LynxesAI-LKEBUAP at REST-MEX 2026: Sentiment Analysis with Synthetic Tourism Reviews using Transformer-Based Conditional Generation and Filtering

Victor G. Morales-Murillo[1,2,*,†], Pierre Delice[3,†], David Pinto[4,†] and Karina Cancino-Villatoro[2,†]

[1]Universidad de las Américas Puebla (UDLAP), Puebla, México
[2]Universidad Politécnica de Tapachula (UPTAP), Chiapas, México
[3]Universidad de Guadalajara (UdeG), Guadalajara, México
[4]Dirección de Innovación y Transferencia de Conocimiento (DITCO), Benemérita Universidad Autónoma de Puebla (BUAP), Puebla, México

Abstract

This paper presents the participation of the LynxesAI-LKEBUAP team in the REST-MEX 2026. The objective is to generate a synthetic corpus of 3,000 Spanish tourism reviews that can be effectively used to train sentiment classification models. Our approach is based on instruction-based fine-tuning of the multilingual text-to-text transformer mT5 using prompt-response pairs derived from the original tourism reviews. We investigate three conditional generation strategies that differ in their decoding settings and the incorporation of a sentiment-based filtering stage. The proposed filtering strategy employs a domain-adapted RoBERTa classifier to select the generated review that best matches the target sentiment polarity from multiple candidates. Experimental results show that this strategy achieved the best performance among the proposed configurations in both the internal and official evaluations. Our best submission obtained a Track Score of 0.5085, a Macro-F1 score of 0.5515, and ranked third in the Synthetic Sentiment Modeling task of REST-MEX 2026.

Keywords

Sentiment Analysis, Synthetic Data Generation, Data Augmentation, Data-centric AI

1. Introduction

Sentiment analysis, also known as opinion mining, is a core subfield of natural language processing (NLP) within artificial intelligence (AI) due to the massive amount of data generated in the digital age, including a large volume of online comments and reviews about products and services [1]. Billions of people express their opinions on social networks, streaming platforms, e-commerce websites, and other online services. Furthermore, more than 6 billion people use the Internet, two-thirds of the global population engage with social media each month, and over 1 billion individuals use generative artificial intelligence (GenAI) [2]. These trends represent an important opportunity to automatically monitor public opinion and support decision-making in several industries [3].

The tourism industry is one of the most important economic sectors worldwide, contributing substantially to gross domestic product (GDP) and creating millions of direct jobs. In countries such as Mexico, tourism represents a key driver of economic growth and employment generation. Most global travelers utilize online platforms for trip planning, where their final decisions are usually influenced by the shared experiences of other travelers expressed in online reviews. In this sense, companies and governments can develop new tourism strategies by analyzing opinions of tourists [4].

IberLEF 2026, September 2026, León, Spain
*Corresponding author.
†These authors contributed equally.
vg055@hotmail.com (V. G. Morales-Murillo); padelice@gmail.com (P. Delice); david.pinto@correo.buap.mx (D. Pinto); investigacion.posgrado@uptapachula.edu.mx (K. Cancino-Villatoro)
https://github.com/Victor055/ (V. G. Morales-Murillo)
0000-0002-6786-9232 (V. G. Morales-Murillo); 0000-0002-4170-4405 (P. Delice); 0000-0002-8516-5925 (D. Pinto); 0000-0002-7467-9082 (K. Cancino-Villatoro)

Despite significant advances in Large Language Models (LLMs) trained on massive datasets, the success of NLP models remains dependent on a large volume of high-quality data [5]. However, obtaining sufficient high-quality training data continues to be a complex issue across multiple domains. Furthermore, the availability of large supervised datasets has been limited by data scarcity, high annotation costs, and privacy constraints [6]. Synthetic data generation has emerged as an alternative to tackle these challenges, where additional training instances are artificially produced within the emerging field of data-centric AI [7].

This paper describes the participation of LynxesAI-LKEBUAP team in the sentiment analysis task at Rest-Mex 2026 [8], organized within the Iberian Languages Evaluation Forum (IberLEF 2026) [9]. Unlike previous editions that focused on direct review classification, this edition focuses on the generation of synthetic tourist opinions and the evaluation of their effectiveness for downstream sentiment analysis.

Our goal is to perform instruction-based fine-tuning of the multilingual text-to-text transformer mT5 using prompt-response pairs derived from tourism reviews. We then design three experiments to generate distinct outputs for Rest-Mex 2026. The first experiment aims to balance coherence and diversity, whereas the second focuses on increasing diversity, as the shared task penalized reviews that were highly similar to those in the training set. In the final experiment, we employ a domain-adapted RoBERTa classifier from Rest-Mex 2023 to filter the generated reviews according to the target polarity.

The rest of this paper is structured as follows: Section 2 presents the related work. Section 3 describes the Rest-Mex 2026 shared task. Section 4 presents our data analysis and model fine-tuning procedures. Section 5 reports the experiments and results. Finally, Section 6 presents the conclusions and outlines future work.

2. Related Work

In recent years, studies [10] [6] [5] have demonstrated that data augmentation can improve sentiment and emotion analysis in low-resource Spanish datasets. Techniques such as text transformations, back-translation, and generative adversarial networks have enhanced the performance of deep learning and transformer-based models. These results motivate the use of synthetic data generation strategies in sentiment analysis.

Previous editions of Rest-Mex have progressively expanded the scope of sentiment analysis tasks in the tourism domain. Rest-Mex 2022 [11] focused on the joint classification of sentiment polarity and the type of tourist attraction from 43,150 TripAdvisor reviews collected between 2002 and 2021. In Rest-Mex 2023 [12], the task was extended to include country identification, covering reviews from Mexico, Cuba, and Colombia and increasing the corpus size to 359,565 opinions. More recently, Rest-Mex 2025 [13] replaced country classification with the identification of one of the 40 most popular Mexican magical towns, requiring systems to jointly predict sentiment polarity, attraction type, and the specific destination from a corpus of 297,217 tourist reviews.

3. Task Description

Unlike others editions, REST-MEX 2026 introduces a fully generative sentiment analysis task. Participants must develop a model capable of generating a synthetic corpus of 3,000 original tourist reviews about Mexican magical towns. Each generated review must be entirely original, avoiding any repetition or paraphrasing of the provided real reviews, while resembling authentic tourist narratives in style and content. Additionally, each review must be assigned a sentiment polarity label:

- `Very Positive`: 5
- `Positive`: 4
- `Neutral`: 3
- `Negative`: 2
- `Very negative`: 1

The generated reviews are then used as the only training data for a sentiment classifier, which is evaluated on a hidden set of real tourist opinions using a weighted macro F1 score that emphasizes minority sentiment classes. Therefore, the central challenge is to generate synthetic reviews that enable a classifier trained solely on artificial data to generalize effectively to real-world tourist opinions.

- `Task Definition:` "Given a limited collection of real tourist reviews written in Spanish, each participating team must develop a generative model capable of producing a synthetic corpus of 3,000 original tourist opinions related to Mexican magical towns".

The official evaluation metric is the Polarity Score (RP), which was specifically designed to assess sentiment classification performance under severe class imbalance, a common characteristic of real-world tourist review datasets. In tourism-related corpora, the majority of reviews are highly positive, whereas negative and very negative opinions occur much less frequently. Consequently, conventional evaluation measures may produce overly optimistic results by being dominated by the majority classes and failing to adequately reflect the model's ability to identify minority sentiments. To address this issue, the RP metric assigns greater importance to underrepresented classes, providing a more balanced and fair evaluation of sentiment classification systems across all polarity levels.

$$RP(k) = \frac{\sum i \epsilon C((1 - \frac{TC_i}{TC})F_i(k)}{\sum i \epsilon C(1 - \frac{TC_i}{TC})}, C = \{1, 2, 3, 4, 5\}. \quad (1)$$

- `C:` The set of sentiment polarity classes.
- `Fi(k):` The F1-score obtained by system k for polarity class i.
- `TCi:` The total number of real reviews belonging to class iii in the evaluation corpus.
- `TC:` The total number of reviews across all polarity classes in the real corpus.

4. Methodology

This section presents the proposed methodology, which consists of four main stages: data analysis, data preprocessing, instruction-based fine-tuning, and conditional synthetic review generation.

4.1. Data Analysis

The organizers of REST-MEX 2026 provide a dataset containing 208,051 opinions, which constitutes the training set for participating teams, while the remaining 30% of the data is reserved for the test set. The dataset consists of the following columns:

- `Title:` The title that the tourist assigned to their opinion.
- `Review:` The opinion issued by the tourist.
- `Polarity:` The label representing the sentiment polarity of the opinion.
- `Town:` The town where the review is focused.
- `Region:` The region (state in Mexico) where the town is located.
- `Type:` The type of place the review refers to.

We perform a data analysis to examine the distribution of the dataset across several key features, including polarity, type, and magical town. As shown in Table 1, the polarity distribution reveals a highly imbalanced dataset, where positive polarities constitute the majority class, while negative polarities are underrepresented. This severe class imbalance poses a significant challenge for generating synthetic reviews with negative polarities.

As indicated in Table 2, the type distribution exhibits a lower degree of imbalance than the polarity distribution, with restaurant reviews representing the majority class, followed by reviews of tourist attractions and hotels.

Table 1
Data Distribution by Polarity

Polarity	Instances
5	136561
4	45034
3	15519
2	5496
1	5441

Table 2
Data Distribution by Type

Type	Instances
Restaurant	86720
Attractive	69921
Hotel	51410

As illustrated in Table 3, the distribution of reviews across magical towns is highly imbalanced, as the dataset contains forty towns with substantially different numbers of reviews. This imbalance poses a significant challenge for synthetic text generation, since the model may tend to generate reviews primarily for the most represented towns. Such behavior increases the risk of producing repetitive or overly similar reviews to those in the original corpus, which is penalized by the evaluation metric.

4.2. Instruction-Based Fine-Tuning of the Text-to-Text Transformer

We employ instruction-based fine-tuning to adapt a multilingual text-to-text Transformer for the generation of synthetic tourism reviews. Unlike conventional language model fine-tuning, the proposed approach represents each training instance as an instruction-response pair, enabling the model to learn conditional text generation from structured prompts. The prompts specify the desired review attributes, including sentiment polarity, review type, region, and magical town, while the target output corresponds to the original human-written review. During the preprocessing stage, each review was reformulated as an instruction-response pair. An example is presented below.

- `Polarity` : 5
- `Type` : Restaurant
- `Region` : Nayarit
- `Town` : Sayulita
- `Input Text (Promt)` : Escribe una reseña turística con una opinión {Polarity} sobre un {Type} ubicado en {Town}, {Region}. La reseña debe describir la experiencia del visitante de forma natural y realista.
- `Target Text (Review)` : Excelente lugar para comer y pasar una buena noche!!! El servicio es de primera y la comida exquisita!!!

Our approach is based on mT5 [14], a multilingual encoder-decoder Transformer that extends the Text-to-Text Transfer Transformer (T5) architecture to 101 languages. While T5 demonstrated the effectiveness of the unified text-to-text framework across a wide range of English NLP tasks, mT5 extends this capability through pre-training on the multilingual Common Crawl (mC4) corpus. Unlike encoder-only models, mT5 formulates every task within a unified text-to-text framework, representing both the input and output as text sequences. Since mT5 is pre-trained using a self-supervised objective rather than task-specific supervision, it requires fine-tuning to adjust its knowledge to downstream tasks. The multilingual encoder-decoder architecture and unified text-to-text formulation make mT5 particularly suitable for the proposed instruction-based fine-tuning strategy.

Table 3
Data Distribution by Magical Town

Magical Town	Instances
Tulum	45345
Isla_Mujeres	29826
San_Cristobal_de_las_Casas	13060
Valladolid	11637
Bacalar	10822
Palenque	9512
Sayulita	7337
Valle_de_Bravo	5959
Teotihuacan	5810
Loreto	5525
TodosSantos	4600
Patzcuaro	4454
Taxco	4201
Tlaquepaque	4041
Ajijic	3752
Tequisquiapan	3627
Metepec	3532
Tepoztlan	3445
Cholula	2790
Tequila	2650
Orizaba	2521
Izamal	2041
Creel	1786
Ixtapan_de_la_Sal	1696
Zacatlan	1602
Huasca_de_Ocampo	1509
Mazunte	1466
Xilitla	1458
Atlixco	1444
Malinalco	1429
Bernal	1252
Tepotzotlan	1013
Cuetzalan	996
Chiapa_de_Corzo	960
Parras	953
Dolores_Hidalgo	909
Coatepec	818
Cuatro_Cienegas	788
Real_de_Catorce	760
Tapalpa	725

After constructing the instruction-response pairs, both the input instructions and the target reviews were tokenized using the `google/mt5-base` tokenizer. The input instructions were tokenized with a maximum sequence length of 128 tokens, whereas the target reviews were tokenized with a maximum length of 256 tokens. Sequences exceeding these limits were truncated, while shorter sequences were padded to the corresponding maximum length. To ensure that padding tokens did not contribute to the training objective, padding token identifiers in the target sequences were replaced with 100, allowing the loss function to ignore them during optimization.

The dataset was divided into training and validation subsets using a stratified 90/10 split based on the polarity label to preserve the original sentiment distribution across both subsets. The mT5 model was fine-tuned using the Hugging Face `Seq2SeqTrainer` framework on Google Colab Pro+ with an NVIDIA A100 GPU. Training was conducted for three epochs using a learning rate of 5e-5, a batch

size of 8, weight decay of 0.01, and gradient accumulation over four steps. Model checkpoints were evaluated at the end of each epoch, and the checkpoint with the lowest validation loss was selected for synthetic review generation. Considering the computational cost of large-scale generation-based evaluation and the objective of REST-MEX 2026, model selection was based on training and validation loss rather than automatic generation metrics such as BLEU or ROUGE. Since the goal of the shared task is to generate a synthetic corpus that supports downstream sentiment classification, the final assessment of the generated reviews relied on a sentiment classification model.

4.3. Conditional Synthetic Review Generation

We designed three generation strategies based on the fine-tuned mT5 model. The first strategy employs conservative generation settings to prioritize coherence and fluency. The second aims to promote greater lexical diversity and originality in the generated reviews, as REST-MEX 2026 explicitly requires fully synthetic reviews and penalizes copying, paraphrasing, or reusing real opinions. Furthermore, the third strategy combines conditional generation with sentiment-based filtering, where three candidate reviews are generated for each condition and a sentiment classifier selects the most suitable output. These strategies were designed to evaluate different generation settings with respect to review quality, diversity, and sentiment consistency.

5. Experiments and Results

This section presents the experimental setup and the results obtained by the proposed generation strategies. All experiments were performed using the same set of generation conditions. Specifically, the original training corpus was stratified according to sentiment polarity, review type, and magical town, and approximately 1.46% of the instances from each group were randomly sampled. This procedure preserved the original data distribution while producing approximately 3,000 representative conditions for synthetic review generation.

The first generation strategy employed conservative decoding parameters to prioritize coherence and fluency in the generated reviews. A relatively low temperature (0.6), together with (top_p=0.85) and (top_k=35), was used to limit randomness during text generation and encourage more consistent outputs. Additionally, a repetition penalty of (1.15) was applied to reduce repeated expressions. The maximum output length was set to 120 tokens. This configuration was implemented to produce realistic and coherent tourism reviews while preserving some degree of lexical diversity.

The second generation strategy employed more exploratory decoding parameters to promote diversity and originality in the generated reviews. A higher temperature (0.7), together with (top_p=0.90) and (top_k=50), was used to increase the variety of candidate tokens considered during text generation. Additionally, a repetition penalty of 1.20 and a no-repeat-ngram constraint of size two were applied to prevent repetitive expressions and reduce similarity to previously generated text. The maximum output length was maintained at 120 tokens. This configuration was designed to generate more diverse and original tourism reviews.

The third generation strategy used the same decoding configuration as the second experiment, but incorporated a sentiment-based filtering step. For each generation condition, three synthetic candidate reviews were produced using the fine-tuned mT5 model. Then, each candidate was evaluated with a RoBERTa-based sentiment classifier [15] previously adapted to the tourism domain and fine-tuned on the REST-MEX 2023 sentiment analysis task. The classifier assigned a probability score to each candidate according to the target polarity, and the review with the highest probability for the expected sentiment label was selected as the final output. This strategy aimed to improve sentiment consistency while preserving the diversity-oriented generation settings of the second experiment.

Table 4 summarizes the generation settings used in the three experiments. While Experiments 1 and 2 differ in their decoding parameters, Experiment 3 extends the second configuration by generating multiple candidates and applying RoBERTa-based sentiment filtering.

Table 4
Generation settings for the three proposed experiments.

Parameter	Exp. 1	Exp. 2	Exp. 3
Temperature	0.6	0.7	0.7
Top-p	0.85	0.90	0.90
Top-k	35	50	50
Repetition Penalty	1.15	1.20	1.20
Generated Candidates	1	1	3
Sentiment Filtering	None	None	RoBERTa

Table 5
Results for the three experiments.

Experiment	F1 (1)	F1 (2)	F1 (3)	F1 (4)	F1(5)	Macro-F1
1	0.492776	0.000000	0.431156	0.305465	0.853210	0.416521
2	0.398551	0.003617	0.413542	0.360717	0.848229	0.404931
3	0.600904	0.223776	0.447946	0.477752	0.841351	0.518346

Table 6
Official REST-MEX 2026 results.

Submission	Track Score	Macro-F1
LynxesAI-LKEBUAP-Run3-mt5-base-finetuned-configuration-2-roberta-filtering	0.5084593819	0.5515316424
LynxesAI-LKEBUAP-Run1-mt5-base-finetuned-configuration-1	0.4611706559	0.5105650788
LynxesAI-LKEBUAP-Run2-mt5-base-finetuned-configuration-2	0.437903523	0.4910894284

Each generation strategy produced a synthetic corpus consisting of 3,000 reviews together with their corresponding sentiment polarity labels. Following the REST-MEX 2026 evaluation protocol, each synthetic corpus was used as the training set to fine-tune a sentiment classifier based on the domain-adapted RoBERTa model `vg055/roberta-base-bne-finetuned-TripAdvisorDomainAdaptation`. The evaluation set corresponded to the 10% validation subset previously extracted from the original REST-MEX 2026 training corpus using stratified sampling based on sentiment polarity. Fine-tuning was performed for two epochs using a learning rate of 2e-5, a batch size of 16, and a weight decay of 0.01. The final evaluation reports the F1-score for each sentiment class together with the macro-F1 score, which was used to assess the effectiveness of the generated synthetic corpus for downstream sentiment classification. The results obtained by the three generation strategies are presented in Table 5.

As shown in Table 5, the experiment 3 achieved the best performance, obtaining a macro-F1 score of 0.5183. This result suggests that the RoBERTa-based filtering strategy improved the quality of the generated synthetic corpus by selecting reviews that better matched the target polarity. In contrast, Experiments 1 and 2 achieved lower macro-F1 scores of 0.4165 and 0.4049, respectively. Although the second configuration promoted greater diversity and originality, the increased randomness may have introduced sentiment inconsistencies that negatively affected downstream classification performance. The filtering strategy incorporated in Experiment 3 mitigated this issue and produced the most effective synthetic corpus for sentiment classification.

Each experiment was submitted as an independent run to the official REST-MEX 2026 evaluation. Table 7 summarizes the official results obtained by the three submitted runs. Run 3, which incorporates RoBERTa-based sentiment filtering, achieved the best performance, obtaining a Track Score of 0.5085 and a Macro-F1 score of 0.5515. These results indicate that sentiment-based filtering improved the quality of synthetic reviews for downstream sentiment classification. Furthermore, Run 3 ranked third in the official Synthetic Sentiment Modeling task of REST-MEX 2026.

6. Conclusions

This work presented an instruction-based framework for generating synthetic tourism reviews for REST-MEX 2026. The experimental results showed that increasing the diversity of the generated reviews alone was not sufficient to improve the performance of downstream sentiment classification. In contrast, incorporating a domain-adapted RoBERTa classifier as a sentiment-based filtering stage produced a synthetic corpus that more accurately preserved the target sentiment, yielding the best performance among the proposed generation strategies. These findings suggest that sentiment consistency is a key factor in improving the effectiveness of synthetic data for sentiment analysis.

As future work, we plan to investigate domain adaptation techniques for mT5 using tourism reviews to further improve the quality and sentiment consistency of the generated synthetic corpus. We also intend to compare the proposed approach with instruction-tuned multilingual generative models, such as FLAN-T5, to further evaluate their effectiveness for conditional synthetic review generation.

7. Acknowledgments

The authors acknowledge the support of Universidad de las Américas Puebla (UDLAP), Universidad Politécnica de Tapachula (UPTAP), Universidad de Guadalajara (UdeG), and Benemérita Universidad Autónoma de Puebla (BUAP) during the development of this work.

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT (OpenAI) to improve the English writing and for technical assistance during software development and model implementation. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of this publication.

References

[1] M. Birjali, M. Kasri, A. Beni-Hssane, A comprehensive survey on sentiment analysis: Approaches, challenges and trends, Knowledge-Based Systems 226 (2021) 107134. URL: https://www.sciencedirect.com/science/article/pii/S095070512100397X. doi:https://doi.org/10.1016/j.knosys.2021.107134.

[2] S. Kemp, Digital 2026 Global Overview Report, Technical Report, Kepios, We Are Social and Meltwater, 2025. URL: https://datareportal.com/reports/digital-2026-global-overview-report.

[3] Y. Mao, Q. Liu, Y. Zhang, Sentiment analysis methods, applications, and challenges: A systematic literature review, Journal of King Saud University - Computer and Information Sciences 36 (2024) 102048. URL: https://www.sciencedirect.com/science/article/pii/S131915782400137X. doi:https://doi.org/10.1016/j.jksuci.2024.102048.

[4] M. Álvarez Carmona, R. Aranda, A. Y. Rodríguez-Gonzalez, D. Fajardo-Delgado, M. G. Sánchez, H. Pérez-Espinosa, J. Martínez-Miranda, R. Guerrero-Rodríguez, L. Bustio-Martínez, Ángel Díaz-Pacheco, Natural language processing applied to tourism research: A systematic review and future research directions, Journal of King Saud University - Computer and Information Sciences 34 (2022) 10125–10144. URL: https://www.sciencedirect.com/science/article/pii/S1319157822003615. doi:https://doi.org/10.1016/j.jksuci.2022.10.010.

[5] L. Long, R. Wang, R. Xiao, J. Zhao, X. Ding, G. Chen, H. Wang, On LLMs-driven synthetic data generation, curation, and evaluation: A survey, in: L.-W. Ku, A. Martins, V. Srikumar (Eds.), Findings of the Association for Computational Linguistics: ACL 2024, Association for Computational Linguistics, Bangkok, Thailand, 2024, pp. 11065–11082. URL: https://aclanthology.org/2024.findings-acl.658/. doi:10.18653/v1/2024.findings-acl.658.

[6] M. Nadăş, L. Dioşan, A. Tomescu, Synthetic data generation using large language models: Advances in text and code, IEEE Access 13 (2025) 134615–134633. URL: http://dx.doi.org/10.1109/ACCESS.2025.3589503. doi:10.1109/access.2025.3589503.

[7] D. Zha, Z. P. Bhat, K.-H. Lai, F. Yang, Z. Jiang, S. Zhong, X. Hu, Data-centric artificial intelligence: A survey, ACM Comput. Surv. 57 (2025). URL: https://doi.org/10.1145/3711118. doi:10.1145/3711118.

[8] M. Álvarez Carmona, Díaz-Pacheco, R. Aranda, M.-M. Reyes-Sierra, A. Y. Rodríguez-González, L. Bustio-Martínez, V. Herrera-Semenets, Overview of rest-mex at iberlef 2026: Research on synthetic text modeling for mexican magical towns, Procesamiento del Lenguaje Natural 77 (2026).

[9] A. Bonet-Jover, J. González-Barba, L. Chiruzzo, Overview of IberLEF 2026: Natural Language Processing Challenges for Spanish and other Iberian Languages, in: Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2026), co-located with the 42nd Conference of the Spanish Society for Natural Language Processing (SEPLN 2026), CEUR-WS.org, 2026.

[10] R. G. Benítez, A. S. Navarrete, C. Vidal-Castro, C. Martínez-Araneda, Guide for the application of the data augmentation approach on sets of texts in spanish for sentiment and emotion analysis, PLoS ONE 19 (2024). URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85205146749&doi=10.1371%2fjournal.pone.0310707&partnerID=40&md5=1ce4498a86d6af6893903c06d3d5f2ac. doi:10.1371/journal.pone.0310707, cited by: 5; All Open Access, Gold Open Access, Green Open Access.

[11] M. Álvarez-Carmona y Ángel Díaz-Pacheco y Ramón Aranda y Ansel Y. Rodríguez-González y Daniel Fajardo-Delgado y Rafael Guerrero-Rodríguez y Lázaro 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) 289–299. URL: http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/6449.

[12] M. Ángel Álvarez-Carmona y Ángel Díaz-Pacheco y Ramón Aranda y Ansel Yoan Rodríguez-González y Victor Muñiz-Sánchez y Adrián Pastor López-Monroy y Fernando Sánchez-Vega y Lázaro Bustio-Martínez, Overview of rest-mex at iberlef 2023: Research on sentiment analysis task for mexican tourist texts, Procesamiento del Lenguaje Natural 71 (2023) 425–436. URL: http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/6572.

[13] M. Álvarez-Carmona y Ángel Díaz-Pacheco y Ramón Aranda y Ansel Y. Rodríguez-González y Lázaro Bustio-Martínez y Vitali Herrera-Semenets, Overview of rest-mex at iberlef 2025: Researching sentiment evaluation in text for mexican magical towns, Procesamiento del Lenguaje Natural 75 (2025) 499–511. URL: http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/6771.

[14] L. Xue, N. Constant, A. Roberts, M. Kale, R. Al-Rfou, A. Siddhant, A. Barua, C. Raffel, mT5: A massively multilingual pre-trained text-to-text transformer, in: K. Toutanova, A. Rumshisky, L. Zettlemoyer, D. Hakkani-Tur, I. Beltagy, S. Bethard, R. Cotterell, T. Chakraborty, Y. Zhou (Eds.), Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Association for Computational Linguistics, Online, 2021, pp. 483–498. URL: https://aclanthology.org/2021.naacl-main.41/. doi:10.18653/v1/2021.naacl-main.41.

[15] V. G. Morales-Murillo, H. Gómez-Adorno, D. Pinto, I. A. Cortés-Miranda, P. Delice, Lke-iimas team at rest-mex 2023: Sentiment analysis on mexican tourism reviews using transformer-based domain adaptation, in: Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2023) co-located with the Conference of the Spanish Society for Natural Language Processing (SEPLN 2023), volume 3496, CEUR Workshop Proceedings (CEUR-WS.org), 2023. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85175333990&partnerID=40&md5=c69922284f684aa47aa32847f11e4d73.