

DSVS at REST-MEX 2025: A Multitask Approach for Sentiment, Place Type, and Town Classification in Spanish Reviews

David Vázquez-Santana^{1,*}, Sergio Damián-Sandoval¹, Cornelio Yáñez-Márquez¹ and Edgardo Felipe-Riverón¹

¹Centro de Investigación en Computación (CIC), Instituto Politécnico Nacional (IPN), México

Abstract

This paper presents a multitask learning approach for analyzing user-generated reviews from Mexican *Pueblos Mágicos* (Towns), as part of the REST-MEX 2025 shared task. We address three interrelated classification subtasks: (1) sentiment polarity on a 1-to-5 scale, (2) categorization of place type into *Restaurant*, *Attractive*, or *Hotel*, and (3) identification of the town being referenced among 40 possible classes. To tackle these tasks, we fine-tuned the pretrained *pysentimiento/robertuito*-base-cased model, originally trained on social media content in Spanish. Our training pipeline incorporates curriculum learning by presenting data in stages based on class frequency, and uses Automatic Weighted Loss to balance task priorities dynamically throughout training. Despite achieving promising results in sentiment and place type classification, the model exhibited poor performance in the town classification task. We analyze potential causes for this failure and outline directions for further experimentation to improve this subtask. Our findings highlight both the opportunities and limitations of multitask learning when applied to real-world, imbalanced, user-generated data in Spanish.

Keywords

Multitask Learning, Sentiment Analysis, Curriculum Learning, Automatic Weighted Loss, Spanish Language Processing, Roberta-based Models

1. Introduction

Reviews have always played a crucial role in the development of various economic activities, as they provide valuable information for potential customers such as geographic identification, spatial context, landmarks, and details about the activities or products offered, including their costs. In the digital era, sharing reviews has become even more relevant. Users can easily publish their experiences through social media platforms, official websites, and travel forums. In this way, reviews influence perception, decision-making, and design of commercial strategies [1, 2, 3].

In particular, the tourism sector has been significantly transformed by the widespread availability of user-generated content on travel platforms such as TripAdvisor, Booking, and Google Maps. These reviews not only guide prospective travelers but also serve as a valuable data source for understanding behavioral patterns, satisfaction levels, and areas of improvement in tourism services [4, 5, 6].

In Mexico, tourism represents one of the most relevant economic activities [7], both due to its contribution to the gross domestic product (GDP) [8] and its social and cultural impact [9]. Within this context, the federal program *Pueblos Mágicos* was created by the Mexican Secretariat of Tourism in 2001 [10], with the primary objective of promoting the integral development of communities with exceptional historical, cultural, and natural attributes. Currently, these so-called Magical Towns receive thousands of national and international visitors each year [7]. Analyzing tourist perception is essential to evaluate the success of the program, identify areas of improvement, and promote sustainable, high-quality tourism.

IberLEF 2025, September 2025, Zaragoza, Spain

✉ dvazquez2019@cic.ipn.mx (D. Vázquez-Santana); sdamians2019@cic.ipn.mx (S. Damián-Sandoval); cyanez@cic.ipn.mx (C. Yáñez-Márquez); edgardo@cic.ipn.mx (E. Felipe-Riverón)

🌐 <https://github.com/Hiram02> (D. Vázquez-Santana); <https://github.com/sdamians> (S. Damián-Sandoval)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



Figure 1: Representative images of three *Pueblos Mágicos*: Cholula, San Miguel de Allende and Bernal.

1.1. Task Description

In this work, we address a natural language processing (NLP) shared task focused on the automatic analysis of reviews about tourist destinations in Mexico. The main goal is to classify reviews related to the *Pueblos Mágicos* across three subtasks:

1. **Sentiment polarity:** This subtask involves evaluating the overall opinion expressed in each review, assigning a score from 1 (very negative) to 5 (very positive).
2. **Type of tourist site:** The review may refer to a hotel, a restaurant, or a tourist attraction. The goal is to classify the review into one of these three categories.
3. **Corresponding /textitPueblo Mágico:** This subtask consists of identifying which of the 60 possible *Pueblos Mágicos* the review refers to.

1.2. Motivation

Automatically classifying this type of information brings value both to academia and to the tourism sector. The widespread use of social media and travel platforms generates large volumes of information that can be mined to extract insights such as recurring topics, visitors’ sentiments, emerging trends, and general criticisms or service issues. This is especially relevant in regions where tourism-related activities represent a significant portion of the local economy, as is the case for the *Pueblos Mágicos*. In such contexts, reviews can help monitor service quality and user satisfaction.

From a scientific perspective, the proposed task contributes to advancements in text analysis and NLP[11, 12, 13]. In particular, the REST-MEX 2025 shared task [14] provides a framework for exploring how these methods can be applied to tourist reviews. This task is part of IberLEF 2025 [15], a congress that promotes research in natural language processing through shared tasks.

Furthermore, the tourism domain introduces specific linguistic and semantic challenges, such as subjective expressions, metaphors, and emotionally charged descriptions, making this a difficult yet valuable problem for the development of robust and adaptable models.

1.3. Relevance and Challenges of the Task

This task combines multiple subproblems in NLP and machine learning. Sentiment analysis on an ordinal scale requires models capable of capturing subtle variations in emotional expressions, which is more complex than binary classification. Additionally, identifying the type of tourist site demands contextual understanding, as users do not always explicitly state whether they are referring to a

hotel, restaurant, or some other attraction. This semantic ambiguity requires the use of sophisticated classification techniques and often benefits from multi-task learning approaches [16].

Finally, assigning the correct Magical Town may seem trivial when geographical metadata is clearly defined. However, in many cases, this information is ambiguous, incomplete, or even contradictory. Solving this subproblem involves geographic disambiguation and inference based on the review's textual context.

Overall, this task represents a comprehensive case study that combines ordinal classification, multi-class categorization, and inference over informal, subjective user-generated text. It highlights the need for intelligent systems capable of understanding and extracting structured insights from unstructured data in real-world tourism contexts. The source code for this work can be found at https://github.com/sdamians/DSVS_REST-MEX2025.

2. Related Work

In recent years, the automatic analysis of tourist reviews has gained interest in the fields of natural language processing (NLP) and tourism research. Tasks such as sentiment polarity, classification of the type of tourist site, and identification of the geographic location have been explored to extract useful information from user-generated content.

2.1. Sentiment Polarity in Tourism

Sentiment polarity aims to detect the emotions or opinions expressed in text. This task has been widely applied to tourism, especially in reviews of hotels and restaurants. One of the first works in this area was done by Hu and Liu [17], who developed methods to find positive and negative opinions in product reviews. Later, Ye et al. [18] applied machine learning to classify sentiment in hotel reviews, showing that this type of analysis can help improve tourism services.

More recently, researchers have used deep learning models such as LSTM and BERT to improve the accuracy of sentiment analysis in tourism texts written in different languages [19, 20]. Some studies also use ordinal scales (for example, from 1 to 5 stars) instead of just positive or negative labels [21]. Classifying sentiments on this type of scale is more complex because it requires detecting small differences in tone and emotions.

2.2. Classification of Tourist Site Type

Another important task is to classify the type of place mentioned in each review. For example, whether the text refers to a hotel, a restaurant, a museum, an archaeological site, a beach, or any other attraction. This task requires understanding the context because users usually do not directly mention the name of the place or category.

Some approaches use supervised learning models trained on labeled data to solve this problem [22]. Others combine keyword-based methods with deep learning to improve the accuracy in classifying different types of tourist sites [23]. Multi-task learning models have also been used successfully to predict sentiment and places at the same time, helping the model learn shared patterns from both tasks [24].

2.3. Geographic Identification and Disambiguation

Identifying the geographic location described in a review is also a difficult task, especially when users do not mention the name of the town or other clear clues. Geographic disambiguation techniques are often used to address this problem. These techniques rely on lists of place names and contextual clues in the text [25].

In tourism, inferring geographic information from reviews can be useful for applications such as planning travel routes [26, 27]. However, there is little research focused on this task in the context

of Spanish-language reviews, particularly those related to Mexican destinations such as the *Pueblos Mágicos*.

2.4. Learning Strategies for the Shared Task

The integration of multiple tasks into a single model, as done in this work, requires learning strategies that can improve convergence and performance. One widely used approach is *Multi-Task Learning* (MTL), where a single model is trained to solve several tasks simultaneously. This aims to reach better generalization through knowledge sharing among related tasks [28, 29].

Another useful strategy is *Curriculum Learning*, which involves organizing the training data in a meaningful order, for example, by starting with easier examples and gradually introducing more difficult ones, which attempts to mimic human learning and can help NLP models converge faster [30].

Finally, we incorporate an *Automatic Weighted Loss* scheme in order to let the model dynamically adjust its weight loss function. This method allows the model to focus more on tasks that are harder or with less certainty, improving the balance across multiple objectives [31].

In summary, although many studies have analyzed sentiment polarity, classified types of location, or identified locations individually, few have tackled all three aspects at once. This work contributes by addressing these tasks together, using real reviews on tourist destinations in Mexico. This approach helps us explore new challenges and develop solutions that take into account the specific language and cultural context.

3. Dataset and Preprocessing

We used a dataset containing **207,873** Spanish-language reviews related to tourist experiences in *Pueblos Mágicos* (Magical Towns) across Mexico. This dataset was provided as part of the shared task and was processed by removing duplicate entries to ensure data quality and better balance. For all three subtasks, the class labels were numerically encoded to facilitate model training. Sentiment scores are represented as integers from 0 to 4 (in this case, we ignored the ordinal meaning of these classes), place types were mapped to values {0: Restaurant, 1: Attractive, 2: Hotel}, and towns were encoded as integer values from 0 to 39.

3.1. Text Preprocessing

All reviews were subjected to standard text normalization, including:

- Unicode normalization to standardize accented characters and the letter ñ.
- Removal of special characters, such as punctuation and non-alphabetic symbols.

3.2. Curriculum-Based Dataset Partitioning

To support a curriculum learning strategy, the dataset was divided into five curriculum-based subsets, using the task with the largest number of classes (Town) as the guiding reference. These subsets were constructed based on the distribution of reviews per town, ensuring that each subset remained representative and excluded towns whose sample counts deviated by more than three standard deviations from the global distribution.

This curriculum-based partitioning did not significantly affect the class distribution for the other two subtasks (sentiment and place type), allowing consistent multitask training across all stages. The original distribution for the training dataset for each subtask is shown in Figure 2.

4. Methodology

We based our architecture on the pre-trained language model `pysentimiento/robertuito` basecased [32], a Spanish RoBERTa variant fine-tuned in social media data. We hypothesized that

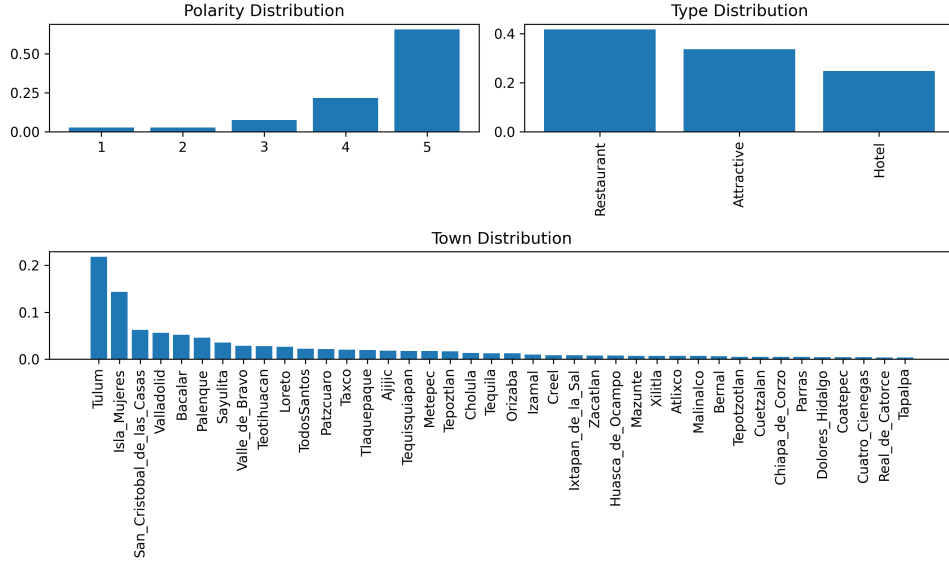


Figure 2: Distribution of instances across polarity scores, place types, and towns.

this model would be particularly well suited to the characteristics of the task, as the reviews in our dataset are user-generated and often contain informal language. Additionally, the model is capable of recognizing offensive or colloquial expressions, which may be informative for the sentiment analysis subtask.

To handle the multitask setup, we implemented a shared encoder with three task-specific heads corresponding to sentiment polarity classification, place type classification, and town classification. Training was guided by the Automatic Weighted Loss strategy, which dynamically adjusts the contribution of each task’s loss during training, allowing the model to prioritize tasks based on their difficulty and learning dynamics.

The training followed a curriculum learning approach. The dataset was divided into five subsets, as described in Table 1, and these were introduced sequentially every three epochs. The curriculum began with subsets containing reviews from the most frequently mentioned towns and progresses toward those with fewer examples, encouraging the model to generalize from abundant to scarce data distributions.

Table 1

F1-score per task and submission. The overall score is the unweighted average of the three subtasks.

Subdataset	Number of instances	Number of towns
subdataset 1	75165	2
subdataset 2	97705	15
subdataset 3	25833	13
subdataset 4	6897	7
subdataset 5	2273	3

Because the encoder model accepts a maximum input length of 128 tokens, reviews that exceed this limit were truncated. This ensures uniform input sizes and reduces training complexity.

Each curriculum subset was split into 90% for training and 10% for validation, maintaining the label distribution across subtasks. Furthermore, we apply task-specific class weights to compensate for class imbalances, especially in the town classification task where some towns are underrepresented. These weights are updated dynamically depending on the class distribution present in each training subset. Figure 3 presents the training process used for the shared task.

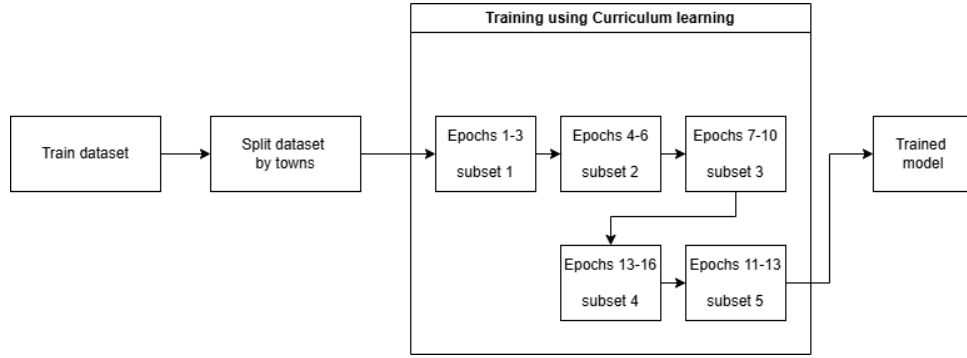


Figure 3: Training process flow. We decided to use curriculum learning so the model learn town information progressively

5. Results

During the training process, we observed that the model consistently struggled with the **town classification** task. Validation scores for this subtask decreased across epochs, suggesting that the model was either overfitting to the training samples or failing to capture distinguishing patterns among the 40 town classes.

We submitted three runs for evaluation on the official test set. Each subsequent submission involved training the model for additional epochs in an attempt to improve its generalization. However, despite this iterative process, the model failed to learn the town classification task, while successfully learning the sentiment analysis and place type classification tasks.

Table 2 summarizes the F1-scores obtained for each subtask and submission, including an overall track score for each submission. As shown in Table 3, although our system did not reach the top 3 in overall performance, it successfully outperformed the baseline, earning an Honorable Mention. This result highlights areas of opportunity for further improvement. One possible direction for future work is to conduct additional experiments that individually address each subtask. This would allow us to better understand the strengths and weaknesses of our approach and determine whether a task-specific strategy could lead to improved results compared to the baseline.

Table 2

F1-score per task and submission. The overall score is the unweighted average of the three subtasks.

Submission	Macro-F1 Polarity	Macro-F1 Type	Macro-F1 Town	Track Score
Run 1	0.4125	0.9317	0.0447	0.3151
Run 2	0.4210	0.9296	0.0408	0.3157
Run 3	0.4170	0.9261	0.0393	0.3130

Table 3

F1-scores of the top three participating systems versus our approach. 'HM' (Honorable Mention) indicates systems that surpassed the baseline performance.

Place	Run	Macro-F1 Polarity	Macro-F1 Type	Macro-F1 Town	Track Score
1st	UDENAR_1	0.6444	0.9877	0.6919	0.7254
2nd	Axolotux_E_T3	0.6395	0.9876	0.6895	0.7225
3rd	Pandas_Rojos_1	0.6864	0.9818	0.6347	0.6464
HM	DSVS	0.4210	0.9296	0.0408	0.3157
	Baseline	0.1583	0.1967	0.0089	0.0900

5.1. Analysis of the Town Classification Task

The poor performance on the town classification subtask may be attributed to several factors:

- **High class imbalance:** Certain towns were significantly underrepresented in the dataset, even after applying class weighting. This imbalance may have hindered the model’s ability to generalize.
- **Semantic overlap:** Reviews for different towns may contain highly similar vocabulary, especially when describing common tourist experiences (e.g., beaches, hotels, local food), which may confuse the model.
- **Suboptimal task balancing:** The failure to effectively adjust the automatic loss weighting or to design an appropriate curriculum learning schedule may have limited the model’s capacity to balance learning across tasks. Alternatively, training separate models per task could have improved performance.
- **Lack of town-specific context:** The model may require auxiliary information (e.g., town descriptions, metadata, or structured features) to effectively distinguish between classes with subtle textual cues.
- **Interference from multitask learning:** Training all subtasks simultaneously may have caused conflicting learning signals, particularly if the objectives of different tasks are not fully aligned. This interference could have negatively impacted the model’s ability to focus on the town classification subtask.

In future iterations, this subtask could benefit from techniques such as town-specific prompts, hierarchical classification, external knowledge integration (e.g., town profiles), or contrastive learning methods to improve class separation in the representation space.

6. Conclusions

In this work, we explored a multitask learning approach for analyzing user-generated reviews from Mexican *Pueblos Mágicos*, addressing three interconnected tasks: sentiment polarity, place type classification, and town identification. By leveraging the pretrained language model *psentimiento/robertuito*-base-cased, along with strategies such as curriculum learning and Automatic Weighted Loss, we aimed to improve generalization across diverse and imbalanced classes.

Our results demonstrate that the model is capable of effectively learning the sentiment and place type tasks, even under a multitask setup. However, the town classification task remained a major challenge, with the model failing to meaningfully distinguish among the 40 town classes. This outcome highlights the difficulty of this subtask, which is likely influenced by vocabulary overlap, class imbalance, and loss of contextual information due to input truncation.

Key insights from this study include:

- Multitask learning can be successfully applied to related classification tasks, especially when there is shared semantic structure.
- Curriculum learning may help when dealing with highly imbalanced datasets, by progressively introducing more difficult or sparse classes.
- Automatic Weighted Loss is a valuable technique for dynamically adjusting learning focus, though it may require further tuning for tasks with high asymmetry in difficulty.
- Complex subtasks like town classification may require additional experimentation and targeted strategies to be effectively solved.

As future work, we plan to continue experimenting with alternative training strategies and task-specific enhancements to improve performance in the town classification subtask, while preserving the gains achieved in the other tasks.

Acknowledgments

This work was done with partial support from the Mexican Government through Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) and Instituto Politécnico Nacional (IPN).

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] J. A. Chevalier, D. Mayzlin, The effect of word of mouth on sales: Online book reviews, *Journal of Marketing Research* 43 (2006) 345–354.
- [2] M. Hu, B. Liu, Opinion mining and summarization of hotel reviews, *Proceedings of the ACM International Conference on Information and Knowledge Management (CIKM)* (2008).
- [3] E. Olmos-Martínez, M. Á. Álvarez-Carmona, R. Aranda, A. Díaz-Pacheco, What does the media tell us about a destination? the cancan case, seen from the usa, canada, and mexico, *International Journal of Tourism Cities* 10 (2023) 639–661. URL: <http://dx.doi.org/10.1108/IJTC-09-2022-0223>. doi:10.1108/ijtc-09-2022-0223.
- [4] R. Filieri, F. McLeay, E-wom and accommodation: An analysis of the factors that influence travelers' adoption of information from online reviews, *Journal of travel research* 53 (2014) 44–57.
- [5] R. Guerrero-Rodríguez, M. A. Álvarez Carmona, R. Aranda, A. P. López-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>.
- [6] 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.
- [7] S. de Turismo de México, Compendio estadístico del turismo en México 2023, <https://datatur.sectur.gob.mx/SitePages/CompendioEstadistico.aspx>, 2023. Disponible en: <https://datatur.sectur.gob.mx/SitePages/CompendioEstadistico.aspx>.
- [8] INEGI, Cuenta satélite del turismo de México 2023, <https://datatur.sectur.gob.mx/SitePages/ProductoDestacado3.aspx>, 2024.
- [9] F. G. Reverté, S. M. Pérez, El impacto cultural y social de los eventos celebrados en destinos turísticos. la percepción desde el punto de vista de los organizadores, *Cuadernos de turismo* (2017) 339–362.
- [10] S. de Turismo, Programa pueblos mágicos, <https://www.gob.mx/sectur/acciones-y-programas/programa-pueblos-magicos>, 2016.
- [11] 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>.
- [12] 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).
- [13] 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, *Procesamiento del Lenguaje Natural* 71 (2023).
- [14] 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.
 - [15] 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.
 - [16] S. Ruder, An overview of multi-task learning in deep neural networks, *arXiv preprint arXiv:1706.05098* (2017).
 - [17] M. Hu, B. Liu, Mining and summarizing customer reviews, in: *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, 2004, pp. 168–177.
 - [18] Q. Ye, Z. Zhang, R. Law, Sentiment classification of online reviews to travel destinations by supervised machine learning approaches, *Expert systems with applications* 36 (2009) 6527–6535.
 - [19] M. Fu, L. Pan, Sentiment analysis of tourist scenic spots internet comments based on lstm, *Mathematical Problems in Engineering* 2022 (2022) 5944954.
 - [20] V. Arefieva, R. Egger, Tourbert: a pretrained language model for the tourism industry. *arxiv prepr, arXiv preprint arXiv:2201.07449* (2022).
 - [21] L. Dong, F. Wei, S. Liu, M. Zhou, K. Xu, A statistical parsing framework for sentiment classification, *Computational Linguistics* 41 (2015) 293–336.
 - [22] Q. Cheng, W. Shi, Hierarchical multi-label text classification of tourism resources using a label-aware dual graph attention network, *Information Processing & Management* 62 (2025) 103952.
 - [23] Q. Li, S. Li, J. Hu, S. Zhang, J. Hu, Tourism review sentiment classification using a bidirectional recurrent neural network with an attention mechanism and topic-enriched word vectors, *Sustainability* 10 (2018) 3313.
 - [24] P. Liu, X. Qiu, X. Huang, Recurrent neural network for text classification with multi-task learning, *arXiv preprint arXiv:1605.05101* (2016).
 - [25] Y. L. Zaila, D. Montesi, Geographic information extraction, disambiguation and ranking techniques, in: *Proceedings of the 9th Workshop on Geographic Information Retrieval*, 2015, pp. 1–7.
 - [26] B. Han, P. Cook, T. Baldwin, Text-based twitter user geolocation prediction, *Journal of Artificial Intelligence Research* 49 (2014) 451–500.
 - [27] S. Roller, M. Speriosu, S. Rallapalli, B. Wing, J. Baldrige, Supervised text-based geolocation using language models on an adaptive grid, in: *Proceedings of the 2012 joint conference on empirical methods in natural language processing and computational natural language learning*, 2012, pp. 1500–1510.
 - [28] R. Caruana, Multitask learning, *Machine learning* 28 (1997) 41–75.
 - [29] R. Collobert, J. Weston, A unified architecture for natural language processing: Deep neural networks with multitask learning, in: *Proceedings of the 25th international conference on Machine learning*, 2008, pp. 160–167.
 - [30] Y. Bengio, J. Louradour, R. Collobert, J. Weston, Curriculum learning, in: *Proceedings of the 26th annual international conference on machine learning*, 2009, pp. 41–48.
 - [31] A. Kendall, Y. Gal, R. Cipolla, Multi-task learning using uncertainty to weigh losses for scene geometry and semantics, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 7482–7491.
 - [32] J. M. Pérez, D. A. Furman, L. Alonso Alemany, F. M. Luque, RoBERTuito: a pre-trained language model for social media text in Spanish, in: *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, European Language Resources Association, Marseille, France, 2022, pp. 7235–7243. URL: <https://aclanthology.org/2022.lrec-1.785>.