Classifying Emotions in Brazilian Stock Market Tweets

Anonymous ACL submission

Abstract

Stock markets play an important role in accelerating economic growth of developing countries like Brazil and, typically, leads to strong emotions in people, which may be reflected in their behaviour at social media like X, old Twitter. At the bright side, analysing these emotions could unveil interesting insights about public perception, potentially leading to more accurate and profitable stock market forecasts. Automatic emotion detection in tweets has been explored by many studies in the past years. State-of-the-art pre-trained language models have also been used to this end. We propose to detect emotion on tweets related to the brazilian stock market, which have a few dedicated researches. We trained BERTimbau (Portuguese version) large and base on a free domain training dataset and tested the models on the target domain (Brazilian stock market), a cross-domain approach. Applying Plutchik's wheel in its basic form, in this work we considers only the four main emotion pairs, to wit, joy $\times$ sadness, anger $\times$ fear, trust $\times$ disgust and surprise $\times$ anticipation. Models performance drops to values ranging from F1 nil, for BERTimbau Large and *Joy* to $F1 = 0.78$, also for BERTimbau Large, but with *Trust*. Results by both BERTimbau Large and Base, after test on the *Free Domain Corpus* test set (same domain used for training) reached almost 100% accuracy for all emotions.

1 Introduction

Artificial Intelligence (AI) conjectures that every aspect of learning or any another feature of intelligence can be precisely described and that a machine can be developed to simulate it (Dick, 2019). In this context, the automatic detection of emotions is one of the research fields that has presented a major challenge to the AI area (Al-Omari et al., 2020).

Human beings express emotions directly or indirectly through speech, facial expressions, gestures or writing. There are then many information sources that can be used to analyse emotions, specifically in texts, such as blogs, newspaper articles, social media posts, etc. (Sailunaz et al., 2018).

With the widespread use of socially-aimed technology over the years, events, news or activities around the world began to be discussed through social media by millions of people (Gaind et al., 2019). X, which was formerly known as Twitter until its recent rebranding, is one of the commonly and widely used resources to this end, encouraging its users to express what they think on a daily basis and real time.

Unsurprisingly, X also plays a major role in discussing stock market moves by its participants. Often, companies publish information that changes share prices, and investors rely on X to seek for opinions and comments, so as to try to figure out what would be a better deal (Simões et al., 2017). Parallel to this, a very common belief is that investor sentiment is one of the most important sources behind market movements. In this sense, although classic financial theory assumes that investors are rational, studies have revealed the significant influence of their irrational behaviour, such as optimistic or pessimistic feelings, among others (Hiew et al., 2019).

Stock markets play an important role in accelerating economic growth (Sharma et al., 2017) and, typically, volatility in such markets leads to strong emotions in people (Liu et al., 2017), which may be reflected in their comments at social media. At the bright side, analysing these emotions could unveil interesting insights about public perception, potentially leading to more accurate and profitable stock market forecasts.

The downside with this procedure lies, however, in the very feature that makes social media interesting for this purpose: its popularity. There simply is a vast amount of data which cannot be analysed quickly enough for such forecasts to be produced. At this point, machine learning methods might come in handy, by automatically detecting the emotions portrayed by many users, through the analysis of their comments on social media, specially on X (Matla and Badugu, 2020).

To do so, one has to rely on some Natural Language Processing (NLP) technique, so as to try to identify the emotion carried by the content of comments by investors. On this regard, BERT is considered state-of-the-art for several NLP tasks and applications (Valdes et al., 2021), given its previous training and possibility to be fine-tuned in different data, in an attempt to transfer the generic linguistic information previously learned to the new task.

Since transfer learning approaches have already shown improvements in general performance on many tasks, and pre-trained language models are valuable solutions especially for languages with few labelled training samples (Souza et al., 2020), this seems to be an alternative to be explored in the case of Portuguese, specially in such a narrow domain such as comments by stock market investors. Within this context, an existing annotated corpus of stock market tweets (Silva et al., 2020) written in Brazilian Portuguese, already annotated with emotions, might come in handy as a basis for determining whether such a model could be used in this task.

In their work, Silva et al. (2020) have taken a cross-domain approach, by training the system in a corpus which was automatically annotated with emotions and then studying how the trained system would perform in their human annotated corpus. Since their focus was on determining whether the syntactic structure of sentences might play a major role in this task, they left unresolved an important gap: the fact that the model developed had not been compared to current distributed representation models, such as BERT for example.

In this work we intend to fill in this gap, by presenting the results obtained by fine-tunning and running BERTimbau (Souza et al., 2020) – a BERT flavour trained specifically in Brazilian Portuguese texts – in the corpus by Silva et al. (2020). The rest of this article is organised as follows. Chapter 2 provides an overview of different approaches, findings and knowledge from existing literature relevant related to the topic of this work. In Chapter 3 we describe the materials and methods used to develop our research: the methodology, the datasets and the experimental setup. Chapter 4 discusses the results of the experimental research. Chapter 5 concludes this paper emphasising the new knowledge that contributed to the field of study and future work.

2 Related Literature

In the stock market context, some researches has explored the relationship between economy and the emotions expressed by people on social media, examining whether these emotions can be influenced by the stock market index or if it is possible to predict fluctuating market through these emotions (*e.g.* (Kang et al., 2017)). Twitter, rebranded to X since 2023, has become widely accepted as the leading platform for this purpose (Michalak, 2020).

Several studies have showed there to be a relationship between emotion and stock market moves (*e.g.* (Liu et al., 2017; Rossouw et al., 2020; Saurabh and Dey, 2020; Kang et al., 2017; Bhatia et al., 2018; Lazeski, 2020)). These, however, concentrate mainly on the US market, there being only a few dedicated to exploring the context of other markets, such as the Brazilian stock market (*e.g.* (Medeiros and Borges, 2019; Silva et al., 2020)).

Even when it comes to our reference study Silva et al. (2020), which builds on a corpus of stock market tweets written in Brazilian Portuguese, there is still a lack of comparison between obtained results and those by the direct application of current language models, such as BERT's. In this case, the fact that BERT has been used with success in tweets written in different languages and to a broad range of tasks (*e.g.* (Sawhney et al., 2021; Pranesh et al., 2020; Kabir and Madria, 2021; Hassan et al., 2021; Abdelali et al., 2021; Ahmad et al., 2021; Chiril et al., 2022; De Bruyne et al., 2022)) serves as an indication of its probable suitability to the stock market domain too.

From this perspective, we propose in this work a research task to detect emotions using BERTimbau, a BERT-based model pre-trained on textw written in Brazilian Portuguese, which has achieved state-of-the-art performance in many NLP tasks, including emotion detection (*cf.* (Hammes and Freitas, 2021)).

3 Materials and Methods

An important aspect to be considered for automatic emotion detection are the emotion models that delimit the classification process (Graterol et al., 2021). Emotion models are fundamental to convey affective meaning in a readable way for humans and computers (Horvat et al., 2022) and define how one emotion differs from another (Acheampong et al., 2021).

Emotion models can be divided in two large groups: discrete and dimensional (Yang et al., 2021). While in psychology there are several theories about the representation of emotions, within NLP two stand out as most commonly used: Ekman's basic emotions (discrete model) and Plutchik's wheel of emotions (multidimensional model) (Graterol et al., 2021).

Plutchik's Wheel of Emotions comprises eight basic emotions: joy, trust, fear, surprise, sadness, disgust, anger and anticipation (Zad et al., 2021). The model is represented through four opposing axes pairs, along which emotions are defined as different points, as illustrated in Figure 1. New emotions can be defined on the basis of different combinations of these basic emotions, although much of the extant work on practical emotion detection considers only a small subset of this group (Graterol et al., 2021).

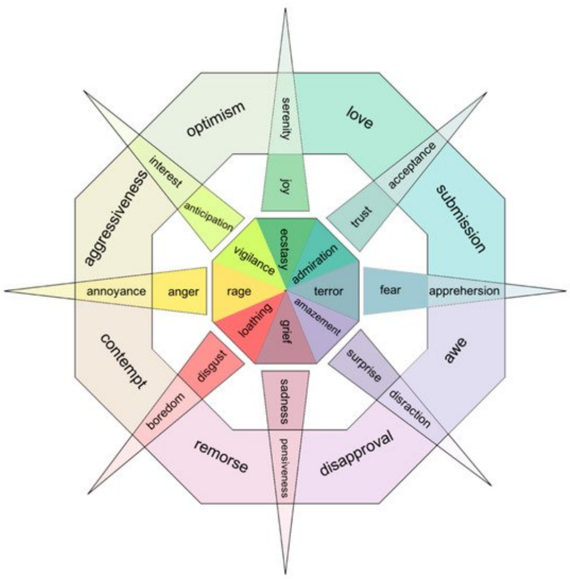

Figure 1: Plutchik's Wheel of Emotions (Zhou et al., 2020)

For the sake of comparison, we decided to use the work by Silva et al. (2020) as reference, thereby applying Plutchik's wheel in its basic form. Hence, in this work we considers only the four main emotion pairs, to wit, joy × sadness, anger × fear, trust × disgust and surprise × anticipation.

In machine learning, there usually is an issue regarding labelled data, given is limited amount, specially regarding human produced data, whose production is usually time consuming and expensive (Willemink et al., 2020). To aid with this problem, transfer learning comes in handy, by relying on very cheap unstructured data that is available online to pre-train general models, which can then be applied to more scarce labelled data (Diethe et al., 2015).

The most popular approach taken to make this transfer is pre-refinement followed by fine-tuning. Basically, it consists of two training steps applied sequentially. First, a general purpose model is trained on unstructured data. Then the model is transferred and continues its training with data labelled according to the target task. This has been reported to achieve state-of-the-art performance in numerous applications (Han et al., 2021).

In this work we have elected BERTimbau (Almeida Neto et al., 2021) as our pre-trained model. The fact that BERTimbau was originally trained in one of the (if not the) largest open corpus written in Brazilian Portuguese –brWaC, which contains 2.68 billion tokens from 3.53 million web pages, along with and 200,000 random articles from Brazilian Wikipedia and a generated vocabulary of 30,000 unique words (Souza et al., 2019), makes it a natural candidate to this task.

3.1 Data

In order to verify the possibility of cross-domain learning, in this research we work with two different data sets, the *Free Domain Corpus* and the *Stock Market Domain Corpus*, both described in Silva et al. (2020). The *Stock Market Domain Corpus.* comprises 4,517 non-repeated tweets mentioning any of the stocks that build the IBOVESPA index (the main index of B3, the Brazilian Stock Exchange market), collected between March and May 2014. In this corpus, tweets were manually annotated by a total of 442 volunteers according to the Plutchik's wheel of emotions, so as to guarantee that each tweet was annotated by at least 3 volunteers.

During the annotation, and as a way to prevent tweets from being assigned opposing emotions, annotators had to classify each tweet according to each of the emotional axes. Hence, for each axe they had to pick either of its opposite emotions, or classify it as neutral regarding that emotion pair. A tweet was considered neutral only if it was annotated as neutral for all emotion pairs.

Along with these categories, there was also an *I don't know* option to annotators, so proper neutral tweets could be told apart from cases where annotators where in doubt. As a result, 240 (5.3%) out of the 4,517 tweets were discarded because they were marked as *I don't know* in all of their emotion pairs by the majority of users. The final corpus delivers, then, a total of 4,277 tweets, labelled, for each emotion pair, with the emotion assigned by the majority of at least three annotators. Table 1 summarises these figures, by presenting the amount of tweets in each of these groups. The distribution of emotions per pair can be seen in Table 2.

At least one emotion assigned	2,340
Neutral	779
Inconclusive (all emotion pairs)	184
Inconclusive + "I don't know" + Neutral	974
Total	**4,277**

Table 1: Distribution of tweet annotations.

Emotion Pair	**Total**	**Emotion**	**N°**
Joy vs Sadness	3,307	Sadness	437
		Joy	531
		Neutral	2,339
Trust vs Disgust	3,088	Disgust	709
		Trust	851
		Neutral	1,528
Anger vs Fear	3,510	Fear	232
		Anger	256
		Neutral	3,022
Anticipation vs Surprise	2,913	Surprise	496
		Anticipation	656
		Neutral	1,761

Table 2: Number of labelled tweets per emotion in each final corpus.

For the cross-domain part of the research, we also rely on the *Free Domain Corpus* built by Silva et al. (2020). This corpus was automatically collected using Twitter API3, by fetching tweets written in Portuguese that contained hashtags naming the emotions from Plutchik's wheel (*i.e.* joy, sadness, anger, fear, confidence, disgust, surprise and anticipation) and labelling them with the emotions (*i.e.* hashtags) found. In total, 230,857 not repeated tweets were collected from September 2015 to October 2016, almost 54 times the amount of tweets of the human annotated corpus.

It is important to notice, at this point, that the *Free Domain Corpus*, as the name implies, has not focused on any particular domain. Instead, Twitter's API was set to retrieve any tweets presenting any of the sought emotions. Table 3 shows the distribution of tweets in this corpus associated to each of Plutchik's emotions. As it turns out, there is a prevalence of *SAD* hashtags in the corpus, followed by *JOY*. Together, these correspond to over 77% of the corpus[1].

Annotated Emotions Count	
Emotion	**Count**
SAD	109,768
JOY	71,004
FEA	27,488
DIS	9,207
TRU	5,192
SUR	5,049
ANG	4,006
ANT	1,326

Table 3: Number of annotated tweets per emotion.

3.2 Experimental Setup

BERTimbau-cased was used in our experiments, in both its Large and Base versions. The hyperparameters were adjusted according to Ferreira[2]. Settings were then as follows:

- nclasses = 3;
- nepochs = 5;
- batch_size = 8;
- batch_status = 32;
- learning_rate = 1e-5; and
- early_stop = 2.

The only hyperparameter we changed was max_length (the maximum length of the tokenised text), which was set to 240. For this parameter, we also tested max_length = 200 in BERTimbau base.

All models were trained using the training dataset which is composed of a subset randomly extracted from the *Free Domain Corpus*, containing

[1] The total amount of tweets in Table 3 is higher than 230,857 because some tweets shown more than one emotion.

[2] Available at https://www.youtube.com/watch?v=GncyWR-dYW8: "Aula 7.6: HuggingFace: BERT para Classificação de Sequência | Linguística Computacional"

80% of their original tweets, with the other 20% building our test dataset.

Each model was trained per emotion pair, in a total of four models: joy vs sadness, trust vs disgust, anger vs fear and anticipation vs surprise.

The models were tested in two different test datasets. The first test dataset consists of 20% of the tweets in the *Free Domain Corpus*, randomly chosen (these tweets were not used for training). The second test dataset consisted of the *Stock Market Domain* corpus. The distribution of emotion labels across both test sets can be seen in Table 4. We calculated the precision, recall and F1-Score for each emotion.

Emotion	**Free Domain Corpus**	**Stock Market Corpus**
Joy	14,373 (31.6%)	531 (12.4%)
Sadness	22,167 (48%)	437 (10.2%)
Disgust	1,848 (4%)	709 (16.6%)
Trust	1,093 (2.4%)	851 (29.9%)
Anger	770 (1.7%)	256 (6.0%)
Fear	5,436 (11.8%)	232 (5.4%)
Surprise	999 (2.2%)	496 (11.6%)
Anticipation	280 (0.6%)	656 (15.3%)
Total	**46,171**	**4,277**

Table 4: Number of annotated tweets per emotion.

4 Results and Discussion

Surprisingly, fine-tuning both versions of BERTimbau (Large and Base) at the *Free Domain Corpus* training set led to nearly perfect results at the train set for the same domain, as illustrated in Table 5. Such a puzzling result can only be explained on the simplicity of the resulting corpus, which included the very tag used to classify it (since it is common to use the hashtag as a surrogate for some word in the text). This might have given the model the very clue it needed for such a high performance.

When tested in the *Stock Market Domain Corpus*, however, the tables turn considerably, and performance drops to values ranging from F1 nil, for BERTimbau Large and *Joy* to $F1 = 0.78$, also for BERTimbau Large, but with *Trust*. Table 6 details these results.

BERTimbau large outperformed the others BERTimbau bases for 3 emotions: Trust (78%), Anticipation (6%) and Surprise (61%). Next the difference between their best results: Trust (14%), Anticipation (4%) and Surprise (1%).

These results were obtained with $max\_length = 240$. By setting it to 200 one gets the results in Table 7. As it turns out, performance at some emotions, such as *Joy* and *Anger*, have raised considerably, whereas *Trust* have dropped.

BERTimbau base fine-tuned with $max\_length = 200$ outperforms BERTimbau large and base fine-tuned with $max\_lenght = 240$ for 4 emotions: Joy (51%), Disgust (67%), Anger (40%) and Fear (68%). The difference between their best results is Joy (7%), Disgust (5%), Anger (11%) and Fear (1%).

As it seems, changing the text length does affect performance, although it is not so clear how. To answer this question, we will have to carry out more experimentation, beyond what these preliminary results indicate.

Peculiarly, we could see impressive results where BERTimbau base with any configuration had a better performance than BERTimbau large. For example, *Sad* emotion results tied for both max_length settings (200 and 240) in the BERTimbau base, but the results still better than BERTimbau large (the difference was only 5%).

In this context, the worst BERTimbau large result was for the *Joy* emotion, the accuracy was zero. We can see 100% precision but zero for recall, meaning the model misclassifying all negative instances as positive.

Related literature Silva et al. (2020) trained with SVM with Tree Kernel model, wich obtained better results for two emotions: Anger (41%) and Anticipation (23%). However the Anticipation emotion had a very low F1 score. One possible reason may be the samples for the training dataset were the lowest among all emotions with 1,046 samples.

Also, building a reference corpus with distant supervision the way described in (Silva et al., 2020) may not be the best approach, leading to some distorted results within the corpus, with not much to carry out to other domains. Still, the system did not perform so badly for some emotions, which indicates this could be a direction to be pursued.

5 Conclusion

In this research, we set out to determine how a Large Language Model like BERTimbau would behave in a cross-domain situation where the source corpus was automatically annotated using a distant supervision technique. As it turns out, results by

	BERTimbau Base			BERTimbau Large		
Emotion	**Prec.**	**Recall**	**F1**	**Prec.**	**Recall**	**F1**
Sad	1.00	1.00	1.00	1.00	1.00	1.00
Joy	1.00	1.00	1.00	1.00	1.00	1.00
Trust	1.00	1.00	1.00	1.00	1.00	1.00
Disgust	1.00	1.00	1.00	1.00	1.00	1.00
Anger	0.99	0.99	0.99	1.00	0.99	1.00
Fear	1.00	1.00	1.00	1.00	1.00	1.00
Anticipation	1.00	1.00	1.00	1.00	1.00	1.00
Surprise	1.00	1.00	1.00	1.00	1.00	1.00

Table 5: Bertimbau large and base results after testing in the source domain (Free Domain) test set

	BERTimbau Base			BERTimbau Large		
Emotion	**Prec.**	**Recall**	**F1**	**Prec.**	**Recall**	**F1**
Sad	0.52	0.93	**0.67**	0.45	1.00	0.62
Joy	0.83	0.30	**0.44**	1.00	0.00	0.00
Trust	0.62	0.49	0.55	0.68	0.89	**0.78**
Disgust	0.51	0.63	0.57	0.80	0.51	**0.62**
Anger	0.88	0.17	**0.29**	0.68	0.16	0.26
Fear	0.52	0.97	**0.67**	0.50	0.92	0.65
Anticipation	0.80	0.01	0.02	0.81	0.03	**0.06**
Surprise	0.43	1.00	0.60	0.44	0.99	**0.61**

Table 6: Results after testing BERTimbau large and base in the target domain test set (Stock Market)

Emotion	**Prec.**	**Recall**	**F1**
Sad	0.54	0.89	**0.67**
Joy	0.80	0.38	**0.51**
Trust	0.74	0.57	0.64
Disgust	0.59	0.76	**0.67**
Anger	0.80	0.27	**0.40**
Fear	0.53	0.93	**0.68**
Anticipation	0.43	0.00	0.01
Surprise	0.43	0.99	0.60

Table 7: Results after testing BERTimbau base fine-tuned `max_length` =200 in the target domain test set.

both BERTimbau Large and Base, after test on the *Free Domain Corpus* test set (same domain used for training) reached almost 100% accuracy for all emotions.

This, in turn, has raised some concerns regarding whether the models have actually learned something or just memorised some existing pattern in the corpus. The fact that the hashtags used for classifications were part of the text furnished the clues that a memorisation effect might be taking place.

This suspect became stronger after the running of these trained models at the *Stock Market Domain Corpus*, where we could observe a very pronounced drop in performance for all emotions, some reaching as bad a result as a nil F1 score. When changing the maximum length of the input string, from 240 to 200, one sees some increased values e some decreasing. It is still not clear how this hyperparameter might help in this task, something we leave for future investigation. We also propose, for future work, to train the models in the stock market domain and evaluate the results.

References

Ahmed Abdelali, Sabit Hassan, Hamdy Mubarak, Kareem Darwish, and Younes Samih. 2021. Pre-training BERT on arabic tweets: Practical considerations. *CoRR*, abs/2102.10684.

Francisca Adoma Acheampong, Henry Nunoo-Mensah, and Wenyu Chen. 2021. Transformer models for text-based emotion detection: a review of bert-based approaches. *Artificial Intelligence Review*, 54:5789–5829.

Zishan Ahmad, Asheesh Kumar, Asif Ekbal, and Pushpak Bhattacharyya. 2021. Emotion driven crisis response: A benchmark setup for multi-lingual emotion analysis in disaster situations. In *2021 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8.

Hani Al-Omari, Malak A. Abdullah, and Samira Shaikh. 2020. Emodet2: Emotion detection in english textual dialogue using bert and bilstm models. In *2020 11th International Conference on Information and Communication Systems (ICICS)*, pages 226–232.

Antônio Manoel dos Santos Almeida Neto, Bruna Almeida Osti, Camila de Araujo Azevedo, Fernando Antonio Asevedo Nobrega, Henrico Bertini Brum, Luis Heriberto Osinaga Peinado, and Marciele Bittencourt. 2021. Sidi-nlp-team at idpt2021: Irony detection in portuguese 2021.

Ashish Bhatia, Hani Hagras, and Jason J. Lepley. 2018. Machine learning approach to extracting emotions information from open source data for relative forecasting of stock prices. In *2018 10th Computer Science and Electronic Engineering (CEEC)*, pages 142–147.

Patricia Chiril, Endang Wahyu Pamungkas, Farah Benamara, Véronique Moriceau, and Viviana Patti. 2022. Emotionally informed hate speech detection: a multi-target perspective. *Cognitive Computation*, 14(1):322–352.

Luna De Bruyne, Pepa Atanasova, and Isabelle Augenstein. 2022. Joint emotion label space modeling for affect lexica. *Computer Speech Language*, 71:101257.

Stephanie Dick. 2019. Artificial intelligence. *Harvard Data Science Review*, 1(1). Https://hdsr.mitpress.mit.edu/pub/0aytgrau.

Tom Diethe, Niall Twomey, and Peter Flach. 2015. Bayesian active transfer learning in smart homes. In *ICML Active Learning Workshop*, volume 2015.

Thiago Castro Ferreira. Aula 7.6: Huggingface: Bert para classificação de sequência | linguística computacional.

Bharat Gaind, Varun Syal, and Sneha Padgalwar. 2019. Emotion detection and analysis on social media. *ArXiv*, abs/1901.08458.

Wilfredo Graterol, Jose Diaz-Amado, Yudith Cardinale, Irvin Dongo, Edmundo Lopes-Silva, and Cleia Santos-Libarino. 2021. Emotion detection for social robots based on nlp transformers and an emotion ontology. *Sensors*, 21.

Luiz Hammes and Larissa Freitas. 2021. Utilizando bertimbau para a classificação de emoções em português. In *Anais do XIII Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana*, pages 56–63.

Wenjuan Han, Bo Pang, and Yingnian Wu. 2021. Robust transfer learning with pretrained language models through adapters.

Sabit Hassan, Shaden Shaar, and Kareem Darwish. 2021. Cross-lingual emotion detection. *CoRR*, abs/2106.06017.

Joshua Zoen-Git Hiew, Xin Huang, Hao Mou, Duan Li, Qi Wu, and Yabo Xu. 2019. Bert-based financial sentiment index and lstm-based stock return predictability. *arXiv: Statistical Finance*.

M. Horvat, A. Stojanović, and Ž. Kovačević. 2022. An overview of common emotion models in computer systems. In *2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO)*, pages 1008–1013.

Md. Yasin Kabir and Sanjay Madria. 2021. Emocov: Machine learning for emotion detection, analysis and visualization using covid-19 tweets. *Online Social Networks and Media*, 23.

Y. Kang, J. Wang, Y. Wang, S. Angsuesser, and T. Fei. 2017. Mapping the sensitivity of the public emotion to the movement of stock market value: A case study of manhattan. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W7:1213–1221.

Dejan Lazeski. 2020. Stock market prediction: A multi-class classification on emotions and sentiment analysis for tweets and news headlines. Master's thesis, Univerzita Karlova, Fakulta sociálních věd.

Clare H. Liu et al. 2017. Applications of twitter emotion detection for stock market prediction. Master's thesis, Massachusetts Institute of Technology.

Suhasini Matla and Srinivasu Badugu. 2020. Emotion detection framework for twitter data using supervised classifiers. *Data Engineering and Communication Technology*, pages 565–576.

Murilo Medeiros and Vinicius Borges. 2019. Tweet sentiment analysis regarding the brazilian stock market. In *Anais do VIII Brazilian Workshop on Social Network Analysis and Mining*, pages 71–82.

Joanna Michalak. 2020. Does pre-processing affect the correlation indicator between twitter message volume and stock market trading volume? *Ekonomia i Prawo. Economics and Law*, 19:739–755.

Raj Ratn Pranesh, Ambesh Shekhar, and Mehrdad Farokhnejad. 2020. CMTA: A framework for multilingual COVID-19 tweet analysis.

S Rossouw, Talita Greyling, and D Steyn. 2020. Sentiment, emotions and stock market predictability in developed and emerging markets.

Kashfia Sailunaz, Manmeet Dhaliwal, Jon Rokne, and Reda Alhajj. 2018. Emotion detection from text and speech: a survey. *Social Network Analysis and Mining*, 12:8.

Samant Saurabh and Kushankur Dey. 2020. Unraveling the relationship between social moods and the stock market: Evidence from the united kingdom. *Journal of Behavioral and Experimental Finance*, 26:100300.

Ramit Sawhney, Harshit Joshi, Alicia Nobles, and Rajiv Ratn Shah. 2021. Tweet classification to assist human moderation for suicide prevention. In *Proc Int AAAI Conf Weblogs Soc Media*, volume 15, pages 609–620.

Ashish Sharma, Dinesh Bhuriya, and Upendra Singh. 2017. Survey of stock market prediction using machine learning approach. *2017 International conference of Electronics, Communication and Aerospace Technology (ICECA)*, 2:506–509.

Fernando José Vieira Silva et al. 2020. *Cross-Domain Emotion Detection in Tweets: detecção de emoções inter-domínio em tweets*. Ph.D. thesis.

Carlos Simões, Rui Neves, and Nuno Horta. 2017. Using sentiment from twitter optimized by genetic algorithms to predict the stock market. In *2017 IEEE Congress on Evolutionary Computation (CEC)*, pages 1303–1310.

Fábio Souza, Rodrigo Nogueira, and Roberto Lotufo. 2020. Bertimbau: Pretrained bert models for brazilian portuguese. In *Intelligent Systems*, pages 403–417, Cham.

Fábio Souza, Rodrigo Frassetto Nogueira, and Roberto de Alencar Lotufo. 2019. Portuguese named entity recognition using BERT-CRF. *CoRR*, abs/1909.10649.

Alberto Valdes, Jesus Lopez, and Manuel Montes. 2021. UACH-INAOE at SMM4H: a BERT based approach for classification of COVID-19 Twitter posts. In *Proceedings of the Sixth Social Media Mining for Health (#SMM4H) Workshop and Shared Task*, pages 65–68, Mexico City, Mexico.

Martin J Willemink, Wojciech A Koszek, Cailin Hardell, Jie Wu, Dominik Fleischmann, Hugh Harvey, Les R Folio, Ronald M Summers, Daniel L Rubin, and Matthew P Lungren. 2020. Preparing medical imaging data for machine learning. *Radiology*, 295(1):4–15.

Chunfeng Yang, Jiajia Lu, Qiang Wu, and Huiyu Chen. 2021. Research progress of speech emotion recognition based on discrete emotion model. In *Journal of Physics: Conference Series*, volume 2010.

Samira Zad, Maryam Heidari, James H Jr Jones, and Ozlem Uzuner. 2021. Emotion detection of textual data: An interdisciplinary survey. In *2021 IEEE World AI IoT Congress (AIIoT)*, pages 0255–0261.

Feng Zhou, Shu Kong, Charless C. Fowlkes, Tao Chen, and Baiying Lei. 2020. Fine-grained facial expression analysis using dimensional emotion model. *Neurocomputing*, 392:38–49.