Application of Multitask BERT Models for the Classification of Sentiments, Types of Places, and Magical Towns in Spanish Tourist Reviews

Danileth Almanza-Gonzalez[1,*], Jairo Serrano Castañeda[1,†], Juan Carlos Martinez-Santos[1,†] and Edwin Puertas[1,†]

[1]*Universidad Tecnológica de Bolívar, Cartagena, Colombia*

Abstract

This article presents a system based on multiclass and multitask BERT models developed in the context of the IberLEF 2025 competition, specifically for the task "Sentiment Analysis and Magical Towns Detection." The objective was to simultaneously address three key functions in Spanish tourist reviews: sentiment polarity prediction, classification of the type of place (hotel, restaurant, or attraction), and identification of the mentioned Magical Town. The proposed architecture integrates advanced preprocessing techniques, data balancing strategies (RandomOverSampler, Condensed Nearest Neighbour, and SMOTE), and the explicit activation of attention layers to enhance the model's performance and interpretability. Despite being trained with a reduced sample, the system achieved competitive results, standing out, particularly in classifying the type of place. The results demonstrate the potential of pre-trained language models for complex tasks in the tourism domain.

Keywords

Sentiment Analysis, Tourist Reviews, BERT Model, Natural Language Processing (NLP), Tourism

1. Introduction

Feelings are emotional manifestations that arise from situations and moods experienced by human beings, directly influencing their decisions and behaviors. These emotions emerge from constant interaction with the environment and lived experiences, thus becoming a valuable source of information. Understanding and analyzing these feelings is particularly relevant in tourism, as it allows for adapting the tourism offer according to the emotional influences that motivate tourists [1]. Nowadays, technological advancements have significantly enhanced individuals' interaction with various digital platforms specialized in tourism, such as TripAdvisor and social media. These platforms provide valuable information about user behavior, primarily through their reviews. Natural language processing (NLP) and artificial intelligence (AI) make it possible to perform sentiment analysis on these texts. It allows for identifying the emotional polarity expressed in the opinions and recognizing the type of places under review. It provides key data to understand user behavior during their trips [2].

With the emergence and development of smart cities and the growing adoption of data-driven approaches, platforms like TripAdvisor have become fundamental tools for tourism studies. These data analysis-based strategies enable the planning and promotion of the tourism offer according to travelers' specific preferences and interests. In particular, sentiment analysis provides relevant information about tourists' satisfaction levels, which is essential for improving the quality of the tourism experience. However, despite the increasing volume of data available on these platforms, one of the main challenges lies in correctly interpreting such data to obtain accurate and valuable information for decision-making.

It highlights the need for models that not only detect the feelings expressed in reviews but also identify the context in which they occur, such as the type of place visited and its geographical location, key aspects for better understanding the tourism experience [3].

Given this context, the "Sentiment Analysis and Magical Towns Detection Task Iberlef 2025" aims to infer the sentiment polarity in reviews, classifying them on a scale from 1 to 5. It also seeks to identify the type of place referred to in each review. This aspect contributes to understanding the activities and actions carried out by tourists during their travel experiences and determining the Magical Town in Mexico to which each comment belongs. The data used in this task come from reviews obtained from the TripAdvisor platform [4][5]. In this regard, the research developed for this task used BERT-based models specifically trained for sentiment analysis. We applied advanced data processing techniques and different balancing strategies to significantly improve the performance and accuracy of the implemented models. As a result, the proposal reached 46th place in the task's official competition, representing considerable participation. The repository is available via the following link (available after reviewing.)
[1]

2. Related Work

In recent years, various studies in the field of tourism have focused on applying sentiment analysis techniques and natural language processing, aiming to understand travelers' preferences, emotions, and behaviors based on digital content such as online reviews. These studies have made it possible to classify the emotional polarity of comments, identify the type of places visited, and extract relevant contextual information to improve decision-making and personalize tourism services. Within the framework of Rest-Mex 2023, a study analyzed sentiments in TripAdvisor reviews, classifying polarity (1–5), type of place, and associated country, and grouped tourism news into four topics. They used transformer-based models such as RoBERTa and BETO, along with oversampling and back-translation techniques, achieving significant performance improvements [6, 7, 8]. Another work proposed a system of transformer classifiers: for polarity, they used cascaded binary classifiers, and for the type of attraction and country, multiclass models with BERT and RoBERTa; BETO obtained the best result (F1 = 0.719) [9]. Finally, a third study applied RoBERTa and phonesthetic embeddings in financial sentiment analysis, using an SVM classifier to predict polarity toward economic entities, highlighting the potential of sentiment analysis in other domains [10].

On the other hand, a study proposes a predictive model of tourist destinations based on the interests and comments of Iranian tourists using sentiment analysis through text and data mining. They used the CRISP-DM methodology, applying clustering with X-means and classification with Decision Trees. The model achieved an accuracy of 95%, and they integrated it as a recommendation system for travel agencies [11]. An article proposes a model incorporating social media sentiment analysis into business strategic planning, specifically to construct a Balanced Scorecard (BSC) automatically. The authors used lexical approaches (VADER) and machine learning (SVM and Naive Bayes) to classify opinions from TripAdvisor and Facebook. The model was applied to a tourist restaurant, facilitating strategic decisions [12]. In a study, more than 12,000 TripAdvisor reviews about tourist attractions in Guanajuato, Mexico, were analyzed using NLP techniques such as Mutual Information and the Jaccard coefficient. The objective was to identify recurring themes according to the polarity of the experience. The main negative themes detected were cleanliness and prices, with the Mummy Museum, the Alley of the Kiss, and the Hidalgo Market being the worst-rated sites [13].

Another study proposes a comprehensive deep learning-based method (CDLM-ITT) to identify topics and sentiments in tourism texts. It uses web scraping, pre-trained embeddings, UMAP, clustering, and the GPT-3 model to generate topic descriptions and lexical sentiment analysis. It was applied to over 3,000 news articles about Cancún in U.S. and Canadian media, successfully detecting relevant topics such as tourism [3]. An article proposes a method to improve the accuracy of sentiment analysis in tourist reviews through a feature selection technique based on linguistic rules. They used POS tagging,

---

[1]https://github.com/VerbaNexAI/IberLEF2025/tree/main

n-grams, and statistical filters such as Information Gain, Chi-Square, and Gini Index, combined with a majority voting technique (MVT), achieving an accuracy of 94.7% [14]. Likewise, a study proposes a sentiment analysis method based on Formal Concept Analysis (FCA) to build polarity dictionaries automatically adapted to specific texts. It uses document-term matrices, formal concept extraction, and polarity weight calculation for each term, comparing its performance with standard dictionaries. The results show that the dictionaries generated with FCA outperform traditional approaches in AUC and accuracy [15].

3. Data

IberLEF 2025 provides the data used in this study as part of the Sentiment Analysis and Magical Towns Detection task. This task aims to infer, from tourist reviews, three key elements: (1) the polarity of the expressed sentiment, classified on a scale from 1 (very negative) to 5 (very positive), based on the original score given by the tourist; (2) the type of place mentioned in the review hotel, restaurant, or tourist attraction based on relevant textual features; and (3) the Magical Town referred to in the review. The data come from the TripAdvisor platform. The training set includes 208,051 instances with the following columns: title, review content, polarity, name of the Magical Town, and the corresponding region of the state of Mexico. In total, it considers 40 locations distributed across 19 regions. The test set, in turn, consists of 89,166 reviews, each with the following columns: identifier (ID), the title assigned by the tourist, and the review text.

4. Architecture
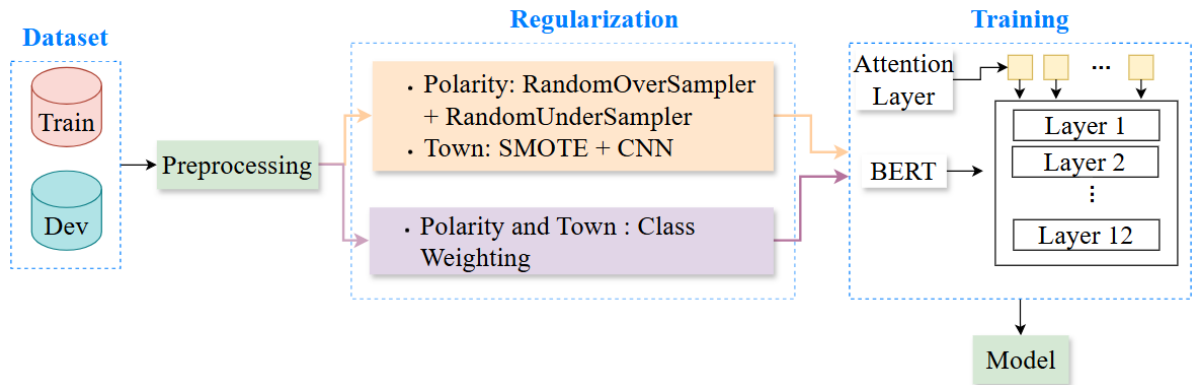
Figure 1: Architecture system.

In the methodology developed for training and validating the model, we implemented various techniques, ranging from text pre-processing to final evaluation of results. This integration enabled the simultaneous handling of multiple classification tasks based on Spanish tourist reviews. Figure 1 shows the general process of the proposed system architecture, which includes specific modules for text cleaning and tokenization, the configuration of the multi-task pipeline with BERT models, the application of data balancing techniques, the activation of attention layers for interpretive analysis, and the evaluation of performance using standard metrics.

4.1. Preprocessing

We carried out text pre-processing using the TextProcessing class, which applies cleaning and normalization techniques for natural language in Spanish. Initially, characters are normalized using Unicode data to remove accents and non-ASCII symbols. Subsequently, we removed punctuation marks, special symbols, numbers, and other irrelevant patterns using regular expressions. In addition, we replaced

emojis and URLs with generic tags such as [EMOJI] and [URL]. We also converted the text to lowercase, and we removed unnecessary spaces. These transformations allow the text to be standardized and reduce informational noise. The result is clean and structured text, ready to be processed by deep learning models such as BERT. This pre-processing improves the quality of the textual representation and, therefore, the model's performance.done

4.2. Tokenization and Data Preparation

In the initial phase of the process, we carried out tokenization and data preparation using custom classes to generate datasets and dataloaders compatible with transformer models, specifically BERT. We defined two specific classes: SentimentDataset for the polarity task and TextDataset for the site type and magical town tasks. Both classes implemented methods to load text from specific columns, tokenize using BERT tokenizer (AutoTokenizer and BertTokenizerFast), truncate or pad sequences to a fixed length (max_length=500), and return tensors ready for training. Using the create_dataloader method simplified the creation of batches (batch_size=8) that were later fed into the models during training and evaluation.

4.3. Pipeline Configuration and Hyperparameter Tuning

We developed a multi-task pipeline, consisting of three independent estimators based on the pre-trained BERT model (dccuchile/bert-base-spanish-wwm-uncased) specialized for Spanish, which simultaneously tackled the classification tasks of polarity, site type, and magical town. The hyperparameters remained consistent across the three estimators, setting a learning rate 2e-5, regularization via weight_decay of 0.01, a batch size of 8, and a fixed maximum sequence length of 500 tokens. We conducted a specific training experiment to determine the optimal number of epochs, evaluating from epoch 1 to epoch 4. During this evaluation, we carefully analyzed the performance evolution of the model using standard metrics such as precision, recall, and F1-score.

4.4. Regularization via Data Balancing Techniques

Given the imbalanced distribution in the data, we applied various balancing techniques to improve model performance in the polarity and magical town classification tasks. In the first methodology, oversampling was initially performed using the RandomOverSampler method for the polarity task, thus increasing the representation of minority classes. Subsequently, we applied RandomUnderSampler to slightly reduce the number of instances of the majority classes, aiming for a more balanced distribution. For the magical town prediction task (Town), a more complex strategy was employed, which first used CondensedNearestNeighbour undersampling to remove redundant examples and then applied SMOTE (Synthetic Minority Oversampling Technique) to generate additional synthetic examples, thus strengthening the representation of less frequent classes. In an alternative second methodology, we used a technique based on assigning class weights via the class_weight parameter, adjusted proportionally according to the frequency of each class in the Polarity and Town tasks. This technique allows the model to penalize errors in underrepresented classes more heavily during training.

4.5. Activation of Attention Layers in the Models

A notable feature implemented in this work was the explicit activation of the BERT model's attention layers during training and inference. Specifically, during the initialization of the custom estimators (BetoEstimator and BertTokenClassifier), the parameter output_attentions=True was set within the AutoConfig and BertConfig classes, respectively. This configuration enabled the model to provide final predictions (logits) and to obtain detailed attention matrices reflecting the importance assigned to each token by each model layer. These matrices are essential for subsequent interpretability analysis and understanding of the model's decisions. In parallel, we also trained a variant of the model where the attention layers were not activated (output_attentions=False). When comparing both variants in terms

of performance, we observed that explicitly activating the attention layers allowed the model to achieve better results.

4.6. Architecture of the BERT Models Used

We based the core of the developed system on pre-trained BERT-type models. Specifically, we used the Spanish variant dccuchile/bert-base-spanish-wwm-uncased, an optimized version with whole-word masking that enhances the understanding of the Spanish linguistic context. We configured AutoModelForSequenceClassification to classify up to 5 classes for the polarity task. In contrast, BertForTokenClassification was used and adapted for specific token classification for the type and magical town tasks, enabling attention and hidden state outputs. Both implementations leveraged BERT representational capacity to capture deep semantic information from Spanish text.

4.7. Training and Evaluation of the Multi-Task Pipeline

The core stage consisted of training the multi-task pipeline that integrated models for the three defined tasks: polarity, site type, and magical town. Each task was trained separately within the pipeline, using previously balanced and prepared data. Once we completed training, predictions were made on the test set, generating specific categorical results for each task. The dataset was split into training (70%) and test (30%) sets to ensure proper model evaluation. Before this split, we took a random sample of 10,000 records from the original dataset. Additionally, we performed basic text preprocessing using a custom class called TextProcessing, which was configured for the Spanish language since the data is in Spanish.

4.8. Calculation and Interpretation of Performance Metrics

We used specific performance evaluation metrics for each task within the pipeline. For polarity, we adapted the original labels to a range from 0 to 4 to facilitate the calculation of standard metrics such as precision, recall, and F1-score. We followed similar procedures in the tasks of type and magical town. This quantitative analysis confirmed the overall effectiveness of the model. It allowed the identification of potential areas for improvement in future development iterations.

5. Experiments Conducted and Training

First, we implemented a dual balancing strategy that combined oversampling and undersampling techniques for the Polarity and City (Pueblo Mágico) tasks, activating the attention layer. Specifically, for the Polarity task, RandomOverSampler and RandomUnderSampler were applied. In contrast, we used a more robust approach with SMOTE (k=2) and Condensed Nearest Neighbour (k=3) for the City task. As shown in Table 1, progressive improvement in performance over the four training epochs. For example, in the City task, the Macro F1 score significantly increased from 0.3479 in Epoch 1 to 0.4910 in Epoch 4, highlighting the positive impact of this strategy. Similarly, the Polarity and Type tasks improved, reaching Macro F1 scores of 0.5600 and 0.9611 in Epoch 4. It demonstrates that combining these balancing techniques with attention layer activation is highly effective, especially in multiclass contexts with imbalanced distributions.

To evaluate the influence of the attention layers, the previous experiment was replicated by explicitly turning off this functionality in the BERT model. As shown in Table 2, this modification negatively impacted performance, particularly in the City task, where the Macro F1 score dropped to 0.3462, similar to the value observed in the first epoch of the model with attention activated. This decline demonstrates that although balancing techniques are essential, attention enables the model to capture more complex contextual relationships between tokens, thereby improving prediction quality. In the Polarity task, the Macro F1 score dropped to 0.5348, while the Type task was barely affected, maintaining a high score of 0.9630. This analysis confirms that attention layer activation improves overall performance

and is particularly critical for classification tasks with many imbalanced classes, such as Pueblo Mágico classification.

Subsequently, we explored a second methodology based on the class weighting technique (class_weight) applied to the Polarity and City tasks, with the attention layer activated. As detailed in Table 3, the results were mixed. While performance in the Polarity task remained competitive Macro F1 = 0.5798 and even showed a slight improvement compared to the model with dual balancing, the City task experienced a significant drop in Macro F1 score to 0.3856. It suggests that although class weighting may be a practical solution for moderately imbalanced scenarios, it is insufficient for multiclass tasks with highly unequal distributions, such as the classification of the 40 municipalities. The Type task, for its part, remained robust with a Macro F1 score of 0.9528, reaffirming that this classification is less sensitive to changes in the balancing method.

Comparing the three analyzed methodologies, it concluded that the combination of oversampling and undersampling techniques and attention layer activation offers the best overall performance, especially for the most challenging task: the classification of Pueblos Mágicos. While class weighting maintained competitive metrics in the Polarity task, it limited effectiveness in scenarios with high-class imbalance. On the other hand, disabling attention led to a generalized drop in performance, reinforcing its importance in BERT-based models. With attention activated, we obtained the best results using a hybrid balancing strategy (SMOTE + CNN and ROS + RUS). It allowed the model to learn richer and more generalizable representations, especially in challenging multiclass classification and extreme imbalance contexts.

Table 1

Results with balancing technique RandomOverSampler + RandomUnderSampler (Polarity), SMOTE + CNN (Town) and the attention layer activated at different epochs.

Epochs	Polarity			Type			Town		
	Macro Prec.	Macro Recall	Macro F1	Macro Prec.	Macro Recall	Macro F1	Macro Prec.	Macro Recall	Macro F1
1	0.5320	0.5189	0.5201	0.9664	0.9624	0.9642	0.5546	0.3040	0.3479
2	0.5471	0.5348	0.5467	0.9607	0.9550	0.9573	0.5875	0.3389	0.4181
3	0.5620	0.5389	0.5501	0.9601	0.9622	0.9600	0.5540	0.3367	0.3456
4	0.5826	0.5474	0.5600	0.9604	0.9624	0.9611	0.6019	0.4494	0.4910

Table 2

Results with balancing technique and attention deactivated.

Epochs	Polarity			Type			Town		
	Macro Prec.	Macro Recall	Macro F1	Macro Prec.	Macro Recall	F1	Macro Prec. .	Macro Recall	Macro F1
4	0.5601	0.5164	0.5348	0.9641	0.9619	0.9630	0.5541	0.3004	0.3462

Table 3

Results with Class Weighting technique (Polarity and Town) and the attention layer activated.

Epochs	Polarity			Type			Town		
	Macro Prec.	Macro Recall	Macro F1	Macro Prec.	Macro Recall	F1	Macro Prec.	Macro Recall	Macro F1
4	0.5868	0.5813	0.5798	0.9511	0.9550	0.9528	0.4051	0.4187	0.3856

6. Results

The results obtained on the test set, as shown in Table 4, were reported by the organizers of the Sentiment Analysis and Magical Towns Detection task at IberLEF 2025, following the evaluation of the predictions generated by the developed system. Our team, VerbaNexAI, achieved an overall Track Score of 0.4521, with distinct metrics for each subtask: a Macro F1 of 0.1818 for the Polarity task, 0.9043 for the Place Type task, and 0.4817 for the Magical Town task. This performance reflects the impact of the varying difficulty levels inherent to each task, with place type classification being the easiest and polarity and magical town classification significantly more complex. It is essential to highlight that we trained the system using a sample of 10,000 instances randomly selected from the total training set, which initially consisted of 208,051 records, due to limitations in computational resources, which restricted the ability to perform large-scale training using the entire dataset. Despite this limitation, the results in the place type task were highly satisfactory, with a performance close to 0.90, demonstrating that even with a reduced subset, the model effectively captured the textual features identifying whether a review refers to a hotel, restaurant, or attraction.

In contrast, the metrics obtained for the Polarity and Magical Town tasks reflect the challenges inherent to these domains. In the case of polarity, we attributed the low Macro F1 score of 0.1818 to the high subjectivity in sentiment interpretation and the class imbalance where most opinions are positive. Although the performance was better at 0.4817 for the Magical Town task, it still faces the difficulty of dealing with 40 unequally distributed classes, which requires more aggressive balancing strategies and a greater generalization capacity of the model. Overall, the results indicate that while the model performs well in tasks with well-defined and less imbalanced classes, such as the Type task, its performance could improve substantially if trained with the full dataset and more robust computational resources, especially in the more complex tasks.

Table 4
Results on the test set

Team	Track Score	Macro F1 (Polarity)	Macro F1 (Type)	Macro F1 (Town)
VerbaNexAI	0.4521	0.1818	0.9043	0.4817

7. Conclusion

This work presented a system based on multiclass and multitask BERT models to address sentiment analysis and the detection of Pueblos Mágicos in Spanish tourist reviews. We integrated various preprocessing techniques, data-balancing methods, and the activation of attention layers through a modular architecture, allowing the simultaneous handling of three key tasks: predicting sentiment polarity, classifying the type of place, and identifying the Pueblo Mágico. The design of a multitask pipeline proved effective in extracting deep, contextualized semantic representations of the text, even with a reduced sample of the original training set. The experiments carried out highlighted the importance of employing combined balancing strategies, such as oversampling with RandomOverSampler, undersampling with Condensed Nearest Neighbour, and synthetic data generation with SMOTE, especially in tasks with severe class imbalance like Pueblo Mágico classification. Likewise, we verified that the explicit activation of attention layers improves performance in Macro Precision and Macro F1 and provides interpretability to the model, allowing visualization of which parts of the text influence the classifier decisions.

The results obtained on the official IberLEF 2025 test set reflect both the potential and the challenges of the proposed approach. While the task of classifying the type of place achieved a remarkable performance (Macro F1 = 0.9043), the polarity (Macro F1 = 0.1818) and Pueblo Mágico (Macro F1 = 0.4817) tasks evidenced the inherent complexity of these categories, derived from factors such as language subjectivity, extreme class-distribution imbalance, and semantic diversity among tourist

regions. We could mitigate these limitations in future research by using the complete dataset and more intensive training with greater computational resources. Finally, this research demonstrates that using pre-trained language models, advanced text-processing techniques, and learning strategies can offer practical solutions for complex tasks within tourism. Future work plans to explore even more robust architectures, such as multilingual models or adaptations with more recent transformers. These advances would enable constructing more precise, equitable, and culturally sensitive recommendation and analysis systems for the tourism industry.

Acknowledgment

Declaration on Generative AI

We declare that the present manuscript has been written entirely by the authors and that no generative artificial intelligence tools were used in its preparation, drafting, or editing.

References

[1] M. Liu, Y. Shibuya, Y. Sekimoto, Emotions, behaviors and places: Mapping sentiments with behaviors in japanese tweets, Cities 155 (2024) 105449. URL: https://www.sciencedirect.com/science/article/pii/S0264275124006632. doi:https://doi.org/10.1016/j.cities.2024.105449.

[2] S. Cvetojevic, H. H. Hochmair, Modeling interurban mentioning relationships in the u.s. twitter network using geo-hashtags, Computers, Environment and Urban Systems 87 (2021) 101621. URL: https://www.sciencedirect.com/science/article/pii/S0198971521000284. doi:https://doi.org/10.1016/j.compenvurbsys.2021.101621.

[3] Á. Díaz-Pacheco, R. Guerrero-Rodríguez, M. Á. Álvarez-Carmona, A. Y. Rodríguez-GonzÁlez, R. Aranda, A comprehensive deep learning approach for topic discovering and sentiment analysis of textual information in tourism, Journal of King Saud University - Computer and Information Sciences 35 (2023) 101746. URL: http://dx.doi.org/10.1016/j.jksuci.2023.101746. doi:10.1016/j.jksuci.2023.101746.

[4] J. Á. González-Barba, L. Chiruzzo, S. M. Jiménez-Zafra, Overview of IberLEF 2025: Natural Language Processing Challenges for Spanish and other Iberian Languages, in: Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2025), co-located with the 41st Conference of the Spanish Society for Natural Language Processing (SEPLN 2025), CEUR-WS. org, 2025.

[5] M. Á. Álvarez-Carmona, Á. Díaz-Pacheco, R. Aranda, A. Y. Rodríguez-González, L. Bustio-Martínez, V. Herrera-Semenets, Overview of rest-mex at iberlef 2025: Researching sentiment evaluation in text for mexican magical towns, volume 75, 2025.

[6] M. Á. Álvarez-Carmona, Á. Díaz-Pacheco, R. Aranda, A. Y. Rodríguez-González, L. Bustio-Martínez, V. Muñis-Sánchez, A. P. Pastor-López, F. Sánchez-Vega, Overview of rest-mex at iberlef 2023: Research on sentiment analysis task for mexican tourist texts, 2023.

[7] M. Á. Álvarez-Carmona, Á. Díaz-Pacheco, R. Aranda, A. Y. Rodríguez-González, D. Fajardo-Delgado, R. Guerrero-Rodríguez, L. Bustio-Martínez, Overview of rest-mex at iberlef 2022: Recommendation system, sentiment analysis and covid semaphore prediction for mexican tourist texts, Procesamiento del Lenguaje Natural 69 (2022).

---

[2]https://github.com/VerbaNexAI

[8] M. Á. Álvarez-Carmona, R. Aranda, S. Arce-Cárdenas, D. Fajardo-Delgado, R. Guerrero-Rodríguez, A. P. López-Monroy, J. Martínez-Miranda, H. Pérez-Espinosa, A. Rodríguez-González, Overview of rest-mex at iberlef 2021: Recommendation system for text mexican tourism, Procesamiento del Lenguaje Natural 67 (2021). doi:https://doi.org/10.26342/2021-67-14.

[9] P. Mirabal, S. HernÃₐndez-Alvarado, J. I. Abreu Salas, Analyzing sentiment, attraction type, and country in spanish language tripadvisor reviews using language models, 2023-09-26.

[10] J. Cuadrado, E. Martinez, J. C. M. Santos, E. Puertas, Team utb-nlp at finances 2023: Financial targeted sentiment analysis using a phonestheme semantic approach, in: IberLEF@SEPLN, 2023. URL: https://ceur-ws.org/Vol-3496/finances-paper4.pdf.

[11] B. Sohrabi, I. Raeesi Vanani, N. Nasiri, A. Ghassemi Rudd, A predictive model of tourist destinations based on tourists' comments and interests using text analytics, Tourism Management Perspectives 35 (2020) 100710. URL: https://www.sciencedirect.com/science/article/pii/S2211973620300775. doi:https://doi.org/10.1016/j.tmp.2020.100710.

[12] J. R. Grande-RamÃₙrez, E. RoldÃₐn-Reyes, A. A. Aguilar-Lasserre, U. JuÃₐrez-MartÃₙnez, Integration of sentiment analysis of social media in the strategic planning process to generate the balanced scorecard, Applied Sciences 12 (2022) 12307. URL: https://doi.org/10.3390/app122312307. doi:10.3390/app122312307.

[13] R. Guerrero-Rodriguez, M. A. Ãₐlvarez Carmona, R. Aranda, A. P. LÃₛpez-Monroy, Studying online travel reviews related to tourist attractions using nlp methods: the case of guanajuato, mexico, Current Issues in Tourism 26 (2023) 289–304. URL: https://doi.org/10.1080/13683500.2021.2007227. doi:10.1080/13683500.2021.2007227. arXiv:https://doi.org/10.1080/13683500.2021.2007227.

[14] N. Saraswathi, T. Sasi Rooba, S. Chakaravarthi, Improving the accuracy of sentiment analysis using a linguistic rule-based feature selection method in tourism reviews, Measurement: Sensors 29 (2023) 100888. URL: https://www.sciencedirect.com/science/article/pii/S2665917423002246. doi:https://doi.org/10.1016/j.measen.2023.100888.

[15] M. Ojeda-HernÃₐndez, D. LÃₛpez-RodrÃₙguez, Ãₐngel Mora, Lexicon-based sentiment analysis in texts using formal concept analysis, International Journal of Approximate Reasoning 155 (2023) 104–112. URL: https://www.sciencedirect.com/science/article/pii/S0888613X23000130. doi:https://doi.org/10.1016/j.ijar.2023.02.001.