ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 11 Issue 01 January 2024

BERT for Satisfaction Analysis of Consumer Dispute Resolution

[1] Daniel G. Silva, [2] William B. A. M. Betker, [3] Daniel P. Gonçalves, [4] Ugo S. Dias

[1][2][3][4] University of Brasília, School of Technology, Department of Electrical Engineering, Brasília, Brazil
E-mail: [1] danielgs@unb.br

Abstract*— Several consumer protection entities have their own online platforms to collect consumers' experience and to mediate the resolution of conflicts between them and the suppliers. In Brazil, one of the most relevant tools for this purpose is the Consumidor.gov.br platform: it enables the communication of consumers with product/service suppliers in order to solve conflicts/disputes between the parts; at the end of the mediation process, the consumer writes his/her final opinion about the service provided and about the fulfillment of the initial expectations. This work proposes to build a sentiment analysis model through Transfer Learning, performing the fine-tuning process of the BERT model via a training set which aggregates well-known public data-sets of customer reviews in Portuguese, which is subsequently evaluated in the task of sentiment analysis of Consumidor.gov.br complaints, through a test set exclusively labeled for this purpose. The model is deployed in Senacon data analysis environment and is able to perform sentiment analysis of the finalized disputes on the platform, on a daily basis, with expected average F1-score of 0.72.

Index Terms*— Sentiment Analysis, Consumer Dispute Resolution, BERT, Transfer Learning

I. INTRODUCTION

In the context of the Digital Transformation, companies always try to keep up with the technological advances, whether in physical commerce, whether through the purchase and sale of products and services via social networks or via *e-commerce* platforms. After a purchase, customers can express personal opinions about their experience in the form of text and media published and shared through different platforms. Furthermore, several consumer protection entities have their own online platforms to collect consumers' experience and to mediate the resolution of conflicts between them and the suppliers.

One of the most relevant tools in Brazil for this purpose is the Consumidor.gov.br [1] platform, a service of the National Consumer Secretariat (Senacon - *Secretaria Nacional do Consumidor*) of the Ministry of Justice and Public Security (MJSP - *Ministério da Justiça e Segurança Pública*). This Secretariat was created in 2012 and its attributions are defined by the Consumer Protection Code and by the Decree 2,181/97, encompassing the planning, elaboration, coordination and execution of the National Consumer Relations Policy [2].

The *Consumidor.gov.br* platform enables the communication of consumers with product/service suppliers in order to solve conflicts/disputes between the parts. One of the system's functionalities allows the consumer, at the end of the mediation process with the company, to write his/her final opinion about the service provided and about the fulfillment of the initial expectations, when the complaint was submitted. Furthermore, the platform is jointly monitored by other Consumer Rights institutions, such as Public Defenders, Public Prosecutors, Regulatory Agencies and by all society in potential, since the information of each process composes a public database, available to any interested party, regardless of request, in open format, in accordance with the guidelines of access to information and active transparency, treated in Brazilian Presidency Decree 8,777 [3] and Law 12,527 [4].

The platform has a quite representative universe of complaints, in terms of coverage. It was launched in June 2014, since then, it has already registered more than 6.3 million complaints by the end of 2022, with a base of 4.2 million registered users and more than 1,275 accredited companies. Since 2020, more than 1 million complaints are processed annually [5]. This success results in technological challenges. The growth of its database and the maturation of its processes culminated in a great need for the automation of analyzes and the optimization of workflows. Naturally, several applications of automatic knowledge extraction can be designed. Due to the large amount of data, it is necessary to build automated applications capable of extracting relevant information from data.

An important information that can be automatically analyzed is the consumers' feedback about the services provided by the companies. The *Consumidor.gov.br* platform allows the consumer to register his/her overall perception about the service received when the complaint/dispute process is finalized. This sort of data offers an opportunity to employ Natural Language Processing (NLP) to infer the polarity of feelings and, consequently, to measure, even indirectly, the general opinion about the services provided by the companies.

However, there are some caveats: (*i*) the large volume of complaints, (*ii*) the impossibility of a thorough labeling process of these data, and (*iii*) the need for the model's prediction of new complaints to be performed in near real-

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 11 Issue 01 January 2024

time. Therefore, it is necessary to consider modern Machine Learning techniques for this challenge. Hence, this work proposes to build a sentiment analysis model through Transfer Learning [6], performing the fine-tuning process of the Bidirectional Encoder Representations from Transformers (BERT) model via a training set which aggregates well-known public data-sets of customer reviews in Portuguese, which is subsequently validated in the task of sentiment analysis of *Consumidor.gov.br* complaints, through a test set exclusively labeled for this purpose. Finally, the validated model is deployed in Senacon data analysis environment, becoming able to perform sentiment analysis of the finalized disputes on the platform, on a daily basis.

The rest of this work is organized as follows: Section II presents the related work; Section III describes the process of data preparation, followed by training, validation and deployment of the model in a production environment; Section IV presents the results obtained; and Section V presents the conclusions and future work.

II. RELATED WORK

One can find several works about Sentiment Analysis, since it is a problem that has been studied for a long time, within the area of NLP. Such scenario is also true to the analysis of customers' sentiment about products and services they have purchased. However, the sentiment analysis regarding the resolution (or not) of a customer dispute is a problem that is still little explored. Thus, this section analyzes recent works that involve (*i*) sentiment analysis in big data, (*ii*) consumer disputes resolution or (*iii*) the use of Transformer models for this task.

The author of [7] performs a systematic review of articles on sentiment analysis in big data. Based on different filters, 23 relevant articles are selected to analyze. The advantages and disadvantages of each technique are studied and their key issues are emphasized. Among the conclusions, the author points out that a better analysis of textual big data in terms of sentiment increases efficiency, flexibility and intelligence for a business to obtain promising business ideas.

The authors of [8] analyze the influence of consumer product reviews on new product development strategies. They use big data strategies in a dataset of 3 million online reviews obtained from a mobile app. The study finds out that the volume of reviews has a curvilinear relationship with customer responsiveness. Moreover, this responsiveness has a curvilinear relationship with product performance. The work contributes by demonstrating the influence of companies' ability to use online customer reviews and their impact on product performance.

The authors of [9] use data from social networks comments to study public opinion regarding the resolution of purchase conflicts on China's largest used goods e-commerce platform, through techniques of word frequency analysis, thematic analysis and sentiment analysis. The study reveals that disputes are mainly focused on return and refund issues, and indicates that the platform needs to improve the system design and increase the service's resource channels.

The authors of [6] study the BERT model for sentiment analysis in the Portuguese language. The work compares (*i*) different ways of aggregating the feature vector produced at the output of the model, (*ii*) its variants for Brazilian Portuguese and for multiple languages, and (*iii*) the combination of different datasets for training and testing. The results indicate the BERTimbau model as the best variant of BERT for the referred task and show that models trained with a dataset different from the test set still present satisfactory results.

The aforementioned works show that the application of Transformer models for sentiment analysis is a recent perspective. The same can be said about the use of big data for this task. However, the literature on sentiment analysis in the context of consumer disputes resolution is quite scarce and, to the best of our knowledge, non-existent when it comes to the Portuguese language and general platforms/services of consumer disputes resolution, such as the *Consumidor.gov.br* platform. Moreover, there are no works that use Transformer models for this specific context. Such gaps indicate the research direction of this work, as will be presented in more detail in the subsequent sections.

III. MATERIALS AND METHODS

It was already mentioned, in the previous section, that the Consumidor.gov.br platform is a tool for mediation of disputes between consumers and companies. The consumer registers his/her complaint and the company has up to 10 days to respond. After the response, the consumer has the opportunity to comment on the response, classify the complaint as Resolved or Not Resolved and, finally, indicate the degree of satisfaction with the service received [5].

This work focus on analyzing the sentiment associated with the final comment left by the consumer. Since the number of complaints is very large, it is not possible to manually label all of them. Thus, we adopt the strategy of labeling a sample of comments to be the test set and training the BERT model through public datasets of product reviews by consumers. Then, the model's performance is evaluated on the test set and a complexity optimization is performed via the quantization of the learned parameters, in order to speed up the inference time, which should enable the model operation in a production environment. Next, we describe each of these steps.

A. Consumidor.gov.br Database

The database systems staff of MJSP provided a read-only copy of the platform's database for the development of the solution proposed in this work. The available information are:

- number of protocol;
- situation of the complaint;

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 11 Issue 01 January 2024

- consumer's report;
- supplier's response;
- solution status indicator;
- consumer's final evaluation text.

Due to the unstructured nature of the original data and their massive volume, an Extract, Transform and Load (ETL) process is performed as follows: the complaints records are extracted from the database through the Azure DataFactory tool, which allows data integration from different sources. The data is then transferred to a Data Lake [10] divided into different zones. The initial storage zone is called RawZone, where the data are stored in their original form, without transformations.

Then, data is processed through the DataBricks tool, where the necessary cleanings are performed for: removal of personal data, in accordance with the Brazilian General Data Protection Law (LGPD - *Lei Geral de Proteção de Dados*) [11]; deduplication of records and application of algorithms to ensure data integrity. After this step, the data is stored in the TrustedZone, a data repository where treatments have already been performed to ensure data quality, such that they can be considered accurate and reliable.

Finally, processing is performed to provide the data in its final format, tailored for the sentiment analysis application. After an enrichment process that involves steps such as joining related tables and pre-processing the text of the complaints, data reaches the RefinedZone, a zone that contains refined and ready-to-consume data: either to firstly extract the necessary data to create the test set that will validate the model training, or to apply the trained model to new data in MJSP BI environment.

B. Test Set Creation

The labeling process of the test set was performed by three volunteers, who previously defined a set of criteria to assign a category to each record. During this preparation phase, some complaints were evaluated together to clarify possible doubts. The labeling consisted of reading the texts written by the consumers, when the complaint process was finalized, and identifying the sentiment that best suited the text: satisfied, dissatisfied or neutral. The neutral sentiment was used when it was not possible to identify in the consumer's text any minimum expressiveness of sentiment. Table I presents examples of labeled texts for each sentiment.

After a first round of labeling, the complaints that were not classified with the same sentiment by all the volunteers were separated. A new round was then started only with these records, in which the volunteers had access to the labels that were previously given and had to, together, define which one best suited. If it was not possible to reach a consensus, the text was discarded, i.e. it would not enter the final set of labeled examples.

When the labeling process ended, a total of 496 labeled complaints was achieved, as shown in Table II. The distribution of the classes is unbalanced, with a predominance of dissatisfied consumers. This is a common scenario in the context of customer reviews, since dissatisfied consumers tend to express their dissatisfaction more than satisfied consumers express their satisfaction.

C. Training and Validation Set

As previously mentioned, a thorough labeling of Consumidor.gov.br platform data is not possible, due to the large volume of complaints. Therefore, we propose to use other sources of data, related to consumer sentiment analysis, to compose the training and validation sets. Hence, in this work, we use three public datasets of product reviews by consumers, in Portuguese: Olist [12], Buscapé [13] and B2W [14].

Table I. Test set examples.

Sentiment	Example
Satisfied	*"Problema resolvido com rapidez e eficiência. Obrigado."* (Problem solved quickly and efficiently. Thank you.)
Dissatisfied	*"O objeto foi entregue com muito tempo de atraso e só consegui ser respondida após o mesmo ser entregue."* (The item was delivered very late and I was only able to receive a response after it was delivered.)
Neutral	*"Não consigo ler a resposta do fornecedor mas consegui as passagens."* (I can't read the supplier's answer but I got the tickets.)

Table II. Test set class distribution.

Class	Quantity	Percentage
Satisfied	145	29.2%
Dissatisfied	244	49.2%
Neutral	107	21.6%

The Brazilian E-Commerce Public Data-set by Olist [12] contains 100,000 orders from 2016 to 2018, which were placed in several Brazilian e-commerce marketplaces. Each order has a diverse set of information such as location, status, price. For the sake of this work, we consider only the final evaluation texts of the orders, issued by the consumers, which comprises around 30,000 examples.

The Buscapé data-set [13] is composed of more than 80,000 product reviews made by consumers in September 2013, captured from Buscapé[1], one of the most important price comparison websites in Brazil.

The B2W-Reviews01 [14] dataset is composed of more than 130,000 product reviews made by consumers between January and May 2018, collected from an important Brazilian e-commerce site, Americanas[2].

[1] https://www.buscape.com.br

[2] https://www.americanas.com.br

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 11 Issue 01 January 2024

All these datasets are public and contain, besides a text review, the consumer's evaluation of the product, in the form of a score ranging from 1 to 5. In order to use them for the task of sentiment analysis of the *Consumidor.gov.br* complaints, we perform a conversion of the score into the three sentiment classes used in this work, as follows: scores 1 and 2 are considered as "Dissatisfied", score 3 is considered as "Neutral" and, finally, scores 4 and 5 are considered as "Satisfied". The aggregation of the three datasets results in a total of 259,107 examples, which are randomly divided into training (70%) and validation (30%) sets, respectively.

Table III presents the class distribution in the training and validation sets. The distribution is unbalanced, with a predominance of examples of the "Satisfied" class.

D. Model

BERT is a language model based on the Transformer architecture [15], designed in 2019 for numerous NLP applications, including sentiment analysis [16]. It is a multi-layer bidirectional Transformer encoder: it is fed with the input text at once, instead of the traditional sequential form of recurrent architectures, while explores consolidated approaches designed from previous strategies, namely the attention mechanism and representation of words in latent spaces. The base model consists of a stack of 12 Transformer encoders. Each Transformer block has 6 identical layers, each of which has two sub-layers, the first being a multi-head self-attention mechanism and the second being a simple fully connected feed-forward network. A residual connection is then applied around each of the two sub-layers, followed by a Layer Normalization technique.

Table III. Training and validation sets class distribution.

Class	Training	Validation
Satisfied	121,680 (67.1%)	52,107 (67.0%)
Dissatisfied	37,831 (20.9%)	16,145 (20.8%)
Neutral	21,863 (12.0%)	9,481 (12.2%)
Total	**181,374 (100%)**	**77,733 (100%)**

Using a language model like BERT consists of two steps, namely pre-training and fine-tuning. In pre-training, the model undergoes an unsupervised training in two generic tasks. In fine-tuning, the model is initialized with the pre-trained parameters, which continue to be adjusted using labeled data from downstream tasks, such as sentiment analysis [6]. BERT is available in two sizes: $\text{BERT}_{\text{BASE}}$, with 110 million parameters, and $\text{BERT}_{\text{LARGE}}$, with 340 million.

Brazilian researchers recently performed the pre-training of BERT models for the Portuguese language, calling these variants BERTimbau [17]. To build them, they used the brWaC corpus [18], so far, the largest open corpus of texts in Portuguese. BERTimbau is also available in two sizes, Base and Large, respectively pre-trained on the Base and Large versions of BERT, through the same tasks.

The BERT model is available in the HuggingFace[3] library, which provides a Python interface for the use of several NLP models, including BERT and its variants. The fine-tuning performed for this work considers the BERTimbau model, in its Base version, and is described in the next section.

1) *Fine-tuning*

Due to the complexity of BERTimbau, the fine-tuning for the classification task was performed on a dedicated machine with an Nvidia GTX 1080Ti GPU with 11 Gigabytes of GDDR5X memory. Moreover, the design of the Transformer architecture demands the documents (sequences of tokens) to be truncated to a fixed maximum size.

In order to select an appropriate maximum size, the length of the sequence of tokens is computed for each entry in the training set and a histogram is built, as shown in Fig. 1. Thus, it was defined as a criterion to choose the sequence length that encompasses the 95th percentile of the set: in this case, the approximate value that meets is 128. From there, all documents are truncated to up to 128 tokens.

A batch size of 83 examples was used, in order to fit the training convergence with the maximum use of GPU memory. The training was performed for 10 epochs, with the final model being the one from the epoch in which the best F1-Score value was obtained in the validation set. The cross-entropy function was used as the cost function, with a soft-max output layer.

After the end of the fine-tuning process, the model is validated on the test set, which was manually labeled as described in previous sections. The assessment is performed via the calculation of Precision, Recall and F1-Score metrics.

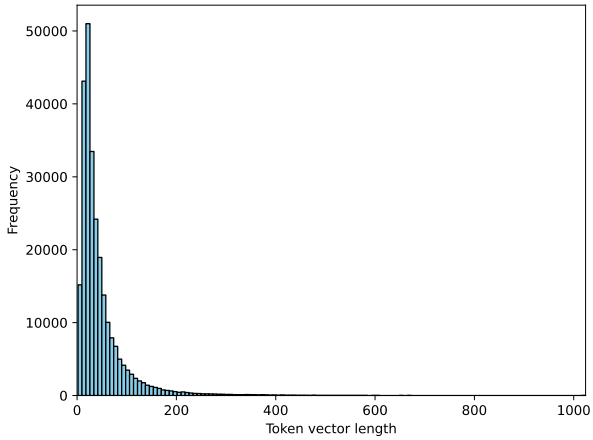


Figure 1. Documents size (number of tokens) histogram.

IV. RESULTS AND DISCUSSION

The fine-tuning took 355 minutes to complete. The best

[3] https://huggingface.co/

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 11 Issue 01 January 2024

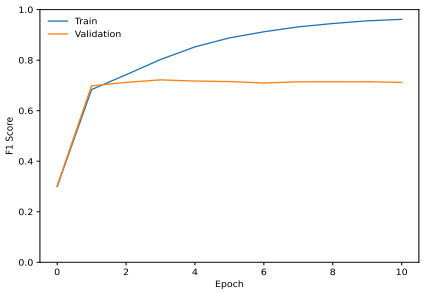
F1-Score value was obtained in the 3rd epoch, with a value of 0.7221. After that, the F1-Score of the validation set starts to decline, indicating a potential situation of over-fitting, as shown in Fig. 2a. A similar result occurs for the Accuracy, as shown in Fig. 2b.

With the model fine-tuned, the next step is to validate it on the test set. The results are shown in Table IV.

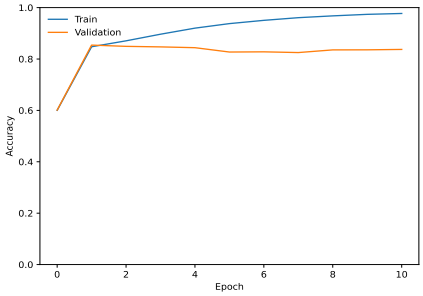
Despite a good value of average F1-Score, the model presents a lower performance for the Neutral class, which can be explained by the low amount of training data belonging to this class. However, we can consider as positive the results achieved specifically for the Satisfied and Dissatisfied classes, especially when we recall that the training set was not built from the Consumidor.gov.br platform, but from data of consumers in similar, but distinct, circumstances from the target platform, in which they are making comments about the resolution (or not) of a consumer dispute with a certain company.

It is also important to note the slightly superior performance of the model for the Dissatisfied class. It is precisely this class that is usually first analyzed by Senacon, since prioritizing the analysis of complaints from dissatisfied consumers with the process on the platform may indicate improvements in the secretariat's policies.

Finally, Fig. 3 presents the confusion matrix, for a better visualization of the model's classification capability.



(a)



(b)

Figure 2. Model F1-Score (a) and Accuracy (b) curves during the fine-tuning process.

Table IV. Test set results. Average statistics are weighted according to the proportion of the class in the test set.

Class	Precision	Recall	F1-Score
Dissatisfied	87%	77%	82%
Neutral	40%	45%	42%
Satisfied	75%	83%	79%
Average	**73%**	**72%**	**72%**

A. Model Deployment

After the model is trained and validated, the next step is to integrate it into the Senacon data analysis environment and make the results available through Business Intelligence Dashboards.

The main objective, in this phase, consists of using the model to calculate metrics associated to consumer satisfaction with companies and present results that are always up to date.

The files containing the trained model and its structure were converted to the ONNX Run-time format[4], so that the inference process could be accelerated to a viable execution time, considering the size of the platform's database. Then, they can be loaded by the scripts that process the data from the BI dashboard within the Azure platform.

Each evaluation text of a consumer's complaint over the last 5 years within the platform was classified in a process that took about 35 hours to complete, due to the large number of records in the historical database. Then, an incremental update process was implemented, executed daily in an automated way, classifying only the new evaluation texts of recently finalized complaints. This process, in turn, is executed in a few minutes and ensures that the data used in decision making reflects recent periods.

The probabilities produced by the model, for each class, are used to visualize the results. Since the individual analysis of the texts would not have practical interest for the Senacon technical staff, the complaints are grouped and the average statistics of each class are presented in a tab of the BI Dashboard, built with the PowerBI tool. Fig. 4 presents an example of visualization of the results (the names of the suppliers are anonymized). Note that it is possible to analyze texts grouped by supplier, economic groups or market segments, which will eventually help to identify suppliers that are associated to a higher degree of dissatisfaction with the resolution of complaints, for example.

[4] https://onnxruntime.ai/index.html

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 11 Issue 01 January 2024

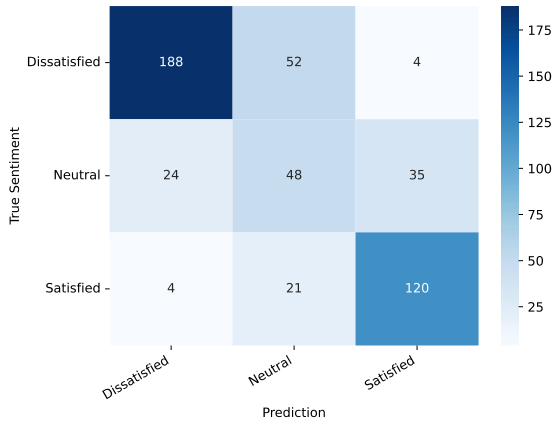


Figure 3. Confusion Matrix over the test set.

Figure 4. Dashboard for visualization of the metrics calculated via the sentiment analysis model.

V. CONCLUSION

This work presented a solution proposal for the sentiment analysis of the final evaluations of the complaints against companies, registered by consumers in the *Consumidor.gov.br* platform. The solution uses the well-known language model based on the Transformer architecture, BERT, which went through the fine-tuning process for the task of text classification into three classes: Satisfied, Neutral and Dissatisfied. Specifically, the variant of the BERTimbau model was used, pre-trained for the Portuguese language.

Due to the fact that the platform has a base of complaints that grows daily in the order of thousands and that does not have any previous labeling, a fine-tuning strategy of the model was designed via a data-set formed by public databases of consumers' product and service reviews, which was then validated through a test set manually labeled, this one, particularly, with the texts of consumer reviews of the Consumidor.gov.br platform.

The results indicate a satisfactory performance of the model, with an average F1-Score of 72%, average Precision of 73% and average Recall of 72% in the test set. Finally, the validated model was incorporated into the internal data analysis environment of Senacon, which enables the visualization of daily updated metrics of consumer satisfaction with the companies, in a timely manner for decision making.

As future work, we intend to improve the sentiment analysis model with the inclusion of more training data, in order to improve the overall performance, especially for the Neutral class. It is also important to perform a feasibility study about the use of more recent language models, with performance even superior to BERT in several natural language processing tasks.

REFERENCES

[1] Senacon, "Plataforma - Consumidor.gov.br," [Online]. Available: https://www.consumidor.gov.br/. [Accessed december 2023].

[2] Senacon, "Senacon - Ministério da Justiça e Segurança Pública," 2021. [Online]. Available: https://www.gov.br/mj/pt-br/assuntos/seus-direitos/consumidor. [Accessed december 2023].

[3] Presidência da República, "Decreto Nº 8.777, Institui a Política de Dados Abertos do Poder Executivo federal," 2021. [Online]. Available: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/d8777.htm. [Accessed december 2023].

[4] Presidência da República, "Lei Nº 12.527," 2021. [Online]. Available: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm. [Accessed december 2023].

[5] Senacon, "Boletim Consumidor.gov.br," Ministério da Justiça e Segurança Pública, Brasília, 2023.

[6] F. D. Souza e J. B. d. O. e. S. Filho, "BERT for Sentiment Analysis: Pre-trained and Fine-Tuned Alternatives," em Computational Processing of the Portuguese Language, Fortaleza, 2022.

[7] M. Hajiali, "Big data and sentiment analysis: A comprehensive and systematic literature review," Concurrency and Computation: Practice and Experience, 2020.

[8] S. Zhou, Z. Qiao, Q. Du, G. A. Wang, W. Fan e X. Yan, "Measuring Customer Agility from Online Reviews Using Big Data Text Analytics," Journal of management information systems, pp. 510-539, 2018.

[9] Y. Liu e Y. Wan, "Consumer Satisfaction with the Online Dispute Resolution on a Second-Hand Goods-Trading Platform," Sustainability, 2023.

[10] "An Introduction to Data Lake," i-manager's Journal on Information Technology, 2016.

[11] Presidência da República, "Lei 13.709, Lei Geral de Proteção de Dados Pessoais (LGPD)," 2018. [Online]. Available: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. [Acessed december 2023].

[12] Olist e A. Sionek, "Brazilian E-Commerce Public Dataset by Olist," Kaggle, 2018. [Online]. Available: https://www.kaggle.com/dsv/195341. [Acessed december 2023].

[13] N. Hartmann, L. Avanço, P. Balage, M. Duran, M. das Graças Volpe Nunes, T. Pardo e S. Aluísio, "A Large Corpus of Product Reviews in Portuguese: Tackling Out-Of-Vocabulary Words," em Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14), Reykjavik, 2014.

[14] B2W, "B2W-Reviews01," GitHub, 2018. [Online]. Available: https://github.com/americanas-tech/b2w-reviews01.

[15] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser e I. Polosukhin, "Attention is All you Need," em Advances in Neural Information Processing Systems, Long Beach, 2017.

[16] J. Devlin, M.-W. Chang, K. Lee e K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," em 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), Minneapolis, 2019.

[17] F. Souza, R. Nogueira e R. Lotufo, "BERTimbau: Pretrained BERT Models for Brazilian Portuguese," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 403--417, 2020.

[18] J. A. Wagner Filho, R. Wilkens, M. Idiart e A. Villavicencio, "The brWaC Corpus: A New Open Resource for Brazilian Portuguese," em Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), Miyazaki, 2018.