seus métodos tradicionais atuais de classificação de análise de sentimento?

Ainda são poucos os métodos para análise de sentimentos na língua portuguesa, com evidente preferência pelos léxicos. Quanto ao uso de ontologias, os testes com a SenticNet demonstraram que o *framework* de *sentic computing* ainda está longe de seu potencial de execução quando comparado com os métodos mais tradicionais. Porém, partindo de premissas, tanto sob a ótica linguística, quanto da teoria da computação voltada para a sub-área de PLN e AS, a metodologia tem bastante fundamentação. Sendo assim, é importante observar as futuras possíveis contribuições da *sentic computing* para o campos da linguística computacional, principalmente com o advento de novos recursos e inovações em modelos já consolidados como os LLMs e GPTs.

O teste com o ChatGPT, da OpenAI [OpenAI, 2022], ainda que um simples piloto foi importante para evidenciar que, conforme observado por Cambria and White [2014], tarefas de PLN e AS podem ter melhor desempenho com ferramentas mais robustas, porém os experimentos aqui discutidos indicaram que, atualmente, uma abordagem ontológica sem um domínio específico pode não ser a melhor solução. Ao contrário da MFOEM, a SenticNet não é exatamente uma ontologia de análise de sentimentos, mas tem sua base em onto-pares originados dos conceitos primitivos presentes num dado *corpus* de palavras de domínio aberto e utilizar uma ontologia sem domínio é algo relativamente complexo, já que ontologias tratam sempre da representação formal de algo.

Há pouco material no estado-da-arte no Brasil a respeito de ontologias de análise de sentimentos e *sentic computing*, sendo a maioria das pesquisas voltadas para léxicos. Esta pesquisa traz novo material para a literatura, com uma revisão de possíveis métodos com base em ontologias que pode auxiliar o desenvolvimento de novas pesquisas no Brasil. Além disso, apresentamos um novo *corpus* paralelo inédito, versionado em inglês, de um conjunto de dados de AS, em português.

Tais conjuntos de dados podem ser utilizados em testes comparativos para avaliações de novas ferramentas de AS que surjam para a língua portuguesa, permitindo uma comparação mais justa com outros recursos desenvolvidos apenas para o inglês e trazendo maior equilíbrio nas verificações e testes. Porém na atual conjuntura, não há necessidade do desenvolvimento de uma ontologia voltada apenas para a classificação de sentimento, sendo os modelos de aprendizado profundo, como os LLMs, ou mesmo léxicos tradicionais, suficientes para a solução de tais problemas, ainda que haja espaço para melhorias. É importante destacar que, até o momento, ainda que os LLMs tenham grande capacidade de inferência e geração de textos, o cenário de PLN e AS segue com os mesmos problemas já observados em diversos estudos: a dificuldade do acompanhamento da evolução da linguagem humana em tempo hábil, gerando ambiguações ou erro de classificação em textos com gírias, jargões e excesso de linguagem coloquial, conforme relatado por Jung et al. [2017].

Este estudo evoca novas preocupações e estudos no campo acadêmico brasileiro de PLN e AS, uma vez que ferramentas multilíngues e traduções automáticas, como as presentes na SenticNet, não oferecem um desempenho adequado para conjuntos de dados com grandes variações linguísticas como *tweets*. Para trabalhos futuros, esperamos desenvolver um protótipo de uma ontologia de sentimento baseada na pesquisa de Smith [2019], e Rosa et al. [2019], bem como realizar novos experimentos com a SenticNet 6 e com a Senpy. Pretendemos, ainda, realizar novas pesquisas sobre a WordNet em OWL e WordNet Affect, a fim de avaliar a possibilidade do mapeamento da WordNet Affect BR para OWL, pensando num relacionamento ou derivação da WordNet BR, caso esta já esteja em OWL.

Devido à carência de ontologias com foco em análise de sentimento na língua portuguesa, pretendemos fazer uma tradução da MFOEM [Smith, 2019]. Ademais, temos como objetivo executar uma análise comparativa com diferentes métodos de AS, *Bag of Words* (BOW), técnicas de aprendizado por comitês com léxicos, bem como modelos de aprendizado de máquina, como BERT (*Bidirectional Encoder Representations for Transformers* [Devlin et al., 2019], RoBERTa (*Robustly Optimized BERT Pretraining Approach* [Liu et al., 2019], LLM, na língua portuguesa. Para tanto, será necessário tempo para o desenvolvimento dos diferentes modelos de AM.

Um *benchmark* comparativo dos principais *frameworks* consolidados no mercado também será elaborado, bem como a avaliação de diferentes métodos de tradução e melhor desempenho em desambiguação. No que concerne aos testes com os modelos de LLM, visamos dar início a pesquisas com arquiteturas de grafos, como Graph RAG e conjuntos de pré-treino com base em ontologias e avaliar o desempenho dessas classificações.

Referências Bibliográficas

Ali, F., Kwak, D., Khan, P., Islam, S. R., Kim, K. H., and Kwak, K. (2017). Fuzzy ontology-based sentiment analysis of transportation and city feature reviews for safe traveling. *Transportation Research Part C: Emerging Technologies*, 77:33–48.

Allen, J. F. (1988). *Natural Language Understanding*. Benjamin-Cummings Publishing Company, São Francisco, Estados Unidos da América.

Almeida, M., Pinto, C., Figueira, H., Mendes, P., and Martins, A. (2015). Aligning opinions: Cross-lingual opinion mining with dependencies. In *Annual Meeting of the Association for Computational Linguistics - ACL*.

Amin, M. M., Cambria, E., and Schuller, B. W. (2023). Will affective computing emerge from foundation models and general artificial intelligence? a first evaluation of chatgpt. *IEEE Intelligent Systems*, 38(02):15–23.

Arendt, D., Shaw, Z., Shrestha, P., Ayton, E., Glenski, M., and Volkova, S. (2021). Crosscheck: Rapid, reproducible, and interpretable model evaluation. In *Proceedings of the Second Workshop on Data Science with Human in the Loop: Language Advances*. Association for Computational Linguistics.

Azevedo, G. d., Pettine, G., Feder, F., Portugal, G., Schocair Mendes, C. O., Castaneda Ribeiro, R., Mauro, R. C., Paschoal Júnior, F., and Guedes, G. (2021a). Mqd-1465. `https://github.com/LaCAfe/MQD-1465`. Accessed: 27 Mar. 2024.

Azevedo, G. d., Pettine, G., Feder, F., Portugal, G., Schocair Mendes, C. O., Castaneda Ribeiro, R., Mauro, R. C., Paschoal Júnior, F., and Guedes, G. (2021b). NAT: Towards an emotional agent. In *2021 16th Iberian Conference on Information Systems and Technologies (CISTI)*, pages 1–4.

Baccianella, S., Esuli, A., and Sebastiani, F. (2010). SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In Calzolari, N., Choukri, K., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S., Rosner, M., and Tapias, D., editors, *Proceedings of the*

Seventh International Conference on Language Resources and Evaluation (LREC'10), Valletta, Malta. European Language Resources Association (ELRA).

Borth, D., Ji, R., Chen, T., Breuel, T., and Chang, S.-F. (2013). Large-scale visual sentiment ontology and detectors using adjective noun pairs. In *Proceedings of the 21st ACM International Conference on Multimedia*, MM '13, page 223–232, Nova Iorque, Estados Unidos da América. Association for Computing Machinery.

Borth, D., Ji, R., Chen, T., Breuel, T., and Chang, S.-F. (2014). Visual Sentiment Ontology. `http://visual-sentiment-ontology.appspot.com/`. Acessado: 14 fev. 2021.

Broden, T. F. (2018). Semiologia/semiótica em saussure e jakobson: conceitos, filiações, debates. *Revista do GELNE*, 19:299–309.

Brum, H. and Volpe Nunes, M. d. G. (2018). Building a sentiment corpus of tweets in Brazilian Portuguese. In Calzolari, N., Choukri, K., Cieri, C., Declerck, T., Goggi, S., Hasida, K., Isahara, H., Maegaard, B., Mariani, J., Mazo, H., Moreno, A., Odijk, J., Piperidis, S., and Tokunaga, T., editors, *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japão. European Language Resources Association (ELRA).

Cambria, E. and Hussain, A. (2012). *Sentic Computing: Techniques, Tools, and Applications*. SpringerBriefs in Cognitive Computation. Springer, Dordrecht, Países Baixos.

Cambria, E. and Hussain, A. (2015). *Sentic computing: A Common-Sense-Based Framework for Concept-Level Sentiment Analysis*. Socio-Affective Computing. Springer International Publishing, Cham, Suíça, 1 edition.

Cambria, E., Li, Y., Xing, F. Z., Poria, S., and Kwok, K. (2020). Senticnet 6: Ensemble application of symbolic and subsymbolic ai for sentiment analysis. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, CIKM '20, page 105–114, Nova Iorque, Estados Unidos da América. Association for Computing Machinery.

Cambria, E., Liu, Q., Decherchi, S., Xing, F., and Kwok, K. (2022). SenticNet 7: A commonsense-based neurosymbolic AI framework for explainable sentiment analysis. In Calzolari, N., Béchet, F., Blache, P., Choukri, K., Cieri, C., Declerck, T., Goggi, S., Isahara, H., Maegaard, B., Mariani, J., Mazo, H., Odijk, J., and Piperidis, S., editors, *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 3829–3839, Marseille, França. European Language Resources Association.

Cambria, E., Poria, S., Gelbukh, A., and Thelwall, M. (2017). Sentiment analysis is a big suitcase. *IEEE Intelligent Systems*, 32(6):74–80.

Cambria, E. and White, B. (2014). Jumping NLP curves: A review of natural language processing research [review article]. *IEEE Computational Intelligence Magazine*, 9(2):48–57.

Carneiro, E. (2022). senticnet-6-test-2022. `https://github.com/EricaCQ/senticnet-6-test-2022`. Accessed: 17 mar. 2024.

Carneiro, E., Guedes, G. P., and Belloze, K. T. (2022). A review of ontology-based approaches for sentiment analysis: Possible improvements on the brazilian affective computing scenario. In de Oliveira, J. M. P., Schmidt, D., Garcia, L. F., and Felipe, E. R., editors, *Proceedings of the XV Seminar on Ontology Research in Brazil (ONTOBRAS 2022) and VI Doctoral and Masters Consortium on Ontologies (WTDO 2022)*, volume 3346 of *CEUR Workshop Proceedings*, pages 127–136, online, Brasil. CEUR-WS.org.

Carvalho, F., Junior, F. P., Ogasawara, E., Ferrari, L., and Guedes, G. (2023). Evaluation of the brazilian portuguese version of linguistic inquiry and word count 2015 (BP-LIWC2015). *Language Resources and Evaluation*, 58(1):203–222.

Carvalho, F., Rodrigues, R., Santos, G., Cruz, P., Ferrari, L., and Guedes, G. (2019). Evaluating the Brazilian Portuguese version of the 2015 LIWC Lexicon with sentiment analysis in social networks. In *Anais do VIII Brazilian Workshop on Social Network Analysis and Mining*, pages 24–34, Porto Alegre, RS, Brasil. SBC.

Carvalho, F., Santos, G., and Guedes, G. P. (2018). AffectPT-br: an affective lexicon based on LIWC 2015. In *2018 37th International Conference of the Chilean Computer Science Society (SCCC)*, pages 1–5.

Carvalho, P. and Silva, M. (2015). SentiLex-PT: Principais características e potencialidades. *Oslo Studies in Language*, 7.

Ceci, F., Gonçalves, A., and Weber, R. (2016). A model for sentiment analysis based on ontology and cases. *IEEE Latin America Transactions*, 14(11):4560–4566.

Chen, Y. and Skiena, S. (2014). Building sentiment lexicons for all major languages. In Toutanova, K. and Wu, H., editors, *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 383–389, Baltimore, Estados Unidos da América. Association for Computational Linguistics.

Cruz, P., Rodrigues, R., Belloze, K., and Guedes, G. (2017). Uma revisão sistemática sobre léxicos afetivos para o português do Brasil. *Nuevas Ideas en Informatica Educativa*, 13:539 – 544.

da Silva Conrado, M., Felippo, A. D., Salgueiro Pardo, T. A., and Rezende, S. O. (2014). A survey of automatic term extraction for brazilian portuguese. *Journal of the Brazilian Computer Society*, 20(1).

de Kok, S., Punt, L., van den Puttelaar, R., Ranta, K., Schouten, K., and Frasincar, F. (2018). Review-level aspect-based sentiment analysis using an ontology. In *Proceedings of the 33rd Annual ACM Symposium on Applied Computing*, SAC '18, page 315–322, Nova Iorque, Estados Unidos. Association for Computing Machinery.

de Saussure, F. (2008). *Curso de linguística geral*. Editora Cultrix, São Paulo, SP, Brasil, 27 edition.

de Souza, K. F., Pereira, M. H. R., and Dalip Hasan, D. (2017). Unilex: Método léxico para análise de sentimentos textuais sobre conteúdo de tweets em português brasileiro. *Abakós*, 5(2):79–96.

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In Burstein, J., Doran, C., and Solorio, T., editors, *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Di Felippo, A. (2010). The TermiNet project: an overview. page 92–99.

Distante, D., Faralli, S., Rittinghaus, S., Rosso, P., and Samsami, N. (2021). DomainSenticNet: An ontology and a methodology enabling domain-aware sentic computing. *Cognitive Computation*, 14(1):62–77.

Dragoni, M., Poria, S., and Cambria, E. (2018). OntoSenticNet: A commonsense ontology for sentiment analysis. *IEEE Intelligent Systems*, 33(3):77–85.

EMBRAPA (2022). e-termos. `https://www.etermos.cnptia.embrapa.br/index.php`. Accessed: 27 Jul. 2022.

Esuli, A. (2019). Sentiwordnet. `https://github.com/aesuli/SentiWordNet`.

Esuli, A. and Sebastiani, F. (2006). SENTIWORDNET: A publicly available lexical resource for opinion mining. In Calzolari, N., Choukri, K., Gangemi, A., Maegaard, B., Mariani, J., Odijk, J., and Tapias, D., editors, *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)*, Genoa, Italy. European Language Resources Association (ELRA).

Filho, P. P. B., Pardo, T. A. S., and Aluísio, S. 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*.

Fondazione Bruno Kessler (2009a). Wordnet affect. https://wndomains.fbk.eu/wnaffect.html. Accessed: 26 Apr. 2023.

Fondazione Bruno Kessler (2009b). Wordnet domains. https://wndomains.fbk.eu/. Accessed: 26 Abr. 2023.

Freitas, C. (2013). Sobre a construção de um léxico da afetividade para o processamento computacional do português. *Revista Brasileira de Linguística Aplicada*, 13(4):1031–1059.

Freitas, L. A. and Vieira, R. (2013). Ontology based feature level opinion mining for portuguese reviews. In *Proceedings of the 22nd International Conference on World Wide Web*, WWW '13 Companion, page 367–370, Nova Iorque, Estados Unidos da América. Association for Computing Machinery.

Freitas, L. A. D. and Vieira, R. (2019). Ontology-based feature-level sentiment analysis in Portuguese reviews. *International Journal of Business Information Systems*, 32(1):30–55.

Fundação para a Computação Científica Nacional (2022). Linguateca. https://www.linguateca.pt/. Accessed: 27 Jul. 2022.

García-Díaz, J. A., Cánovas-García, M., and Valencia-García, R. (2020). Ontology-driven aspect-based sentiment analysis classification: An infodemiological case study regarding infectious diseases in Latin America. *Future Generation Computer Systems*, 112:641–657.

Google (2024). Cloud translation. https://cloud.google.com/translate/docs/basic/translate-text-basic?hl=pt-br. Acessado: 24 mar. 2024.

Gruber, T. (2009). Ontology. http://tomgruber.org/writing/ontology-definition-2007.htm. Acessado: 18 fev. 2021.

Gruber, T. R. (1993). Towards Principles for the Design of Ontologies Used for Knowledge Sharing. In Guarino, N. and Poli, R., editors, *Formal Ontology in Conceptual Analysis and Knowledge Representation*, Deventer, Países Baixos. Kluwer Academic Publishers.

Guarino, N., Oberle, D., and Staab, S. (2009). What is an ontology? In Staab, S. and Studer, R., editors, *Handbook on Ontologies*, pages 1–17. Springer Berlin Heidelberg, Berlin, Alemanha.

Han, J., Kamber, M., and Pei, J. (2011). *Data Mining: Concepts and Techniques*. Morgan Kaufmann, Oxford, Reino Unido, 3 edition.

Han, L., Jones, G., and Smeaton, A. (2020). AlphaMWE: Construction of multilingual parallel corpora with MWE annotations. In Markantonatou, S., McCrae, J., Mitrović, J., Tiberius,

C., Ramisch, C., Vaidya, A., Osenova, P., and Savary, A., editors, *Proceedings of the Joint Workshop on Multiword Expressions and Electronic Lexicons*, pages 44–57, online. Association for Computational Linguistics.

Hastings, J., Ceusters, W., Smith, B., and Mulligan, K. (2011). The emotion ontology: Enabling interdisciplinary research in the affective sciences. In Beigl, M., Christiansen, H., Roth-Berghofer, T. R., Kofod-Petersen, A., Coventry, K. R., and Schmidtke, H. R., editors, *Modeling and Using Context*, pages 119–123, Berlin, Alemanha. Springer Berlin Heidelberg.

Hitzler, P. (2021). A review of the semantic web field. *Communications of the ACM*, 64(2):76–83.

Hu, M. and Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '04, page 168–177, New York, Estados Unidos da América. Association for Computing Machinery.

Hutto, C. and Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1):216–225.

IBGE (2020). IBGE divulga estimativa da população dos municípios para 2020. `https://agenciadenoticias.ibge.gov.br/agencia-sala-de-imprensa/2013-agencia-de-noticias/releases/28668-ibge-divulga-estimativa-da-populacao-dos-municipios-para-2020`. Acessado: 22 nov. 2020.

Instituto de Letras da UFRGS (2022). TextQuim. `http://www.ufrgs.br/textecc/textquim/`. Accessed: 27 Jul. 2022.

Jiang, X. and Tan, A.-H. (2005). Mining ontological knowledge from domain-specific text documents. In *Fifth IEEE International Conference on Data Mining (ICDM'05)*, pages 4 pp.–.

Jiang, Y. E., Liu, T., Ma, S., Zhang, D., Sachan, M., and Cotterell, R. (2022). A bilingual parallel corpus with discourse annotations. `https://arxiv.org/abs/2210.14667`. arXiv:2210.14667.

Jung, H., Park, H.-A., and Song, T.-M. (2017). Ontology-based approach to social data sentiment analysis: Detection of adolescent depression signals. *Journal of Medical Internet Research*, 19(7):e259.

Jurafsky, D. and Martin, J. (2008). *Speech and Language Processing*. Pearson, Upper Saddle River, Estados Unidos da América, 2 edition.

Karthik, R. and Ganapathy, S. (2021). A fuzzy recommendation system for predicting the customers interests using sentiment analysis and ontology in e-commerce. *Applied Soft Computing*, 108:107396.

Kasama, D. Y., Zavaglia, C., and de Barcellos Almeida, G. M. (2011). Do termo à estruturação semântica: representação ontológica do domínio da nanociência e nanotecnologia utilizando a estrutura quali. *Linguamática*, 2(3):43–58.

Kontopoulos, E., Berberidis, C., Dergiades, T., and Bassiliades, N. (2013). Ontology-based sentiment analysis of Twitter posts. *Expert Systems with Applications*, 40(10):4065–4074.

Kristensen, C. H., Gomes, C. F. d. A., Justo, A. R., and Vieira, K. (2011). Normas brasileiras para o affective norms for english words. *Trends in Psychiatry and Psychotherapy*, 33(3):135–146.

Leiter, C., Zhang, R., Chen, Y., Belouadi, J., Larionov, D., Fresen, V., and Eger, S. (2024). ChatGPT: A meta-analysis after 2.5 months. *Machine Learning with Applications*, 16:100541.

Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Springer International Publishing, Cham, Suíça.

Liu, L., Nie, X., and Wang, H. (2012). Toward a fuzzy domain sentiment ontology tree for sentiment analysis. In *2012 5th International Congress on Image and Signal Processing*, pages 1620–1624.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., and Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach.

Longhi, M. T. (2011). *Mapeamento de aspectos afetivos em um ambiente virtual de aprendizagem*. PhD thesis, UFRGS.

Machado, A., Longhi, M., Nunes, M. A. S. N., and Pardo, T. (2015). Personalitatem lexicon: Um léxico em português brasileiro para mineração de traços de personalidade em textos. In *Simpósio Brasileiro de Informática na Educação*, SBIE. Sociedade Brasileira de Computação - SBC.

Menezes, S. F. and Silva, M. A. (2018). Língua e linguagem para Saussure: Um percurso de elaboração dos conceitos no 1º, 2º e 3º curso. *PERcursos Linguísticos*, 8(20):64–74.

Microsoft (2024). Microsoft Bing. https://www.bing.com/translator. Acessado: 8 abr. 2024.

Miller, G. A. (1995). WordNet: a lexical database for English. *Commun. ACM*, 38(11):39–41.

Moura, M. F., dos Santos, F. F., and Rezende, S. O. (2022). An experimental comparison of label selection methods for hierarchical document clusters. `http://sites.labic.icmc.usp.br/toptax/`. Accessed: 27 Jul. 2022.

Musen, M. A. (2015). The protégé project: a look back and a look forward. *AI Matters*, 1(4):4–12.

Musto, C., Semeraro, G., and Polignano, M. (2014). A comparison of lexicon-based approaches for sentiment analysis of microblog. *CEUR Workshop Proceedings*, 1314:59–68.

Narayanasamy, S. K., Srinivasan, K., Mian Qaisar, S., and Chang, C.-Y. (2021). Ontology-enabled emotional sentiment analysis on covid-19 pandemic-related twitter streams. *Frontiers in Public Health*, 9.

Nascimento, R., Carvalho, F., and Guedes, G. (2019). Identificando sintomas depressivos: um estudo de caso no Youtube. In *Anais do VIII Brazilian Workshop on Social Network Analysis and Mining*, pages 119–130. SBC, Porto Alegre, RS, Brasil.

Nascimento, R. d. S. (2021). Uma análise comparativa entre léxicos afetivos em português. mestrado, CEFET/RJ, Rio de Janeiro, RJ, Brasil.

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.

Nguyen, D. N., Phan, T. T., and Do, P. (2021). Embedding knowledge on ontology into the corpus by topic to improve the performance of deep learning methods in sentiment analysis. *Scientific Reports*, 11(1).

Niles, I. and Pease, A. (2003). Linking lexicons and ontologies: Mapping wordnet to the suggested upper merged ontology. In Arabnia, H. R., editor, *Proceedings of the International Conference on Information and Knowledge Engineering. IKE'03, June 23 - 26, 2003, Las Vegas, Nevada, USA, Volume 2*, pages 412–416. CSREA Press.

Olson, D. L. and Delen, D. (2008). Performance evaluation for predictive modeling. In *Advanced Data Mining Techniques*, pages 137–147. Springer Berlin Heidelberg, Berlin, Alemanha.

OntoLP (2008). OntoLP portal de ontologias. `https://ontolp.inf.pucrs.br/Recursos/downloads-ontolpplugin.php`. Acessado: 22 nov. 2020.

OpenAI (2022). ChatGPT apr 23 version. `https://chat.openai.com/`. Acessado: 23 jun. 2023.

OpenDover (2022). OpenDover website. `https://opendover.nl/`. Acessado:10 jun. 2022.

Pang, B., Lee, L., and Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. In *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP 2002)*, page 79—86. Association for Computational Linguistics.

Parrott, W. G., editor (2000). *Emotions in social psychology*. Key Readings in Social Psychology. Psychology Press, Londres, Reino Unido.

Pasqualotti, P. and Vieira, R. (2008). WordnetAffectBR: uma base lexical de palavras de emoções para a língua portuguesa. *RENOTE*, 6.

Peñalver-Martinez, I., Garcia-Sanchez, F., Valencia-Garcia, R., Rodrıguez-García, M. Á., Moreno, V., Fraga, A., and Sánchez-Cervantes, J. L. (2014). Feature-based opinion mining through ontologies. *Expert Systems with Applications*, 41(13):5995–6008.

Pennebaker, J., Booth, R.J. ANN Boyd, R., and Francis, M. (2015). Linguistic inquiry and word count: Liwc2015. `www.liwc.net`.

Pereira, D. A. (2020). A survey of sentiment analysis in the portuguese language. *Artificial Intelligence Review*, 54(2):1087–1115.

Picard, R. W. (2000). *Affective Computing*. MIT Press, Londres, Reino Unido.

Pino, D. H. P. (2021). Aspectos semânticos da terminologia do biodiesel. Master's disseertation, UFSCAR, São Carlos, SP, Brasil.

Plutchik, R. (1980). *Emotion: A Psychoevolutionary Synthesis*. Longman Higher Education, Harlow, Reino Unido.

Poria, S., Cambria, E., Winterstein, G., and Huang, G.-B. (2014). Sentic patterns: Dependency-based rules for concept-level sentiment analysis. *Knowledge-Based Systems*, 69:45–63.

PortLEX (2022). LIWC. Acessed: 26 Jul. 2022.

Qihuang, Z., Ding, L., Liu, J., Du, B., and Tao, D. (2023). Can ChatGPT understand too? A comparative study on ChatGPT and fine-tuned BERT. arXiv:2302.10198.

Rachid, B. A., Azza, H., and Henda, B. G. (2018). Sentiment analysis approaches based on granularity levels. In *Proceedings of the 14th International Conference on Web Information Systems and Technologies - WEBIST*, pages 324–331. INSTICC, SciTePress.

Reis, J., Gonçalves, P., Araújo, M., Pereira, A., and Benevenuto, F. (2015). Uma abordagem multilíngue para análise de sentimentos. In *Anais do IV Brazilian Workshop on Social Network Analysis and Mining*, page , Porto Alegre, RS, Brasil. SBC.

Rosa, R. L., Schwartz, G. M., Ruggiero, W. V., and Rodríguez, D. Z. (2019). A knowledge-based recommendation system that includes sentiment analysis and deep learning. *IEEE Transactions on Industrial Informatics*, 15(4):2124–2135.

Rossetti, A. and Gaspari, F. (2017). Modelling the analysis of translation memory use and post-editing of raw machine translation output: A pilot study of trainee translators ' perceptions of difficulty and time effectiveness. In Hansen-Schirra, S., Czulo, O., and Hofmann, S., editors, *Empirical modelling of translation and interpreting*, pages 41–67. Language Science Press, Berlim, Alemanha.

Sánchez-Rada, J. F., Iglesias, C. A., Corcuera, I., and Araque, (2016). Senpy: A pragmatic linked sentiment analysis framework. In *2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, pages 735–742.

Saranya, K. and Jayanthy, S. (2017). Onto-based sentiment classification using machine learning techniques. In *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, pages 1–5.

SenticNet (2012a). Sentic apis. `https://sentic.net/api/`. Acessado:15 jun. 2022.

SenticNet (2012b). Sentic computing. `https://sentic.net/computing/`. Acessado:15 jun. 2022.

SenticNet (2022). Senticnet. `https://sentic.net/about/`. Accessed: 26 Jul. 2022.

SenticNet (2024). Senticnet. `https://sentic.net//`. Accessed: 20 Fev. 2024.

Silva, E., Malheiros, Y., Nunes, R. T. A., Antunes, I. L., and do Rêgo, T. G. (2019). Um conjunto de dados extraído do twitter para análise de sentimentos na língua portuguesa. In *XII Symposium in Information and Human Language Technology and Collocates Events*, Salvador, Brasil. Sociedade Brasileira de Computação.

Smith, B. (2018). The ontology of emotions (abril 2018). `https://www.youtube.com/watch?v=BYAT8vHZ8ik`. Acessado:20 January. 2022.

Smith, B. (2019). Introduction to basic formal ontology. `https://www.youtube.com/watch?v=p0buEjR3t8A`. Acessado: 1o Maio 2021.

Socher, R., Chen, D., Manning, C. D., and Ng, A. (2013). Reasoning with neural tensor networks for knowledge base completion. In Burges, C., Bottou, L., Welling, M., Ghahramani, Z., and Weinberger, K., editors, *Advances in Neural Information Processing Systems*, volume 26. Curran Associates, Inc.

Soliman, A. (2023). An unsupervised linguistic-based model for automatic glossary term extraction from a single pdf textbook. *Education and Information Technologies*, 28(12):16089–16125.

Souza, E., Alves, T., Teles, I., Oliveira, A. L. I., and Gusmão, C. (2016). TOPIE: An open-source opinion mining pipeline to analyze consumers' sentiment in brazilian portuguese. In *Computational Processing of the Portuguese Language*, page 95–105. Springer International Publishing.

Souza, M., Vieira, R., Busetti, D., Chishman, R., and Alves, I. M. (2011a). Construction of a Portuguese opinion lexicon from multiple resources. In *Proceedings of the 8th Brazilian Symposium in Information and Human Language Technology*.

Souza, M., Vieira, R., Busetti, D., de Oliveira Chishman, R. L., and da Rosa Alves, I. M. (2011b). Construction of a portuguese opinion lexicon from multiple resources. In *Brazilian Symposium in Information and Human Language Technology*.

Stavrianou, A., Andritsos, P., and Nicoloyannis, N. (2007). Overview and semantic issues of text mining. *SIGMOD Rec.*, 36(3):23–34.

Steinberger, J., Lenkova, P., Kabadjov, M., Steinberger, R., and van der Goot, E. (2011). Multilingual entity-centered sentiment analysis evaluated by parallel corpora. In Mitkov, R. and Angelova, G., editors, *Proceedings of the International Conference Recent Advances in Natural Language Processing 2011*, pages 770–775, Hissar, Bulgária. Association for Computational Linguistics.

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.

The National Center for Biomedical Ontology (2022). Emotion ontology. `http://bioportal.bioontology.org/ontologies/MFOEM`. Acessado: 24 mar. 2024.

Thelwall, M. (2017). The heart and soul of the web? sentiment strength detection in the social web with sentistrength. In Holyst, J. A., editor, *Cyberemotions: Collective Emotions in Cyberspace*, pages 119–134. Springer International Publishing, Cham, Suíça.

Thomasson, A. (2019). Categories. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2019 edition.

Tran, H. T. H., Martinc, M., Caporusso, J., Doucet, A., and Pollak, S. (2023). The recent advances in automatic term extraction: A survey. `https://arxiv.org/abs/2301.06767`. arXiv:2301.06767.

Vilares, D., Peng, H., Satapathy, R., and Cambria, E. (2018). Babelsenticnet: A commonsense reasoning framework for multilingual sentiment analysis. In *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1292–1298.

W3C (2024a). OWL web ontology language. https://www.w3.org/TR/owl-features/. Accessed: 27 Mar. 2024.

W3C (2024b). W3C. https://www.w3.org/. Accessed: 27 Mar. 2024.

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., and Zhou, D. (2023). Chain-of-thought prompting elicits reasoning in large language models. https://doi.org/10.48550/arXiv.2201.11903. arXiv:2201.11903v6.

World Population Review (2024). Twitter users by country 2024. https://worldpopulationreview.com/country-rankings/twitter-users-by-country. Acessado: 04 abr. 2024.

Zehra, S., Wasi, S., Jami, I., Nazir, A., Khan, A., and Waheed, N. (2017). Ontology-based sentiment analysis model for recommendation systems. In *Proceedings of the 9th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*. SCITEPRESS - Science and Technology Publications.

ŽiŽka, J., DaŘena, F., and Svoboda, A. (2019). *Text mining with machine learning: Principles and techniques*. CRC Press, Londres, Reino Unido.