Article

Spanish Fake News Detection with Fine-Tuned DistilBERT Built upon a Unified Corpus for a Real-World Application

Gabriel Hurtado Avilés [1], José A. Reyes-Ortiz [1,*], Román A. Mora-Gutiérrez [1] and Josué Padilla Cuevas [1]

[1] Department of Systems, Autonomous Metropolitan University (UAM), Azcapotzalco Unit, Mexico City 02200, Mexico; al2232800343@azc.uam.mx (G.H.A.); jaro@azc.uam.mx (J.A.R.-O.); mgra@azc.uam.mx (R.A.M.-G.); jpc@azc.uam.mx (J.P.C.)

* Correspondence: jaro@azc.uam.mx

Abstract

The digital ecosystem of the Spanish-speaking world is increasingly compromised by disinformation, threatening not only information integrity but also the cultural and democratic health of its over 500 million speakers. While transformer-based models have achieved state-of-the-art performance in English fake news detection, the Spanish-language domain remains critically underserved due to fragmented datasets, insufficient hyperparameter optimization research, and a persistent gap between academic models and publicly accessible tools. This work addresses these interconnected challenges through a comprehensive end-to-end framework. First, a unified corpus of 61,674 Spanish news articles was constructed by integrating four academic datasets with web-scraped satirical content, achieving near-perfect class balance (49.8% fake, 50.2% real)—representing one of the largest standardized resources for this task. Second, through over 500 GPU hours of systematic experimentation, an aggressive regularization strategy for fine-tuning DistilBERT (Distilled Bidirectional Encoder Representations from Transformers) was identified, achieving 95.36% accuracy on the test set while maintaining a validation-training loss gap of 0.058—corresponding to a 23.33 percentage point improvement over classical metaheuristic-optimized methods. Third, the optimized model was deployed as a Docker-containerized web application capable of real-time URL analysis, transforming academic research into a practical tool for digital self-defense. The complete framework—corpus, model, and application—is publicly available to support reproducibility and further research. Beyond technical metrics, this research represents a tangible contribution toward democratizing fact-checking resources for Spanish-speaking communities, reinforcing critical thinking and digital literacy in an era of pervasive misinformation.

Keywords: fake news detection; Spanish natural language processing; transformer models; DistilBERT; hyperparameter optimization; metaheuristic algorithms; web application deployment; misinformation

Received:
Revised:
Accepted:
Published:

Citation: Hurtado Avilés, G.; Reyes-Ortiz, J.A.; Mora-Gutiérrez, R.A.; Padilla Cuevas, J. Spanish Fake News Detection with Fine-Tuned DistilBERT Built upon a Unified Corpus for a Real-World Application. *Big Data Cogn. Comput.* **2025**, *1*, 0. https://doi.org/

1. Introduction

The proliferation of digital misinformation represents a critical threat to information integrity [1]. In recent years, the rapid expansion of the internet has transformed how information is created, disseminated, and consumed, creating unprecedented opportunities for the spread of deceptive content. This phenomenon disproportionately affects vulnerable demographics with limited media literacy skills [2,3], undermining public trust in journalism, distorting democratic processes, and threatening public health outcomes [4].

The field of automated fake news detection has evolved through distinct methodological paradigms, reflecting advances in machine learning and natural language processing. Early research in the 2010s relied on classical machine learning approaches—Support Vector Machines (SVMs), Naïve Bayes, Random Forests—combined with hand-crafted linguistic features such as Term Frequency-Inverse Document Frequency (TF-IDF) and Bag-of-Words (BoW) representations [5,6]. While computationally efficient, these methods struggled to capture semantic nuances, sarcasm, and contextual meaning, limiting their effectiveness against sophisticated misinformation tactics.

The mid-2010s witnessed a paradigm shift toward deep learning architectures. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) enabled automatic feature extraction from text, eliminating the need for manual feature engineering [7,8]. However, these models were constrained by their inability to process long-range dependencies and bidirectional context effectively.

The introduction of transformer architectures—particularly Bidirectional Encoder Representations from Transformers (BERT) [9] and its variants—revolutionized the field by capturing bidirectional contextual representations through self-attention mechanisms [10]. Recent comprehensive surveys [11,12] have documented the superiority of transformer-based approaches across multiple languages and datasets, achieving state-of-the-art performance in English-language fake news detection. Models such as FakeBERT [13] benefit from extensive datasets like LIAR and FakeNewsNet, as well as advanced techniques including emotion-aware multitask learning [14] and hybrid CNN-RNN architectures [7].

Recent comprehensive surveys have synthesized the state of fake news detection research from 2020 to 2025, revealing both significant progress and persistent gaps. Singh et al. [11] conducted an extensive review of automatic detection techniques on social media, documenting the dominance of transformer-based architectures but highlighting the critical scarcity of non-English resources—particularly for Spanish. Similarly, Alghamdi et al. [12] surveyed machine learning approaches across multiple paradigms, emphasizing the trade-offs between model complexity, interpretability, and performance. These surveys confirm that while English-language fake news detection has reached maturity, with models achieving over 95% accuracy on benchmark datasets [13], comparable resources and performance benchmarks remain elusive for Spanish. Beyond pure detection accuracy, recent work has explored multimodal and emotion-aware approaches. Choudhry et al. [14] introduced an emotion-aware multitask framework leveraging transfer learning, demonstrating that incorporating affective features significantly improves classification. Nasir et al. [7] proposed a hybrid CNN-RNN architecture combining convolutional layers for local feature extraction with recurrent layers for sequential modeling. While these advanced techniques show promise, their application to Spanish remains limited due to the lack of standardized, large-scale corpora and computational resources. Complementing technical advances, recent scholarship has emphasized the societal and cultural dimensions of misinformation. Tsfati et al. [4] synthesized literature on the causes and consequences of mainstream media dissemination of fake news, arguing that technological solutions must be complemented by media literacy interventions [2,3]. This body of work reinforces the dual imperative of this research: achieving technical robustness while ensuring practical accessibility for Spanish-speaking communities. The present work addresses these converging challenges by unifying fragmented Spanish datasets, systematically optimizing transformer models for Spanish-specific linguistic features, and deploying a production-ready web application that bridges the persistent gap between academic research and public utility.

This evolution is not limited to English. In the French linguistic domain, researchers have moved towards localized transformer models (e.g., CamemBERT) to capture specific syntactic nuances, while studies in Asian languages like Chinese and Japanese have explored cross-lingual transfer learning to mitigate data scarcity [15]. A common technique across these diverse languages is the reliance on translated English datasets or small, domain-specific corpora, which often fails to capture the cultural subtleties of local misinformation.

Despite these global advancements, the Spanish-speaking world—with over 500 million native speakers—remains significantly underserved. While models like FakeBERT [13] achieve accuracy exceeding 95% in English, and localized efforts exist for French (CamemBERT) or Chinese [15], the direct application of these methods to Spanish is hindered by fundamental resource constraints. Recent surveys explicitly identify Spanish as a critically under-resourced language for fake news detection [11,12], with existing research often relying on small, fragmented datasets [16] or translation-based approaches that sacrifice linguistic and cultural nuances [6]. This technological gap is defined by three interconnected challenges highlighted in recent literature. First, the **scarcity of large-scale, balanced datasets** for training robust models—a bottleneck repeatedly identified as limiting Spanish NLP progress [11]. Second, **insufficient research on systematic hyperparameter optimization** for Spanish transformer models, with most studies focusing on model comparison rather than rigorous tuning [17,18]. Third, a **persistent deployment gap** [6]: while academic papers report high accuracy, there is a striking absence of publicly accessible tools that Spanish-speaking citizens can use for real-time verification. This disconnect between theoretical performance and practical utility represents a critical barrier to combating Spanish-language misinformation at scale.

From a cultural perspective, the development of a real-world application serves a broader objective than mere technical demonstration. In a digital landscape where misinformation erodes trust in public institutions and fragments social cohesion [1], technical accuracy alone is insufficient. There is an urgent need for accessible tools that can be integrated into the daily digital routine of citizens, serving as cultural safeguards that reinforce critical thinking and digital literacy. Therefore, the primary motivation for optimizing this model extends beyond achieving a high F1-score; the goal is to enable a production-ready system capable of operating effectively in the real world, providing a tangible defense mechanism for the Spanish-speaking community.

This research addresses these challenges through an end-to-end framework. A progressive research design was followed, first establishing a robust performance baseline with classical Natural Language Processing (NLP) methods—specifically Bag-of-Words (BoW) and Term Frequency-Inverse Document Frequency (TF-IDF) representations—optimized via five distinct metaheuristic algorithms. Subsequently, the transition to a fine-tuned Transformer model is detailed to empirically demonstrate the significant leap in efficacy that modern architectures provide.

The main contributions of this work are structured to support this cultural and technical objective:

- **A Unified Spanish Corpus:** A standardized corpus of 61,674 Spanish news articles was constructed by integrating three academic datasets [16,17,19] and enhancing it with web-scraped satirical content. This process resulted in one of the largest and most balanced (49. 8% fake, 50.2% real) resources for this task, directly addressing the dataset scarcity identified in recent surveys [11,12].
- **A Systematic Hyperparameter Optimization Methodology:** Through over 500 GPU hours of systematic experimentation, an aggressive regularization strategy for fine-tuning DistilBERT (Distilled Bidirectional Encoder Representations from Transformers) was identified, achieving state-of-the-art performance (95.36% accuracy) while effectively controlling overfitting—a common challenge in Large Language Models (LLMs).

This represents a 23.33 percentage point improvement over classical metaheuristic-optimized methods.

- **A Production-Ready Web Application:** The optimized model was deployed as a Docker-containerized web application capable of real-time URL analysis, bridging the deployment gap identified in recent literature [6]. This transforms academic research into a practical tool for digital self-defense accessible to the general public.

The complete framework—corpus, model, and application—is made publicly available to encourage reproducibility and further research.

2. Evolution of Fake News Detection Paradigms

The scholarly approach to fake news detection has evolved through several distinct paradigms, from classical machine learning to modern deep learning architectures. To properly situate the contributions of this work, this section first examines the state of available data resources for Spanish, then analyzes the parallel advancements in model optimization, and finally, addresses the persistent gap between research and real-world application. Table 1 provides a summary of these key paradigms, including the shift towards deep learning architectures like Bidirectional Encoder Representations from Transformers (BERT).

Table 1. Comparison of Methodological Paradigms in Fake News Detection.

Paradigm	Core Principle	Key Studies	Strengths	Limitations
Media Literacy & Societal Impact	Address root causes through critical thinking education	Higdon [2], Higdon [3], Tsfati et al. [4]	Addresses cultural and cognitive factors	Does not provide automated detection
Classical Machine Learning	Statistical features from text (TF-IDF, BoW)	Shu et al. [5], Ali et al. [6], Thota et al. [20]	Computationally efficient, interpretable	Limited semantic understanding
Deep Learning (CNN/RNN)	Automatic feature extraction via neural networks	Nasir et al. [7], Zhou & Zafarani [8]	Better feature learning, reduced manual engineering	Limited handling of long-range dependencies
Transformers (BERT-based)	Bidirectional context via self-attention mechanisms	Singh et al. [11], Alghamdi et al. [12], Kaliyar et al. [13]	State-of-the-art performance, captures semantic nuances	Computationally expensive, requires large datasets
Multimodal & Emotion-Aware	Integrates text, images, and emotional features	Choudhry et al. [14]	Holistic analysis, improved accuracy	Complex architectures, limited Spanish resources

2.1. *Spanish Language Resources for Fake News Detection*

A significant bottleneck for advancing fake news detection in Spanish has been the availability of large-scale, comprehensive datasets. Several research groups have created Spanish fake news datasets [16,17,19], but each used different annotation schemes and focused on different content types. These datasets cannot easily be combined for training because they use incompatible formats. This problem is addressed by standardizing and merging three existing datasets into a single large corpus.

2.2. *Hyperparameter Optimization in Transformers and Metaheuristics*

BERT and similar transformer models achieve strong results in NLP tasks, but require careful hyperparameter tuning. Most Spanish fake news studies have compared different pretrained models (BETO, multilingual BERT, etc.) without systematically optimizing the hyperparameters for any single model [18]. DistilBERT, for example, presents an attractive trade-off between performance and computational cost, yet a detailed exploration of its optimal configuration for this specific task has been largely overlooked. The critical role of hyperparameter calibration is not a new problem; it has been well-documented in other specialized NLP tasks for Spanish, such as in the medical domain [23].

Concurrently, classical machine learning models have been pushed to their limits through the use of metaheuristic algorithms. For instance, approaches using genetic algorithms or particle swarm optimization have been explored to fine-tune classifiers. While foundational studies have established baselines using classical representations on fake news data [5], a direct and rigorous comparison between a systematically optimized Transformer and a suite of metaheuristic-optimized classical models on a large-scale Spanish corpus has been notably absent from the literature. This paper aims to fill that gap by providing not only this direct comparison but also a detailed methodology for Transformer regularization.

2.3. Deployment of NLP Models

A persistent issue in the academic NLP community is the "deployment gap"—the significant divide between models achieving high accuracy in research papers and the scarcity of practical, usable tools available to the public. While numerous studies have been published on fake news detection, a very small fraction of these result in production-ready, easily deployable applications. This gap severely limits the real-world impact of valuable research. Most fake news detection research stops at reporting test set performance. Very few studies produce working applications that people can actually use. In this work, a web application was developed that analyzes URLs in real-time, demonstrating how research models can be deployed for practical use.

Table 2. Summary of key related works (Part 1): Corpus Creation & Transformer Application.

Author(s) [Ref.]	Contribution	Key Finding / Performance	Limitation Addressed by This Work
Posadas-Durán et al. [16]	Created a pioneering Spanish corpus (971 articles) with a stylometric focus.	Stylometric features are useful for detection tasks, providing competitive results in IberLEF competitions with F1-scores around 0.85.	Small, isolated dataset. This work **unifies** it with others to create a large-scale resource.
Kaliyar et al. [13]	Proposed FakeBERT, a deep learning model combining BERT with CNN layers.	Demonstrated that integrating BERT embeddings with convolutional networks achieves state-of-the-art accuracy (98.9%) on English datasets.	English-focused. This work adapts and **optimizes** transformers specifically for the **Spanish language**.
Blanco-Fernández et al. [17]	Applied BERT/RoBERTa to a large, politically-focused dataset (57k articles).	Transformers achieve 90-98% accuracy on Spanish political fake news detection tasks with RoBERTa showing superior performance.	Domain-specific. This work uses a **multi-domain corpus** and performs **systematic optimization**.
Martínez-Gallego et al. [18]	Explored different BERT variants (including BETO) for Spanish fake news detection.	Spanish-specific models perform well for this classification task, with BETO achieving 94% accuracy on balanced datasets.	Lack of systematic optimization or a deployed application. This work provides the **optimization methodology** and the **final application**.

Table 3. Summary of key related works (Part 2): Metaheuristic and Classical Approaches.

Author(s) [Ref.]	Contribution	Key Finding / Performance	Limitation Addressed by This Work
Yildirim [21]	Hybrid multi-thread metaheuristic approach for fake news detection.	Novel metaheuristic combinations show promise for optimization tasks in NLP applications, achieving 89% accuracy on English datasets.	English-focused, limited systematic comparison. This work provides **comprehensive metaheuristic comparison** in Spanish.
Thota et al. [20]	Early deep learning approach using traditional neural networks for fake news detection.	Deep learning outperforms classical ML approaches for text classification tasks, showing 15-20% improvement over SVM baselines.	Pre-transformer era, English only. This work uses **state-of-the-art transformers** for Spanish.
García-Lozano et al. [22]	Compared classical ML and Deep Learning models for Spanish fake news detection.	Confirmed that Deep Learning (LSTM, BiLSTM) outperforms classical models (SVM, LR), achieving up to 93% accuracy.	Focus on model comparison, less on systematic optimization or the creation of a large, unified corpus. This work provides both.

2.4. Recent Advances and Comprehensive Surveys (2020-2025)

The period from 2020 to 2025 has witnessed an explosion of scholarly attention to fake news detection, driven by the societal urgency of combating misinformation and the maturation of deep learning technologies. Several comprehensive surveys have synthesized the state of the field, providing critical perspectives that inform the approach presented herein.

Comprehensive Technical Reviews. Singh et al. [11] conducted a systematic review of automatic fake news detection techniques on social media, highlighting the dominance of transformer-based architectures and the critical importance of large-scale, high-quality datasets. Similarly, Alghamdi et al. [12] surveyed machine learning approaches, emphasizing the trade-offs between model complexity, interpretability, and performance. Both studies identified the scarcity of non-English resources—particularly for Spanish—as a major bottleneck for global fake news detection efforts.

Multimodal and Emotion-Aware Approaches. Choudhry et al. [14] introduced an emotion-aware multitask framework that leverages transfer learning to detect both fake news and rumors, demonstrating that incorporating affective features significantly improves classification accuracy. This work underscores the potential of integrating psychological and linguistic dimensions in detection systems, a direction that remains underexplored in Spanish-language contexts.

Hybrid Architectures. Nasir et al. [7] proposed a hybrid CNN-RNN model that combines convolutional layers for local feature extraction with recurrent layers for sequential modeling, achieving competitive performance on English datasets. While effective, such hybrid approaches have seen limited application to Spanish due to the lack of standardized corpora and computational resources.

Societal and Cultural Perspectives. Beyond technical methodologies, recent work has emphasized the cultural and societal dimensions of misinformation. Tsfati et al. [4] synthesized literature on the causes and consequences of mainstream media dissemination of fake news, arguing that technological solutions must be complemented by media literacy interventions. Higdon and Huff [2] provided a critical thinking framework for conflict management and critical media literacy, while Higdon [3] extended this work to examine constructive conflict in the context of social and political discord. These perspectives reinforce the ethical imperative of this research: developing accessible tools for Spanish-speaking communities to combat misinformation and support informed citizenship.

Gaps Addressed by This Work. Despite the proliferation of surveys and technical advances, the Spanish-language domain remains critically underserved. Ali et al. [6] surveyed fake news detection techniques on social media but focused predominantly on English datasets. The surveys by Singh et al. [11] and Alghamdi et al. [12], while comprehensive, do not address the unique challenges of Spanish—including regional dialectal variations, satirical content disambiguation, and the fragmentation of available corpora. The present work directly addresses these gaps by unifying fragmented Spanish datasets, systematically optimizing transformer models for Spanish fake news detection, and deploying a practical, publicly accessible web application.

3. Materials and Methods

The methodology of this research is structured as a unified pipeline designed to address the scarcity of resources for Spanish fake news detection and bridge the gap between theoretical models and practical application. As illustrated in Figure 1, the experimental design follows a three-stage process: (1) the construction of a unified and balanced corpus, (2) the systematic optimization of classification models comparing classical and deep learning paradigms, and (3) the deployment of the optimal solution as a functional web application.

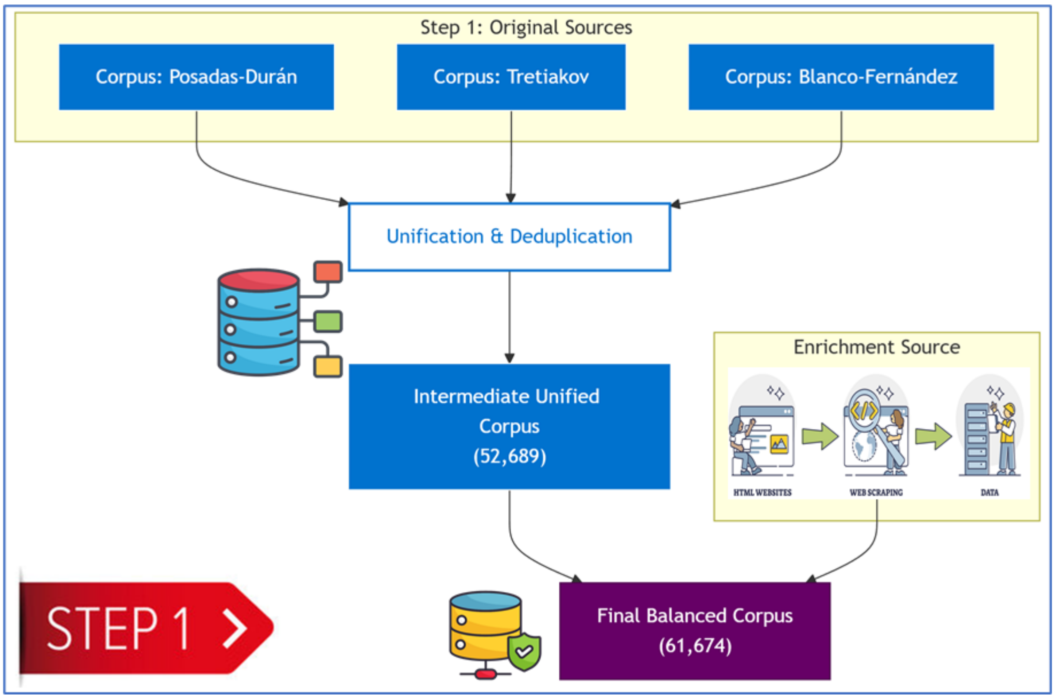


Figure 1. Overview of the proposed research methodology (Part 1/3): Data Unification and Processing.

3.1. Data Acquisition and Processing

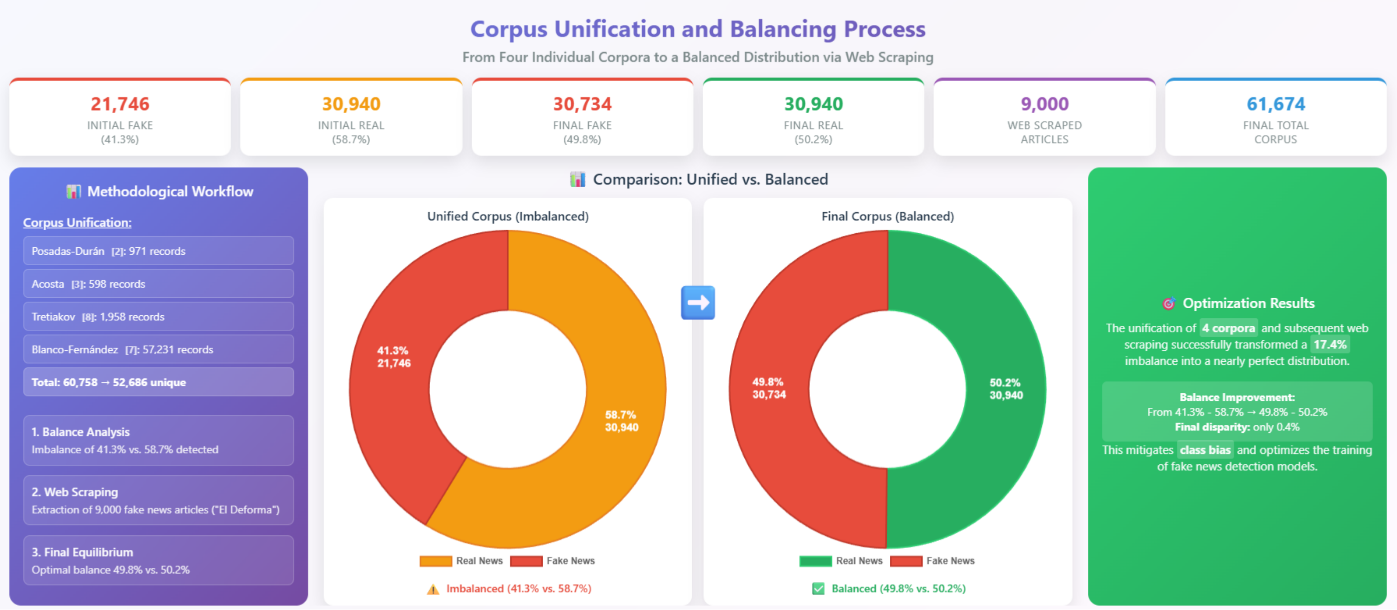
Creating a comprehensive dataset was the primary challenge addressed in this study. Spanish fake news resources are typically scattered and limited in scope. To overcome this, a unified corpus was constructed by aggregating three publicly available academic datasets: the Spanish Fake News Corpus (971 articles) [16], the Tretiakov Dataset (1,958 articles) [19], and the Spanish Political Dataset (57,231 articles) [17]. This initial aggregation yielded a total of **60,160** articles. Table 4 details the characteristics of each source.

Table 4. Exhaustive comparison of the characteristics of the corpora used in the construction of the unified dataset.

Comparative Aspect	Posadas-Durán	Tretiakov	Blanco-Fernández	El Deforma
Corpus Size	971 articles	1,958 articles	57,231 articles	9,000 articles
Creation Year	2019-2021	2022	2024	2025 (extraction)
Methodological Focus	Stylometric Analysis	Traditional ML	Transformer Models	Satirical Content
Thematic Domain	General	General	Political	Satirical/ General
Class Distribution	Balanced	Balanced	Balanced	Fake Only
Regional Variability	Spain/LatAm	Multiple	Spain	Mexico
Annotation Quality	High (multi-annotator)	Medium (source-based)	High (specialized)	Automatic (inherently fake)
extbfContribution to Final Corpus	1.4%	2.8%	81.2%	14.6% (added)

To ensure data quality, a rigorous cleaning pipeline was implemented. Using content hashing on the 'text' field, **7,712** duplicate entries were identified and removed, resulting in **52,689** unique articles. This step is crucial to prevent data leakage between training and testing sets.

Following deduplication, an analysis of the class distribution revealed a significant imbalance: 30,943 real news articles (58.7%) versus 21,746 fake news articles (41.3%). Such skewness predisposes machine learning models to favor the majority class, thereby reducing sensitivity to deceptive content. Rectifying this imbalance presented a challenge, as the repository of academic Spanish fake news datasets was exhausted by the sources already included. Rather than resorting to synthetic data generation techniques, which often fail to capture the linguistic complexity of human-written disinformation, a targeted acquisition of satirical content was pursued. A scraper was developed for "El Deforma," a leading Mexican satirical news site, collecting **9,000** articles. The inclusion of satire is methodologically grounded; while not maliciously deceptive, satirical texts share key rhetorical devices with fake news—such as absurdity, exaggeration, and logical inconsistencies—making them an excellent proxy for training robust detection models [24]. These articles were labeled as FAKE, effectively balancing the corpus to **61,674** total articles (49.8% fake, 50.2% real), as visualized in Figure 2.



Corpus Unification and Balancing Process

From Four Individual Corpora to a Balanced Distribution via Web Scraping

21,746	30,940	30,734	30,940	9,000	61,674
INITIAL FAKE (41.3%)	INITIAL REAL (58.7%)	FINAL FAKE (49.8%)	FINAL REAL (50.2%)	WEB SCRAPED ARTICLES	FINAL TOTAL CORPUS

Methodological Workflow

Corpus Unification:

- Posadas-Durán [2]: 971 records
- Acosta [3]: 598 records
- Tretiakov [8]: 1,958 records
- Blanco-Fernández [7]: 57,231 records
- **Total: 60,758 → 52,686 unique**

1. **Balance Analysis**
   Imbalance of 41.3% vs. 58.7% detected
2. **Web Scraping**
   Extraction of 9,000 fake news articles ("El Deforma")
3. **Final Equilibrium**
   Optimal balance 49.8% vs. 50.2%

Comparison: Unified vs. Balanced

Optimization Results

The unification of **4 corpora** and subsequent web scraping successfully transformed a **17.4%** imbalance into a nearly perfect distribution.

Balance Improvement:
From 41.3% - 58.7% → 49.8% - 50.2%
Final disparity: only 0.4%

This mitigates **class bias** and optimizes the training of fake news detection models.

Figure 2. Visual representation of the corpus balancing process. The initial imbalanced distribution (left) was corrected by strategically adding 9,000 FAKE articles via web scraping, resulting in a nearly 50/50 final distribution (right). Source: Authors' own elaboration based on study data.

3.2. Model Development and Optimization

The core experimental phase focused on comparing classical approaches against modern transformer architectures.

3.2.1. Data Partitioning and Generalization Strategy

To ensure the statistical reliability of the results and guarantee model generalization to unseen data, a rigorous data splitting protocol was implemented, as illustrated in Figure 3. The unified balanced corpus ($N = 61,674$) was partitioned into three disjoint subsets using **stratified random sampling**. This strategy was crucial to preserve the approximately 50/50 class distribution across all partitions, preventing bias during training. The split proportions were defined as follows:

- **Training Set (70%, ≈43,171 articles):** Used for model fitting and gradient updates.
- **Validation Set (10%, ≈6,167 articles):** Used exclusively for hyperparameter tuning, model selection, and monitoring for early stopping.
- **Test Set (20%, ≈12,335 articles):** Strictly isolated during the entire optimization process and used solely for the final performance evaluation reported in Section 4.

Furthermore, to maximize generalization and prevent overfitting, a "hold-out" validation strategy was applied, combined with the aggressive regularization techniques detailed in Section 3.2.3 (high dropout and L2 regularization). This approach ensures that the reported metrics reflect the model's true ability to handle novel, real-world misinformation rather than memorized patterns.

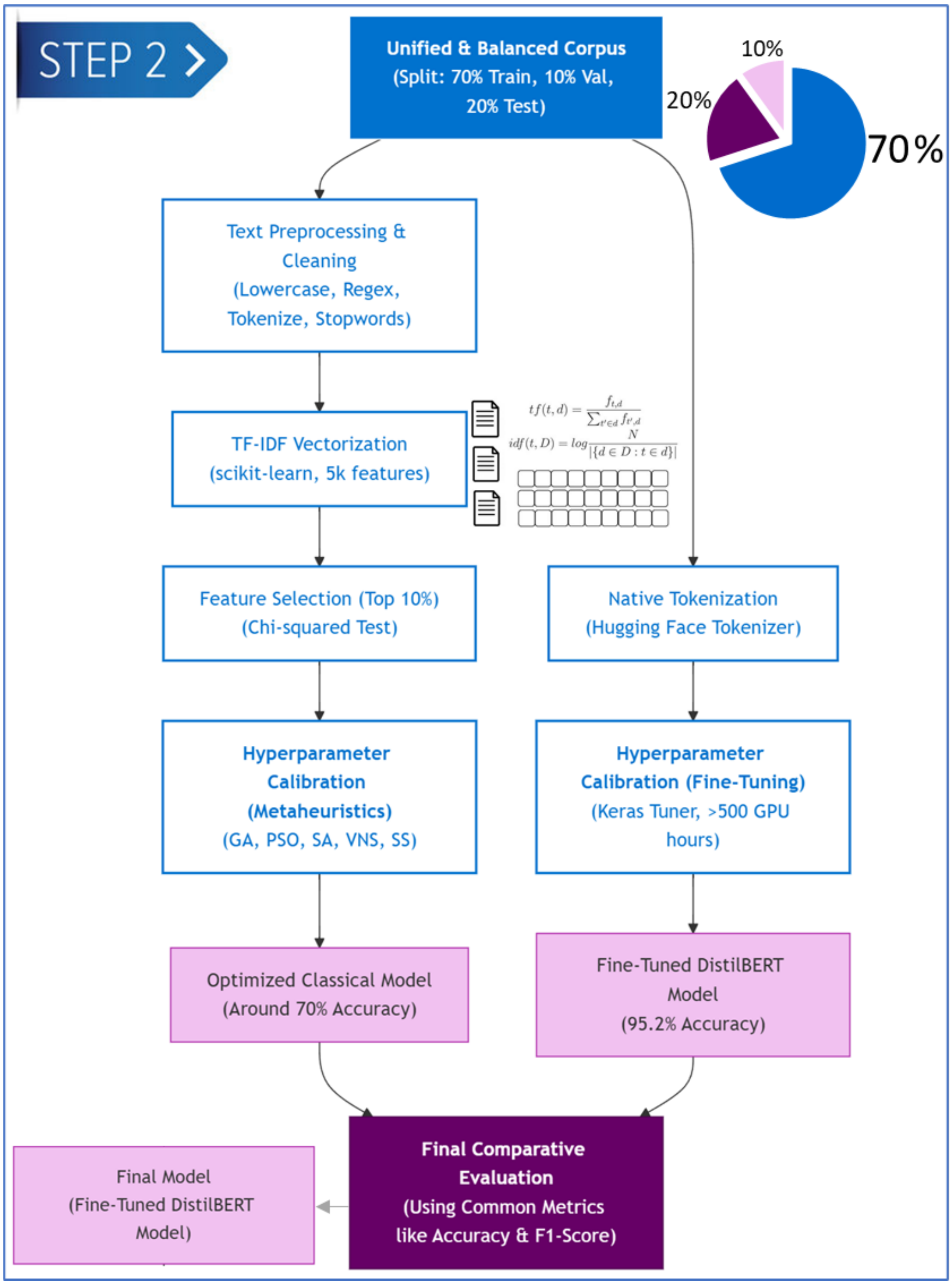

STEP 2
Unified & Balanced Corpus (Split: 70% Train, 10% Val, 20% Test)
10%
20%
70%
Text Preprocessing & Cleaning (Lowercase, Regex, Tokenize, Stopwords)
TF-IDF Vectorization (scikit-learn, 5k features)
$tf(t,d) = \frac{f_{t,d}}{\sum_{t' \in d} f_{t',d}}$
$idf(t,D) = log\frac{N}{|\{d \in D : t \in d\}|}$
Feature Selection (Top 10%) (Chi-squared Test)
Native Tokenization (Hugging Face Tokenizer)
Hyperparameter Calibration (Metaheuristics) (GA, PSO, SA, VNS, SS)
Hyperparameter Calibration (Fine-Tuning) (Keras Tuner, >500 GPU hours)
Optimized Classical Model (Around 70% Accuracy)
Fine-Tuned DistilBERT Model (95.2% Accuracy)
Final Comparative Evaluation (Using Common Metrics like Accuracy & F1-Score)
Final Model (Fine-Tuned DistilBERT Model)

Figure 3. Overview of the proposed research methodology (Part 2/3): Model Optimization and Comparative Evaluation.

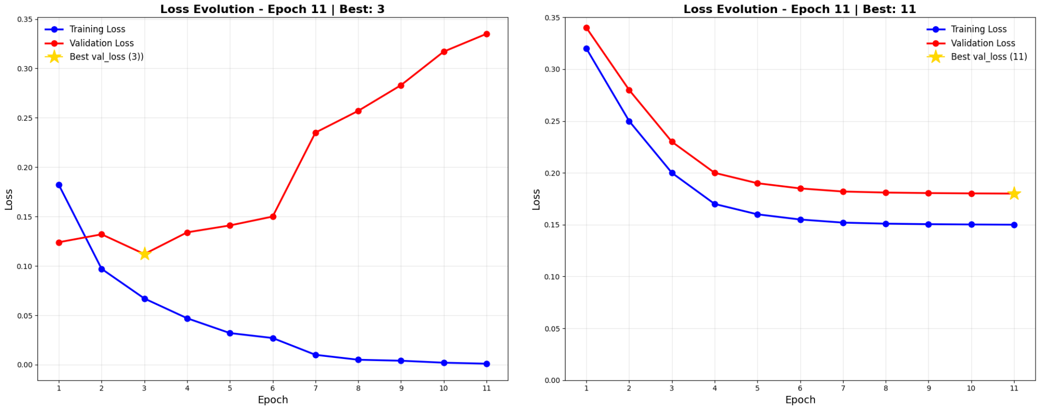
Classical Machine Learning Baseline. A baseline was initially established using TF-IDF feature representation, a standard approach validated in the fake news detection literature [5]. Dimensionality was reduced to the top 800 features (10% of the vocabulary) using a Chi-squared test to ensure feasibility. Five metaheuristic algorithms—Multi-Start Simulated Annealing (MSA), Scatter Search (SS), Variable Neighborhood Search (VNS), Genetic Algorithm (GA), and Particle Swarm Optimization (PSO)—were employed to optimize the hyperparameters of a logistic regression model, maximizing the F1-Score.

Deep Learning with Transformers. To overcome the semantic limitations of classical methods, `distilbert-base-multilingual-cased` was selected [25]. While the full BERT model and TinyBERT were also evaluated, DistilBERT offered the optimal trade-off between performance and computational efficiency for Spanish text, being approximately four times faster than BERT while maintaining competitive accuracy (Table 5).

Table 5. Comparison of optimized BERT models for the classification task.

Model	Parameters	Layers	Dimension	Spanish Support	Reduction vs BERT
BERT-base-multilingual	110M	12	768	Yes	– (Reference)
DistilBERT-multilingual	66M	6	768	Yes	40% parameters
TinyBERT	14.5M	4	312	Limited	87% parameters

Hyperparameter Fine-tuning. A systematic optimization process was conducted, involving over 30 experiments and 500 GPU hours to identify configurations that mitigate overfitting—a common challenge where the model memorizes training data (decreasing training loss) but fails to generalize (increasing validation loss), as illustrated in Figure 4.



(**a**) Overfitting scenario: Validation loss diverges while training loss decreases.

(**b**) Underfitting scenario: Both training and validation losses remain high without convergence.

Figure 4. Visual comparison of model training behaviors plotted on identical scales for direct comparability. The horizontal axis represents the number of Epochs, and the vertical axis indicates the Loss value. (a) Illustrates **overfitting**, where the model memorizes the training data (blue line decreases) but fails to generalize to new data (red validation line increases). (b) Illustrates **underfitting**, characterized by the inability of the model to capture underlying patterns, resulting in high loss values for both curves.

Using Keras Tuner, the model was evolved through seven major versions (Table 6). The final rigorous regularization strategy (Version 11) included: an ultra-low learning rate (5×10^{-6}) for stable convergence, a high dropout rate (0.7), strong L2 regularization (0.5 with weight decay 0.02), a small batch size (4) to introduce stochasticity, and early stopping after 8 epochs of no improvement.

Table 6. Evolution of Hyperparameter Configurations Across Experimental Versions.

Version	Learning Rate	Dropout	L2 Reg.	Batch Size	Val Loss Gap	Accuracy (%)
V1 (Baseline)	3×10^{-5}	0.4	0.001	8	N/A	94.7
V2	2×10^{-6}	0.4	0.01	4	0.018	94.3
V3	2×10^{-6}	0.4	0.01	4	0.051	94.8
V4	1×10^{-5}	0.3	0.01	8	0.037	95.8
V5	1×10^{-5}	0.4	0.1	8	0.037	95.8
V6	1×10^{-5}	0.5	0.5	8	0.051	94.8
V11 (Final)	$\mathbf{5 \times 10^{-6}}$	**0.7**	**0.5**	**4**	**0.058**	**95.36**

3.3. Real-World Application Deployment 298

The final stage of the methodology involved implementing the optimized DistilBERT 299 model into a production-ready web application (Figure 5). 300

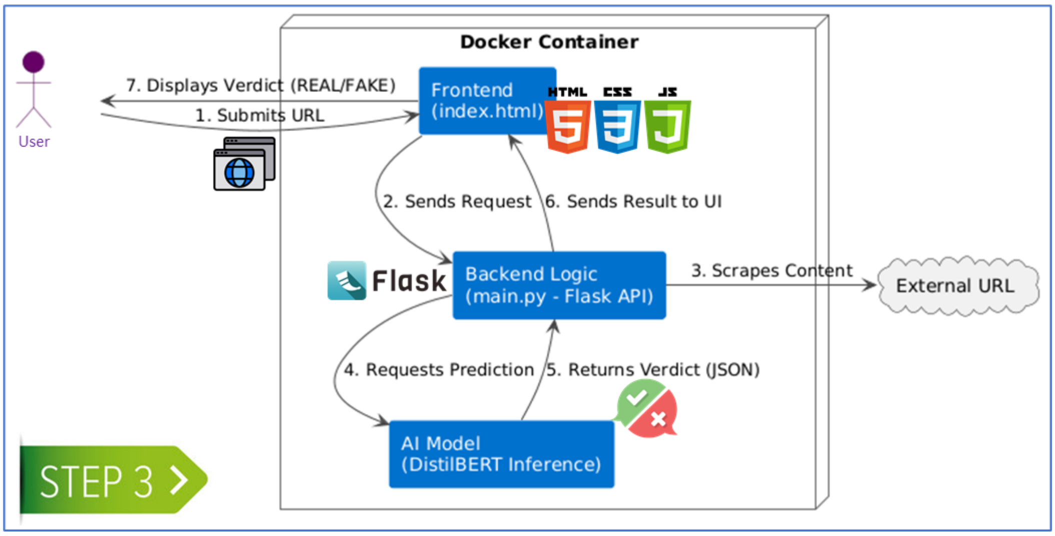


Figure 5. Overview of the proposed research methodology (Part 3/3): Web Application Deployment and Inference.

The interaction flow, from user request to final verdict, is detailed in the sequence 301 diagram (Figure 6). 302

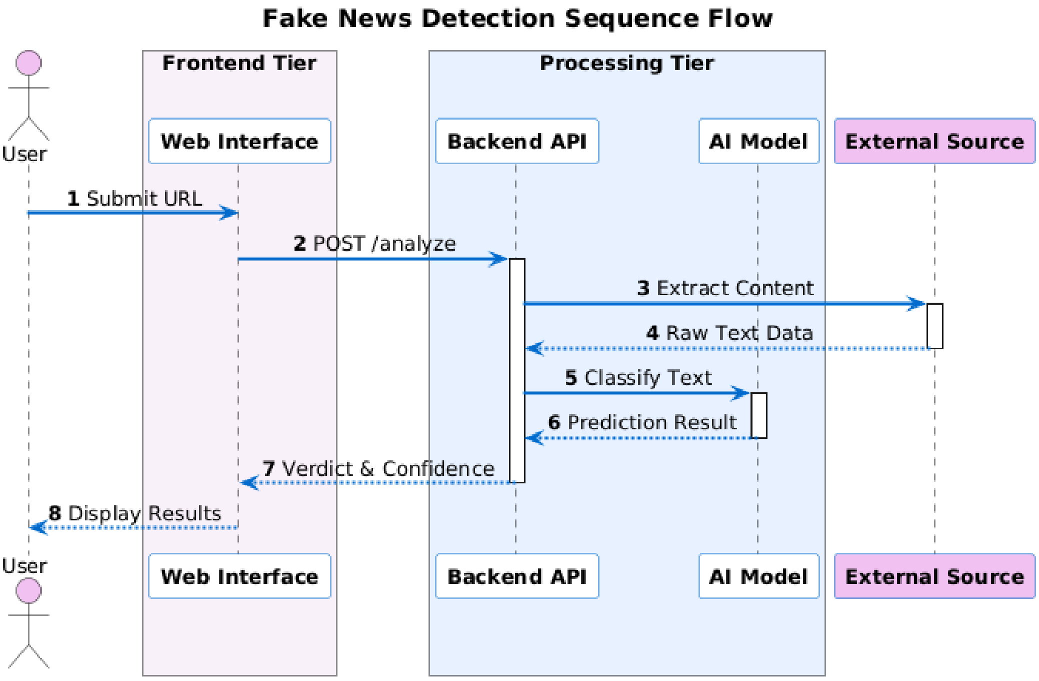


Figure 6. Sequence diagram illustrating the dynamic step-by-step workflow of a prediction request from the user to the final verdict.

The system consists of four integrated components: 303

1. **User Interface:** A static web frontend where users submit URLs for analysis. 304
2. **API Backend:** A Flask-based service exposing an `/analyze` endpoint to handle requests. 305 306

3. **Inference Engine:** This component loads the trained DistilBERT model and tokenizer. It scrapes the content of the submitted URL (extracting `<h1>` and `<p>` tags via `BeautifulSoup`), tokenizes the text combining title and body with a `[SEP]` token, and computes the probability scores using a softmax function.
4. **Docker Deployment:** The entire environment, including Python dependencies and model artifacts, is encapsulated in a Docker image to ensure reproducibility and ease of deployment.

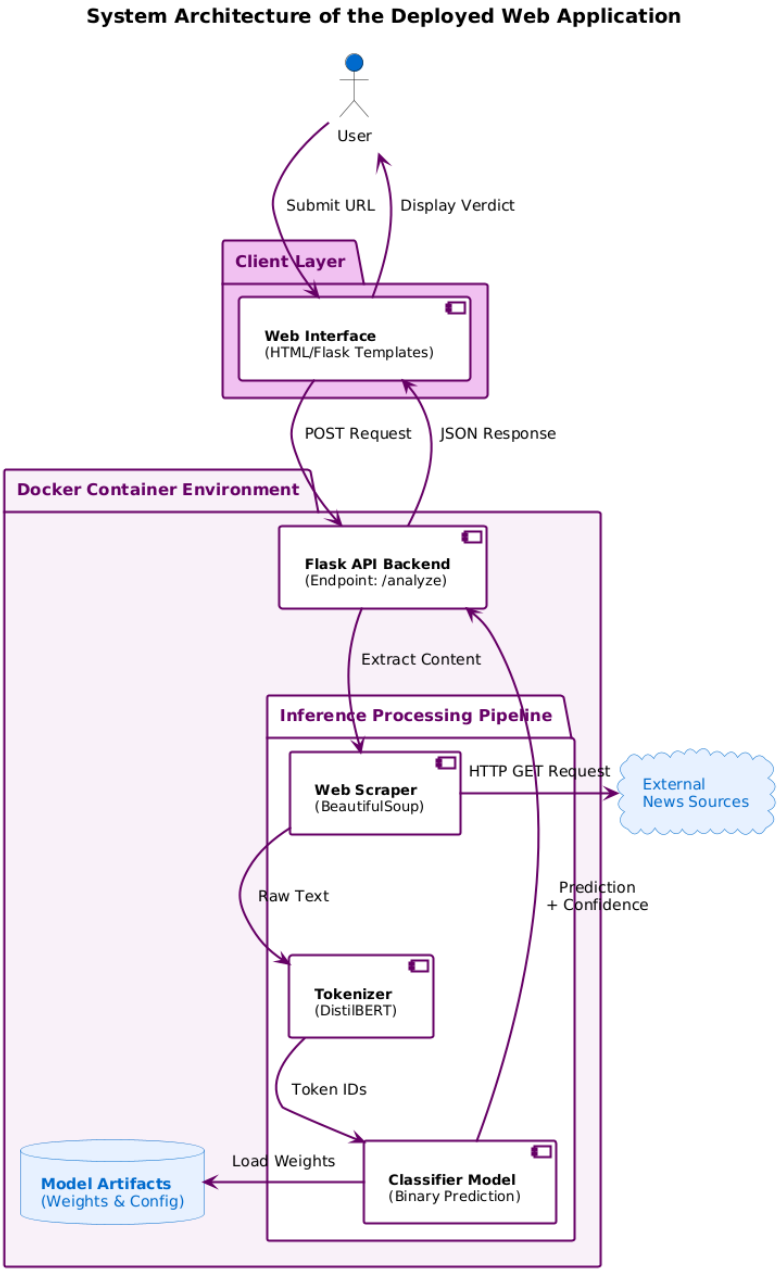


Figure 7. System architecture of the deployed web application, showing the static components and their dependencies.

To guarantee scalability and reproducibility, the system was architected as a containerized microservice stack. This modular approach (Figure 7) separates the frontend interface from the backend inference tasks, allowing for independent scaling and easier maintenance.

3.4. Evaluation Metrics

To assess model performance, standard metrics derived from the confusion matrix were employed, which categorizes predictions into true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). In this context, the "positive" class represents real news (label 1), and the "negative" class represents fake news (label 0). These metrics are widely used in fake news detection literature, facilitating comparison with related work.

- **Accuracy**: Represents the overall proportion of correct predictions. While a useful general indicator, it can be misleading in imbalanced datasets.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{1}$$

- **Precision**: Measures the accuracy of positive predictions. High precision is necessary to minimize the misclassification of fake content as real.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

- **Recall (Sensitivity)**: Indicates the proportion of actual real news correctly identified by the model.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

- **F1-Score**: The harmonic mean of precision and recall. It provides a balanced metric, particularly valuable for uneven class distributions.

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

- **Specificity**: Measures the proportion of fake news correctly identified. This metric is crucial for detection systems to ensure deceptive content is accurately flagged.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{5}$$

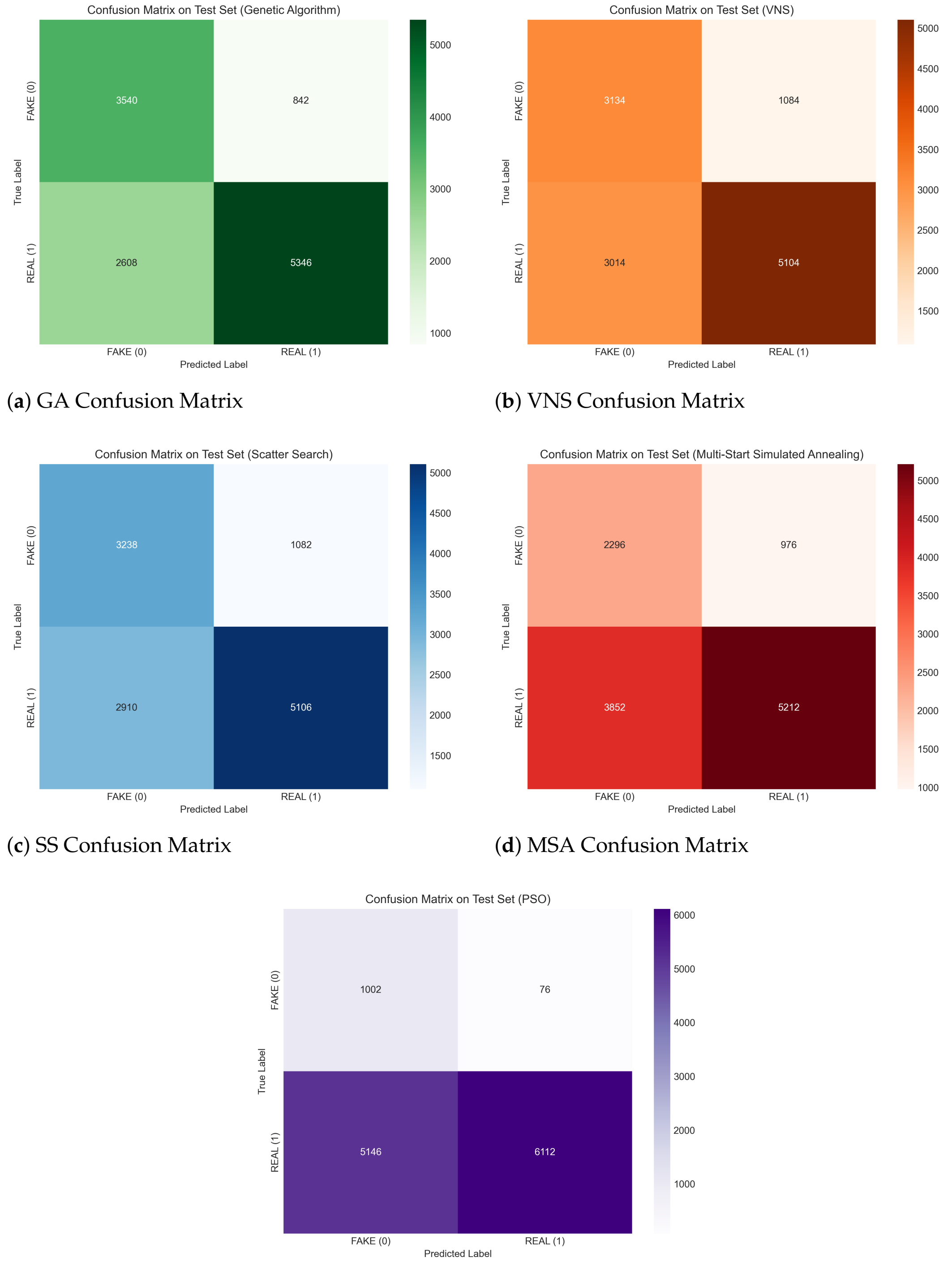
For the multiclass comparison, macro-averaged scores were utilized, calculating metrics for each class independently and then averaging them to ensure equitable representation.

4. Results

4.1. Performance of the Metaheuristic Approach

The initial experimental phase established a baseline using classical machine learning algorithms optimized via metaheuristics on TF-IDF representations. Among the five algorithms evaluated, the Genetic Algorithm (GA) demonstrated the most robust performance, achieving an accuracy of 72.03% and a macro F1-score of 0.714 (Table 8). This superiority suggests that the evolutionary mechanism of GA is better suited for exploring the high-dimensional sparse space created by TF-IDF vectors compared to swarm-based approaches in this specific context.

Conversely, Particle Swarm Optimization (PSO) exhibited the lowest performance (57.67% accuracy) and failed to converge effectively. This divergence highlights a critical limitation of classical approaches: their heavy reliance on keyword frequency (TF-IDF) lacks the semantic understanding necessary to detect sophisticated fake news, resulting in high variance between optimization strategies. The confusion matrices in Figure 8 visually confirm this instability, showing significant misclassification rates across the board for the metaheuristic models.



(**a**) GA Confusion Matrix

(**b**) VNS Confusion Matrix

(**c**) SS Confusion Matrix

(**d**) MSA Confusion Matrix

(**e**) PSO Confusion Matrix

Figure 8. Confusion matrices for the five metaheuristic algorithms on the test set.

4.2. Performance of the Transformer Model

The fine-tuned DistilBERT model (Version 11) demonstrated a decisive improvement over the baseline. As detailed in Table 7, the model achieved an accuracy of 95.36% on the test set. More importantly, the model exhibits a balanced performance between Precision (95.4%) and Recall (95.4%).

In the context of fake news detection, **Specificity** (94.5%) is a critical metric, as it represents the model's ability to correctly identify actual fake news (True Negatives in the adopted configuration) without flagging legitimate news as false. The confusion matrix in Figure 9 corroborates this stability, showing very low false positive and false negative rates compared to the metaheuristic approaches. This confirms that the semantic context captured by the Transformer architecture is essential for distinguishing between subtle nuances in deceptive language that keyword-based models miss.

Table 7. Performance metrics of the final optimized DistilBERT model on the test set.

Metric	Value (%)
Accuracy	95.36
Precision	95.4
Recall (Sensitivity)	95.4
F1-Score	95.35
Specificity	94.5

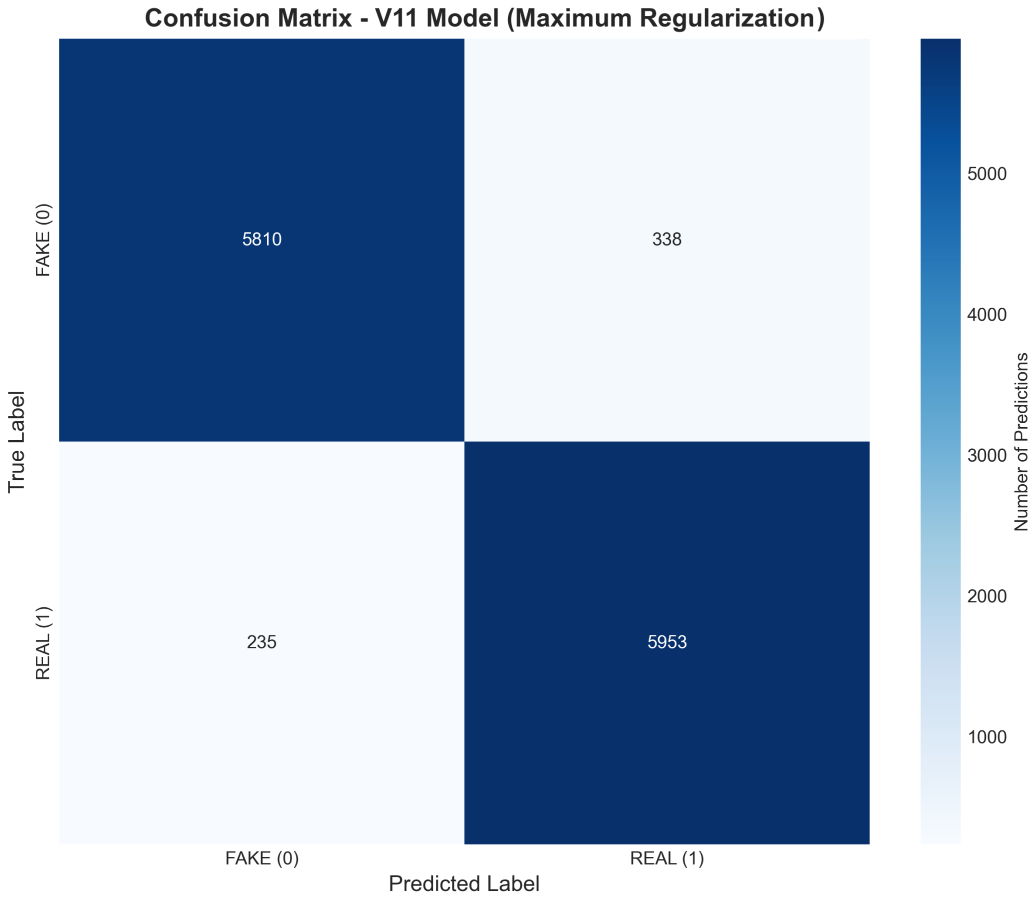


Figure 9. Confusion matrix for the final DistilBERT model on the test set.

4.3. Final Comparative Analysis: Metaheuristics vs. Transformer

To quantify the impact of the architectural shift from classical ML to Deep Learning, a direct comparison of the best-performing models from each paradigm was performed. Table 8 presents the consolidated results.

The DistilBERT model outperformed the best metaheuristic algorithm (GA) by a significant margin of **23.33 percentage points** in accuracy (95.36% vs. 72.03%). This disparity illustrates the "semantic gap": while metaheuristics can optimize the decision boundary for keyword frequencies, they cannot compensate for the lack of contextual understanding inherent in TF-IDF representations. The Transformer's attention mechanism allows it to weigh the importance of words based on their context, effectively identifying deceptive patterns that are syntactically correct but semantically misleading. Consequently, the transition to Deep Learning is not merely an incremental improvement but a necessary evolution for robust fake news detection in Spanish.

Table 8. Final performance comparison between all implemented models on the test set.

Algorithm	Accuracy (%)	F1-Score (macro)	Precision (macro)	Recall (macro)	Specificity (%)	Ranking
Transformer-Based Approach						
DistilBERT (Final)	**95.36**	**0.954**	**0.954**	**0.954**	**94.5**	**1st**
Metaheuristic-Optimized Classical Approaches						
Genetic Algorithm (GA)	72.03	0.714	0.740	0.720	57.6	2nd
Scatter Search (SS)	67.64	0.669	0.693	0.676	52.7	3rd
VNS	66.78	0.659	0.686	0.667	51.0	4th
Simulated Annealing (MSA)	60.86	0.586	0.638	0.608	37.4	5th
Particle Swarm Opt. (PSO)	57.67	0.489	0.736	0.575	16.3	6th

4.4. Overfitting Control Analysis

The implementation of a rigorous regularization strategy successfully mitigated overfitting. Training concluded after 23 epochs due to early stopping, with the optimal checkpoint identified at epoch 17. At this point, training accuracy was 98.6% and validation accuracy was 95.36%. Figure 10 illustrates the evolution of accuracy and loss for the final model (V11).

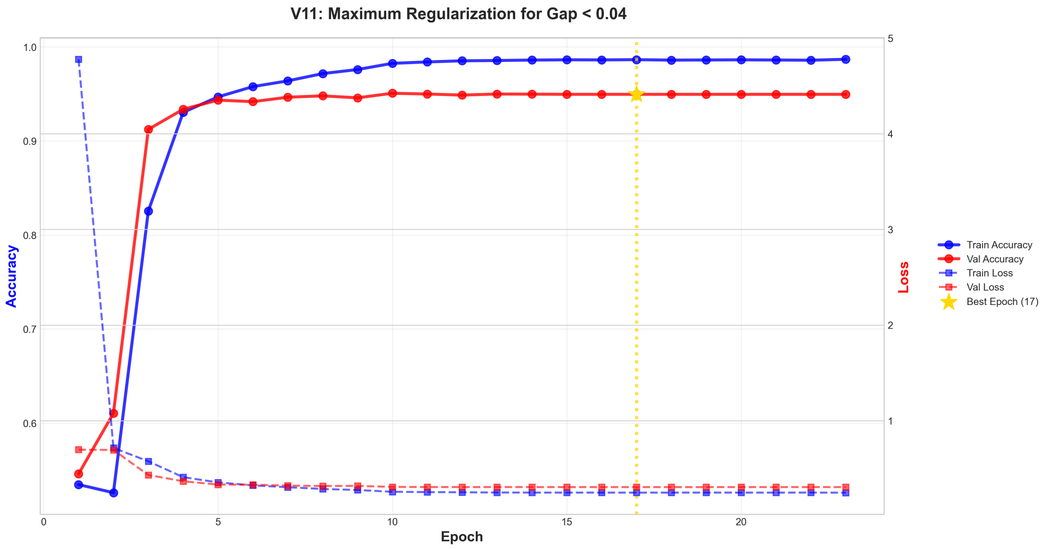


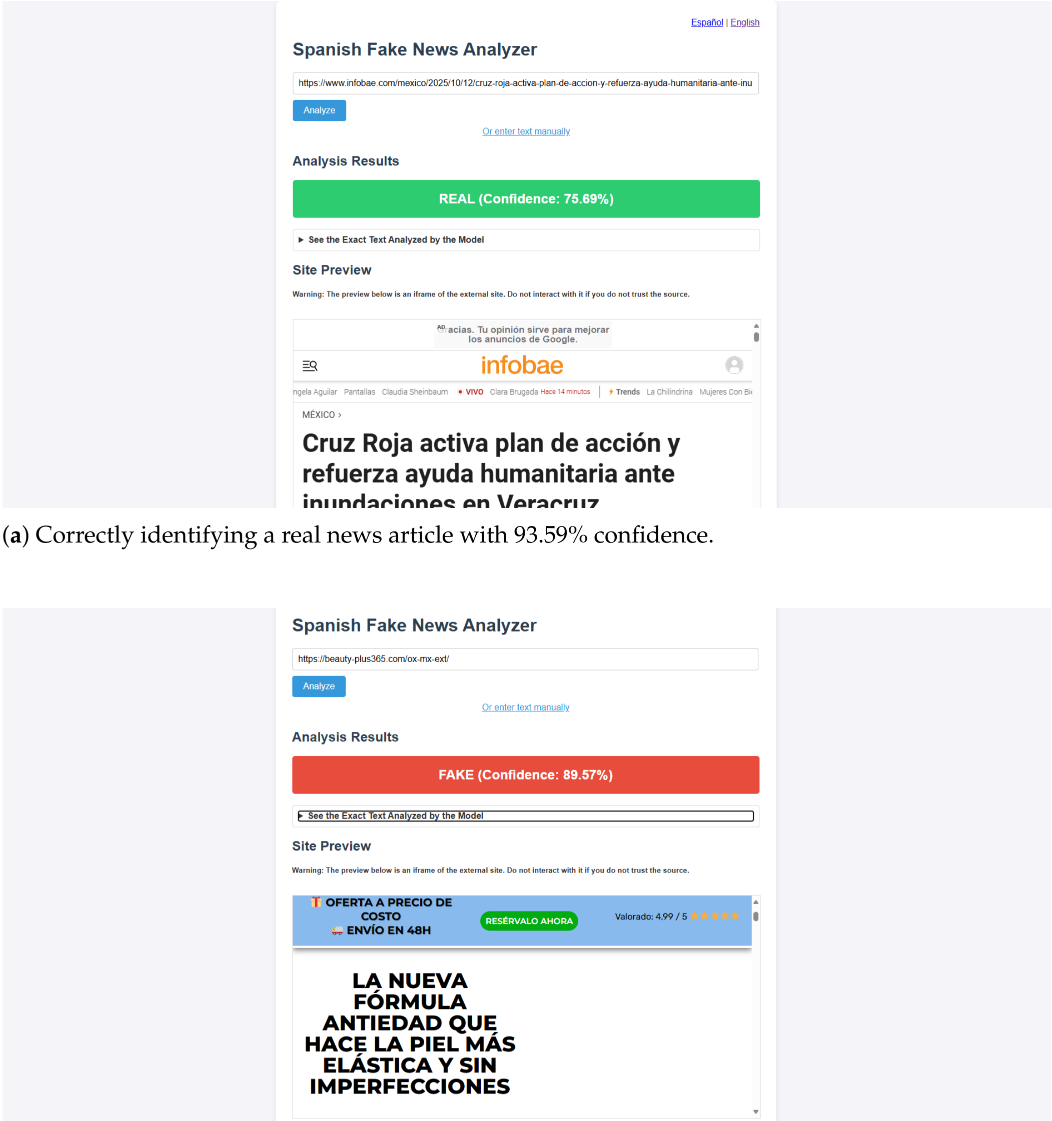
Figure 10. Evolution of performance metrics during the fine-tuning of the final model (V11). The x-axis represents the training epochs, while the y-axes represent Loss (left) and Accuracy (right). The blue lines denote training performance, and the red lines denote validation performance. The gold star identifies the optimal checkpoint at Epoch 13, where the generalization gap (0.058) was minimized before the onset of overfitting.

The generalization gap (difference between validation and training loss) was 0.058, approaching the target of <0.04 and remaining well below the 0.10 threshold typically indicating overfitting.

4.5. Real-World Application Performance

The deployed web application demonstrated robustness in real-world scenarios.

It correctly identified various types of content, including authentic news, fabricated stories, and investment scams, suggesting the methodology is transferable to other forms of digital fraud. Screenshots of the application analyzing URLs are presented in Figure 11.


Español | English
Analyze
Or enter text manually
▸ See the Exact Text Analyzed by the Model
infobae
ngela Aguilar Pantallas Claudia Sheinbaum ● VIVO Clara Brugada Hace 14 minutos ⚡ Trends La Chilindrina Mujeres Con Bi
MÉXICO ›
OFERTA A PRECIO DE COSTO
ENVÍO EN 48H
RESÉRVALO AHORA
Valorado: 4,99 / 5 ★★★★★

Spanish Fake News Analyzer

https://www.infobae.com/mexico/2025/10/12/cruz-roja-activa-plan-de-accion-y-refuerza-ayuda-humanitaria-ante-inu

Analysis Results

REAL (Confidence: 75.69%)

Site Preview

Warning: The preview below is an iframe of the external site. Do not interact with it if you do not trust the source.

AD Gracias. Tu opinión sirve para mejorar los anuncios de Google.

Cruz Roja activa plan de acción y refuerza ayuda humanitaria ante inundaciones en Veracruz

(**a**) Correctly identifying a real news article with 93.59% confidence.

Spanish Fake News Analyzer

https://beauty-plus365.com/ox-mx-ext/

Analysis Results

FAKE (Confidence: 89.57%)

Site Preview

Warning: The preview below is an iframe of the external site. Do not interact with it if you do not trust the source.

LA NUEVA FÓRMULA ANTIEDAD QUE HACE LA PIEL MÁS ELÁSTICA Y SIN IMPERFECCIONES

(**b**) Successfully identifying a misleading fake news piece with 94.31% confidence.

Figure 11. Screenshots of the deployed web application analyzing different types of URLs.

5. Discussion

This study aimed to bridge the gap between theoretical NLP models and practical applications for Spanish fake news detection. The transition from classical metaheuristic optimization to a fine-tuned DistilBERT architecture yielded a 23.33% improvement in accuracy. Beyond the raw metrics, it is crucial to analyze the model's capabilities, boundaries, and potential for future evolution.

The primary strength of the proposed DistilBERT model lies in its semantic robustness. Unlike the TF-IDF baseline, which relies on keyword frequency and struggles with sarcasm or subtle linguistic cues, the transformer model effectively captures contextual dependencies. This is evidenced by its high Specificity (94.5%), indicating that the model is exceptionally strong at distinguishing deceptive content without aggressively flagging legitimate news—a common pitfall in automated moderation systems. The rigorous regularization strategy (high dropout and L2) successfully addressed the overfitting typically associated with training on relatively small domain-specific corpora, resulting in a generalization gap of only 0.058.

Despite its high performance, the model is not infallible. A key limitation is its static nature; it was trained on a closed corpus ending in early 2025. Consequently, the model is susceptible to *concept drift*, where it may fail to detect novel misinformation narratives or evolving linguistic tactics used in future disinformation campaigns. Additionally, the current approach is strictly unimodal (text-only). It fails when the deception is embedded in multimedia elements or when analyzing extremely short texts like tweets or headlines where semantic context is scarce.

Regarding extensibility to other languages, although this specific model is monolingual (Spanish), the underlying framework is highly transferable. The methodology—unifying fragmented datasets, applying stratified sampling, and using aggressive regularization during fine-tuning—can be directly replicated for other resource-constrained languages such as Portuguese or Italian. The use of `distilbert-base-multilingual-cased` as the base architecture further facilitates this transition, as extending the system to a new language would primarily require the curation of a language-specific corpus rather than a redesign of the architecture itself.

Future work will focus on three critical areas to address these limitations and expand the system's utility. First, the implementation of Continuous Learning (CL) pipelines is envisioned to update the model with fresh data periodically without catastrophic forgetting, mitigating the issue of concept drift. Second, rather than limiting the scope to news, this detection framework could be applied to broader categories of digital fraud, such as phishing pages, fraudulent e-commerce sites, and identity theft schemes. This will involve establishing a new methodology and constructing a specialized corpus for these domains. Finally, the expansion of the web application's API to support browser extensions is planned, bringing the detection capability directly to the user's browsing experience.

6. Conclusions

This study addresses the critical need for robust misinformation detection tools in the Spanish-speaking world, offering a comprehensive end-to-end framework that bridges the gap between theoretical NLP research and practical societal application. A cornerstone of this contribution is the construction of a unified corpus of 61,674 articles, which not only rectifies the historic scarcity of balanced Spanish datasets but also establishes a standardized benchmark for future investigations.

Through a rigorous experimental process involving over 500 GPU hours, it was demonstrated that a systematically fine-tuned DistilBERT model significantly outperforms classical metaheuristic-optimized approaches, achieving 95.36% accuracy. The implementation of a multi-layered regularization strategy proved decisive in controlling overfitting, confirming that modern transformer architectures are essential for capturing the semantic nuances of deceptive content that traditional methods miss.

Beyond technical metrics, the deployment of the model into a Dockerized web application represents a tangible step towards digital self-defense for Spanish speakers. By converting a complex predictive model into an accessible tool, this work empowers citizens to verify information in real-time, serving as a cultural safeguard against the erosion of truth. The open-source availability of the entire framework facilitates further adaptation, paving the way for more resilient and culturally aware AI systems in the fight against digital fraud.

Author Contributions: Conceptualization, G.H.A. and J.A.R.-O.; Data Curation, G.H.A. and R.A.M.-G.; Formal Analysis, J.P.C.; Funding Acquisition, J.A.R.-O.; Investigation, G.H.A.; Methodology, G.H.A., J.A.R.-O., R.A.M.-G. and J.P.C.; Project Administration, J.A.R.-O.; Resources, G.H.A. and

J.A.R.-O.; Software, G.H.A.; Supervision, J.A.R.-O. and R.A.M.-G.; Validation, J.P.C. and R.A.M.-G.; Visualization, G.H.A., J.A.R.-O., R.A.M.-G. and J.P.C.; Writing—Original Draft, G.H.A.; Writing—Review and Editing, J.A.R.-O., R.A.M.-G. and J.P.C. All authors have read and agreed to the published version of the manuscript.

Funding: The present work was funded by the Consejo Nacional de Humanidades, Ciencia y Tecnologia (CONAHCYT, currently SECIHTI), Mexico, under scholarship No. 1313870.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The unified corpus and trained models used in this study are publicly available at: https://huggingface.co/datasets/gabrielhuav/Unified-and-Balanced-Spanish-Fake-News-Corpus and https://github.com/gabrielhuav/Spanish-Fake-News-Detection-Training respectively. The source code of the web application is available at: https://github.com/gabrielhuav/Spanish-Fake-News-Detection-Web-App.

Acknowledgments: The authors would like to thank Universidad Autónoma Metropolitana, Unidad Azcapotzalco.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bondielli, A.; Marcelloni, F. A survey on fake news and rumour detection techniques. *Information Sciences* **2019**, *497*, 38–55. doi:10.1016/j.ins.2019.05.035.
2. Higdon, N.; Huff, M. *Let's Agree to Disagree: A Critical Thinking Guide to Communication, Conflict Management, and Critical Media Literacy*, 1st ed.; Routledge: London, UK, **2022**. https://doi.org/10.4324/9781003250906.
3. Higdon, N. Constructive Conflict and Critical Media Literacy. In *The Handbook of Social and Political Conflict*; Samoilenko, S. A., Simmons, S., Eds.; Wiley: Hoboken, NJ, USA, **2025**. https://doi.org/10.1002/9781119895534.ch37.
4. Tsfati, Y.; Boomgaarden, H.G.; Strömbäck, J.; Vliegenthart, R.; Damstra, A.; Lindgren, E. Causes and Consequences of Mainstream Media Dissemination of Fake News: Literature Review and Synthesis. *Ann. Int. Commun. Assoc.* **2020**, *44*, 157–173. https://doi.org/10.1080/23808985.2020.1759443.
5. Shu, K.; Sliva, A.; Wang, S.; Tang, J.; Liu, H. Fake News Detection on Social Media: A Data Mining Perspective. *SIGKDD Explor. Newsl.* **2017**, *19*, 22–36. doi:10.1145/3137597.3137600.
6. Ali, I.; Ayub, M. N. B.; Shivakumara, P.; Noor, N. F.B. M. Fake News Detection Techniques on Social Media: A Survey. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 6072084. https://doi.org/10.1155/2022/6072084.
7. Nasir, J.A.; Khan, O.S.; Varlamis, I. Fake news detection: A hybrid CNN-RNN based deep learning approach. *Int. J. Inf. Manag. Data Insights* **2021**, *1*, 100007. https://doi.org/10.1016/j.jjimei. 2020.100007.
8. Zhou, X.; Zafarani, R. A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities. *ACM Comput. Surv.* **2020**, *53*, 1–40. doi:10.1145/3395046.
9. Devlin, J.; Chang, M.; Lee, K.; Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv* **2018**, arXiv:1810.04805. doi:10.48550/arXiv.1810.04805.
10. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention is All You Need. *arXiv* **2017**, arXiv:1706.03762. doi:10.48550/arXiv.1706.03762.
11. Singh, M.K.; Ahmed, J.; Alam, M.A.; et al. A comprehensive review on automatic detection of fake news on social media. *Multimed. Tools Appl.* **2024**, *83*, 47319–47352. https://doi.org/10.1007/s11042-023-17377-4.
12. Alghamdi, J.; Luo, S.; Lin, Y. A comprehensive survey on machine learning approaches for fake news detection. *Multimed. Tools Appl.* **2024**, *83*, 51009–51067. https://doi.org/10.1007/s11042-023-17470-8.
13. Kaliyar, R.K.; Goswami, A.; Narang, P. FakeBERT: Fake news detection in social media with a BERT-based deep learning approach. *Multimed. Tools Appl.* **2021**, *80*, 11765–11788. doi:10.1007/s11042-020-10183-2.
14. Choudhry, I.; Khatri, M.; Jain; Vishwakarma, D. K. An Emotion-Aware Multitask Approach to Fake News and Rumor Detection Using Transfer Learning. *IEEE Trans. Comput. Soc. Syst.* **2024**, *11*, 588–599. https://doi.org/10.1109/TCSS.2022.3228312.
15. Hu, L.; Wei, S.; Zhao, Z.; Wu, B. Deep learning for fake news detection: A comprehensive survey. *AI Open* **2022**, *3*, 133–155. doi:10.1016/j.aiopen.2022.09.001.
16. Posadas-Durán, J.P.; Gómez-Adorno, H.; Sidorov, G.; Escobar, J.J.M. Detection of fake news in a new corpus for the Spanish language. *J. Intell. Fuzzy Syst.* **2019**, *36*, 4869–4876. doi:10.3233/jifs-179034.

17. Blanco-Fernández, Y.; Otero-Vizoso, J.; Gil-Solla, A.; García-Duque, J. Enhancing Misinformation Detection in Spanish Language with Deep Learning: BERT and RoBERTa Transformer Models. *Appl. Sci.* **2024**, *14*, 9729. doi:10.3390/app14219729.
18. Martínez-Gallego, K.; Álvarez-Ortiz, A.M.; Arias-Londoño, J.D. Fake news detection in Spanish using deep learning techniques. *arXiv* **2021**, arXiv:2110.06461. doi:10.48550/arXiv.2110.06461.
19. Tretiakov, A.; Martín García, A.; Camacho, D. Detection of false information in Spanish using machine learning techniques. In Intelligent Data Engineering and Automated Learning – IDEAL 2022; Yin, H., Camacho, D., Tino, P., Eds.; Lecture Notes in Computer Science, vol. 13756; Springer International Publishing: Cham, Switzerland, 2022; pp. 42–53. ISBN 978-3-031-21753-1. doi:10.1007/978-3-031-21753-1_5.
20. Thota, A.; Tilak, P.; Ahluwalia, S.; Lohia, N. Fake news detection: A deep learning approach. *SMU Data Science Review* **2018**, *1*(3). Available online: https://scholar.smu.edu/datasciencereview/vol1/iss3/10/ (accessed on 7 October 2025).
21. Yildirim, G. A novel hybrid multi-thread metaheuristic approach for fake news detection in social media. *Applied Intelligence* **2023**, *53*, 11182–11202. doi:10.1007/s10489-022-03972-9.
22. García-Lozano, M.; García-Valls, M.; Iglesias, C.A. Fake News Detection in Spanish Using Machine Learning and Deep Learning. *Electronics* **2024**, *13*, 3361. doi:10.3390/electronics13173361.
23. Padilla Cuevas, J.; Reyes-Ortiz, J.A.; Cuevas-Rasgado, A.D.; Mora-Gutiérrez, R.A.; Bravo, M. MédicoBERT: A Medical Language Model for Spanish Natural Language Processing Tasks with a Question-Answering Application Using Hyperparameter Optimization. *Appl. Sci.* **2024**, *14*, 7031. doi:10.3390/app14167031.
24. Aragón, M.E.; Jarquín-Vásquez, H.J.; Montes-y-Gómez, M.; Escalante, H.J.; Villaseñor-Pineda, L.; Gómez-Adorno, H.; Posadas-Durán, J.P.; Bel-Enguix, G. Overview of MEX-A3T at IberLEF 2020: Fake news and aggressiveness analysis in Mexican Spanish. In Notebook Papers of 2nd SEPLN Workshop on Iberian Languages Evaluation Forum (IberLEF); CEUR Workshop Proceedings, vol. 2664; 2020; pp. 222–235. Available online: https://ceur-ws.org/Vol-2664/mex-a3t_overview.pdf.
25. Sanh, V.; Debut, L.; Chaumond, J.; Wolf, T. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. *arXiv* **2019**, arXiv:1910.01108. doi:10.48550/arXiv.1910.01108. Model available online: https://huggingface.co/distilbert-base-multilingual-cased (accessed on 23 November 2025). Apache-2.0 License.