Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 8s (2025)

A Novel Hybrid Feature Engineering Approach for Enhanced Sentiment Analysis

M. Amala Jeyaseeli[1], Heren Chellam G[2].

[1]Research Scholar, (Reg. No. 22211172282007), Rani Anna Government College for Women

[2]Associate Professor, Rani Anna Government College for Women

[1&2]Affliated to Manonmaniam Sundaranar University, Tirunelveli-627012

Author's mail id: [1] amalafast@mail.com, [2] hchellam@gmail.com

Article History:
Received: 15-06-2025
Revised: 19-07-2025
Accepted: 10-08-2025

Abstract:

Sentiment analysis, a critical subfield of Natural Language Processing (NLP), involves identifying and categorizing sentiments expressed in textual data. However, analyzing human language is inherently challenging because of its complexity and ambiguity. In this study, we propose a hybrid approach that integrates Latent Dirichlet Allocation (LDA) for feature extraction with Non-Negative Matrix Factorization (NMF) for dimensionality reduction to enhance the sentiment classification performance. The effectiveness of the proposed method is evaluated using two datasets to assess its robustness and generalizability. experimental results demonstrate that the hybrid LDA–NMF model achieves high accuracy scores, with performance improvements reaching up to 93% on one dataset and 79% on the other, significantly outperforming standalone LDA and NMF-based approaches. These findings highlight the potential of hybrid feature engineering techniques to transform complex textual data into more discriminative representations, thereby substantially improving sentiment analysis outcomes.

Keywords: Sentiment Analysis, Natural Language Processing, Latent Dirichlet Allocation, Non - negative Matrix Factorization, Machine Learning.

I Introduction

With the rapid growth of digital communication and widespread use of the Internet, social media platforms such as Twitter, Facebook, and online forums have become major sources of user-generated content. These platforms enable individuals to express their opinions, emotions, and perspectives on various topics, resulting in a vast amount of unstructured textual data [8]. Extracting meaningful insights from these data poses a significant challenge in the field of natural language processing (NLP). One of the key areas of NLP research is stance-based sentiment analysis, which involves determining both the sentiment (positive or negative) and stance (support or oppose) of an author toward a specific target entity. The accuracy of stance-based sentiment analysis largely depends on the effectiveness of feature extraction techniques that transform raw text into structured representations suitable for machine learning models. The development of robust feature extraction methods is crucial for capturing the linguistic and contextual cues necessary for accurate stance classification. This is particularly important for applications such as public opinion mining, social media monitoring, and policy decision making.

Sentiment analysis presents several challenges:

- **Mixed Sentiment Expression:** The same text may contain both positive and negative sentiments, making classification difficult.
- **Contextual Dependence:** The sentiment of a text can change depending on its surrounding context and discourse structure.
- **Implicit Sentiment:** Some opinions are expressed indirectly, requiring models to infer sentiment beyond explicit word choices.
- **Lexical Ambiguity:** Words and phrases may have different meanings based on context, leading to potential misclassification.
- **Use of Figurative Language:** Sarcasm, irony, and metaphor significantly affects sentiment interpretation and often misleads traditional classifiers.

To address these challenges, effective feature extraction plays a significant role in sentiment analysis by transforming unstructured text into structured numerical representations suitable for machine-learning models. Among the various feature extraction methods, semantic feature extraction is particularly valuable because it captures the underlying meaning and contextual associations of words, rather than relying solely on surface-level representations. Semantic feature extraction enables models to identify latent themes in a text, distinguish implicit sentiment expressions, and handle lexical ambiguity more effectively. For instance, word embedding and topic modelling techniques can help extract meaningful semantic structures and improve sentiment classification accuracy.

In this study, we propose a hybrid feature extraction technique that integrates LDA for topic-based representation and NMF for dimensionality reduction. LDA identifies latent topics within text, enabling a deeper understanding of semantic structures, whereas NMF reduces dimensionality [6][7][14], eliminating noise and improving feature interpretability. The extracted features are then used to train multiple classifiers to optimize the sentiment prediction. By leveraging the strengths of both topic modeling and matrix factorization techniques, our approach enhances model performance, leading to improved sentiment classification accuracy. We evaluated our proposed method on a large-scale Twitter sentiment dataset, demonstrating significant improvements over the existing models.

The main focuses of this study are as follows:

1. **Hybrid Feature Extraction for Stance-Based Sentiment Analysis:** We propose a hybrid feature extraction technique that combines LDA for semantic representation and NMF for dimensionality reduction, thereby enhancing interpretability and sentiment classification performance.
2. **Improved Sentiment Classification Performance:** Our method outperforms traditional feature extraction approaches such as bag-of-words (BoW), Term Frequency TF-IDF, Word2Vec, LDA, NMF) across multiple classifiers, achieving higher accuracy and better generalization on a large-scale

Twitter sentiment dataset.

3. **Reduction in Classification Errors:** The proposed approach effectively reduces misclassification errors by capturing deeper semantic structures, leading to improved precision, recall, and F1-score in stance-based sentiment analysis.

The remainder of this paper is organized as follows. Section II provides an overview of related work and discusses advancements in topic modeling, feature engineering, and sentiment analysis of Twitter data. Section III outlines the proposed methodology and describes the approaches and techniques used in the study. Section IV presents the experimental setup and results, including an analysis of the experiments conducted. Section V summarizes the key findings and contributions of this study.

II Related Work

2.1 Feature Extraction Techniques for Sentiment Analysis

Early sentiment analysis models primarily relied on statistical techniques such as Bag of Words (BoW), Term Frequency-Inverse Document Frequency (TF-IDF), and N-grams, which represent text as word frequency vectors but do not capture contextual relationships [1][2][3]. Ahuja et al. [1] demonstrated that TF-IDF improves classification performance by 3–4% compared to N-gram models on the SS-Tweet dataset, while Zhan [5] observed that although TF- IDF achieves 99% accuracy on training data, it tends to overfit, whereas Word2Vec generalizes better on unseen data. To overcome the limitations of frequency-based methods, researchers have introduced word embeddings such as Word2Vec, GloVe, and FastText, which capture semantic relationships between words [2][3]. Kalaivani et al. [2] showed that BoW lacks semantic understanding, whereas word embedding enhances contextual representation. Semary et al. [3] found that TF-IDF and Word2Vec outperformed BoW in sentiment classification on Amazon and Twitter datasets. Despite TF-IDF achieving high accuracy on structured datasets, such as Amazon reviews, Semary et al. [3] reported inconsistent performance on Twitter data, which Zhan [5] attributed to Word2Vec's ability to generalize better with larger datasets. In addition to feature extraction, feature selection techniques, such as Document Frequency (DF), Information Gain (IG), and Chi-Square Statistics (CSS), help refine the extracted features by filtering out irrelevant information [4]. Hung et al. [4] showed that combining TF–IDF with feature selection improves the classification efficiency by reducing the computational load while maintaining accuracy.

2.2 Topic Modeling Approaches in NLP

Topic modeling is a fundamental technique in natural language processing (NLP) that is used to uncover latent thematic structures in large text corpora. Among the various methods explored, Latent Dirichlet Allocation (LDA) is widely adopted as a probabilistic generative model that represents documents as mixtures of latent topics, with each topic characterized by a distribution of words [6][7]. Twil et al. [6] applied LDA to sentiment analysis in TripAdvisor reviews, demonstrating its ability to extract latent topics for aspect-based sentiment analysis, while Rahmadan et al. [7] utilized LDA on Twitter data to extract topics related to disaster events, showcasing its capability to handle large-scale social media text. Several variants have been introduced to extend the functionality of LDA. The Joint Sentiment- Topic (JST) model integrates sentiment labels with topic modeling, allowing for simultaneous sentiment and topic extraction [9]. Akhmedov et al. [9] developed an LDA-based topic/document/sentence (TDS) model, incorporating sentence-level sentiment classification to refine sentiment polarity detection. Another extension combined LDA with lexicon-based sentiment analysis, as demonstrated by Rahmadan et al. [7], who integrated sentiment lexicons with LDA to improve sentiment classification in disaster-related discussions. Additionally, hierarchical topic models such as hierarchical LDA (hLDA) organize topics into structured hierarchies, enabling a more detailed representation of topic relationships and subtopics [9].

2.3 Dimensionality Reduction Techniques in NLP

Dimensionality reduction techniques are essential in natural language processing (NLP) to manage high-dimensional text data by transforming feature spaces while retaining key information. Principal Component Analysis (PCA) and Singular Value Decomposition (SVD) are widely used linear techniques that project high-dimensional text features into lower-dimensional spaces [12][13]. PCA identifies the principal components that capture the most variance in the data, and is commonly applied in document clustering and text feature selection [12][15]. Mohamed et al. [15] demonstrated its effectiveness in improving clustering quality and computational efficiency in Arabic text classification. SVD, particularly in Latent Semantic Analysis (LSA), is employed to reduce noise and extract latent topics, and Muningsih et al. [17] compared its performance against PCA in NLP clustering tasks. Beyond linear methods, Non-Negative Matrix Factorization (NMF) has gained prominence for feature reduction by factorizing data into interpretable components [14]. It has been applied in sentiment analysis, topic modeling, and document classification [10][16], with Das et al. [16] highlighting its suitability for sparse, non-negative datasets. Unlike PCA, NMF preserves part-based representations, making it beneficial for semantic feature extraction [10].

Recent advancements have introduced deep learning-based dimensionality reduction techniques, including auto-encoders (AE) and transformer-based embedding, which capture nonlinear relationships in text data [13][14]. AE models are often combined with traditional methods such as PCA and NMF for improved feature selection, whereas transformer-based approaches, as explored by García et al. [13], have been applied in multilingual settings to optimize the processing speed without compromising accuracy. These deep-learning approaches offer new possibilities for dimensionality reduction in NLP by refining feature representations beyond conventional statistical techniques.

2.4 Hybrid Feature Extraction in NLP

Hybrid feature extraction techniques in natural language processing (NLP) integrate multiple approaches to enhance text representation by combining statistical, syntactic, and deep-learning-based methods [21]. Various studies have explored hybrid models for tasks such as text classification, stance detection, and emotion recognition, leveraging word embeddings, lexical features, and dimensionality reduction techniques to capture both syntactic and semantic aspects of language [20]. Ahanin et al. [20] proposed a hybrid model for emotion classification that integrates Word2Vec embeddings with TF-IDF and part-of-speech (POS) tagging, demonstrating improved classification accuracy. Similarly, Kim and Gil [20] utilized Latent Dirichlet Allocation (LDA) for topic modeling combined with TF-IDF weighting to refine feature selection strategies in research paper classification. Dimensionality reduction plays a key role in hybrid feature extraction by optimizing the computational efficiency while preserving meaningful information [18]. Researchers have combined Principal Component Analysis (PCA) and Non-Negative Matrix Factorization (NMF) to improve feature interpretability [18]. Zoya et al. [20] highlighted that hybrid models outperform existing traditional feature techniques, leading to better sentiment analysis results. These hybrid methodologies contribute to more effective text classification and analysis in NLP.

2.5 Stance-Based Sentiment Analysis Studies

Stance detection is a key area in natural language processing (NLP) that focuses on identifying whether a text expresses support, opposition, or neutrality toward a specific target entity, unlike sentiment analysis, which requires contextual interpretation [22]. Various machine-learning approaches have been explored for stance detection, including feature-based and deep-learning-based methods. Al-Ghadir et al. [22] proposed a feature selection method combining term frequency-inverse document frequency (TF- IDF) with sentiment features, achieving an F1-score of 76.45% on the SemEval-2016 Task 6 dataset. Their evaluation included K-nearest neighbor (KNN) variants and Support Vector Machines (SVMs), with weighted KNN performing best for stance classification. Deep-learning models, particularly those using self-attention mechanisms such as transformers, have also been widely adopted for stance detection. Alturayeif et al. [23] conducted a systematic review that highlighted the increasing use of few-shot and multitask learning methods to enhance stance detection models.

Stance detection is closely related to sentiment analysis; however, the two tasks differ fundamentally. Whereas sentiment analysis categorizes text as positive, negative, or neutral, stance detection assigns labels of favor, against, or none, making it a more nuanced task [25]. For instance, a text with positive sentiment may still express opposition to a particular entity (e.g., "I'm so happy that the bill was rejected"). Küçük and Arıcı [26] demonstrated that joint stance and sentiment analysis of COVID-19 vaccine-related tweets provided a richer contextual understanding and improved sentiment classification. Stance detection has significant applications in various domains, including fake news detection, public opinion analysis, and social media monitoring, particularly in debates surrounding political issues, healthcare policies, and controversial social movements [23],[26]. These applications underscore the importance of developing robust stance detection models that are capable of capturing both linguistic and contextual nuances in text.

2.6 Research Gap

Despite advancements in stance-based sentiment analysis, existing methods still face challenges in terms of feature extraction, topic modeling, dimensionality reduction, hybrid approaches, and stance detection. These challenges impact classification performance, computational efficiency, and contextual understanding. Table 1 summarizes these research gaps and provides a clear foundation for addressing these limitations and guiding future improvements.

Table 1. Research Gaps

Category	Limitations
Feature Extraction	Techniques like word embeddings improve semantic representation but often lead to high-dimensional feature spaces, increasing computational inefficiencies.
Topic Modeling	Methods like LDA uncover latent themes but struggle with short-text data and require better optimization for feature interpretability.
Dimensionality Reduction	PCA, SVD, and NMF reduce feature space but fail to preserve contextual meaning, focusing more on numerical decomposition than semantic coherence.
Hybrid Feature Extraction	Hybrid methods enhance NLP classification but often emphasize word embeddings and statistical selection, overlooking semantic optimization and feature reduction.
Stance-Based Sentiment Analysis	Stance detection is hindered by short-text limitations and reliance on large annotated datasets, affecting scalability across domains.

The identified limitations highlight the need for techniques that optimize the semantic representation and computational efficiency. Existing methods offer insights, but lack a balance between semantic coherence, feature interpretability, and model efficiency in stance-based sentiment classification.

This paper proposes a topic-based feature extraction approach combined with dimensionality reduction to

- Preserve semantic meaning while reducing feature dimensionality.
- Refine stance classification through improved feature structures.
- Lower computational costs without compromising accuracy.

By integrating these improvements, our approach enhances stance detection and sentiment classification, thereby ensuring both interpretability and performance.

III Proposed Methodology

In this paper, we propose a hybrid methodology that combines LDA for topic modeling with NMF for dimensionality reduction to enhance sentiment analysis. The proposed sentiment classification framework incorporates a systematic pipeline consisting of dataset selection, preprocessing, hybrid feature extraction, machine-learning-based classification, and performance evaluation. The following subsections outline each stage in detail: The

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 8s (2025)

architecture of the proposed methodology is illustrated in Figure 1.

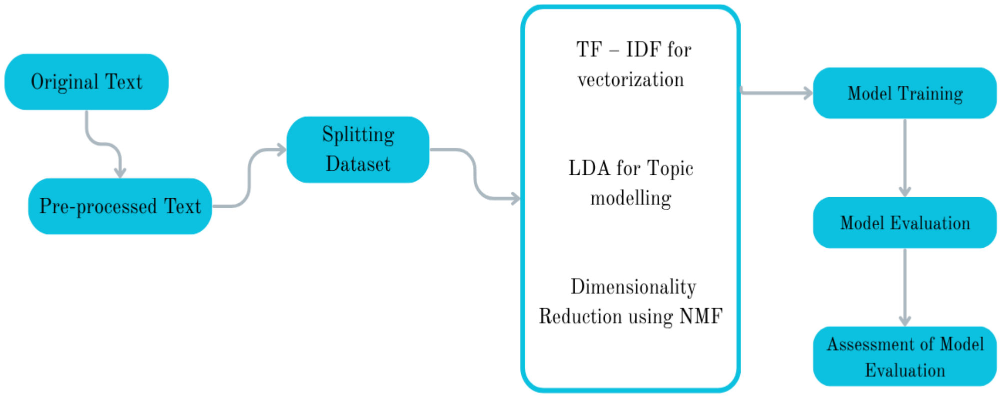

Figure 1. Architecture of the proposed hybrid LDA-NMF methodology for sentiment classification.

3.1 Dataset Description

Two publicly available sentiment datasets were used to evaluate the performance and generalizability of the proposed approach. The first dataset consists of labeled tweets related to Roe vs. Wade, with a total of 10804 samples divided into Pro-Choice and Pro-Life. The second dataset comprises 9940 IMDB movie reviews annotated with Positive and Negative sentiment labels. Summary statistics, including the total samples and class distributions, are provided in Table 1.

Table 2. Dataset Description

Dataset	Total Samples	Training Samples	Testing Samples	Classes/ Labels	
Roe Vs Wade Tweets	10804	8643	2161	Prochoice – 0 Pro-life - 1	0 - 5397 1 - 5407
IMDB Movie Review	9940	7952	1988	Negative – 0 Positive - 1	0 – 5007 1 - 4933

a) Data Pre-processing

- **Text Cleaning:** The text was pre-processed by converting it to lowercase, removing URLs, and eliminating non-word characters (such as special symbols and numbers) to reduce noise and ensure uniformity.
- **Tokenization**: The text was tokenized into individual words using the nltk.tokenize method, which served as the foundation for further processing.
- **Stop word Removal:** Commonly occurring words, such as "the," "is," and "on," were removed, as they do not contribute significant meaning for classification tasks.

- **Lemmatization**: Reduced words to their **base or dictionary form** (e.g., "running" → "run") to maintain linguistic consistency.
- **Data splitting:** A standard train–test split of 80:20 was applied to both the datasets. A summary of the training and testing divisions is provided in Table 1.

3.2 Feature Extraction Process

Feature extraction plays a significant role in sentiment classification by transforming raw text into structured representations that machine-learning models can process effectively. Extracting meaningful features enhances classification accuracy, while reducing computational complexity. In natural language processing (NLP), textual features can be categorized into three primary types, each associated with a specific feature extraction technique.

Table 3. Feature Types and Their Computational Techniques for Text Representation

Feature Type	Description	Techniques Used
Linguistic Feature	Capture surface-level textual patterns and statistical word distributions.	Bag of Words (BoW), TF-IDF, Hashing Vectorizer
Syntactic Features	Derived from grammatical structures (POS tagging, dependency parsing, and sentence structure analysis).	POS tagging, Dependency Parsing, N-grams
Semantic Features	Capture latent meaning, contextual relationships, and topic distributions, offering deeper insights into textual intent.	Latent Dirichlet Allocation (LDA), Word2Vec, Non-negative Matrix Factorization (NMF)

Among these, semantic features are particularly significant in stance-based sentiment analysis for the following reasons.

1. Contextual Dependence: Stance is often implicit and requires models to infer hidden relationships between words and topics.
2. Limitations of Statistical Approaches: Traditional methods, such as TF-IDF and n-grams, fail to capture the conceptual meaning of text, focusing only on frequency-based representations.
3. Topic-Driven Sentiment Differentiation: Grouping words into latent topics provides a more structured representation of user opinions, aiding stance differentiation, even when sentiment expressions are similar.

To assess the effectiveness of different feature extraction techniques, this study evaluated multiple methods, including BoW, TF-IDF, Word2Vec, LDA, NMF, and Hashing Vectorizer. Based on the experimental results, a hybrid LDA- NMF approach was proposed to enhance semantic feature extraction and dimensionality reduction for stance-based sentiment classification.

3.2.1 Hybrid Feature Extraction Approach:

This study proposes a hybrid feature extraction approach that integrates Latent Dirichlet Allocation (LDA) for topic-based semantic feature extraction and Non-Negative Matrix Factorization (NMF) for dimensionality reduction. The primary objective of this methodology is to extract a compact and meaningful set of semantic features to enhance classification accuracy while reducing computational complexity.

a) LDA for Topic Feature Extraction:

LDA is a probabilistic generative model that represents a document as a mixture of latent topics, with each topic being distributed over words. The LDA process is defined as follows.

- Document-Topic Distribution: Each document d is represented as a distribution over K topics:

$$\theta_d \sim Dir(\alpha)$$

where α is the Dirichlet prior for topic distribution.

- Topic-Word Distribution: Each topic is z represented by a distribution over V words:

$$\phi_z \sim Dir(\beta)$$

where β is the Dirichlet prior for the word distribution.

- Generative Process: For each word $w_{d,n}$ in document d:
  1. Sample a topic $z_{d,n}$ from the topic distribution θ_d.
  2. Sample a topic $w_{d,n}$ from the topic-word distribution ϕ_z .

This process produces a topic-word matrix ϕ, where each topic is represented by a distribution of words. In this study, ten topics were extracted, each represented by seven top words, forming a topic-word matrix $\phi \in \mathrm{R}^{\mathbf{10\times7}}$.

b) NMF for Dimensionality Reduction:

After extracting topics using LDA, NMF was applied to reduce the dimensionality of the topic-word matrix while preserving meaningful topic structures. NMF factorizes the non-negative matrix ϕ into two lower-rank non-negative matrices W and H, such that:

$$\phi \approx W.H$$

- $\phi \in \mathrm{R}_{+}^{\mathbf{10\times7}}$ is the topic-word matrix obtained from the LDA.
- $W \in \mathrm{R}_{+}^{\mathbf{10\times}r}$ is the basis matrix, where r is the reduced number of dimensions.
- $H \in \mathrm{R}_{+}^{r\times\mathbf{7}}$ is a coefficient matrix representing the relationship between topics and words.

The final low-dimensional feature set was obtained by solving the following optimization problem:

$$\min_{\{W,H\}|} || \phi - W \cdot H ||_F^2 \text{ subject to } W, H \geq 0$$

where $|| \cdot ||_F$ denotes the Frobenius norm that measures the difference between the original and approximated matrices.

3.3 Classification Model:

The feature sets extracted from various feature extraction techniques serve as inputs for stance-based sentiment classification. To assess the impact of different feature representations, multiple feature extraction methods were applied before implementing the proposed LDA-NMF hybrid approach. These methods encompass both the linguistic and semantic approaches. Table 4 summarizes the approaches and techniques for feature extraction and highlights the key methods utilized in our study, including linguistic and semantic techniques.

After analyzing the classifier performance of these methods, a hybrid feature extraction approach using LDA and NMF was proposed to improve semantic representation and dimensionality reduction.

The extracted feature set from each method, including the proposed LDA-NMF, was then used to train and evaluate the six supervised machine learning classifiers.

- Support Vector Machine (SVM): Constructs a hyperplane that maximizes the margin between classes. The decision function is given by

$$f(x) = w^T x + b$$

where w is the weight vector, x is the input feature vector, and b is a bias term.

Table 4. Feature Extraction Approaches and Techniques Used for Text Classification

Approach	Techniques	Description
Linguistic Approaches	BoW	Represents text using word frequency counts.
	TF-IDF	Weighs words based on importance by considering term frequency and inverse document frequency.
	Hashing Vectorizer	Uses a fixed-size vector space to encode text, reducing memory usage.
Semantic Approaches	Word2Vec	Captures word relationships based on contextual similarity.
	LDA	Identifies latent topics within text.
	NMF	captures hidden patterns in text for better classification,

		reduces high-dimensional feature spaces while preserving structure

- K-Nearest Neighbours (KNN): Classifies a new sample based on the majority vote of its k nearest neighbors using the Euclidean distance:

$$d(x_i, x_j) = \sqrt{\sum_{m=1}^{n} (x_{i,m} - x_{j,m})^2}$$

- Logistic Regression (LR): The sigmoid function is used to estimate the probability of a sample belonging to a class:

$$P(y = 1|x) = \frac{1}{1 + e^{-(w^T x + b)}}$$

- Decision Tree (DT): Constructs a tree in which decisions are made by selecting features that minimize entropy:

$$H(S) = -\sum_{i=1}^{n} p_i log_2 p_i$$

where $H(S)$ is the entropy, and p_i is the probability of class i in data set S.

- Random Forest (RF): Builds multiple decision trees and combines their outputs using majority voting.

$$\hat{y} = \frac{1}{T}\sum_{t=1}^{T} h_t(x)$$

where $h_t(x)$ is the prediction from tree t, and T is the number of trees.

- Gradient Boosting (GB): Sequentially improve the weak models by minimizing the loss function $L(y, f(x))$ using gradient descent:

$$f_m(x) = f_{m-1}(x) + \gamma h_m(x)$$

where $h_m(x)$ is the weak learner, and γ is the learning rate.

Each classifier was trained on the extracted feature set and evaluated using the unseen data. Performance was measured through accuracy, precision, recall, and F1-score.

3.4 Model Evaluation:

This study employed standard evaluation metrics to systematically analyze the performance of six classifiers trained on features extracted using existing techniques and the proposed hybrid technique.

- Accuracy: Accuracy measures the proportion of correctly classified instances to the

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 8s (2025)

total instances

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

where TP (True Positive), TN (True Negative), FP (False Positive), and FN (False Negative) represent classification outcomes.

- Precision: Measures how many predicted positive instances are truly positive.

$$Precision = \frac{TP}{TP + FP}$$

- Recall: Measures how many actual positive instances were correctly classified.

$$Recall = \frac{TP}{TP + FN}$$

- F1-Score: A balanced measure that calculates the harmonic mean of Precision and Recall.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

The classification results were compared across six classifiers (SVM, KNN, LR, DT, RF, and GB) using these metrics. The best-performing model was identified based on the highest F1-score and the overall balanced performance.

Algorithm: Hybrid LDA-NMF for Feature Extraction

Input: Text dataset D

Output: Low-dimensional feature set F

Step 1: Pre-processing:

- Convert text to lowercase
- Remove URLs, non-words, and stop words
- Tokenize text, and apply lemmatization.

Step 2: Splitting the Dataset &Vectorization

- Split the dataset into training, validation and testing sets.
- Convert text into numerical representations using TF-IDF

Step 3: LDA for Topic Feature Extraction:

- Define the number of topics K.
- Apply LDA to D, generating:

- Topic-word matrix ϕ, where ϕ_{ij} represents the probability of word i in topic j.

- Document-topic matrix θ, where θ_{ij} represents the probability of topic j in document i.

Step 4: NMF for Dimensionality Reduction:

- Use the topic-word matrix ϕ as input for NMF.

- Decompose ϕ into two non-negative matrices:

- Basic matrix $W \in R_{+}^{mxr}$, where m is the number of documents and r is the reduced dimensionality.
- Coefficient matrix H $\in R_{+}^{rxn}$, where n is the vocabulary size.

- Optimize W and H by minimizing the reconstruction error:

$$\min_{\{W,H\}} || \phi - W \cdot H ||_{\mathrm{F}}^{2} \text{ subject to W, H} \geq 0$$

where $|| \cdot ||_F$ is the Frobenius norm.

Step 5: Feature Set Output:

- Use coefficient matrix H as the final low-dimensional feature set F for classification.

Step 6: Classification & Testing: Train machine learning models using feature set F.

IV Result and Discussion

This section presents the experimental results and performance evaluation of various feature extraction techniques and machine-learning classifiers for sentiment classification. Experiments were conducted on two datasets—Roe vs. Wade Tweets and IMDB Movie Reviews—to assess the effectiveness, scalability, and generalizability of the proposed hybrid feature extraction technique across two textual domains. All experiments were performed in the Google Colab environment using Python-based libraries, such as scikit-learn, TensorFlow, and NLTK, for preprocessing, feature extraction, model training, and evaluation.

To transform raw textual data into structured representations, multiple feature extraction techniques were applied, including BoW, TF-IDF, Word2Vec, LDA, NMF, Hashing Vectorizer, and the proposed LDA–NMF hybrid approach. These features were evaluated using six supervised classifiers: SVM, KNN, logistic regression, decision tree, Random Forest, and gradient boosting. This setup allows a comprehensive comparison of traditional, semantic, and hybrid feature representations across classifiers, highlighting the performance and robustness of the proposed method.

4.1 Accuracy Comparison of Feature Extraction Techniques

The comparative performance of the six existing feature extraction techniques and the proposed hybrid method is summarized in Table 2 and visualized in Figure 1,2. Across both datasets, the proposed technique consistently achieved the highest accuracy with all classifiers, indicating its ability to capture richer semantic and contextual features from textual data.

Table 5. Classification Accuracy (%) of Various Feature Extraction Techniques Across Multiple Classifiers on Different Datasets

	Roe v Wade						**IMDB**					
	SVM	**KNN**	**LR**	**DT**	**RF**	**GB**	**SVM**	**KNN**	**LR**	**DT**	**RF**	**GB**
BoW	85.98	79.22	86.21	84.36	87.78	87.37	63.58	54.68	64.69	54.48	60.71	60.06
Tf-IDF	86.44	80.01	85.05	84.73	87.46	87.04	63.73	58.35	62.78	54.02	61.27	62.98
Hash	82.83	80.10	82.14	77.74	83.11	83.53	58.90	53.62	60.16	51.91	59.96	61.57
LDA	78.57	81.49	76.26	85.65	87.97	86.35	63.63	58.55	63.68	56.49	61.32	63.33
NMF	88.71	86.81	87.27	81.49	86.86	87.88	57.60	52.82	58.70	50.60	55.53	57.75
W2V	87.97	82.92	88.85	78.53	84.73	85.89	56.29	53.97	56.04	51.21	53.17	56.74
Hybrid	**93.06**	**92.27**	**92.97**	**93.01**	**93.06**	**93.11**	**79.18**	**73.79**	**79.18**	**67.15**	**77.67**	**78.47**

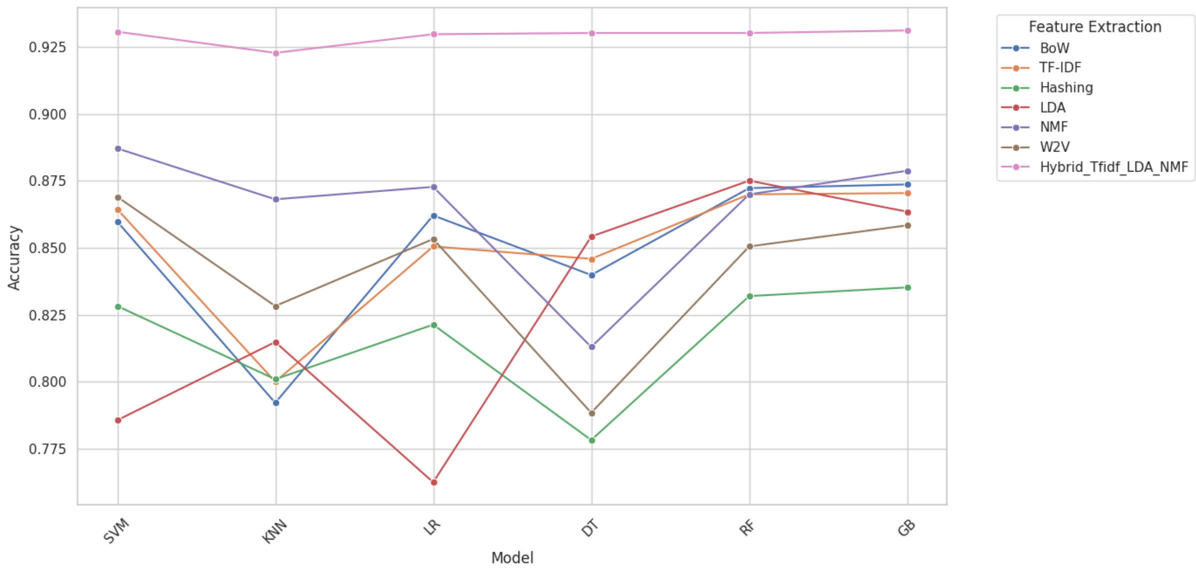

Figure 2. Accuracy comparison of feature extraction techniques across six classifiers on Roe vs. Wade dataset

Accuracy Performance of Feature Extraction Techniques across Six Machine Learning Classifiers: Roe vs. Wade Tweet Dataset Figure 1. The image displays a line graph showing the classification accuracy of different feature extraction methods using six supervised classifiers. The accuracy of the methods is shown on the y-axis ranging from 75 % to 95 %. Seven different feature extraction methods are mentioned on the x-axis: BoW, TF-IDF, Hash, LDA, NMF, W2V and the Hybrid. Six types of supervised classifiers are mentioned: support vector machine (SVM), K-nearest neighbors or KNN, logistic regression or LR, decision tree or DT, random forest or RF, and gradient boosting or GB. Different colored lines and markers represent different classifiers. The Hybrid LDA-NMF method showed the maximum accuracy across all classifiers, ranging between 92.27% and 93.11%. Thus, it performs much better than the other feature-extraction methods. The chart shows that the ensemble methods (RF and GB) generally perform better than other classifiers using classical features, whereas the hybrid technique increases the performance of all classifiers. This graph confirms that the hybrid feature extraction approach is accurate and can be used for the stance-based sentiment classification of social media text.

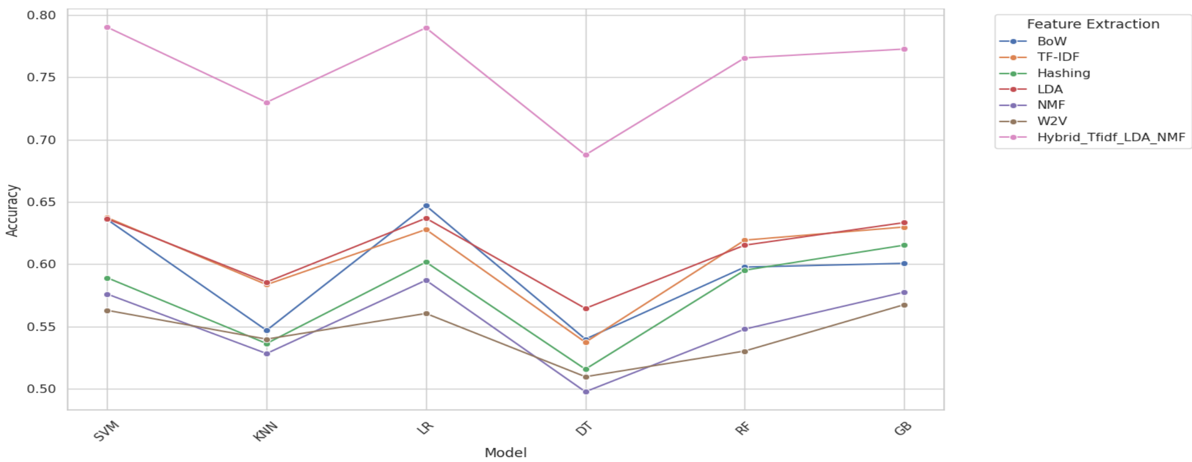

Figure 3. Accuracy comparison of feature extraction techniques across six classifiers on IMDB dataset.

A visual representation of the feature extraction method performance across the six machine learning classifiers on the IMDB movie review dataset is illustrated in Figure 2. The line graph shows the classification accuracy along the y-axis in the range of 50–80 on the x-axis for seven feature extraction methods (BoW, TF-IDF, Hash, LDA, NMF, W2V, and Hybrid) evaluated over six supervised classifiers (SVM, KNN, LR, DT, RF, and GB). Each classifier trajectory is represented by a colored line marked by data points. The Hybrid LDA-NMF technique considerably improves performance. All classifier accuracies show 67.15% to 79.18% results. It outperformed the baseline techniques by 5% to 20% points in overall results. It's worth mentioning that standard frequency-based methods (BoW, TF-IDF) yield moderate performance (54-65%), whereas standalone semantic methods (LDA, NMsF, W2V) show mixed results. The hybrid method produced the highest accuracy with all classifier architectures, indicating that it was able to identify the semantic and contextual features among longer text documents more effectively for movie sentiment analysis.

4.2 Correlation Analysis

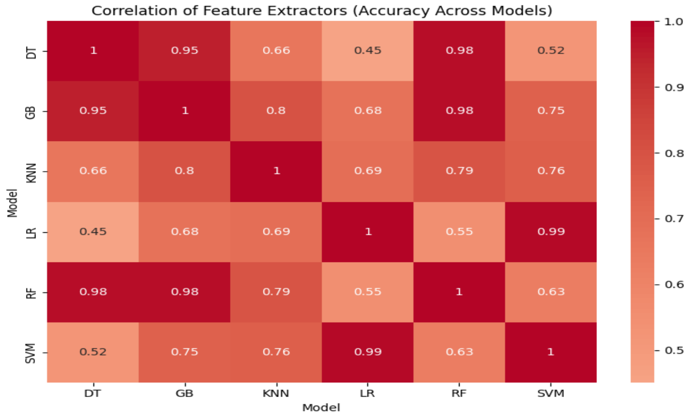

Figure 4. Correlation heat-map of classifier performances on Roe vs. Wade dataset

This is a heat map displayed in Figure 3, which shows that the classifiers performed well in relation to one another. The heatmap shows the Pearson correlation coefficients (from 0 to 1) of the six machine learning classifiers (SVM, LR, KNN, DT, RF, and GB) assessed on the extracted features. The gradient indicates the strength of the correlation: light yellow, low correlation; dark purple, very strong correlation (1). The values are indicated for each cell. Strong positive correlations (>0.95) were observed between the tree-based ensemble methods, DT, GB, and RF. This indicates that these methods follow similar patterns of classification. The linear models, LR and SVM, are highly correlated at 0.99, indicating that they have similar decision boundaries. KNN is correlated with other classifiers (0.66-0.80) because it is instance-based. The correlation between DT and LR is 0.45, showing that they are fundamentally different classification approaches. Correlation analysis exhibits distinct classifier clusters and shows the necessity of comparing multiple model architectures to properly judge the feature extraction quality in stance-based sentiment classification.

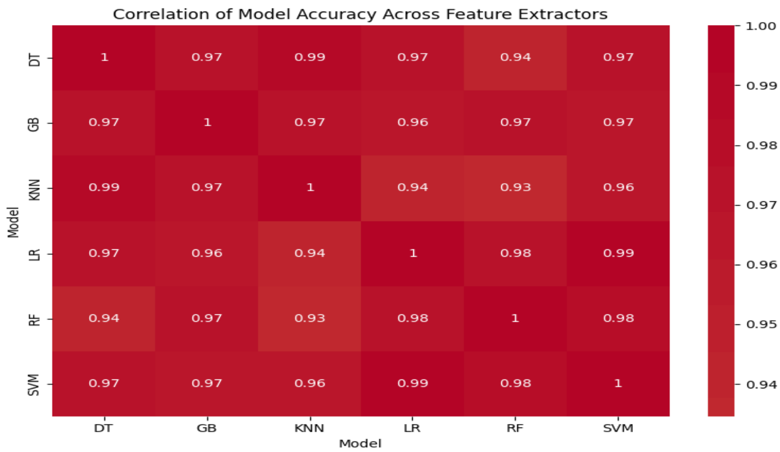

Figure 5. Correlation heat-map of classifier performances on IMDB Movie Review dataset

The performances given by the different classifiers are compared to show their correlation, as shown in Figure 4. The heat map is symmetrical. The heat map shows the correlations between the SVM, LR, KNN, DT, RF, and GB classifiers. The correlation is given on a color scale ranging from light to dark purple. This shows that the correlation is between zero and one. Models based on trees (DT, GB, RF) have very high inter-correlations (0.94-0.97) in the movie review corpus. The linear classifier (LR and SVM) models are highly consistent (0.98-0.99). The KNN correlation with most models is fairly high at 0.93-0.96 but is lower than other correlations with models. The overall high correlation values indicate that most classifiers agree on processing longer-form review texts. However, subtle differences in performance indicate the benefits of multi-classifier evaluations. This study provides insights into classifier behavior consistency and shows the robustness of the extracted features under different learning paradigms in document-level sentiment analysis.

4.3 Visualization of Extracted Features

The visual representation of features extracted by the proposed method (Figure 6 and 7) shows distinct and compact clusters, confirming the enhanced discriminability compared to traditional methods.

Figure 6 illustrates the topic-word structure generated using the proposed hybrid LDA-NMF technique. In the diagram, each large blue circle represents a topic, whereas the smaller grey circles denote the most important words associated with that topic. The connecting lines show the strength of the relationship between a word and a topic: solid lines indicate strong, high-confidence associations, whereas dotted lines represent weaker or shared connections across multiple topics. This figure demonstrates how the hybrid technique formed clear groupings of meaningful words for each topic. The clustering pattern shows that the model effectively identifies dominant keywords and captures subtle associations that may not be detected using a single method. Overall, Figure 6 shows how the hybrid approach builds richer and more interpretable feature representations for a given dataset.

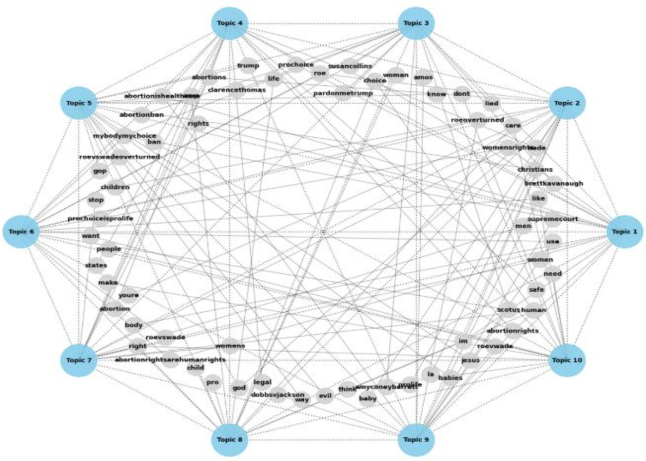

Figure 6. Extracted features from Roe vs wade dataset

Figure 7 further illustrates the capability of the hybrid LDA-NMF model to capture the relationships between topics and their associated words. Similar to Figure 6, each large blue circle corresponds to a topic, and the gray circles represent the keywords contributing to that topic. The solid lines indicate strong connections, whereas the dotted lines indicate weaker or overlapping relationships between words and topics. This figure highlights the robustness of the hybrid model by illustrating how it organizes semantically important words into a compact and meaningful structure. Compared to individual LDA or NMF outputs, the hybrid representation reveals more detailed interactions and shared patterns across topics. Figure 2 reinforces that the hybrid technique extracts deeper semantic relationships and provides a more comprehensive understanding of the topic–word structure.

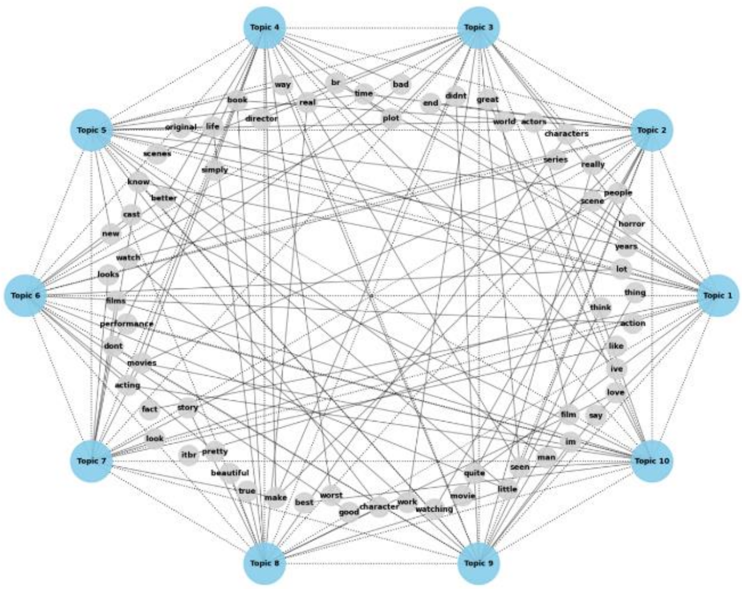

Figure 7. Extracted features from IMDB datase

4.5 Misclassification Analysis

Misclassification patterns (Figures 8–10) reveal that the proposed feature extractor minimizes errors for both classes, reflecting improved generalization in the Roe vs. Wade dataset.

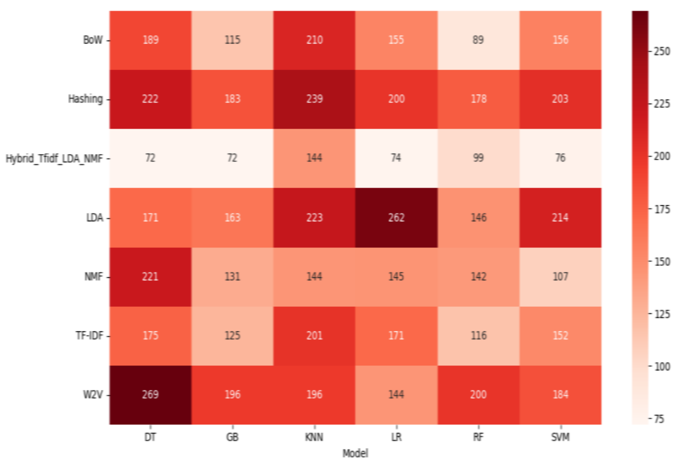

Figure 8. Misclassification Analysis for Roe vs Wade Dataset for class 0

Figure 8 displays the patterns of incorrect classification in the Roe vs. Wade dataset for the Pro-Choice class 0. According to the line graph, the total number of misclassified pro-choice samples (y-axis) for every feature extraction method (x-axis: BoW, TF-IDF, Hash, LDA, NMF, W2V, and Hybrid) and for six classifiers, represented by different colors. Of all the classifiers, the hybrid approach yielded the lowest misclassification counts across all classifiers. In addition, the classification of the class is better in the case of the hybrid approach compared with other methods.

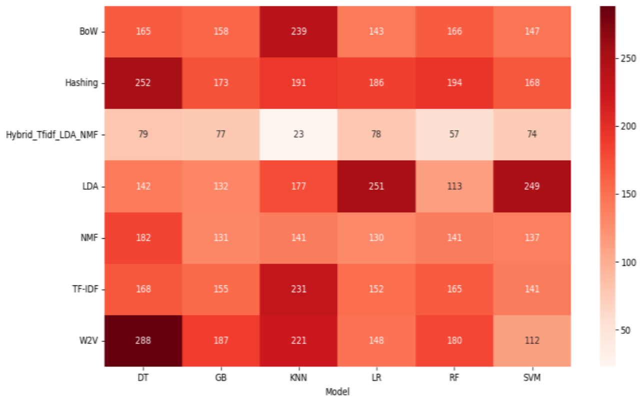

Figure 9. Misclassification Analysis for Roe vs Wade Dataset for class 1

Misclassification patterns for the Pro-Life class (Class 1) across different feature extraction techniques on the Roe vs. Wade dataset. The line graph illustrates the misclassification frequencies (y-axis) for the Pro-Life stance samples across seven feature extraction methods (x-axis) and six classifiers. The Hybrid LDA-NMF technique exhibits minimal error rates across all classifier architectures, indicating robust generalization for both stance categories.

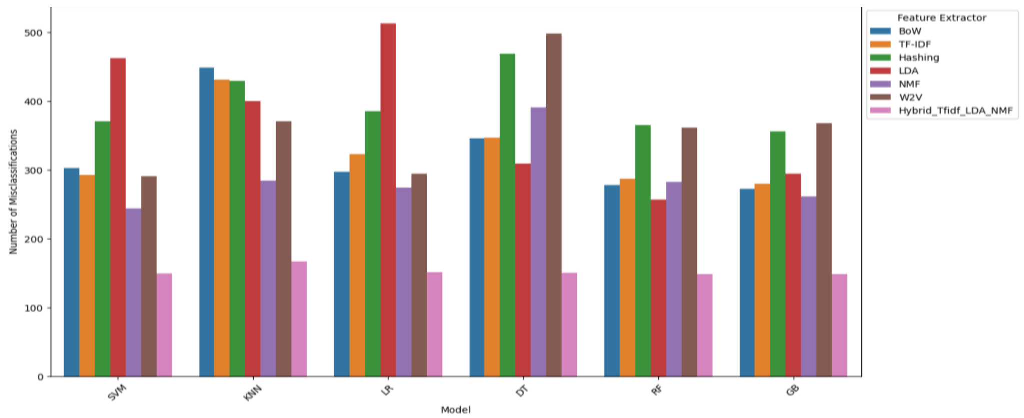

Figure 10. Aggregate misclassification comparison across feature extraction methods on the Roe vs Wade dataset.

The grouped bar chart presents the total misclassification counts for each feature extraction technique across the six classifiers (color-coded bars). Each cluster of bars represents a different classifier (SVM, KNN, LR, DT, RF, GB), allowing direct comparison of the feature extraction performance. The Hybrid approach demonstrated the lowest cumulative misclassification errors across all classifiers, with significant reductions compared with the baseline methods. This comprehensive analysis confirms the effectiveness of the proposed techniques in minimizing both Type I and Type II errors, thereby improving the overall classification reliability of stance-based sentiment analysis.

Misclassification for IMDB Movie Review Dataset:

Misclassification patterns (Figures 11–13) revealed that the proposed feature extractor minimized errors for both classes, reflecting improved generalization in the IMDB Dataset.

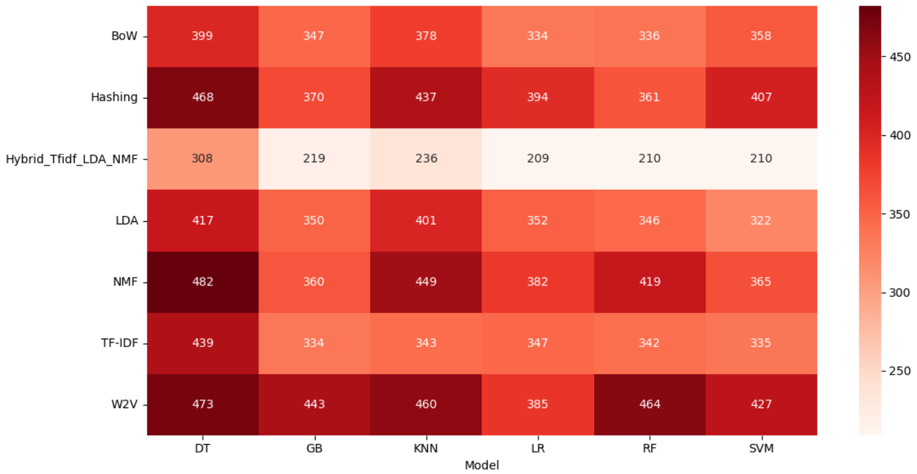

Figure 11. Misclassification Analysis for IMDB Dataset for class 0

Figure 11 shows the misclassification distribution for negative reviews (Class 0) across feature extraction methods on the IMDB dataset. The line graph plots the number of incorrectly classified negative reviews (y-axis) against the feature extraction techniques (x-axis) for the six classifiers. The Hybrid LDA-NMF approach achieves substantially lower misclassification rates than traditional methods, indicating enhanced sensitivity for detecting negative sentiments.

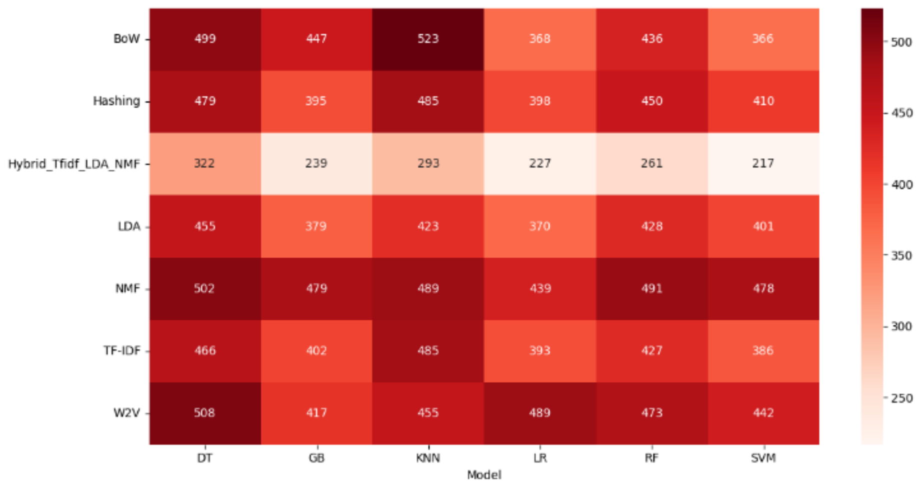

Figure 12.Misclassification Analysis for IMDB Dataset for class 1

Figure 12 shows the misclassification distribution for positive reviews (Class 1) across feature extraction methods on the IMDB dataset. The line graph illustrates the error patterns (y-axis) across the seven feature extraction techniques (x-axis) for the six classifiers. The proposed hybrid method demonstrated consistent superiority in correctly identifying positive sentiments, with minimal false negative rates across all classifier types.

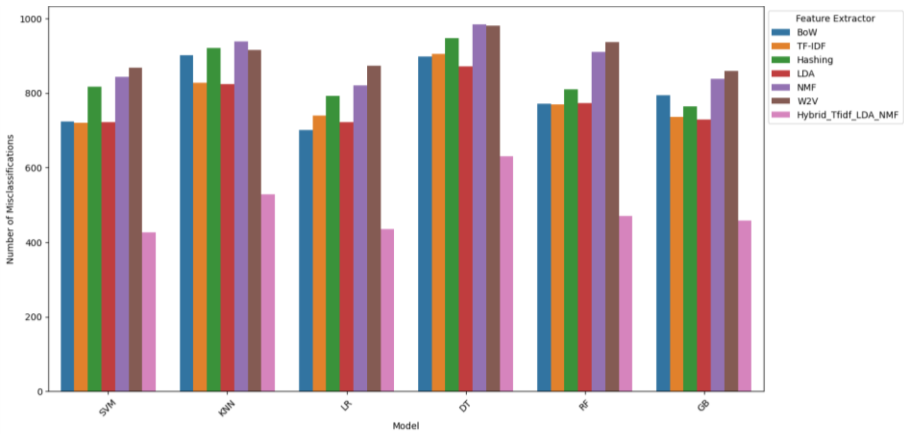

Figure 13. Aggregate misclassification comparison across feature extraction methods on the Roe vs Wade dataset.

Figure 13 shows a comprehensive misclassification comparison using the feature extraction technique on the IMDB dataset. The grouped bar chart visualizes the total classification errors for each method across the six classifiers (represented by color-coded bar clusters). Each group compares the classifier performance for a specific feature extraction technique. The Hybrid approach exhibited the shortest bars across all classifier groups, indicating the lowest total misclassifications. The substantial reduction in errors (15-40% improvement over baseline methods) confirms the hybrid techniques superior generalization capability and robustness for sentiment classification in movie review analysis, validating its effectiveness across diverse classifier architectures and text domains.

4.6 ROC-AUC Analysis

Roe v Wade:

Figure 14 shows the Receiver Operating Characteristic (ROC) curves with Area Under Curve (AUC) scores for all classifiers on the Roe vs. Wade dataset. The multi-panel plot displays ROC curves for six classifiers (SVM, KNN, LR, DT, RF, and GB) with each panel comparing the performance across seven feature extraction methods. The True Positive Rate (sensitivity) was plotted on the y-axis against the False Positive Rate (1-Specificity) on the x-axis, with the diagonal dashed line representing random classification (AUC=0.50). Each feature extraction method is represented by a distinct colored curve with corresponding AUC values in the legend. The proposed Hybrid LDA-NMF approach (typically shown in red or highlighted color) consistently achieves the highest AUC values (>0.93) across all classifiers, with curves positioned closest to the top-left corner, indicating superior discriminative ability. The consistent elevation of hybrid method curves above baseline approaches across diverse classifier architectures demonstrates robust performance in distinguishing between Pro-Choice and Pro-Life stances, validating the effectiveness of the combined topic modelling and dimensionality reduction strategy for binary stance classification tasks.

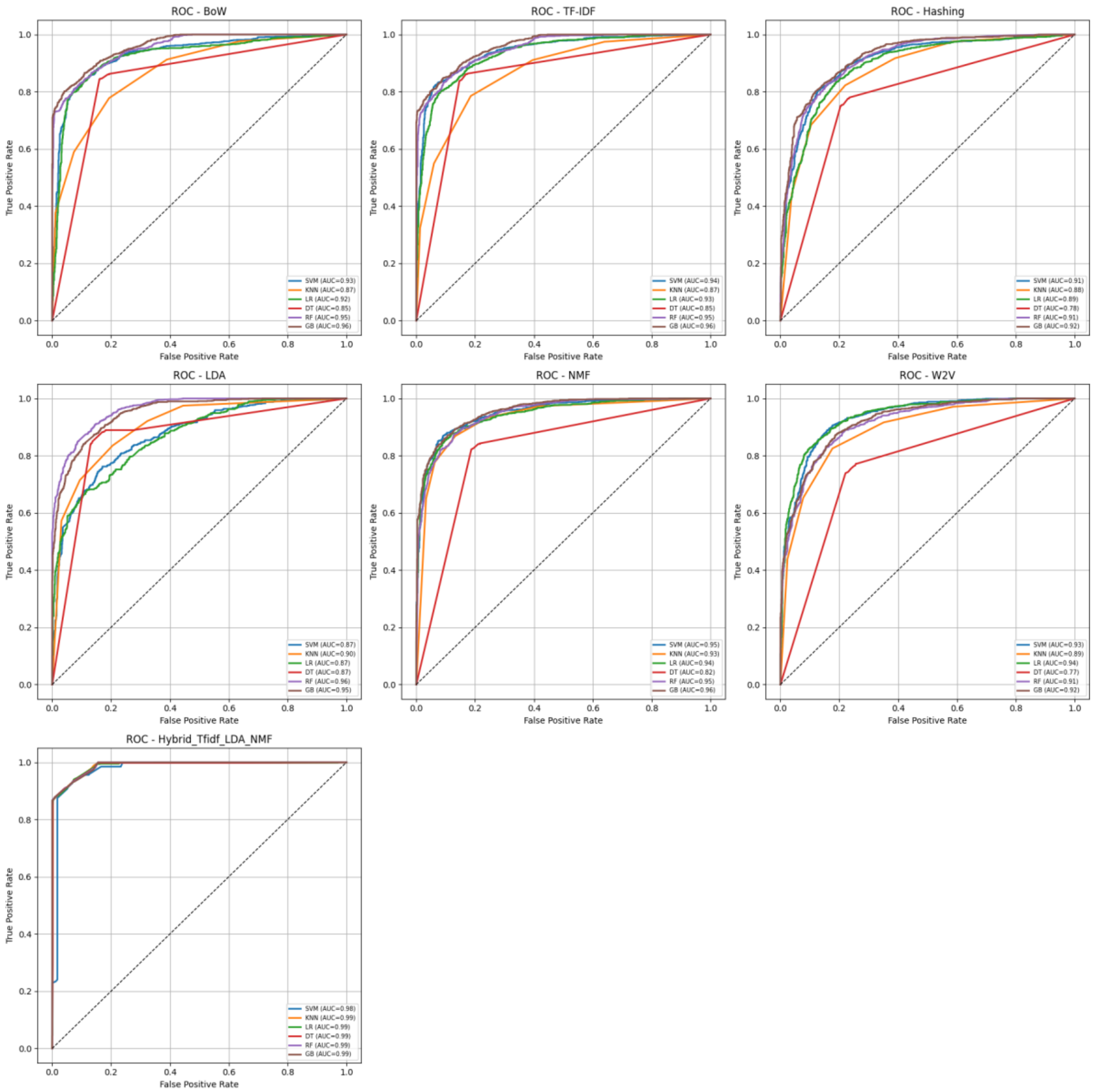

Figure 14. ROC-AUC scores for all classifiers on the Roe v Wade dataset

IMDB:

The ROC curves and AUC metrics for all classifiers were evaluated on the IMDB data. The six panels show the ROC curves of the different feature extraction techniques and types of classifiers depicted in Figure 15, in which the TPR versus FPR curves of the seven methods are displayed. We note that the AUC scores of Hybrid were consistently superior over the baseline methods, where Hybrid ranged between 0.75-0.85 and baselines from 0.55-0.70. The considerable vertical distance between the hybrid method curve and the curves based on traditional approaches suggests that with comparable false-positive rates, we can achieve much higher true-positive rates. This indicates that the hybrid method is more sensitive and specific for sentiment detection. Hybrid methods perform better than linear (SVM, LR) and nonlinear (DT, RF, GB, KNN) classifiers. This shows that hybrid methods are robust and generalizable to document-level sentiment analysis. This thorough ROC analysis clearly shows that the LDA-NMF hybrid feature extraction method is a better alternative that yields superior discriminative ability and classification reliability for various ML architectures on different text domains.

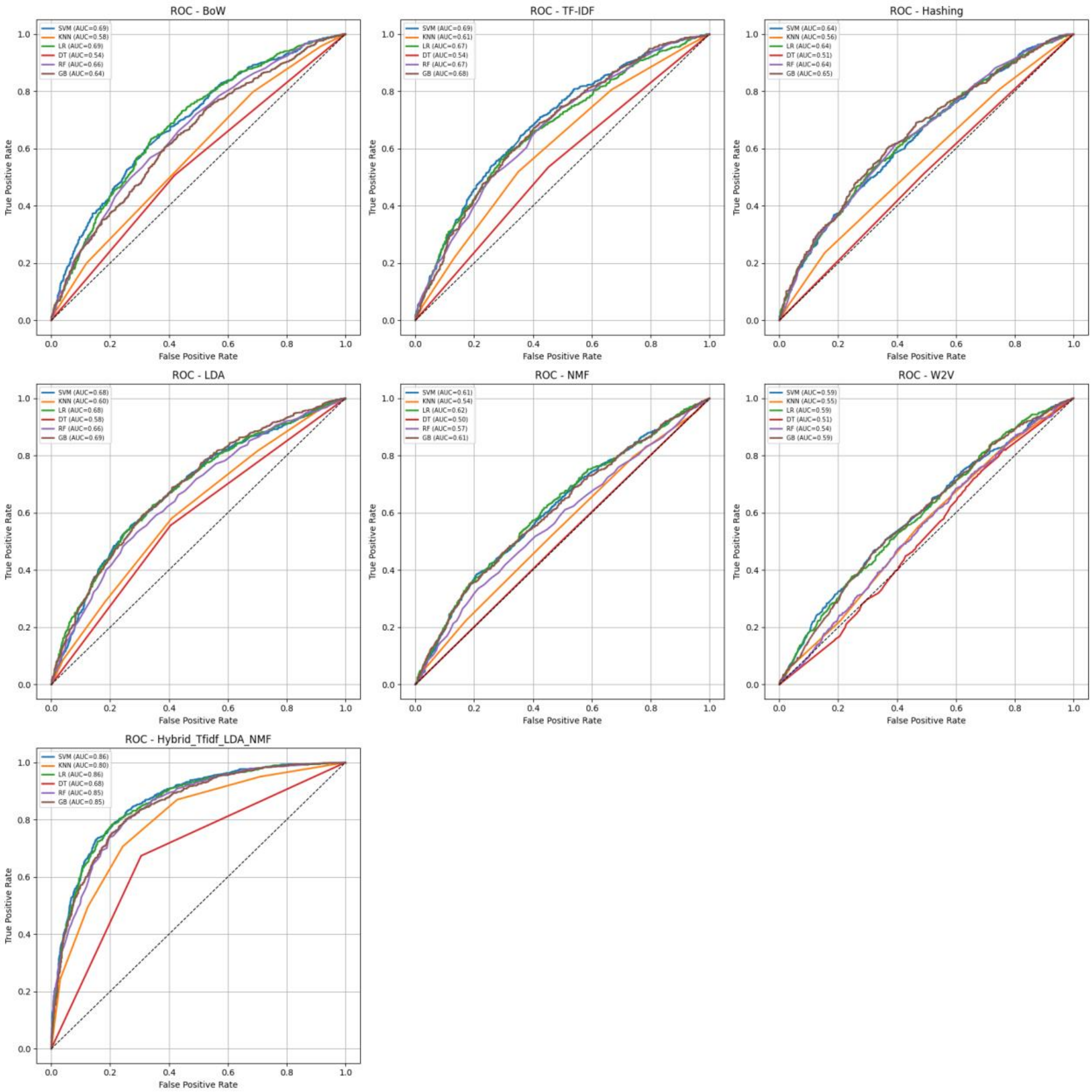

Figure 15. ROC-AUC scores for all classifiers on the IMDB dataset

4.8 Scalability Analysis

Figures 16–17 demonstrate that the training time increased nearly linearly with the dataset size. The proposed technique exhibits stable scalability and maintains consistent performance across varying data volumes.

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 8s (2025)

The six classifiers showed training times as the dataset sizes changed for the Roe vs. Wade dataset. The multipanel line graph depicts the computational efficiency for different data sizes. Each subplot corresponds to a different classifier, namely, SVM, KNN, LR, DT, RF, and GB. The size of the dataset is shown on the x-axis (in number of samples), while the time taken to train (in seconds) is on the y-axis with a logarithmic scale displayed in Figure 16, where each sub-frame of the figure is represented by seven colored lines showing the various feature extraction techniques: BoW, TF-IDF, Hash, LDA, NMF, W2V, and Hybrid. The proposed hybrid approach grows with nearly linear time complexity as the dataset size increases, and it is comparable to or better than the existing methods. The hybrid method curve was stable for all the classifiers. Therefore, we can say that the computational power is not that high an overhead. The computational power was consistent throughout. However, hybrid methods can create more sophisticated semantic representations. According to this analysis, the LDA-NMF technique possesses practical scalability while also achieving superior accuracy, which means that it can be applied successfully in the real world for large-scale social media text analysis together with social media stance detection tasks.

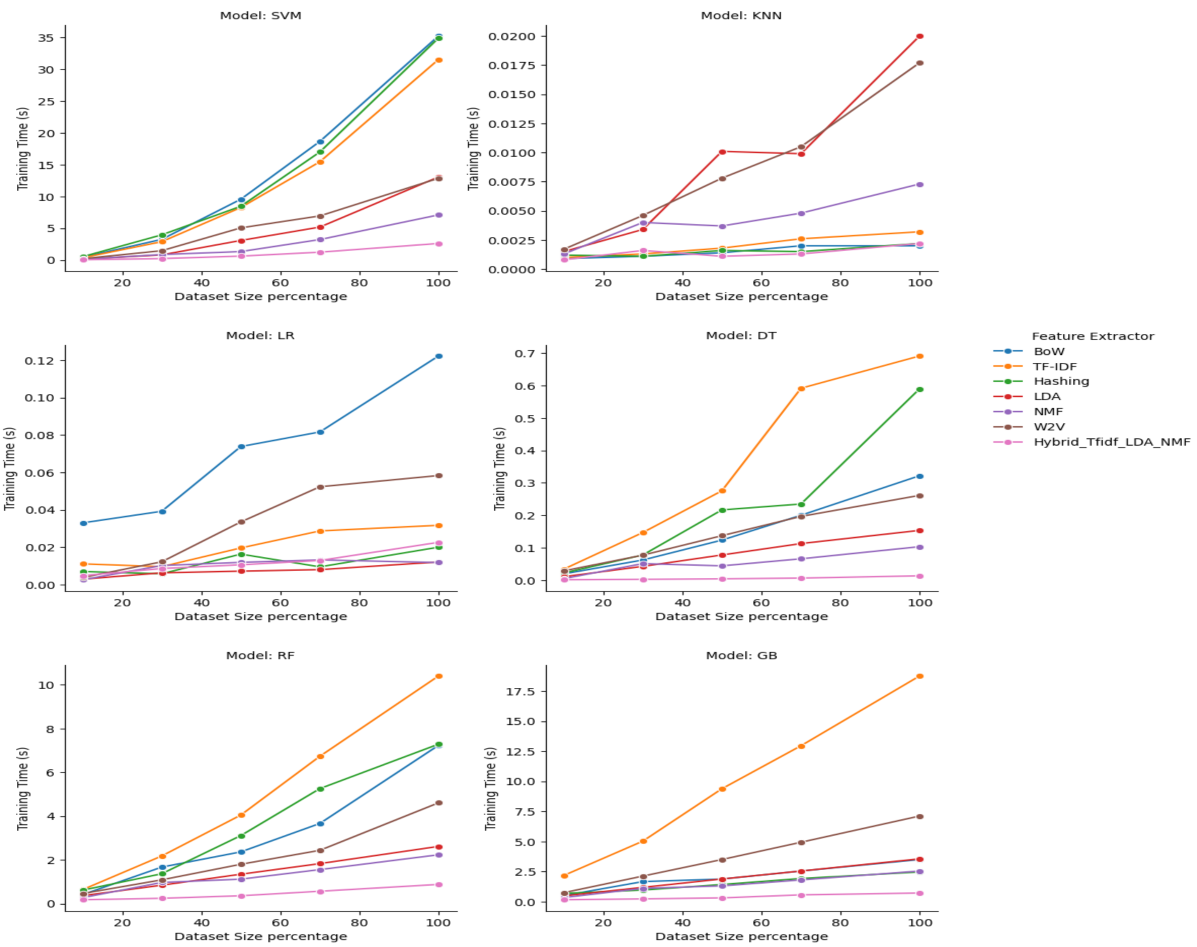

Figure 16. Scalability Analysis for Roe V Wade Dataset

Figure 17 illustrates the impact of the dataset size on the training performance with IMDB movie reviews. The six-panel visualization illustrates the complexity of each classifier, showing how long it takes to train the classifier (y-axis, log scale) as a function of the dataset volume (x-axis). The Hybrid LDA-NMF method exhibits steady and predictable scaling behavior for all classifier architectures, which follows time complexity growth rates similar to those of simple feature extraction methods. The hybrid technique produces richer representations and remains faster than baseline methods. This combines topic modelling and dimensionality-reduction techniques. The slope of the hybrid methods is flat when data volumes are increased, indicating that the option is practically viable in nature. The training time of the methods is not excessive, even when they are complex, such as Gradient Boosting and Random Forest. The scalability study revealed that the proposed hybrid feature extraction technique achieves a proper trade-off between classification accuracy and computational cost, which makes it suitable for production-level sentiment analysis systems that deal with a large volume of text.

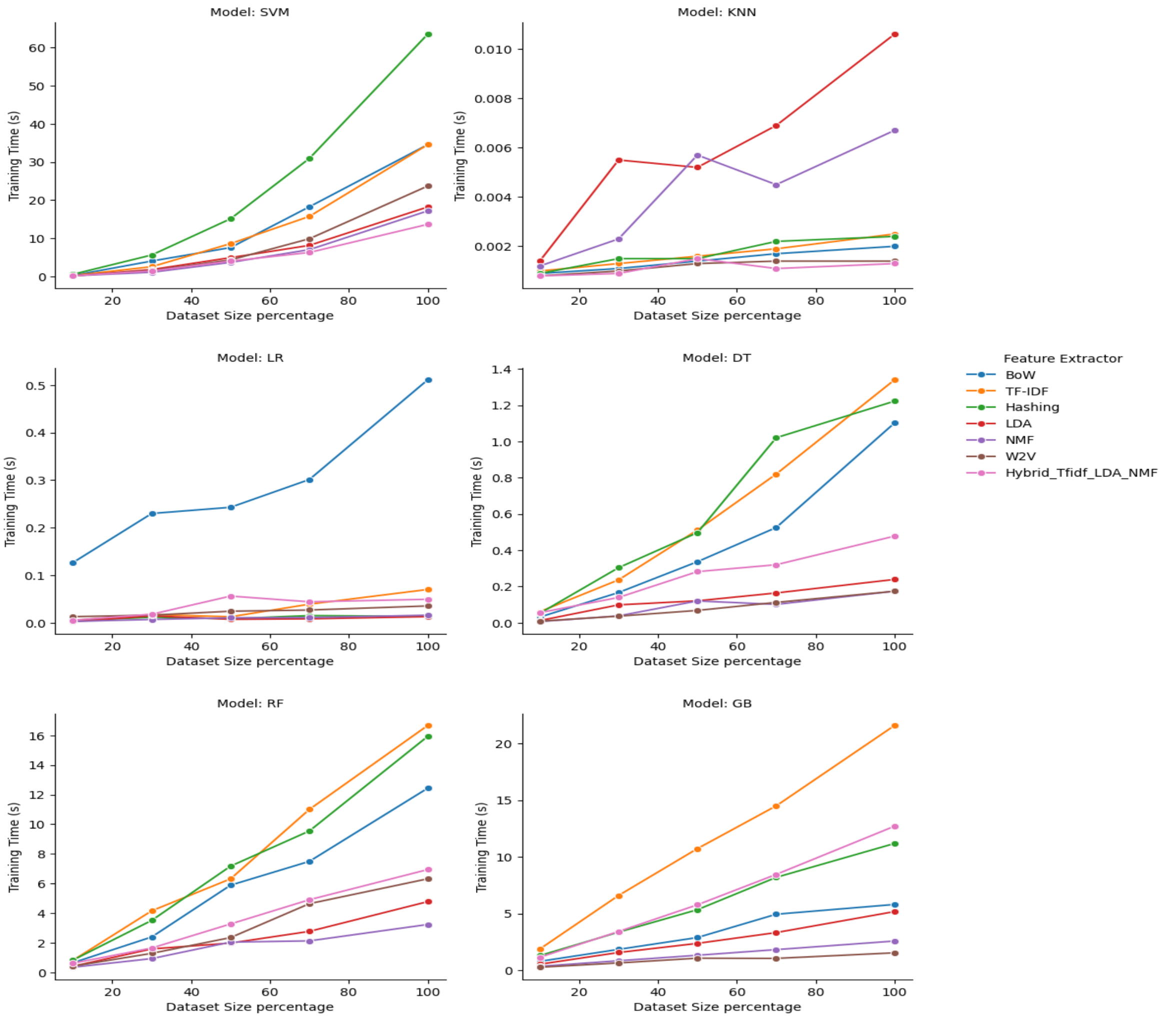

Figure 17. Scalability Analysis for IMDB Dataset

V CONCLUSION

This study proposes a hybrid feature extraction approach integrating LDA for topic-based feature extraction and NMF for dimensionality reduction to enhance stance-based sentiment classification. The experimental results demonstrate that the LDA-NMF hybrid approach outperforms traditional feature extraction techniques (BoW, TF-IDF, Word2Vec, LDA, NMF, and Hashing Vectorizer) by achieving the lowest misclassification rate (1604) and the highest classification accuracy across multiple classifiers. Notably, all classifiers demonstrated improved accuracy when using the proposed hybrid approach compared to existing techniques. For instance, Support Vector Machine (SVM) achieved 94.19% accuracy, while Gradient Boosting (GB) reached 93.71%, reflecting the effectiveness of the LDA-NMF technique in generating refined, low-dimensional semantic representations. This consistent performance improvement across the classifiers further validates the robustness of the proposed approach.

By combining semantic feature extraction with dimensionality reduction, the proposed method effectively addressed feature sparsity, enhanced classifier decision boundaries, and improved sentiment differentiation in stance classification tasks. These findings emphasize the importance of structured topic modeling and feature refinement in natural language processing applications. Furthermore, LDA uncovered hidden themes in the Roe v Wade case, revealing discourses across politics, elections, feminism, bodily autonomy, religion, and human rights. Themes related to children's lives and emotions toward unborn children emerged, reflecting the deep personal and ethical dimensions of the debate. This thematic analysis highlights the multidimensional nature of public opinion, demonstrating how a single case can intersect multiple societal and ideological domains. These insights further reinforce the significance of topic-driven sentiment analysis for understanding the complexity and interconnectivity of public discourse.

However, as the dataset size increases, traditional machine-learning models struggle with scalability. While the LDA-NMF approach proved effective for small-to-moderate datasets, deep learning models can enhance performance on larger datasets by capturing richer semantic representations. Future research can explore hybrid deep-learning models that integrate topic-based feature extraction with advanced neural architectures to improve scalability. Additionally, real-time stance classification of streaming social media data presents an opportunity to enhance dynamic opinion analysis. Finally, cross-domain generalization can extend the LDA-NMF approach beyond stance classification, enabling broader applications in sentiment analysis and NLP-driven opinion mining. The results confirm that the proposed LDA-NMF hybrid approach provides a scalable, efficient, and interpretable solution for stance-based sentiment classification, offering a strong foundation for future advancements in NLP-driven opinion mining.

Reference

[1] R. Ahuja, A. Chugh, S. Kohli, S. Gupta, and P. Ahuja, "The Impact of Features Extraction on the Sentiment Analysis," International Conference on Pervasive Computing Advances and Applications (PerCAA), pp. 341-348, 2019.

[2] K. S. Kalaivani, S. Uma, and C. S. Kanimozhiselvi, "A Review on Feature Extraction Techniques for Sentiment Classification," Proceedings of the Fourth International Conference on Computing Methodologies and Communication (ICCMC 2020), 2020, pp. 679-683, IEEE Xplore, Part Number: CFP20K25-ART; ISBN: 978-1-7281-4889-2.
[3] N. A. Semary, W. Ahmed, K. Amin, P. Pławiak, and M. Hammad, "Enhancing machine learning-based sentiment analysis through feature extraction techniques," PLoS ONE, vol. 19, no. 2, Feb. 2024.
[4] L. P. Hung, R. Alfred, and M. H. Ahmad Hijazi, "A performance comparison of feature selection methods for sentiment classification," in Studies in Computational Intelligence, R. Alfred et al., Eds. Springer Nature, vol. 488, pp. 21–30, 2018.
[5] Z. Zhan, "Comparative Analysis of TF-IDF and Word2Vec in Sentiment Analysis: A Case of Food Reviews," ITM Web of Conferences, vol. 70, p. 02013, 2025.
[6] T. Ali, B. Omar, and K. Soulaimane, "Analyzing tourism reviews using an LDA topic-based sentiment analysis approach," MethodsX, vol. 2022, pp. 1-13, 2022.
[7] M. C. Rahmadan, A. N. Hidayanto, D. S. Ekasari, B. Purwandari, and T. Theresiawati, "Sentiment Analysis and Topic Modelling Using the LDA Method related to the Flood Disaster in Jakarta on Twitter," 2020 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS), Depok, Indonesia, 2020, pp. 1-6, IEEE Explore, 2020.
[8] V. Roseline, G. Heren Chellam, “Role Of Sentiment Analysis Using Deep Learning," International Journal of Emerging Technologies and Innovative Research, ISSN:2349-5162, Vol.6, Issue 2, page no.116-122, February-2019.
[9] F. A. Farkhod, A. Abdusalomov, F. Makhmudov, and Y. I. Cho, "LDA-Based Topic Modeling Sentiment Analysis Using Topic/Document/Sentence (TDS) Model," Applied Sciences, vol. 11, no. 23, p. 11091, 2021.
[10] S. Lee and H.-S. Pang, "Feature Extraction Based on the Non-Negative Matrix Factorization of Convolutional Neural Networks for Monitoring Domestic Activity With Acoustic Signals," IEEE Access, vol. 8, pp. 150177-150186, 2020.
[11] D. Marchette et al., "Enhancing Text Analysis via Dimensionality Reduction," in Proceedings of the 2007 IEEE International Conference on Information Reuse and Integration (IRI), Aug. 2007.
[12] C. O. S. Sorzano, J. Vargas, and A. Pascual-Montano, "A survey of dimensionality reduction techniques,"arXiv:1403.2877, 2014.
[13] A. Huertas-García, A. Martín, J. Huertas-Tato, and D. Camacho, "Exploring Dimensionality Reduction Techniques in Multilingual Transformers," Cognitive Computation, vol. 15, no. 3, pp. 590–612, 2023.
[14] K. M. Monica and R. Parvathi, "Traditional Dimensionality Reduction Techniques using Deep Learning," International Journal of Recent Technology and Engineering (IJRTE), vol. 8, no. 3, pp. 7153, Sep. 2019.
[15] A. A. Mohamed, “An effective dimension reduction algorithm for clustering Arabic text,” Egyptian Informatics Journal, vol. 21, pp. 1–5, 2020.
[16] K. P. Das, J. C. Hapuwitharana, J. Fowler, and S. S. Young, “Non-negative matrix

factorization: A useful method for two-way life tables," Journal of Statistics and Management Systems, 2020. doi: 10.1080/09720510.2020.1756048.

[17] E. Muningsih, H. M. Nur, F. F. D. Imanianwan, S. Saifudin, V. R. Handayani, and F. Endiarto, "Comparative Analysis on Dimension Reduction Algorithm of Principal Component Analysis and Singular Value Decomposition for Clustering," Journal of Physics: Conference Series, vol.1641, no. 1, p. 012101, 2020.

[18] N. Sabri, H. N. Abdull Hamed, Z. Ibrahim, K. Ibrahim, M. A. Isa, and N. M. Diah, "The hybrid feature extraction method for classification of adolescence idiopathic scoliosis using Evolving Spiking Neural Network," Journal of King Saud University – Computer and Information Sciences, vol. 34, pp. 8899–8908, 2022.

[19] Z. O. Zoya, S. Latif, F. Shafait, and R. Latif, "Analyzing LDA and NMF Topic Models for Urdu Tweets via Automatic Labeling," IEEE Access, vol. 9, 2021.

[20] Z. Ahanin, M. A. Ismail, N. S. S. Singh, and A. AL-Ashmori, "Hybrid Feature Extraction for Multi-Label Emotion Classification in English Text Messages," Sustainability, vol. 15, no.12539, 2023.

[21] S.-W. Kim and J.-M. Gil, "Research paper classification systems based on TF-IDF and LDA schemes," Human-Centric Computing and Information Sciences, pp. 1-24, 2019.

[22] A. I. Al-Ghadir, A. M. Azmi, and A. Hussain, "A novel approach to stance detection in social media tweets by fusing ranked lists and sentiments," Information Fusion, vol. 67, pp. 29–40, 2021.

[23] N. Alturayeif, H. Luqman, and M. Ahmed, "A systematic review of machine learning techniques for stance detection and its applications," Neural Computing and Applications, vol.35, pp. 5113–5144, 2023.

[24] F. Arias, M. Z. Núñez, A. Guerra-Adames, N. Tejedor-Flores, and M. Vargas-Lombardo, "Sentiment Analysis of Public Social Media as a Tool for Health-Related Topics," IEEE Access, vol. 10, pp.74850-74872, Jul. 2022.

[25] Ramalho, J. Jorge, and S. Gama, "Representing uncertainty through sentiment and stance visualizations: A survey," Graphical Models, vol. 129, p. 101191, 2023.

[26] D. Küçük and N. Arıcı, "Stance and Sentiment Analysis of Health-related Tweets with Data Augmentation," Journal of Scientific & Industrial Research, vol. 83, no. 4, pp. 381-391, Apr. 2024.