



UNIVERSIDADE FEDERAL DA BAHIA
INSTITUTO DE COMPUTAÇÃO
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO
TRABALHO DE CONCLUSÃO DE CURSO

ANÁLISE DE SENTIMENTO DE DIFERENTES COMUNIDADES
E TRIBOS DE DESENVOLVIMENTO DE SOFTWARE

THIAGO DE OLIVEIRA SANTOS

Salvador - Bahia
8 DE DEZEMBRO DE 2022

ANÁLISE DE SENTIMENTO DE DIFERENTES COMUNIDADES E TRIBOS DE DESENVOLVIMENTO DE SOFTWARE

THIAGO DE OLIVEIRA SANTOS

Trabalho de Conclusão de curso apresentado
como requisito parcial para obtenção do título
de Bacharel em Ciência da Computação.

Orientador: Rodrigo Rocha Gomes e Souza

Salvador - Bahia
8 de dezembro de 2022

ANÁLISE DE SENTIMENTO DE DIFERENTES COMUNIDADES E TRIBOS DE DESENVOLVIMENTO DE SOFTWARE

THIAGO DE OLIVEIRA SANTOS

Trabalho de Conclusão de curso apresentado
como requisito parcial para obtenção do título
de Bacharel em Ciência da Computação.

Banca Examinadora:

Rodrigo Rocha Gomes e Souza
Universidade Federal da Bahia (Orientador)

Gláucya Carreiro Boechat
Universidade Federal da Bahia

Joselito Mota Júnior
Universidade Federal da Bahia

Análise de sentimento de diferentes comunidades e tribos de desenvolvimento de software

THIAGO DE OLIVEIRA SANTOS, Universidade Federal da Bahia, Brasil

Ao usar determinadas linguagens, sistemas e frameworks podemos acabar gostando deles e esse fator pode influenciar na escolha das tecnologias usadas em projetos aos quais o desenvolvedor faz parte. Este artigo tem como objetivo estudar a ocorrência de emoções e sentimentos associados às comunidades de desenvolvedores que usam determinadas linguagens, frameworks e sistemas operacionais e as tecnologias que compõem essas comunidades. Para alcançar esse objetivo, fizemos uso de ferramentas de análise de sentimentos e emoções para analisar o Stack Overflow, um site de perguntas e respostas sobre programação. E chegamos a conclusão que existem emoções e sentimentos associados às comunidades de desenvolvedores que usam determinadas linguagens, frameworks e sistemas operacionais e as tecnologias que compõem essas comunidades, mesmo que em pequenas proporções.

Additional Key Words and Phrases: data sets, análise de emoções, análise de polaridade

ACM Reference Format:

Thiago de Oliveira Santos. 2022. Análise de sentimento de diferentes comunidades e tribos de desenvolvimento de software . 1, 1 (December 2022), 11 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

1 INTRODUÇÃO

Na criação de um projeto de software, uma das primeiras coisas a se levar em consideração é a linguagem, o sistema operacional e frameworks que serão usados nesse projeto. Nessa escolha são levadas em conta as características que estão presentes na linguagem, sistema e framework e se essas características são adequadas para serem usadas no projeto, mas também devem ser levadas em consideração as experiências dos desenvolvedores em relação às tecnologias aplicadas no mesmo.

No processo de decisão de que tecnologias serão usadas nos projetos, os desenvolvedores pesquisam sobre as tecnologias que podem ser usadas e uma forma é a busca de experiência em usar determinadas tecnologias ou problemas ocorridos ao usar determinada linguagem, sistema e frameworks para objetivos parecidos com o projeto em questão. Uma forma de busca é através de sites de perguntas e respostas usados por programadores, como por exemplo o Stack Overflow. Nesse site, além de experiências também podem ser encontradas opiniões e emoções sobre as tecnologias em questão.

Este artigo tem como objetivo fazer uma análise das emoções e sentimentos para descobrir se existe uma relação entre tecnologias – linguagens de programação, frameworks e sistemas operacionais

Author's address: Thiago de Oliveira Santos, Universidade Federal da Bahia, Salvador, Bahia, Brasil.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2022 Association for Computing Machinery.

XXXX-XXXX/2022/12-ART \$15.00

<https://doi.org/10.1145/nnnnnnn.nnnnnnn>

– e as emoções dos desenvolvedores de software. Para fazer essa análise, usamos dados provenientes do Stack Overflow, um site de perguntas e respostas para desenvolvedores. O Stack Overflow é uma plataforma usada principalmente por desenvolvedores para tirar dúvidas ou buscar soluções para problemas. Antes de falarmos sobre as questões de pesquisa usadas para atingir o objetivo falado acima, precisamos definir algumas palavras que serão usadas com frequência neste estudo. A primeira definição é a palavra tribo, que corresponde ao conjunto de desenvolvedores que usam e postam sobre uma linguagem, sistema ou framework estudado, por exemplo a tribo Python ou a tribo Linux.

A palavra comunidade se refere a um grupo de tribos relacionadas a uma função ou objetivo, por exemplo a comunidade de análise de dados é formada por tribos que são usualmente usadas para analisar dados através de estatística, aprendizagem de máquina, visualização, dentre outros. A seguir falaremos sobre a metodologia.

Agora que definimos algumas palavras usadas neste artigo, podemos falar sobre as questões de pesquisa usadas para atingir o objetivo deste artigo que são as seguintes:

- **QP1:** Existe diferença de polaridade de sentimento entre os desenvolvedores de diferentes comunidades de programação?
- **QP2:** Existe diferença de polaridade de sentimento entre os desenvolvedores de diferentes tribos de uma mesma comunidade de programação?
- **QP3:** Existe diferença entre as emoções manifestadas por desenvolvedores de diferentes comunidades de programação?
- **QP4:** Existe diferença entre as emoções manifestadas por desenvolvedores de diferentes tribos de uma mesma comunidade de programação?

2 METODOLOGIA

Inicialmente, foi baixado o data set SOTorrent, contendo dados do Stack Overflow até 2019. Há duas formas de adquirir o dataset: uma forma é através do serviço BigQuery, da Google, e a outra forma é a partir do repositório Zenodo. O método usado para adquirir o dataset foi através do repositório Zenodo. A maioria dos arquivos baixados estão no formato XML e os dados contidos nos arquivos são os mesmos que podem ser obtidas através de consultas na linguagem SQL no serviço BigQuery, ou seja, não existe diferença nos dados entre as duas opções citadas acima.

O SOTorrent é um data set baseado no dump do banco de dados oficial do Stack Overflow, que fornece acesso ao histórico de versões do mesmo, contendo postagens e textos individuais ou blocos de código. O dataset é constituído de todas as tabelas do banco de dados e algumas tabelas adicionais, mas para esse artigo foram usadas as tabelas Posts e Users, que, respectivamente, contêm o conteúdo HTML das postagens do Stack Overflow e os dados relacionados aos seus usuários.

Na figura 1 podemos observar uma representação em diagrama de todas as tabelas que compõem o data set SOTorrent.

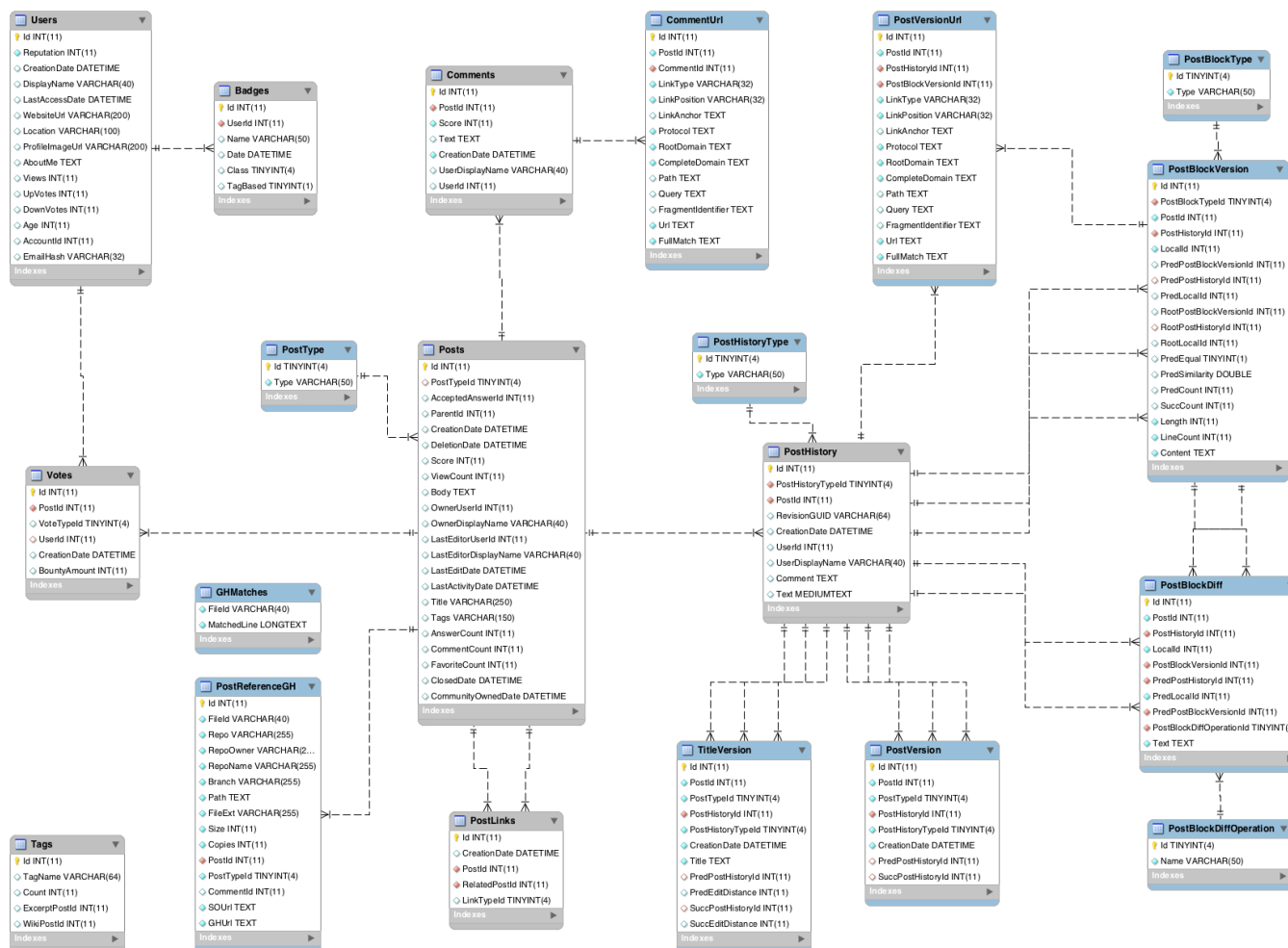


Fig. 1. Representação em diagrama de todas as tabelas que compõe o data set SOTorrent.

No data set existem 45.919.817 postagens no total. As postagens são classificadas em 8 tipos, que são pergunta, resposta, Orphaned tag wiki, Tag wiki excerpt, Tag wiki, Moderator nomination, Wiki placeholder e Privilege wiki. O tipo usado neste artigo foram as perguntas, que representam 18.154.493 de postagens. Esse tipo de postagens foi escolhido, pois supomos que as perguntas poderiam conter mais sinais de sentimentos e emoções que os outros tipos. Após ter o data set, conduzimos a pesquisa em 4 etapas:

- (1) Seleção de dados
- (2) Pre-processamento
- (3) Análise de emoções
- (4) Análise de polaridade

A Figura 2 apresenta um workflow das etapas citadas acima, que são descritas nas próximas seções. Durante o estudo a ferramenta usada nas duas últimas etapas foi o Emotion Mining Toolkit (EMTK) [Calefato et al. 2019], que é uma ferramenta de análise de polaridade e emoções desenvolvida em 2016 e lançada no meio de 2017 como uma aplicação. O EMTK integra duas ferramentas, que são a EmoTxt e Senti4SD,

sendo o EmoTxt a ferramenta que representa o módulo de análise de emoções e o Senti4SD representa o módulo de polaridade. O funcionamento de cada módulo será descrito nas seções em que foram usados.

2.1 Seleção de dados

A seleção de dados tem como objetivo adquirir amostras de postagens pertencentes a cada linguagem, sistema operacional e framework e separá-las por comunidades e tribos para serem analisadas neste artigo. As postagens foram separadas por tribos usando a lista de tags, que são rótulos que ajudam na identificação do assunto ao qual as questões se referem, presente na tabela Posts do data set. Após a separação das postagens por tribos foram selecionadas amostras de 1000 postagens de cada tribo e classificadas em 4 comunidades estudadas neste artigo. As 4 comunidades foram escolhidas devido a sua popularidade e também levamos em conta na escolha as tribos já selecionadas de forma que não existissem comunidades

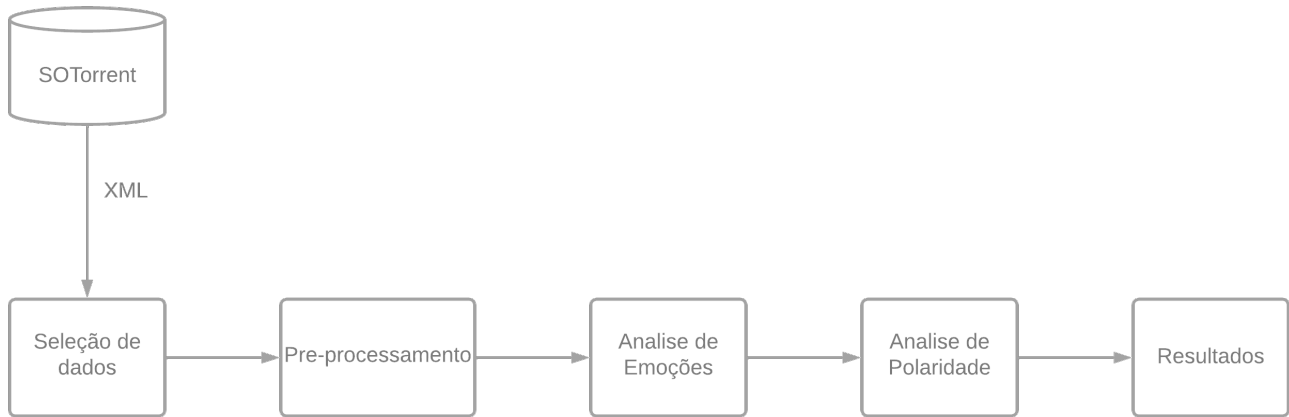


Fig. 2. Workflow de todas as etapas.

com menos de duas tribos a compoendo. As comunidades escolhidas foram:

- Sistema mobile
- Análise de dados
- Programação web
- Sistema operacional

Antes de falarmos sobre as comunidades devemos esclarecer que durante a separação das postagens por tribos levamos em consideração a comunidade à qual cada tribo pertence e tivemos o cuidado de selecionar apenas postagens nas quais houvesse somente uma tag relacionada às tribos pertencentes a uma mesma comunidade. Por exemplo as postagens da tribo Windows são aquelas que possuem a tag windows mas não as tags linux ou macOS.

2.1.1 Sistema mobile. A comunidade de sistema mobile é composta por tribos que representam sistemas operacionais de smartphones e outros dispositivos móveis. As tribos que representam esta comunidade foram selecionadas usando as tags android e ios.

2.1.2 Análise de dados. A comunidade de análise de dados é representada por tribos que são usualmente usadas para analisar dados através de estatísticas, aprendizagem de máquina, visualização, dentre outros. As tribos que compõem esta comunidade são a tribo Python e a tribo R. Para esta comunidade especificamente tivemos que usar tags além de r e python, pois a tribo Python não é somente usada para análise de dados. Por esse motivo, durante a seleção da tribo Python, usamos as tags pandas, scipy, numpy e matplotlib juntamente com a tag python.

2.1.3 Programação web. A comunidade de programação web representa frameworks usados no front-end na programação de sites. Esta comunidade é composta por 3 frameworks, que são o Angular, React e Vue, e para seleção das tribos foram usadas as tags angular, react e vue, respectivamente.

2.1.4 Sistema operacional. A comunidade de sistema operacional representa os sistemas operacionais usados em computadores e

notebooks. As tribos que compõem essa comunidade são Windows, Linux e macOS, que são representadas pelas tags windows, linux e macOS respectivamente.

2.2 Pré-processamento

Foram realizados três processos no pré-processamento dos dados. O primeiro foi a limpeza do atributo Body, presente na tabela Posts do dataset, o segundo foi a criação de 200 arquivos CSV contendo 50 postagens cada no padrão especificado na ferramenta EMTK para passar pela análise de emoções, e o terceiro foi a criação de 10.000 arquivos CSV contendo uma postagem cada no formato especificado pela ferramenta mencionada anteriormente para ser usada na análise de polaridade.

O atributo Body contém o texto da postagem representado em formato HTML. A limpeza do atributo Body foi feita usando o método parse presente no módulo etree da biblioteca xml do Python e tem como objetivos a retirada de elementos desnecessários, como trechos de códigos e links, e transformar em texto o atributo mencionado acima.

O segundo pré-processamento tem como objetivo adequar os dados a serem analisados para o formato exigido pelo módulo de análise de emoções do EMTK e a divisão em 200 arquivos se fez necessária para evitar travamentos do sistema em decorrência do uso de memória maior que o disponível na máquina.

O terceiro pré-processamento tem o objetivo de adequar os dados ao formato exigido pelo módulo de análise de polaridade da ferramenta e sua divisão de arquivos é feita para podermos analisar cada postagem separadamente, pois se a análise for feita sem separar cada postagem não seria possível identificar quais resultados pertencem a cada, devido ao resultado da análise mostrar a polaridade para cada sentença dentro do arquivo.

2.3 Análise de emoções

Na análise de emoções foi usada a ferramenta EMTK [Calefato et al. 2019] para analisar cada uma das 6 emoções, que são amor, surpresa,

alegria, tristeza, raiva e medo. Nesta etapa foi usado o módulo de análise de emoções que é baseado na ferramenta EmoTxt [Calefato et al. 2017a], cujo o meio de classificar uma emoção é usando associação entre palavras e categorias de emoção presentes no WordNet Affect, que é um recurso que contém um conjunto de conceitos afetivos correlacionado com palavras afetivas [Strapparava et al. 2004]. Essa associação foi feita usando tf-idf em cada uma das emoções com base no número de ocorrências das palavras associadas a elas. A ferramenta EmoTxt também leva em consideração polidez e humor durante a análise de emoções, mas que por questões de otimização se tornaram uma feature opcional no módulo de emoções do EMTK. Na análise usada neste estudo também foi usada a feature opcional.

Para cada um dos 200 arquivos CSV criados na etapa de pré-processamento se gerou um resultado para cada análise feita, totalizando 1200 análises.

2.4 Análise de polaridade

Na análise de polaridade foi usado o módulo de polaridade da ferramenta EMTK para analisar cada uma das 10.000 postagens, gerando um resultado para cada análise.

O módulo de polaridade é baseado na ferramenta Senti4SD [Calefato et al. 2017b], feita em Java e R, que foi treinada para analisar canais de comunicação de desenvolvedores de software. A análise deste módulo é baseada em 3 funções diferentes, que são:

- Léxico
- Palavras-chaves
- Semântico

A análise usando a função léxica é feita através do uso de um dicionário no qual cada palavra está associada a um valor inteiro entre 5 e 1 e entre -1 e -5, sendo uma palavra de polaridade positiva quando seu valor está entre 2 e 5, negativa quando está entre -2 e -5 e neutra quando está entre 1 e -1.

A análise usando palavras-chaves usa como features o número de ocorrências dos n-grams que aparecem no documento [Joachims 1998]. Além disso, a ferramenta identifica aspectos de microblogging, como o uso de caixa alta, a ocorrência de palavras alongadas usadas como intensificadores (por exemplo a palavra "óóódio!"), a presença de emoticons positivos e negativos e o uso de gírias que expressam riso.

A análise usando a função semântica identifica a semelhança entre as representações vetoriais dos documentos e vetores protótipos representando as classes de polaridade em um DSM (design structure matrix) [Calefato et al. 2017b]. Os documentos são representados em um DSM como uma soma vetorial de todos os vetores de palavras que ocorre no documento e o vetores protótipos são representações das classes positiva, negativa e neutra em um DSM e um vetor protótipo é a representação vetorial de um perfil lexical que é usado para determinar uma classe de polaridade.

No uso da ferramenta é obrigatório escolher entre quatro opções que são uma das funções faladas anteriormente executada separadamente ou todas as funções combinadas em uma única análise, que foi a opção usada neste artigo, pois devido ao uso de mais de um tipo diferente de feature a análise se torna mais precisa. O resultado da análise é um arquivo CSV que contém o resultado para cada sentença analisada, que pode ser positivo, neutro ou negativo.

3 RESULTADOS

Neste artigo foram analisadas 10.000 perguntas do Stack Overflow. A seguir serão discutidos os resultados obtidos como resposta às questões de pesquisa propostas neste artigo.

3.1 QP1: Existe diferença de polaridade de sentimento entre os desenvolvedores de diferentes comunidades de programação?

Para chegarmos aos resultados, utilizamos os dados obtidos pela análise de polaridade. Foram feitas comparações entre as comunidades e usando o teste qui-quadrado demonstramos que a diferença nas proporções é estatisticamente significativa a um nível de 5% ($p = 3.945 \times 10^{-7}$).

Na Figura 3 podemos observar que a comunidade de análise de dados tende a expressar mais sentimentos positivos do que outras comunidades, e que a comunidade web tende a expressar mais sentimentos negativos. Ao analisarmos a comunidade de análise de dados verificamos que 428 (21,4%) posts referente a essa comunidade possuem sentimentos positivos associados a eles. E ao aplicarmos os mesmos métodos à comunidade de programação web obtivemos resultados parecidos em relação aos sentimentos negativos.

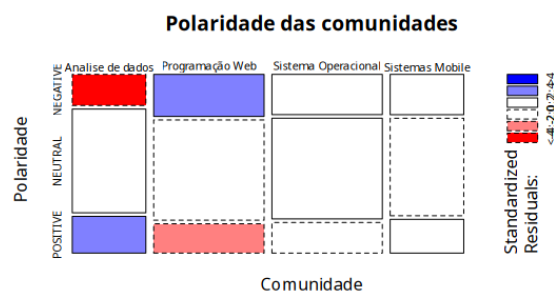


Fig. 3. Gráfico apresentando a polaridade por comunidade. No gráfico podemos observar que a células destacadas pela cor azul e vermelha que respectivamente representam um aumento e diminuição em comparação com o que é esperado no modelo nulo.

3.2 QP2: Existe diferença de polaridade de sentimento entre os desenvolvedores de diferentes tribos de uma mesma comunidade de programação?

Inicialmente, foi feita uma análise dos gráficos comparando as tribos presentes em cada comunidade. Considerando a comunidade de sistemas operacionais na Figura 4 podemos perceber que o sistema macOS tem maior quantidade de postagens com sentimentos negativos e positivos que as outras tribos e também podemos perceber que os sistemas Windows e Linux têm uma quantidade maior de postagens neutras. Aplicando o teste qui-quadrado nas tribos pertencentes à comunidade de sistema operacional obtivemos $p = 0.046$, que significa que as diferenças entre as tribos são estatisticamente significativas. Apesar da diferença ser estatisticamente significativa, a diferença é pequena de um sistema para o outro.

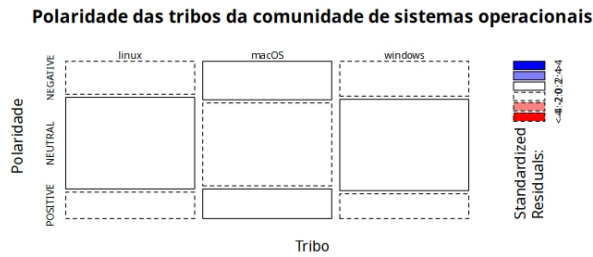


Fig. 4. Gráfico apresentando a polaridade por tribos da comunidade de sistemas operacionais. No gráfico podemos observar que as células pontilhadas representam uma diminuição em comparação com o que é esperado no modelo nulo.

Na comunidade de análise de dados podemos observar, na Figura 5, que a linguagem R tem mais postagens com sentimentos positivos que a linguagem Python e que a linguagem Python tem mais postagens com sentimentos não-positivos. Ao aplicarmos o teste qui-quadrado obtivemos $p = 0.36$ e, assim, podemos dizer que a diferença entre as tribos não é estatisticamente significativa.

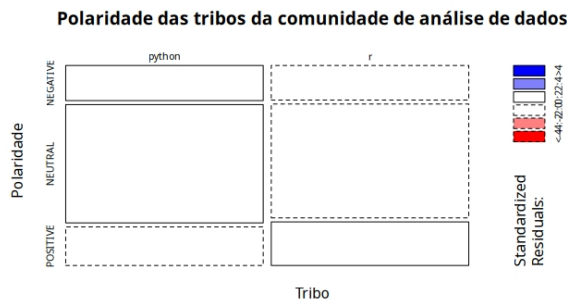


Fig. 5. Gráfico apresentando a polaridade por tribos da comunidade de análise de dados. No gráfico podemos observar que as células pontilhadas representam uma diminuição em comparação com o que é esperado no modelo nulo.

Na comunidade de programação web podemos perceber na Figura 6 que a framework Angular tem maior número de postagens com sentimentos positivos, a framework React com o maior número de postagens com sentimentos negativos e o Vue com a maior número de postagens neutras. Aplicando o teste qui-quadrado temos $p = 0.28$ e assim concluímos que as diferenças entre as tribos desta comunidade não são estatisticamente significativas.

Na comunidade de sistemas mobile podemos observar através da Figura 7 que o sistema Android tem a maior quantidade de postagens com sentimentos positivos e negativos em comparação com o iOS e aplicando teste qui-quadrado obtivemos o $p = 0.03$ que mostra que a diferença entre as duas tribos é estatisticamente significativa.

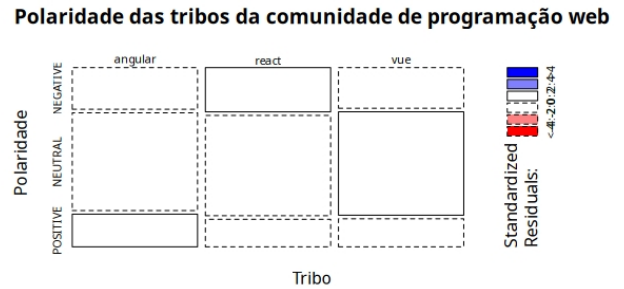


Fig. 6. Gráfico apresentando a polaridade por tribos da comunidade de programação web. No gráfico podemos observar que as células pontilhadas representam uma diminuição em comparação com o que é esperado no modelo nulo.

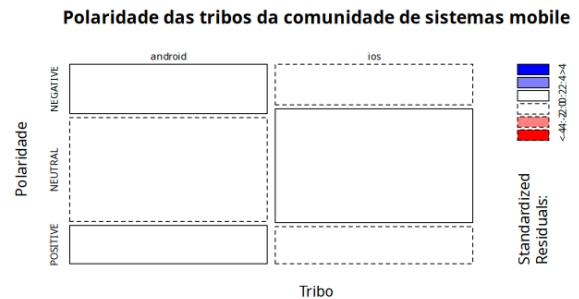


Fig. 7. Gráfico apresentando a polaridade por tribos da comunidade de sistemas mobile. No gráfico podemos observar que as células pontilhadas representam uma diminuição em comparação com o que é esperado no modelo nulo.

3.3 QP3: Existe diferença entre as emoções manifestadas por desenvolvedores de diferentes comunidades de programação?

Para respondermos essa QP, faremos uma análise de cada uma das 6 emoções em relação às comunidades. Na análise da emoção de alegria, podemos observar na Figura 8 que nenhuma das comunidades tende a expressar a emoção de alegria de forma significativa. Aplicando o teste qui-quadrado obtivemos $p = 0.39$ e assim podemos assumir que as diferenças não são estatisticamente significativas.

Na análise da emoção de raiva, podemos perceber na Figura 9 que a comunidade de sistemas mobile tem uma tendência maior de expressar a emoção de raiva e que a comunidade de programação web tem uma tendência menor de expressar a raiva em comparação com as outras comunidades. Quando aplicado o teste qui-quadrado obtivemos $p = 4.548 \times 10^{-5}$ e esse resultado mostra que a diferença é estatisticamente significativa.

Na emoção de tristeza, podemos observar na Figura 10 que não há uma grande diferença entre as comunidades na expressão da emoção

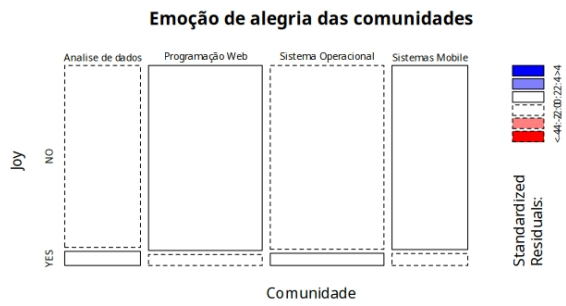


Fig. 8. Gráfico apresentando a emoção de alegria por comunidade. No gráfico podemos observar que as células pontilhadas representam uma diminuição em comparação com o que é esperado no modelo nulo.



Fig. 10. Gráfico apresentando a emoção de tristeza por comunidade. No gráfico podemos observar que as células pontilhadas representam uma diminuição em comparação com o que é esperado no modelo nulo.

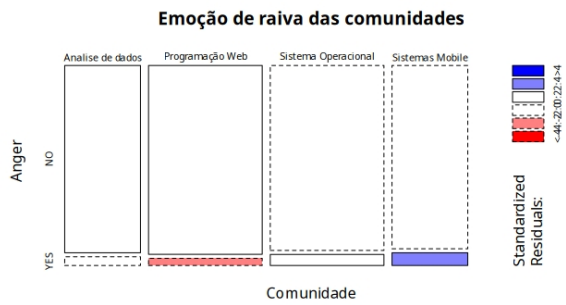


Fig. 9. Gráfico apresentando a emoção de raiva por comunidade. No gráfico podemos observar que as células destacadas pela cor azul e vermelha que respectivamente representam um aumento e diminuição em comparação com o que é esperado no modelo nulo.

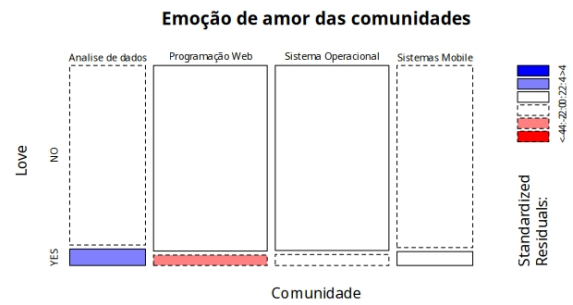


Fig. 11. Gráfico apresentando a emoção de amor por comunidade. No gráfico podemos observar que as células destacadas pela cor azul e vermelha que respectivamente representam um aumento e diminuição em comparação com o que é esperado no modelo nulo.

de tristeza e aplicando teste qui-quadrado temos $p - valor = 0.45$ que nos mostra que a diferença não é estatisticamente significativa.

Na emoção de amor, podemos perceber na Figura 11 que a comunidade de análise de dados tem uma maior tendência de expressar amor e que a comunidade de programação web tem uma tendência menor de expressar amor em comparação com as outras comunidades. Usando o teste qui-quadrado obtivemos $p - valor = 0.0003$, mostrando que a diferença é estatisticamente significativa.

Na emoção de surpresa, podemos observar que a comunidade de análise de dados tende a demonstrar menos expressão de surpresa e a comunidade de sistemas mobile tem uma maior tendência a expressar surpresa em comparação com as outras comunidades, como mostrado na Figura 12. Aplicando o teste qui-quadrado temos $p - valor = 1.017 \times 10^{-7}$ e assim mostrando que a diferença é estatisticamente significativa.

Na emoção de medo, podemos observar na Figura 13 que não há uma diferença significativa entre as comunidades na expressão da emoção de medo e aplicando teste qui-quadrado temos o $p -$

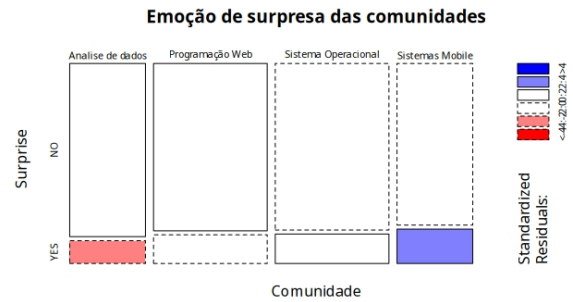


Fig. 12. Gráfico apresentando a emoção de surpresa por comunidade. No gráfico podemos observar que as células destacadas pela cor azul e vermelha que respectivamente representam um aumento e diminuição em comparação com o que é esperado no modelo nulo.

valor = 0.36, que nos mostra que a diferença não é estatisticamente significativa.

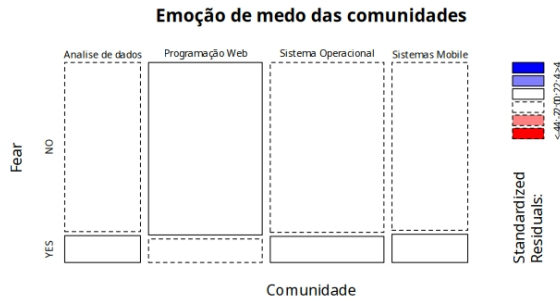


Fig. 13. Gráfico apresentando a emoção de medo por comunidade. No gráfico podemos observar que as células pontilhadas representam uma diminuição em comparação com o que é esperado no modelo nulo.

3.4 QP4: Existe diferença entre as emoções manifestadas por desenvolvedores de diferentes tribos de uma mesma comunidade de programação?

Para chegarmos ao resultado desta questão de pesquisa, usamos o resultado da análise de emoções. Foram feitas comparações entre as tribos pertencentes à mesma comunidade para cada uma das seis emoções que podem ser identificadas pelo EMTK.

Na comunidade de sistemas operacionais, podemos observar na Tabela 1 que a emoção mais encontrada nesta comunidade é a surpresa e que a tribo que mais se destaca nesta emoção é o sistema Linux. Também é perceptível que nenhum dos p-valor tem valor abaixo de 0.05 e assim podemos dizer que nenhuma das análises de emoções desta comunidade é estatisticamente significativa.

Table 1. Tabela apresentando o número de postagens em que foi identificada cada uma das emoções por tribo pertencente à comunidade de sistemas operacionais

Emoção	Windows	Linux	MacOS	p-valor
Joy	59	55	73	0.22
Anger	64	52	53	0.43
Sadness	41	46	42	0.84
Love	55	54	59	0.88
Surprise	150	166	135	0.15
Fear	126	135	139	0.68

Na comunidade de análise de dados, usando a Tabela 2 podemos perceber que a emoção de surpresa é a única que é estatisticamente significativa e que a linguagem Python tem uma quantidade razoavelmente maior de postagens que apresentam a emoção de surpresa que a linguagem R. Também podemos identificar uma diferença razoável na quantidade de postagens que expressam amor em associada à linguagem R, mas não é estatisticamente significativa devido ao seu p-valor.

Table 2. Tabela apresentando o número de postagens em que foi identificada cada uma das emoções por tribo pertencente à comunidade de análise de dados

Emoção	Python	R	p-valor
Joy	72	72	1.00
Anger	48	41	0.51
Sadness	33	35	0.90
Love	73	95	0.09
Surprise	145	88	9.498×10^{-5}
Fear	129	145	0.32

Na comunidade de programação web, podemos perceber na Tabela 3 que a emoção de raiva é a única estatisticamente significativa ($p < 0.05$) e que o framework Vue tem uma quantidade relativamente menor de postagens expressando raiva em comparação com as outros frameworks da mesma comunidade.

Table 3. Tabela apresentando o número de postagens em que foi identificada cada uma das emoções por tribo pertencente à comunidade de programação web

Emoção	Angular	React	Vue	p-valor
Joy	55	59	54	0.88
Anger	45	41	24	0.03
Sadness	37	49	42	0.41
Love	48	56	56	0.65
Surprise	161	150	128	0.10
Fear	112	128	120	0.54

Na comunidade de sistema mobile, podemos observar na Tabela 4 que nenhuma das emoções é estatisticamente significativa.

Table 4. Tabela apresentando o número de postagens em que foi identificada cada uma das emoções por tribo pertencente a comunidade de sistema mobile

Emoção	Android	IOS	p-valor
Joy	56	66	0.40
Anger	56	74	0.12
Sadness	29	43	0.12
Love	78	64	0.26
Surprise	168	183	0.41
Fear	142	146	0.85

3.5 Ameaças à validade

A análise de emoções e sentimentos é uma área complexa e um dos principais problemas para esse tipo de estudo é a validade dos resultados que podem ser questionados por vários fatores. Neste tópico vamos discutir as ameaças à validade dos resultados deste artigo.

Uma das ameaças está relacionada aos dados usados neste estudo e podemos identificar o fato de que os dados terem sido adquiridos

em outubro de 2019. O fato de os dados pertencerem a outubro de 2019 indica que uma quantidade de dados não foram levados em consideração neste estudo, mas por serem dados de 2019 provavelmente não afetarão muito a validade deste estudo, pois o dataset contém uma grande parte dos dados totais e a análise foi feita com linguagens já consolidadas.

Um segundo fator é o fato de que o módulo de polaridade usa as pontuações do léxico do SentiStrength, que é tendencioso para uma classificação negativa, pois não foi criado especificamente voltado para textos voltados para engenharia de software e algumas palavras que são consideradas neutras em textos de engenharia de software são classificadas como negativas, por exemplo a palavra *bug*. Embora seja um fator importante a ser considerado, provavelmente não afetará significativamente a validade deste estudo pelo fato de que a grande maioria das análises é composta por postagens neutras.

Outra ameaça é o fato de usarmos uma única fonte de dados, o Stack Overflow. Por causa deste fato não podemos generalizar os resultados para todos os desenvolvedores de software. Por outro lado, o Stack Overflow é muito popular entre os desenvolvedores de todos os países e comunidades, então provavelmente é representativo da população de desenvolvedores de software.

Outro fator é a generalização dos resultados. A generalização ocorre por causa da análise não cobrir todas as linguagens, frameworks e sistemas pertencentes a cada comunidade. Provavelmente não afetará muito a validade do estudo pelo fato de que estamos estudando uma das sub-comunidades ao qual faz parte da comunidade. Por exemplo a comunidade de programação web é representada por vários subgrupo, como front-end, back-end entre outros, e foi selecionado uma parte da sub-comunidade de front-end para representar a comunidade front-end que por conveniência foi chamada de comunidade de programação web para fins de diferenciação mais precisas entre as comunidades estudadas.

4 TRABALHOS RELACIONADOS

[Ling and Larsén 2018] estudaram que tipo de dado das postagens do Stack Overflow, que são os comentários, respostas e perguntas, continham maior número de sentimentos positivos e negativos e também estudaram a relação entre esses sentimentos e as linguagens de programação. Eles concluíram que existem diferenças substanciais entre as distribuições de sentimento de diferentes tipos de documentos e que existe uma associação entre linguagem de programação e sentimentos. No estudo de [Ling and Larsén 2018] também foram feitas análises de sentimento em linguagens de programação, como dito acima, e dentre as linguagens analisadas somente a linguagem Python é comum entre nossas pesquisas e os resultados foram diferentes em relação a diferença existente entre a proporção de sentimentos negativos e neutros da linguagem Python, sendo que a diferença entre a quantidade de postagens neutras e negativas foram grandes no nosso estudo enquanto a diferença em [Ling and Larsén 2018] foi pequena entre as postagens neutra e negativas. Um ponto a se destacar em nosso estudo é o fato de fizemos comparações entre linguagens de uma mesma comunidade e o fato de existir semelhança entre linguagens que pertencem a uma mesma comunidade foi mencionado no estudo de [Ling and Larsén 2018].

[Cagnoni et al. 2020] fizeram uma análise baseada em emoções de 26 linguagens de programação e concluíram que o uso de um modelo de classificação hierárquico composto por vários classificadores dispostos em diferentes níveis oferece melhores resultados do que usar um único classificador plano ao realizar a detecção de emoções.

5 CONCLUSÃO

Neste estudo, fizemos uma análise de sentimentos e emoções usando a ferramenta EMTK com o objetivo de observar se existem sentimentos e emoções que possam ser associados a determinadas comunidades de linguagens de programação, sistemas operacionais e frameworks.

Foi observado neste estudo que a comunidade de análise de dados tende a expressar mais sentimentos positivos que as demais comunidades, e dentro desta comunidade a tribo da linguagem Python tende a expressar surpresa muito mais que a tribo da linguagem R.

Já a comunidade de programação web tende a expressar mais sentimentos negativos. Em particular, desenvolvedores Angular e React tendem a expressar mais raiva que desenvolvedores Vue.

5.1 Discussão

Podemos observar que a comunidade sistemas mobile tende a expressar mais raiva e surpresa. Dentro desta comunidade, os desenvolvedores Android têm maior tendência de sentimentos positivos e negativos em comparação com os desenvolvedores iOS.

Apesar de a comunidade de programação web tende a expressar mais sentimentos negativos, essa comunidade é a que tende a expressar menos emoções entre todas as comunidades. A comunidade de análise de dados é a comunidade que tende a expressar mais a emoção de amor, o que está de acordo com o fato de ser a comunidade que expressa mais sentimentos positivos. Foi observado também que as emoções de surpresa e medo são identificadas em maior quantidade que as outras emoções em todas as tribos.

6 TRABALHOS FUTUROS

Como trabalho futuro, planejamos repetir o estudo com uma ou mais fontes de dados diferentes. Também planejamos fazer uma análise qualitativa das perguntas para investigar por que certas emoções estão associadas a certas tribos.

REFERENCES

- Stefano Cagnoni, Lorenzo Cozzini, Gianfranco Lombardo, Monica Mordonini, Agostino Poggi, and Michele Tomaiuolo. 2020. Emotion-based analysis of programming languages on Stack Overflow. *ICT Express* 6, 3 (2020), 238–242.
- Fabio Calefato, Filippo Lanubile, Federico Maiorano, and Nicole Novielli. 2017b. Sentiment Polarity Detection for Software Development. *Empirical Software Engineering* (2017). <https://doi.org/10.1007/s10664-017-9546-9>
- Fabio Calefato, Filippo Lanubile, and Nicole Novielli. 2017a. EmoTxt: A Toolkit for Emotion Recognition from Text. In *Proc. of 7th Int'l Conf. on Affective Computing and Intelligent Interaction Workshops and Demos* (San Antonio, TX, USA) (ACII 2017). 79–80. <https://doi.org/10.1109/ACIIW.2017.8272591>
- Fabio Calefato, Filippo Lanubile, Nicole Novielli, and Luigi Quaranta. 2019. Emtk-the emotion mining toolkit. In *2019 IEEE/ACM 4th International Workshop on Emotion Awareness in Software Engineering (SEmotion)*. IEEE, 34–37.
- Thorsten Joachims. 1998. Text categorization with support vector machines: Learning with many relevant features. In *European conference on machine learning*. Springer, 137–142.
- Li Ling and Simon Larsén. 2018. Sentiment analysis on stack overflow with respect to document type and programming language.
- Carlo Strapparava, Alessandro Valitutti, et al. 2004. Wordnet affect: an affective extension of wordnet.. In *Lrec*, Vol. 4. Lisbon, Portugal, 40.