Mutual Information-Guided Prompt Engineering for Synthetic Tourism Review Generation in Mexican Magical Towns

Alejandra Romero-Canton[1,*], Jose Ramon Aranda-Romero[1]

[1]*Secretaría de Educación del Gobierno del Estado de Yucatán, Mérida, México*

Abstract

The generation of high-quality synthetic text has recently emerged as an effective solution for addressing data scarcity, privacy preservation, and severe class imbalance in Natural Language Processing. Within tourism applications, online reviews are typically dominated by highly positive opinions, making the construction of balanced datasets particularly challenging. The REST-MEX 2026 shared task addresses this problem by requiring participants to generate a synthetic corpus capable of training sentiment classifiers that generalize to real-world reviews. This paper proposes an information-theoretic framework that guides Large Language Models (LLMs) through discriminative lexical information automatically extracted from the original corpus. Instead of selecting prompting examples randomly, we first estimate the Mutual Information (MI) between lexical units and class labels, identifying words that best characterize each sentiment polarity and attraction category. These high-information words are subsequently expanded through semantic similarity and incorporated into structured prompts together with contextual metadata such as municipality, attraction type, and target review length. Consequently, the language model receives statistically informative lexical guidance while maintaining sufficient freedom to generate natural and diverse reviews. Unlike previous approaches where Mutual Information has mainly been employed for feature selection in classification tasks, our methodology uses MI as an automatic prompt engineering mechanism for controlled synthetic text generation. Experimental evaluation under the REST-MEX 2026 benchmark demonstrates that incorporating information-theoretic lexical guidance improves lexical diversity, semantic consistency, and downstream sentiment classification performance compared with conventional prompting strategies.

Keywords

Synthetic Text Generation, Prompt Engineering, Mutual Information, Large Language Models, Tourism Reviews, REST-MEX

1. Introduction

Large Language Models (LLMs) have transformed Natural Language Processing by enabling machines to generate coherent, context-aware, and highly fluent text across a wide variety of domains. Rather than relying exclusively on model architecture or parameter scaling, recent studies have shown that the quality of generated text strongly depends on how information is presented to the model. Consequently, prompt engineering has emerged as one of the most influential research topics in modern generative artificial intelligence.

Despite its importance, prompt engineering remains largely heuristic. Most existing approaches rely on manually crafted instructions, expert intuition, or randomly selected few-shot examples. Although these strategies often produce satisfactory results, they rarely exploit one of the most valuable resources available during generation: the statistical knowledge already contained in the training corpus. As a consequence, generated texts frequently overlook discriminative vocabulary associated with particular semantic categories, exhibit repetitive lexical patterns, or fail to adequately represent minority classes.

These limitations become particularly evident in tourism-related applications. Online travel reviews constitute one of the principal sources of information supporting destination management, recommendation systems, customer satisfaction analysis, and tourism intelligence. During the last decade, numerous studies have demonstrated the usefulness of Natural Language Processing for extracting

IberLEF 2026, September 2026, León, Spain

*Corresponding author.

ale_romero85@hotmail.com (A. Romero-Canton); ramon.arac84@gmail.com (J. R. Aranda-Romero)

knowledge from tourism reviews, including sentiment analysis, topic discovery, destination image analysis, and conversational systems [1, 2, 3, 4].

The REST-MEX evaluation campaign has significantly contributed to this research area by providing standardized datasets and evaluation protocols for Spanish-language tourism texts. Earlier editions focused on recommendation systems, sentiment analysis, and polarity classification [5, 6, 7]. More recently, REST-MEX 2026 introduced a substantially different challenge where the objective is no longer to classify existing reviews but to generate a balanced synthetic corpus capable of training classifiers that generalize to real-world opinions [8, 9]. This new formulation shifts the research focus from discriminative modeling toward controllable text generation.

One of the main difficulties in this scenario is ensuring that generated reviews preserve the lexical characteristics associated with each sentiment category. Most prompting strategies describe the desired polarity using natural language instructions such as "generate a positive review" or "write a negative opinion." While intuitive, these descriptions provide only weak lexical guidance and leave the language model responsible for determining which words should characterize each sentiment level.

Interestingly, information-theoretic methods offer an alternative perspective. Mutual Information has been successfully employed to identify the words that best discriminate among sentiment classes, business categories, and geographic locations within tourism corpora. Romero-Cantón and Aranda-Romero proposed an MI-based framework for feature extraction in the REST-MEX 2025 classification task, demonstrating that statistically informative vocabularies improve feature interpretability and contribute to more balanced classifiers [**?**]. However, this lexical information was used exclusively for document classification and has not yet been exploited for synthetic text generation.

In this paper, we argue that discriminative lexical information should guide the generation process itself. Rather than treating Mutual Information as a feature selection mechanism, we reinterpret it as an automatic prompt engineering strategy. The central idea is straightforward: if Mutual Information identifies the vocabulary that best characterizes each sentiment class, then incorporating those words into prompts should encourage language models to generate reviews that more faithfully reflect the linguistic properties of the target category.

To achieve this objective, we first compute Mutual Information scores for every word in the corpus with respect to sentiment polarity and attraction type. The highest-ranked lexical units are subsequently expanded through semantic similarity relationships to improve vocabulary coverage while preserving class specificity. Finally, these information-theoretic lexical profiles are integrated into structured prompts together with metadata describing the municipality, attraction category, and stylistic constraints. The resulting prompts provide both semantic context and statistically grounded lexical guidance, enabling the language model to produce diverse and class-consistent synthetic reviews.

The main contributions of this work are summarized as follows:

- We propose an information-theoretic prompt engineering framework for synthetic tourism review generation.
- We reformulate Mutual Information from a feature selection technique into a lexical guidance mechanism for Large Language Models.
- We introduce lexical profiles that summarize the most discriminative vocabulary associated with each sentiment category.
- We demonstrate that statistically guided prompts improve lexical diversity and downstream sentiment classification under the REST-MEX 2026 evaluation protocol.

The remainder of this paper is organized as follows. Section 2 discusses previous work on information-theoretic feature extraction and synthetic text generation. Section 3 presents the proposed lexical profile construction methodology and the prompt engineering framework. Section 4 describes the experimental setup and evaluation methodology. Experimental results are presented in Section 5, followed by conclusions and future research directions in Section 6.

2. Information-Theoretic Motivation

Large Language Models have demonstrated remarkable capabilities for generating coherent and context-aware text. Nevertheless, the quality of the generated output strongly depends on the information provided within the prompt. Current prompt engineering strategies generally rely on manually designed instructions, handcrafted templates, or randomly selected demonstrations. While these approaches often produce fluent text, they rarely exploit the statistical properties of the training corpus itself. Consequently, language models may overlook lexical patterns that characterize particular semantic categories, especially when generating samples for minority classes.

Information Theory provides a principled mathematical framework for quantifying the dependency between variables. Among its most widely used measures, Mutual Information (MI) estimates the amount of information that one random variable provides about another. In Natural Language Processing, MI has traditionally been employed to identify discriminative words, rank features, estimate semantic relevance, and improve document classification. Because MI measures statistical dependence rather than simple frequency, it is capable of identifying vocabulary that truly characterizes each class instead of merely selecting the most common words.

2.1. Mutual Information in Natural Language Processing

Feature selection has been one of the classical applications of Mutual Information within text mining. Given a collection of labeled documents, MI quantifies how strongly the occurrence of a particular word is associated with a target class. Words appearing uniformly across all classes receive low scores, whereas words concentrated in specific categories obtain significantly larger values.

Several studies have demonstrated the usefulness of Mutual Information for text classification, sentiment analysis, authorship attribution, document retrieval, and semantic representation learning. Unlike frequency-based measures such as TF-IDF, MI explicitly incorporates the class distribution, allowing it to identify vocabulary that contributes to discrimination rather than popularity.

Recent advances in representation learning have not displaced these information-theoretic approaches. On the contrary, they frequently complement deep neural architectures by providing interpretable lexical features that improve model robustness, particularly under highly imbalanced data distributions.

2.2. Mutual Information in Tourism Text Analysis

Within tourism research, statistical feature selection has received comparatively less attention than transformer-based language models. Most recent studies have focused on contextual embeddings, attention mechanisms, and large pre-trained language models for sentiment analysis, destination image analysis, recommendation systems, and topic discovery [1, 2, 4, 3].

An important exception is the work of Romero-Cantón and Aranda-Romero [**?**], who proposed a Mutual Information-driven framework for the REST-MEX 2025 shared task. Their methodology computes normalized Mutual Information scores between lexical units and several tourism-related labels, including sentiment polarity, attraction type, and geographic location. High-scoring words are further enriched through synonym expansion, producing interpretable lexical profiles that improve feature representation for FastText classifiers. Unlike conventional bag-of-words models, this strategy combines statistical relevance with semantic enrichment, allowing minority classes to be represented by more informative vocabularies.

Although their objective was document classification rather than generation, the resulting lexical profiles reveal an interesting property. The extracted vocabularies describe the language naturally associated with each class. Positive reviews tend to contain words related to satisfaction, comfort, cleanliness, or recommendation, whereas negative opinions emphasize service deficiencies, waiting times, costs, or unpleasant experiences. These discriminative lexical distributions constitute valuable prior knowledge that remains largely unexplored within current LLM prompting strategies.

2.3. Prompt Engineering for Synthetic Text Generation

Recent research on synthetic text generation has explored several prompting paradigms, including zero-shot prompting, few-shot prompting, chain-of-thought prompting, persona-based prompting, and retrieval-augmented prompting. Most of these methodologies focus on improving the contextual information presented to the language model by incorporating demonstrations, instructions, or external documents.

Within the REST-MEX 2026 challenge, participating systems have adopted diverse generation strategies. Some approaches rely on semantic prototype retrieval through embedding clustering, while others generate reviews from simulated traveler profiles, synthetic itineraries, or fine-tuned language models [8]. Additional studies have explored parameter-efficient fine-tuning techniques such as QLoRA or fully local language models to preserve privacy during synthetic data generation.

Despite these advances, prompt construction itself remains predominantly empirical. Demonstrations are usually selected randomly, manually curated, or retrieved according to semantic similarity. However, little attention has been devoted to exploiting the lexical statistics already available in the original training corpus. Consequently, prompts frequently omit highly discriminative vocabulary that could guide language models toward generating more class-consistent text.

2.4. Research Gap

The review above reveals an interesting disconnect between two mature research areas. On one hand, Information Theory provides statistically grounded mechanisms for identifying discriminative lexical units within labeled corpora. On the other hand, Large Language Models rely heavily on prompt engineering to control text generation. Surprisingly, these two perspectives have evolved almost independently.

To the best of our knowledge, no previous work has employed Mutual Information as an automatic prompt engineering mechanism for synthetic tourism review generation. Existing MI-based approaches terminate once informative features have been extracted for classification, whereas modern prompting techniques rarely exploit corpus-level lexical statistics beyond simple keyword selection.

This work bridges these two research directions. Rather than using Mutual Information to improve a classifier, we employ it to construct information-rich prompts that guide a Large Language Model during review generation. In this way, discriminative lexical information is transferred from the original corpus into the generation process itself, allowing the model to produce reviews that better preserve the statistical characteristics of authentic tourism opinions.

3. Lexical Prompt Construction Framework

The proposed framework automatically constructs prompts from the statistical properties of the original corpus instead of relying on manually designed templates or randomly selected demonstrations. Figure **??** illustrates the complete pipeline.

The methodology consists of five consecutive stages:

1. Text preprocessing.
2. Mutual Information estimation.
3. Lexical profile construction.
4. Semantic vocabulary expansion.
5. Information-guided prompt generation.

Unlike conventional prompt engineering approaches, every prompt is dynamically generated from lexical information extracted directly from the training corpus.

3.1. Corpus Preprocessing

Given the original corpus

$$D = \{(r_i, p_i, t_i, m_i)\}_{i=1}^{N}, \tag{1}$$

where r_i denotes the review, p_i its polarity, t_i the attraction type, m_i the municipality, every review undergoes several normalization steps before statistical analysis.

The preprocessing stage consists of

1. Lowercasing.
2. Unicode normalization.
3. Removal of duplicated spaces.
4. Tokenization.
5. Lemmatization.
6. Stop-word removal.

Unlike traditional bag-of-words approaches, punctuation is preserved because it contributes to the writing style observed in authentic tourism reviews.

3.2. Mutual Information Estimation

For every token $w \in V$, where V denotes the corpus vocabulary, the dependency between the occurrence of the word and each target class is quantified through Mutual Information.

For a class c, MI is computed as

$$MI(w, c) = P(w, c) \log \frac{P(w, c)}{P(w)P(c)} \tag{2}$$

Words obtaining large positive values are highly representative of the corresponding class, whereas words with values close to zero provide little discriminative information. Unlike previous applications where MI is exclusively employed for feature selection, the proposed methodology interprets MI as a lexical relevance score for prompt construction.

3.3. Lexical Profile Construction

For every polarity-attraction pair (p, t), the words are sorted according to their Mutual Information score, $MI(w, p, t)$. The resulting ordered list

$$L_{p,t} = \{w_1, w_2, \ldots, w_n\}, \tag{3}$$

constitutes the lexical profile associated with that category. Only the highest-scoring words are retained,

$$\Omega_{p,t} = \{w_1, \ldots, w_K\}, \tag{4}$$

where K denotes the maximum vocabulary size.

These lexical profiles summarize the vocabulary naturally associated with each sentiment level without requiring any manual intervention.

3.4. Semantic Vocabulary Expansion

Although Mutual Information identifies highly discriminative words, lexical variability can still be limited because different users frequently describe similar experiences using synonymous expressions.

To increase vocabulary diversity, every representative word is expanded through semantic similarity.

Given a representative word w_i, its semantic neighborhood is defined as

$$\mathcal{N}(w_i) = \{v_1, v_2, \ldots, v_k\}, \tag{5}$$

where every neighbor satisfies

$$Sim(w_i, v_j) > \tau \tag{6}$$

Semantic similarity may be computed using multilingual sentence embeddings or lexical resources such as WordNet. The expanded lexical profile becomes

$$\Omega'_{p,t} = \Omega\_\{p\, t\} \cup \mathcal{N}(\Omega_{p,t}). \tag{7}$$

This enrichment increases lexical diversity while preserving the semantic characteristics of each class.

3.5. Prompt Construction

For every synthetic review, the prompt is automatically assembled using four complementary information sources.

1. Destination metadata.
2. Sentiment polarity.
3. Attraction category.
4. Information-theoretic lexical profile.

The prompt is formally defined as

$$\Pi = \{m, t, p, \Omega'_{p,t}, L\}, \tag{8}$$

where

- m denotes the municipality,
- t the attraction type,
- p the desired polarity,
- $\Omega'_{p,t}$ the expanded lexical profile,
- L the target review length.

Instead of explicitly instructing the language model to employ particular words, the lexical profile is incorporated into the prompt as contextual guidance, allowing the model to naturally integrate discriminative vocabulary during generation.

Given the prompt Π, a Large Language Model generates a synthetic review $s = LLM(\Pi)$. Generation continues until obtaining 600 reviews for every polarity level while maintaining balanced attraction categories. Each review is independently generated, ensuring lexical diversity through stochastic decoding.

4. Information-Theoretic Lexical Profiles

The central hypothesis of this work is that the vocabulary naturally associated with each sentiment category contains valuable information that should guide the generation process. Instead of manually selecting keywords or relying on randomly retrieved examples, we automatically construct lexical profiles that summarize the statistical language associated with every class.

Each lexical profile represents a compact description of the vocabulary that best characterizes a particular sentiment polarity and attraction type. Consequently, the generated prompts become both interpretable and statistically grounded.

4.1. Normalized Mutual Information

Raw Mutual Information values may be biased toward infrequent events. To reduce this effect, every score is normalized according to

$$NMI(w, c) = \frac{MI(w, c)}{\max\limits_{v \in V} MI(v, c)}, \tag{9}$$

where w denotes a vocabulary word, c represents one target class, V is the corpus vocabulary. Consequently, $0 \leq NMI(w, c) \leq 1$. This normalization allows lexical profiles extracted from different classes to become directly comparable.

4.2. Lexical Importance Ranking

After normalization, every class vocabulary is sorted according to $NMI(w, c)$. The resulting ranking

$$R_c = \{(w_1, s_1), (w_2, s_2), \ldots, (w_n, s_n)\}, \tag{10}$$

contains every word together with its normalized information score $s_i = NMI(w_i, c)$. Only the highest-ranked words are retained,

$$R_c^{(K)} = \{w_1, \ldots, w_K\}, \tag{11}$$

where $K = 30$ in our experiments. Unlike frequency-based vocabularies, these words maximize discriminative information rather than occurrence count.

4.3. Category-based Lexical Organization

Rather than directly inserting the ranked words into the prompt, they are organized into semantic categories.

Every representative word is assigned to one of several tourism-related concepts:

- Service
- Facilities
- Food
- Price
- Comfort
- Cleanliness
- Recommendation
- Environment

For example, a positive hotel profile may contain { *comfortable*, *clean*, *friendly*, *excellent*, *recommended*}, whereas a negative restaurant profile may include { *expensive*, *slow*, *cold*, *rude*, *disappointing* }. Grouping words according to semantic concepts facilitates prompt readability while encouraging the language model to naturally incorporate multiple aspects of the tourist experience.

4.4. Lexical Diversity Sampling

Using the same vocabulary repeatedly would inevitably reduce corpus diversity.

Therefore, every generated review samples only a subset of the lexical profile.

Given $R_c^{(K)}$, the sampled vocabulary becomes

$$\Phi_c = RandomSubset(R_c^{(K)}, m), \tag{12}$$

where $m = 10$. Words with larger Mutual Information receive greater sampling probability,

$$P(w_i) = \frac{s_i}{\sum_j s_j}. \tag{13}$$

Consequently, highly representative words appear more frequently while lower-ranked vocabulary still contributes to lexical diversity.

4.5. Prompt Template

Every synthetic review is generated from a structured prompt composed of five information blocks.

1. Municipality.
2. Attraction type.
3. Sentiment polarity.
4. Information-theoretic lexical profile.
5. Stylistic generation instructions.

Formally,

$$Prompt = \{m, t, p, \Phi_c, I\}, \tag{14}$$

where I contains stylistic instructions encouraging

- first-person narration,
- natural Mexican Spanish,
- coherent review length,
- realistic tourism experiences.

Unlike manually designed prompts, the lexical component is automatically generated from the statistical properties of the corpus, making the framework easily transferable to other datasets.

4.6. Illustrative Example

Table 1 presents an example of the proposed prompt construction strategy.

Table 1
Example of information-guided prompt construction.

Target polarity	2
Attraction	Restaurant
Town	Taxco
Lexical profile	slow, expensive, cold, waiting, disappointing
Instruction	Write a natural review using first-person narration.

Instead of explicitly forcing the language model to use every lexical item, the profile acts as contextual guidance. This strategy allows the generated review to remain fluent while naturally incorporating statistically informative vocabulary.

5. Experimental Evaluation

Experiments were conducted using the official REST-MEX 2026 training corpus, which contains more than 208,000 Spanish-language reviews from Mexican Magical Towns. Each review includes the review text, sentiment polarity (1–5), attraction type (Hotel, Restaurant or Attractive), municipality, and state [8].

The objective was to generate a balanced corpus containing exactly 3,000 synthetic reviews following the competition guidelines. Reviews were uniformly distributed across the five sentiment levels and the three attraction categories.

The proposed framework was compared against three representative prompting strategies:

- Zero-shot prompting.
- Random few-shot prompting.
- Semantic prototype prompting.

All generation experiments used identical decoding parameters to ensure fair comparisons.

5.1. Evaluation Metrics

Following the REST-MEX 2026 protocol, every synthetic corpus was used to train a multilingual sentiment classifier that was subsequently evaluated on a hidden collection of authentic reviews.

Performance was measured using

- Track Score,
- Macro-F1,
- Precision,
- Recall.

In addition, we evaluated the quality of the generated corpora using two complementary generation-oriented metrics:

- Lexical Diversity (Type-Token Ratio).
- Semantic Diversity (average cosine distance between Sentence-BERT embeddings).

6. Results

Table 2 summarizes the classification performance obtained by the evaluated prompting strategies.

Table 2
Downstream sentiment classification performance.

Method	Track Score	Macro-F1
Zero-shot Prompting	0.431	0.472
Random Few-shot	0.479	0.508
Semantic Prototypes	0.507	0.529
Proposed Method	**0.521**	**0.546**

The proposed information-theoretic prompting strategy consistently achieved the highest performance across all evaluation metrics. Incorporating discriminative lexical profiles allowed the language model to generate reviews that more accurately reflected the linguistic characteristics of each sentiment class.

Table 3 presents the diversity analysis.

Table 3
Lexical and semantic diversity.

Method	TTR	Cosine Distance
Random Prompting	0.64	0.289
Semantic Prototypes	0.70	0.347
Proposed Method	**0.75**	**0.381**

Table 4
Ablation study.

Configuration	Track Score
Random Prompt	0.479
+ MI Vocabulary	0.503
+ Semantic Expansion	0.513
+ Lexical Profiles	**0.521**

The information-guided prompts generated the most diverse corpus according to both lexical and semantic measures. The proposed lexical profiles encouraged the language model to employ richer vocabulary while preserving coherence across reviews.

To analyze the contribution of each framework component, an ablation study was performed.

The results indicate that each stage contributes positively to the final performance. While Mutual Information alone already improves prompt quality, semantic expansion and structured lexical profiles provide additional gains by increasing vocabulary diversity without sacrificing class consistency.

7. Analysis

The proposed methodology differs from conventional prompt engineering because lexical guidance is derived directly from the statistical structure of the training corpus instead of human intuition.

Interestingly, the generated reviews naturally incorporated representative vocabulary without explicitly forcing the language model to reproduce individual keywords. This suggests that lexical profiles operate as soft semantic constraints rather than rigid lexical templates, allowing the model to preserve fluency while remaining faithful to the linguistic characteristics of each sentiment category.

Another important observation concerns minority sentiment classes. Reviews corresponding to one- and two-star polarities exhibited noticeably greater lexical richness than those generated using conventional prompting strategies. Since the official REST-MEX evaluation metric assigns larger weights to minority classes, this richer representation translated into improved downstream classification performance.

Unlike manually designed prompts, the proposed framework can be automatically transferred to new domains. Computing Mutual Information requires only a labeled corpus, making the methodology applicable to healthcare, finance, education, legal documents, or any other domain where discriminative vocabulary can be statistically estimated.

8. Conclusions

This paper presented an information-theoretic framework for prompt engineering that exploits Mutual Information to guide Large Language Models during synthetic tourism review generation.

Unlike previous applications where Mutual Information has been employed exclusively for feature selection, the proposed methodology reformulates MI as a lexical guidance mechanism capable of automatically constructing prompts from the statistical properties of the training corpus. By combining discriminative vocabularies, semantic expansion, and structured prompt templates, the proposed framework produces synthetic reviews that exhibit higher lexical diversity, stronger semantic consistency, and improved downstream classification performance.

Experimental evaluation under the REST-MEX 2026 benchmark demonstrates that information-theoretic prompting consistently outperforms conventional prompting strategies while remaining fully model-independent. Because lexical profiles are automatically extracted from labeled corpora, the proposed methodology can be readily adapted to other text generation problems without requiring manual prompt engineering.

Future work will investigate contextual lexical profiles conditioned simultaneously on sentiment polarity, attraction type, and municipality. Additionally, adaptive lexical retrieval strategies and graph-based semantic representations will be explored to further improve diversity and realism in synthetic text generation.

Declaration on Generative AI

Generative AI tools were used solely to improve the language, grammar, and readability of the manuscript. The authors reviewed and edited all AI-assisted content and take full responsibility for the final version of the manuscript.

References

[1] M. Á. 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, Á. 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.

[2] A. Diaz-Pacheco, M. Á. Álvarez-Carmona, R. Guerrero-Rodríguez, L. A. C. Chávez, A. Y. Rodríguez-González, J. P. Ramírez-Silva, R. Aranda, Artificial intelligence methods to support the research of destination image in tourism. a systematic review, Journal of Experimental & Theoretical Artificial Intelligence 36 (2024) 1415–1445.

[3] R. Guerrero-Rodriguez, M. Á. 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.

[4] 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.

[5] 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.

[6] 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).

[7] 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).

[8] M. Á. Álvarez-Carmona, Á. Díaz-Pacheco, R. Aranda, 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. Á. 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.