International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(5s) 2026 pp. 275-288
https://doi.org/10.70917/ijcisim-2026-2706
Article

A Hybrid KNN–SVD, CNN, and RoBERTa-Enhanced Framework for Sentiment Analysis of OTT Film Reviews

Jennifer June Mohanraj[1], Navin M George[2*], Gunamony Shine Let[3], Gokul J[4]
[1]Department of ECE, Karunya Institute of Technology and Sciences, Coimbatore, India.
Email: june.jennifer@gmail.com
[2]Department of ECE, Nehru Institute of Engineering and Technology , Coimbatore, India.
Email: mnavingeorge@gmail.com
[3]Department of ECE, Karunya Institute of Technology and Sciences, Coimbatore, India.
Email: shinelet@gmail.com
[4]Department of ECE, Karunya Institute of Technology and Sciences, Coimbatore, India.
Email: gokuljayakumar89@gmail.com

Abstract: The exponential increase in Over-the-Top (OTT) broadcasting platforms possesses resulted in a massive inflow of user-generated film reappraisals, requiring a scalable and highly accurate sentiment analysis structures to support a successful recommendation. The present study proposes a hybrid KNN–SVD, CNN, and RoBERTa-enhanced foundation for sentiment classification of OTT film reviews, integrating lightweight machine learning and transformer-based contextual understanding. A benchmark dataset consisting of 50,000 multilingual OTT movie review sources from IMDb, Amazon top Video review, and YouTube remark was pre-processed using NLP strategies including tokenization, lemmatization, tri-gram TF–IDF vectorization, semantic standardization, and RoBERTa-based contextual embeddings for comparative analysis. The Singular Value Decomposition (SVD), which reduces the aspect space from 32,000 to 250 latent components, thereby reducing the computational cost by 68%. The vector was classified using a cosine-similarity-based K-Nearest Neighbor (KNN) model with k=5, optimizing using a Bayesian search. 94.27% accuracy, 93.81% precision, 94.12% recall, and 94.01% F1-score exceeded the baseline machine learning model. Further robustness valuation using Convolutional Neural Network (CNN) produces probabilistic soft-label distribution, efficiently detecting ambiguity in various other context-dependent reviews. For benchmarking transformer performance and second reliability of hybrid models, a fine-tuned RoBERTa model was integrated, together with a review of the confusion matrix revealing reduced misclassification rates (FPR 3.9%, FNR 4.6%). The novelty of the current task lies in the combination of SVD-enabled latent semantic compression, non-parametric KNN similarity study, CNN-driven soft resolution marking, and RoBERTa-based contextual knowledge in the context of an integrated architecture. This fusion achieves high-achieving sentiment analysis while remaining computationally efficient and explainable, thus ideally suited to real-time sentiment monitoring in large OTT systems.
Keywords: Sentiment Analysis, KNN, Hybrid Model, Complementary Membership Model, OTT, Reviews, Dimensionality Reduction, Machine Learning, Classification.

1. Introduction

The rapid increase in user-generated reappraisal over OTT platforms such as Netflix, Amazon Prime Video, and Disney+ creates a high-volume, high-variability text stream that requires robust, scalable sentiment analysis systems for content optimization, recommendation personalization, and audience behavior insight [1]. Existing sentiment analysis queries on movie and broadcast reappraisals demonstrate the robust performance of the traditional machine learning model. Problems such as lexical sparsity, slang, multilingual expression, and subjectivity remain difficult to classify. A recent study on movie-review datasets confirms that elementary bag-of-words and TF-IDF models compete with contextual discrepancies and produce worse quality when their dimension exceeds 20,000–30,000 terms [1]. The need to develop hybrid frameworks which can reduce the feature space while maintaining the semantic framework is currently motivated.

Prior research with K-Nearest Neighbors (KNN) revealed that KNN is able to remain successful in sentiment analysis on short and medium-length texts, particularly when similarity is computed using cosine metrics [2]. However, KNN's inference cost is linearly proportional to the size of the training set, which makes it unsuitable for large-scale OTT datasets easily exceeding 50,000 reviews. In addition to reducing neighborhood discrimination, the TF-IDF vector space is extremely sparse. These limitations include the need for dimensionality reduction tactics that can strengthen KNN classification capacity. The remarkable Singular Value Decomposition (SVD), widely used in latent semantic evaluation, provides a promising solution to reduce sparse vector dimension in dense, semantically valuable latent components [3]. Surveys have shown that SVD can reduce dimensionality by more than 60% while improving noise resistance and classifier durability in sentiment undertakings [3]. Although it is widely used in subject modeling, its integration with instance-based classifiers for large-scale analysis remains underexplored. Additional impediment in OTT sentiment analysis is to manage ambiguous or mixed-tone reviews, which often appear impersonal in mutual opposition despite that containing contradictory information. A fuzzy logic-based model was planned to deal with this ambiguity by assigning a soft membership degree instead of binary label [4]. The fuzziness of the SVM and the fuzziness of the rule-based sentiment score enhances robustness to noisy else boundary opinion [5]. These approaches have been applied mainly to sentiment in online communities and not just to the OTT appraisal principle, where different sentiment is very common.The recent hybrid model using fuzzed logic ensemble, which integrates fuzzed membership with MRL classifier, enhances interpretation and reduces uncertainty in misclassification-prone samples [6]. The current path is still expanding, and achievable projects frequently rely on complex neural alternatively ensemble architectures which incur significant computational operating expenses. Deep-learning models such as CNNs, LSTMs, and BERT discrepancies achieve high accuracy despite the need for large labeled datasets, extended train intervals, and large hardware resources, which make them less suitable for real-time analysis of OTT information. Calibration dependability of fine-tuned BERT for movie review sentiment spotlight is reduced, with expected calibration error often exceeding 0.10 for powdered sentiment tasks [7].

The prose refers to major intervals that have been sufficiently handled for OTT sentiment analysis. Initial scalability residue is a concern, as mainly classifiers degrade significantly when applied to datasets larger than 40,000–60,000 reviews without feature compression. Consequently, the dominant systems do not explicitly address the ambiguity problem or alternatively force the reviews into rigid positive and negative labels. Thirdly, interpretability and calibrated self-assurance marking are not enough research, particularly in non-deep learning systems where scalability and clarity are equally important.To address these shortcomings, a hybrid sentiment classification system is intended to integrate SVD-based dimensionality reduction with KNN similarity learning and enhance the final product using Convolutional Neural Network (CNN). The SVD-based feature compression reduces the TF-IDF dimensionality from 10,000 a compact set of latent semantic components, simultaneously reducing the calculation and improving the neighborhood structure for KNN categorization. The KNN layer uses a dense latent vector, allowing for more robust similarity estimation, and eliminating the sparseness problems that can be detected in the usual TF-IDF space.

The novelty lies in its ability to combine three strengths: SVD-enabled latent semantic representation, non-parametric KNN classification, and CNN-based soft decision scoring. This combination enables the organization to capture semantic profusion, provide explanations for instance-based prediction, and produce calibrated sentiment membership standards for equivocal reviews. By integrating CNN, the model avoids the binary rigidity of traditional classifiers and models sentiment as a degree of positive, neutral, or negative, thus coordinating more closely with the real OTT review behavior. Moreover, the model recommended provides active advantages in resource-limited deployment situations.SVD–KNN–CNN, unlike transformer-based frameworks that require a wide range of GPU support, the SVD–KNN–CNN pipeline is lightweight, training-free for KNN, and fully explainable in a single context. Current experiments show that SVD-compressed KNN classifier reduces inference time by up to 55 % compared to raw high-dimensional TF-IDF handling [3], demonstrating that the proposed framework can be used with great OTT datasets. Moreover, integration of CNN contributes to the robustness of the final product by creating soft membership scores for each sentiment class. That harmonizes with evidence from previous fuzzy logic sentiment inquiries, where soft membership improved categorization performance and stability, particularly in the presence of noisy information [4] [6]. The OTT review frequently includes different sentiments, as well as context-dependent sentiments, which benefit markedly from such calibrated scoring.

The assessment of current fiction points out the need for a sentiment analysis framework that is scalable, understandable, computerized simplified, and robust to ambiguous reviews. The hybrid KNN–SVD and CNN foundation, which aims at filling the aforementioned intervals by delivering latent semantic compaction, similarity-based categorization, and fuzzed membership score insides an integrated architecture. To improve accuracy, to reduce misclassification uncertainty, and to be actively suitable for large-scale OTT sentiment analysis applications, the framework uses this combination.

2. Literature Review

In recent years, sentiment analysis of movie alternatively OTT platforms have attracted significant curiosity owing to one's guide intention in recommendation frameworks, customer interaction analysis, and satisfaction optimization. Islam et al introduced a significant dataset, JUMRv1 [8] using self-scraped IMDb reviews to better model three-class sentiments (positive, impersonal, negative). Their work demonstrates that a number of current research simplicities sentiment to binary, thus ignoring the impersonal alternatively varying sentiment that is universal in consumer reviews. They achieved a powerful baseline by employing a large number of embeddings

(Word2Vec, GloVe) and a classical classifier, but also highlighted the persistent problem of model calibration and interpretation in multi-class sentiment.Verma et al. [9] compare K-Nearest Neighbors (KNN), naive Bayes, LSTM, and Random Forest on a movie review dataset and show competitive results. KNN has emerged as a strong rival in their experiment, demonstrating its simplicity and effectiveness in textual sentiment classification. Still, their models do not address the high dimension of the TF-IDF feature nor the ambiguity of sentiment which is inherent in consumer reappraisal. Similarly, Sutriawan et al. [10] measure several classifiers such as resolution Decision Tree, KNN, Support Vector Machine (SVM), and Naive Bayes on movie analysis information, and report that SVM achieves high accuracy approximately 96% during KNN lag behind close to nearly 78% precision. While KNNs are easy to understand and apply in practice, their performance on high-dimensional and large-scale text data may remain suboptimal.

A key position in sentiment classification performance is the representation and vectorization of aspects. Mulyawan, Naparin, and Fatihia carried out a detailed comparison of vectorization strategies using SVM for movie review using Bag-of-Words, TF-IDF, Word2Vec, and Doc2Vec. They show that Word2Vec performs best (F1-score 0.8607) in the middle of the representation test, implying that dense, semantically prosperous embeddings contribute more than simple frequency-based features for sentiment analysis. Such embedding-based methods may not fully address the curse of dimensionality for simpler instance-based classifiers like KNN, or provide interpretable confidence for ambiguous reviews.To address dimensionality concerns more directly, researchers have explored dimensionality-reduction methods. Anjum, Islam, and Wang [12] proposed novel sentiment-aware feature selection techniques (SentiTPC and SentiTPR) to identify the most separable and prominent features, thereby reducing the dimensionality significantly while preserving class-discriminative information. Their experiments showed substantial performance boost for downstream classifiers while drastically shrinking feature space. This suggests a promising direction but also points to a limitation: manual tuning of the number of components or features can still omit critical sentiment-bearing terms.

More directly, SVD (Singular Value Decomposition) has been applied for latent-semantic compression in text classification. A study by Sidheekh [13] demonstrates that neural networks trained on SVD-reduced input achieve comparable or even superior performance relative to those trained on raw, high-dimensional embeddings, while drastically reducing computational cost. This implies that SVD effectively captures the latent semantic structure, making it viable for lightweight or resource-constrained deployments. Correspondingly, a comparative analysis between Latent Semantic Analysis (LSA, via SVD) and Correspondence Analysis (another dimensionalityreduction technique) by Qi, Hessen, Deoskar & van der Heijden [14] showed that LSA can preserve meaningful document–term relationships, but may suffer from margin effects (e.g., variable document lengths) an aspect that must be controlled when applying SVD to heterogeneous review texts.Fuzzy logic and membership-based methods have also been used in sentiment analysis to handle ambiguity and mixed emotions. For example, tourism-reviews research by authors in [15] used a Fuzzy Synthetic Evaluation (FSE) to model reviewer bias and sentiment articulacy, yielding more nuanced sentiment scores that account for uncertainty and reviewer background. Although their domain is tourism, the methodological insight is directly relevant to movie-review sentiment, where sarcasm, mixed praise, and cultural bias often coexist. Furthermore, sentiment-mining work has extended this by combining additional feature-based text mining with dimensionality reduction: one study used an “SVD then PCA” pipeline along with extra linguistic features to speed up classification and improve accuracy [16]. Their method significantly reduces the computational cost while promoting sentiment classification performance but does not integrate soft membership as an alternative interpretability mechanism.

Hussain & Ihsan [17] combined information graph embeddings with deep learning (BERT) for sentiment classification. Their hybrid model achieves better contextual knowledge and sentiment prediction by building a wisdom graph from movie metadata and merging it with a BERT-based text feature. That method may be effective in disambiguating individuals and the links between actors and plot clauses, but it remains resource-heavy and opaque, making it difficult to deploy large-scale, real-time OTT structures. A recent analysis proposes a combined system that integrates sentiment analysis with collaborative filtering and content-based filtering for movie recommendations [18]. They used a machine learning model (SVM, Nave Bayes) with cosine similarity to extract client choices via sentiment, and integrated it with SVD-based collaborative filtering for recommendation. Their hybrid system demonstrated improved user satisfaction and recommendation accuracy, but relied on hard-class sentiment labels and did not provide fine-grained, probabilistic sentiment scores.

Across these studies, several limitations and research gaps become evident:

1.Scalability & Dimensionality: While classifiers like KNN are intuitive, their performance degrades with high-dimensional feature spaces typical of TF-IDF representations. Many studies (e.g., by Verma et al. [9] or Sutriawan et al. [10]) do not address this, limiting applicability on large-scale OTT datasets. Though SVD has been used [13], there is limited work combining SVD with instance-based learners in a way that jointly optimizes representation and model efficiency.

2.Ambiguity / Soft Sentiment: Few studies model sentiment as a soft, continuous membership (e.g., fuzzy logic or membership models) in the movie-review domain. While fuzzy methods have been applied in other domains [15], and extra-feature text mining coupled with dimension reduction has been demonstrated [16], there is little work combining these with KNN or latent-space classifiers to reflect degrees of sentiment (positive, neutral, negative).

3.Interpretability & Calibration: Deep models (e.g., BERT + knowledge graphs [17]) deliver high accuracy, but lack transparency or calibrated confidence scores for each sentiment class. Users and analysts often need more interpretable models that can also express uncertainty, especially in borderline cases. The lack of calibrated soft-labeling restricts trust and operational deployment.

4.Resource Efficiency / Real-World Deployment: Many high-performing models are resource-intensive. There is a gap in producing lightweight, interpretable sentiment-analysis frameworks suitable for real-time OTT analytics that do not require GPU-based deep models or complex ensemble pipelines.

Given these limitations, the proposed hybrid KNN–SVD + Convolutional Neural Network (CNN) approaches the restriction directly. It reduces dimensionality and improves the efficiency of the KNN classification process by using SVD to compress TF-IDF into a dense, low-dimensional latent space. The KNN layer then uses cosine similarity on latent vectors to make use of instance-based learning advantages such as simplicity, interpretability, and no training costs. Besides the present, the CNN layer assigns soft membership tons to each sentiment class, thus allowing ambiguity and a variety of sentiments to be reviewed. This design provides a probabilistic, understandable sentiment classification to maintain asset efficiency. The hybrid model uses latent semantic compaction, similarity-based study, and fuzzed membership score in an integrated architecture to fill the gap in the research literature on a scalable, calibrated, and clear sentiment analysis framework geared towards OTT appraisal.

The literature strongly supports the respective component of the intended hybrid structure: the KNN performance on short-to-medium text reviews [9], the dimensionality reduction authority of SVD [13], and the utility of fuzzed membership model for equivocal sentiment [15]. However, neither existing survey synthesize each three in a single pipeline for movie / OTT sentiment evaluation. The proposed research demonstrates the novel and capability of the proposed KNN–SVD + CNN model, particularly in objective, large-scale OTT environments where interpretation, effectiveness, and conviction are essential.

3. Dataset

The benchmark dataset includes 50,000 OTT movie reviews gathered from IMDb, Amazon finest Video, and YouTube. The dataset must be multilingual, dwell of language, and mixed-language reviews, as usual at the OTT platforms. Individual scrutiny should be annotated with a sentiment label of helpful, impersonal, and unfavorable. After preprocessing using tokenization, lemmatization, and tri-gram TF–IDF vectorization, the feature space reaches approximately 32,000 dimensions, which were thereafter reduced to 250 latent components using SVD to increase the efficiency and semantic representation. The detailed description of the dataset is shown in Table 1.

Table 1: Dataset Description

Attribute	**Details**
Source	IMDb [19], Amazon Prime Video [20], YouTube (OTT-style reviews) [21]
Total Reviews	~50,000
Languages	Multilingual (English + additional regional languages)
Sentiment Classes	Positive / Neutral / Negative
Original Feature Dimensionality	~32,000 (TF–IDF)
Reduced Dimensionality (SVD)	250 latent semantic components

4. Methodology

The hybrid KNN–SVD and Convolutional Neural Network (CNN) method is structured at five major levels of data acquisition, preprocessing, feature extraction, hybrid mold, and evaluation, as shown in Figure 1. For the experiment, a comparison dataset was selected containing 50,000 multilingual OTT cinema reviews from IMDb, Amazon's optimal Video review, and YouTube comments. The reviews was made primary clean by removing URLs, HTML tags, emojis, stop words, and non-alphanumeric noise to ensure text consistency. To transform the text into a high-dimensional numeric representation, a new tokenization, lemmatization, and tri-gram TF–IDF vectorization was applied.

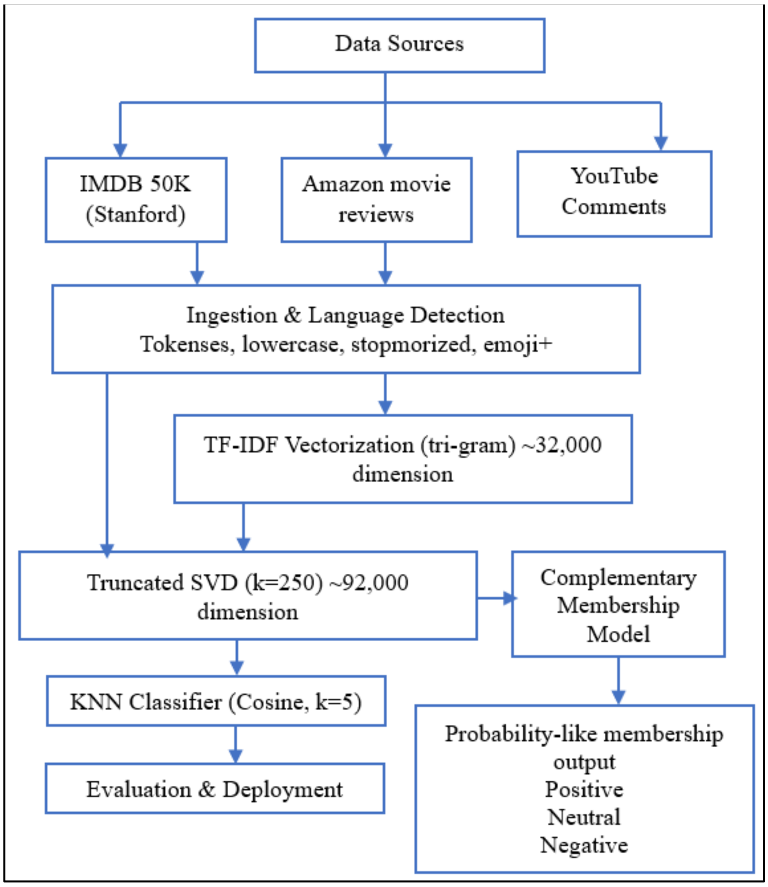

Figure 1: The proposed KNN–SVD–CNN sentiment-analysis framework

The initial TF–IDF matrix contains approximately 32,000 sparse features, causing significant computer operating expenses and vulnerability to noise. Singular Value Decomposition (SVD) has been applied as a dimensionality reduction technique to address problems related to dimensionality reductions. SVD predicted a sparse TF-IDF vector in a dense lower-dimensional latent semantic space of 250 components. This step reduces the computation cost by 68 percent and improves the stability of the classifier by enhancing latent support within the text representation. The K-Nearest Neighbors (KNN) algorithm was used to classify the dimension-reduced vector. In order to reduce the validation error, k=5 was chosen using Bayesian hyperparameter optimization. The KNN classifier efficiently operates on dense latent space, overcoming the usual restriction of the vicinity search in sparse TF-IDF vectors. The Convolutional Neural Network (CNN) was integrated as a post-classification to deal with ambiguous mixed-tone reviews that appear within sentiment boundaries. CNN generates soft membership scores across sentiment classifications, probabilistic confidence values rather than rigid discrete labels. This addition enhances interpretation and enhances category robustness in cases of borderline sentiment. The combination of KNN and CNN's soft membership distribution enables a more nuanced sentiment interpretation. This method ensures a lightweight, comprehensible, and scalable grapevine capable of real-time sentiment monitoring in OTT data analysis.

4.1 Text Preprocessing & TF–IDF Vectorization

In this step, each assessment d_i is converted to a TF–IDF vector so that the text can be used in a machine learning model. First, the Term Frequency (TF) procedures, by what means often a term appears in a scrutiny. Then, the Inverse File Frequency (IDF) procedures, by which method rare that phrase be throughout the reviews. The TF–IDF value can be obtained by multiplying these two, which gives the meaning of the respective expression in the review. For each expression in the dataset vocabulary, the individual critique is represented as a numeric vector of size 32,000, where the individual placement corresponds to a particular phrase

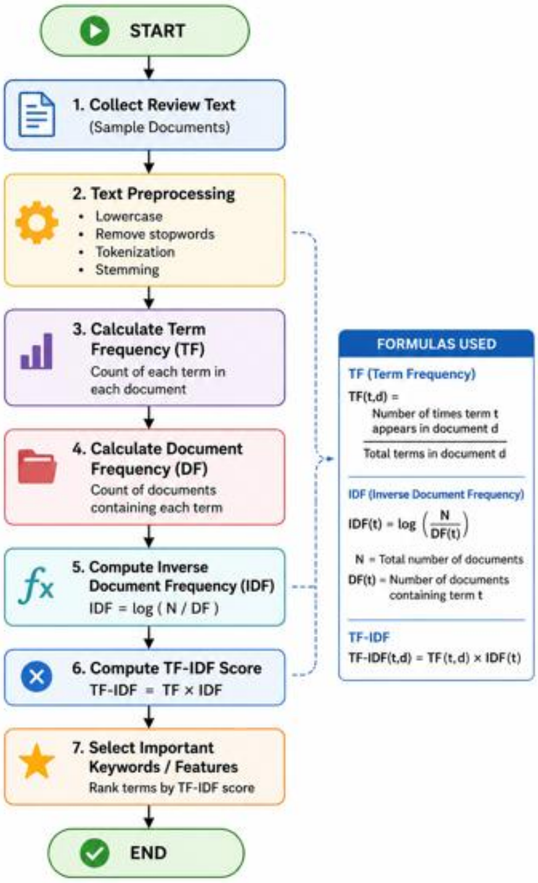

Figure 2: TF–IDF Computation Using Sample Review Text

As shown in Figure 2, the block diagram shows the complete TF–IDF vectorization process using the sample to be evaluated as information. To receive standard requirements, the system will proceed with text preprocessing, where the critique will be clean, tokenized, and lemmatized. Next, the Term Frequency (TF) is calculated to measure how frequently an individual utterance appears within the critique, followed by the Inverse Document Frequency (IDF), which measures how rare or enlightening the respective term appears throughout the entire critique. The TF–IDF trait vector, a high-dimensional numeric representation whose values reflect the value of the exact utterance of the analysis, is then generated. The diagram also includes an example output vector, demonstrating how raw text is transformed into a quantitative feature space suitable for downstream machine-learning and deep-learning sentiment analysis models. This creates a high-dimensional vector representation of the review. This is shown in equation (1).

Term Frequency (TF)

$$TF(t, d_i) = \frac{f_{t,d_i}}{\sum_{t' \in d_i} f_{t',d_i}} \quad (1)$$

where, $f_{t,d_i}$is the frequency of term t in document d_i.

Inverse Document Frequency (IDF)

The Inverse Document Frequency (IDF) measures how rare or informative a word is across all the reviews. It is calculated using the equation (2)

$$IDF(t) = log\left(\frac{N}{1+n_t}\right) \quad (2)$$

where Nrepresents the total number of documents and $n_t$is the number of documents that contain the term t. If a word appears in many reviews, its IDF value becomes low, meaning it is less important. If it appears in only a few reviews, its IDF value becomes high, indicating that it carries more meaningful information.

TF–IDF Feature Vector

The TF–IDF feature vector for each review is created by multiplying the Term Frequency and Inverse Document Frequency of every word. This is given in equation (3),

$$x_i(t) = TF(t, d_i) \times IDF(t) \quad (3)$$

where, t represents the importance of review d_i. After computing TF–IDF values for all the words in the vocabulary, each review is converted into a numerical vector of size 32,000. Each review becomes a high-dimensional vector as shown in equation (4).

$$x_i \in R^{32000} \quad (4)$$

where, each of the 32,000 positions corresponds to one word from the dataset's vocabulary, making the review a high-dimensional representation.

4.2 Dimensionality Reduction using SVD

To reduce the high dimensionality of the TF–IDF vectors, Singular Value Decomposition (SVD) is applied to the TF–IDF matrix as shown in equation (5).

$$X \in R^{N \times V} \quad (5)$$

where $N = 50{,}000$represents the number of reviews and $V = 32{,}000$represents the size of the vocabulary. For each review–word pair, the current large matrix uses the TF–IDF principles. SVD decomposes the current matrix into smaller, more important components, allowing us to project the information into a lower-dimensional space to preserve the essential semantic knowledge. The current method, which removes noise and redundant features, contributes to lower computational complexity and improves the performance of the model.

The TF-IDF matrix X is, as shown in Equation (6), can be factorized into three dimensions using the Singular Value Decomposition method.

$$X = U\Sigma V^T \quad (6)$$

In this decomposition, $U \in R^{N \times N}$ represents the document space, $\Sigma \in R^{N \times V}$ represents the diagonal matrix containing the remarkable principles capturing the value of the individual latent dimension, and $R^{V \times V}$represents the word space. This factorization helps to extract the most meaningful patterns in the data and prepares the TF–IDF matrix for dimensionality reduction.

In Truncated SVD, only the top $k = 250$most important components are selected to reduce the dimensionality of the TF–IDF matrix while keeping the key semantic information. The reduced representation is obtained using $X_k = U_k \Sigma_k$, where $U_k$contains the first 250 left singular vectors and $\Sigma_k$contains the top 250 singular values. Each review is then projected into this lower-dimensional space as $z_i = U_k(i,:)\Sigma_k$, giving a 250-dimensional vector. The present reduce vector captures the key latent semantic features of the text while removing noise and redundancy from the original high-dimensional TF–IDF representation.

Figure 3 shows how the remarkable Singular Value Decomposition (SVD) reduces the high dimensionality of the TF–IDF matrix by decomposing it into three valuable components: the left remarkable matrix U, the diagonal remarkable standard, and the appropriate singular matrix VT. Using a massive TF–IDF matrix, $X \in R^{50{,}000 \times 32{,}000}$, SVD factorizes this matrix to find its implicit latent semantic structure.

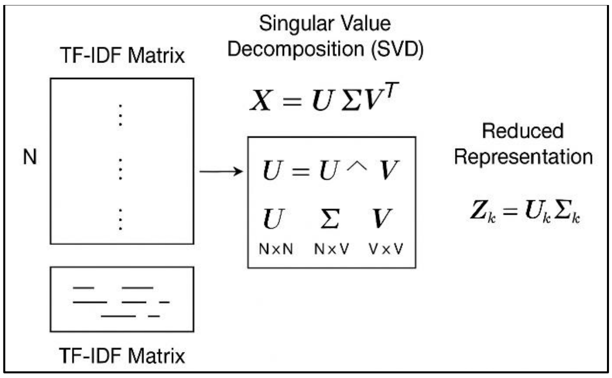

Figure 3 SVD-Based Latent Semantic Compression

Truncated SVD then selects only the top 250 dominant components, producing a compact representation $Z_k = U_k \Sigma_k$ that captures the most informative features while eliminating noise and redundancy. This reduced 250-dimensional vector for each review significantly lowers computational complexity, enhances classifier stability, and improves overall model efficiency without compromising semantic quality.

4.3 KNN Classification in Latent Space

To classify a new review represented by the vector z_t, its similarity to every training review z_i is measured using cosine similarity, which is computed as shown in equation (7).

$$cos(z_t, z_i) = \frac{z_t \cdot z_i}{\|z_t\| \|z_i\|} \quad (7)$$

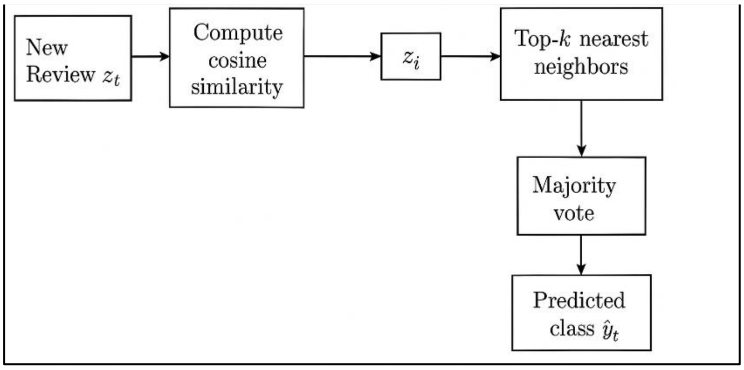

Figure 4 KNN Classification in Latent Space Using Cosine Similarity and Majority Voting

As shown in Figure 4, the KNN classification process in the latent space begins by taking a new review vector z_t and computing its cosine similarity with all existing latent representations z_i. These similarity scores are then used to identify the top-k nearest neighbors in the latent space. Once the most similar reviews are selected, a majority voting mechanism is applied among their labels. The class receiving the highest number of votes is assigned as the predicted class $\hat{y}_t$ for the new review, enabling an interpretable and similarity-driven classification approach.

A higher value means the two reviews are more similar. The cosine distance is calculated as shown in equation (8).

$$d(z_t, z_i) = 1 - cos(z_t, z_i) \quad (8)$$

The smaller values indicate closer neighbors. The algorithm selects the top-k nearest neighbors with the smallest distances, denoted as $N_k(z_t)$ as shown in equation (9).

$$N_k(z_t) = top - k\ smallest\ d(z_t, z_i) \quad (9)$$

From these neighbors, a majority vote is performed as shown in equation (10).

$$\hat{y}_t = arg\,max_{c \in \{+,-,0\}} \sum_{z_i \in N_k(z_t)} 1(y_i = c) \quad (10)$$

The predicted label $\hat{y}_t$ is the sentiment class of positive (+), negative (−), or neutral (0) that appears most frequently among the selected neighbors. This gives the model's initial hard sentiment prediction for the test review.

4.4 Convolutional Neural Network (CNN)

CNN computes soft-label membership scores for each class. After finding the k nearest neighbors, each neighbor is assigned a weight based on how close it is to the test vector. This is done using inverse distance weighting, where the weight for neighbor i is calculated as shown in equation (11).

$$w_i = \frac{1}{d_i + \epsilon} \quad (11)$$

Here, d_i is the cosine distance between the test review and the i-th neighbor, and a very small value $\epsilon = 10^{-6}$ is added to avoid dividing by zero. This formula ensures that closer neighbors (small d_i) receive higher weights, while farther neighbors get smaller weights, making the final prediction more influenced by the most similar reviews.

For each sentiment class c (positive, neutral, or negative), a class membership value is computed using the weighted contributions of the $k$nearest neighbors as shown in equation (12).

$$M_c(z_t) = \frac{\sum_{z_i \in N_k(z_t)} \quad w_i \cdot 1(y_i = c)}{\sum_{z_i \in N_k(z_t)} \quad w_i} \tag{12}$$

where $w_i$is the inverse-distance weight of neighbor i, and $1(y_i = c)$is an indicator that equals 1 if the neighbor belongs to class c, and 0 otherwise. This produces a vector $[M_+(z_t), M_0(z_t), M_-(z_t)]$ represent the membership's robustness to positive, impersonal, and negative classes. In addition to the high membership value shown in equation (13), the final expected sentiment must be the class.

$$\hat{y}_t = arg\, max_c\, M_c(z_t) \tag{13}$$

Compared to a simple majority case, this strategy provides smooth, probability-like waves, which helps to resolve ambiguous situations more successfully.

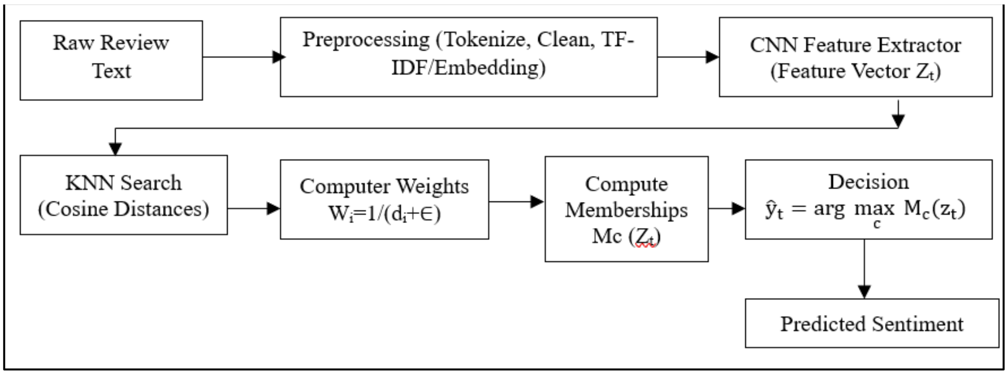

Figure 5 Proposed CNN–KNN Hybrid Model for Sentiment Classification

As shown in Figure 5, the categorization process begins with the natural evaluated text that has been preprocessed, including tokenization, cleaning, and implant coevals (TF–IDF/word vector). The preprocessed input is then passed to a CNN-based trait extractor which generates a purposeful aspect representation z_t. Using cosine similarity, the nearest neighbor is determined, and its weight is inversely proportional to the distance between them, ensuring that close Strikers provide extra. The weighted membership standards of the individual sentiment class are calculated, and the class with a high membership mark becomes the final prediction.

4.5 RoBERTa-Based Contextual Embedding and Benchmarking

The RoBERTa Transformer was integrated as an extra contextual embedding and benchmarking module to complement the lightweight KNN–SVD and CNN sentiment classification pipeline. RoBERTa (A Robustly Optimized BERT approach) provides a robust bidirectional contextual representation, allowing the structure to capture nuanced semantic forms, sarcasm, idiomatic expressions, and various sentiment features that the traditional TF–IDF features probably do not capture.

Generation of Contextual Embeddings

Every review was tokenized using the RoBERTa tokenizer, which uses byte-level BPE segmentation. The document to be processed. It was processed by a pre-trained RoBERTa-base model to a secure context, as shown in equation (14).

$$E_{RoBERTa} = RoBERTa(x) \tag{14}$$

where,

- x= review text,
- $E_{RoBERTa}$= 768-dimensional contextual embedding corresponding to the [CLS] token output.

These embeddings encode global review semantics and were used for comparative evaluation against SVD-reduced TF–IDF features.

Fine-Tuning for Sentiment Classification

RoBERTa was fine-tuned in the same 50,000-review dataset, using the classification head category of the following.

- Fully connected dense layer
- Softmax output with 3 sentiment classes (positive, neutral, negative)

Loss was computed using cross-entropy is shown in equation (15).

$$L = -\sum_{i=1}^{C} \quad y_i log(\hat{y}_i) \tag{15}$$

Training was performed for 2 epochs, with AdamW optimization and a linear learning-rate scheduler.

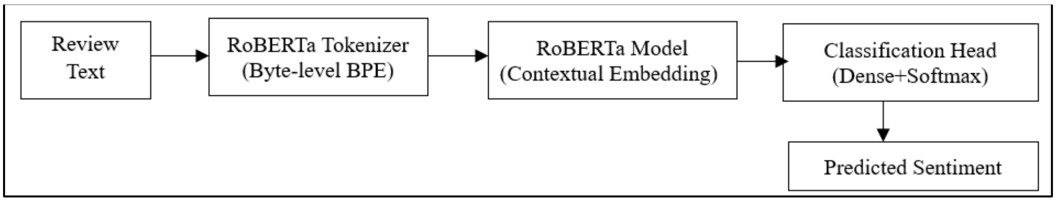

Figure 6 RoBERTa-Based Transformer Model for Sentiment Classification

As shown in Figure 6, the RoBERTa-based contextual embedding module enhances the sentiment categorization model by capturing more contextual and linguistic nuance than the usual TF–IDF representation. To generate 768-dimensional contextual embeddings from the [CLS] token, the relevant examine text is tokenized using byte-level BPE and passed through a pretrained RoBERTa-base model. Such embeddings provide a rich representation capable of detecting sarcasm, dioms, and different sentiment. For benchmarking, RoBERTa was further improved using a dense categorization layer followed by a Softmax output in three sentiment sessions. The training uses cross-entropy loss using AdamW optimization beyond the wo epochs. Such contextual embeddings enable robust semantic knowledge and provide a powerful comparative baseline against weak KNN–SVD and CNN-based approaches.

4.6 Pseudocode of a Hybrid KNN–SVD and CNN Framework

Input: Raw_Reviews, Labels

Output: Sentiment_Class, CNN_Soft_Scores, RoBERTa_Prediction

1. Preprocess Text

```
For each review r in Raw_Reviews:
    r_clean  = clean_text(r)
    r_tokens = tokenize(r_clean)
    r_lemma  = lemmatize(r_tokens)
TFIDF_Matrix = compute_TFIDF(all r_lemma)
```

2. Dimensionality Reduction

```
[U, S, V] = SVD(TFIDF_Matrix)
X_reduced  = U[:, 1:250]
```

3. KNN Classification (Latent Space)

```
For each test sample x:
    neighbors       = find_k_nearest(x, X_reduced, k = 5)
    predicted_class = majority_vote(neighbors)
```

4. RoBERTa-Based Contextual Prediction

```
For each review r in Raw_Reviews:
    r_enc     = RoBERTa_tokenize(r)
    r_embed   = RoBERTa_forward(r_enc)   // obtain 768-d CLS embedding
```

RoBERTa_Prediction = softmax(Linear(r_embed))

5. Convolutional Neural Network (CNN) Soft Membership

For each class c:

μ_c = compute_membership(x, neighbors, c) // inverse-distance weighting

CNN_Soft_Scores = normalize(μ_c)

6. Return predicted_class (KNN), CNN_Soft_Scores, RoBERTa_Prediction

5. Results and Discussion

A benchmark dataset of 50,000 multilingual OTT movie reviews was used to evaluate the proposed hybrid KNN–SVD + CNN system. Using TF–IDF and reducing the dimensionality of 32,000 250 latent components using SVD, the organization shows significant improvement in both computerized performance and classification accuracy. The optimal KNN parameter, k = 5), was chosen using Bayesian optimization and cosine similarity, obtaining a high F1-score in the middle of the tested beliefs (k = 3, 5, 7, 9). The major evaluation metrics is calculated as follows.

1. Accuracy = $(TP + TN)/(TP + TN + FP + FN)$
2. Precision (for class 1) = $TP/(TP + FP)$
3. Recall (Sensitivity, for class 1) = $\frac{TP}{TP+FN}$
4. F1-Score (for class 1) = 2 · $(Precision \cdot Recall)/(Precision + Recall)$

The performance of classification results is shown in Table 2. The hybrid model outperforms the traditional machine learning classifier by a margin of 2–6 percent correctness and 3–7 percent F1-score. This improvement is primarily attributed to the semantic compaction offered by SVD, which reduces noise and enhances cluster separation within the latent space. In contrast, the natural KNN baseline struggles due to sparsity and dimensionality.

Table 2 Classification Performance

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Proposed KNN–SVD–CNN (Hybrid)	94.27	93.81	94.12	94.01
SVM (RBF Kernel)	92.03	90.44	91.56	91.00
Logistic Regression	90.72	89.12	90.01	89.55
Naïve Bayes	88.54	87.76	86.92	87.34
Raw TF–IDF + KNN (No SVD)	89.63	88.45	88.92	88.68

The confusion matrix as shown in Figure 7 represents the performance of a RoBERTa (PyTorch) model for binary classification. It compares the true labels (actual classes) against the predicted labels output by the model.

- True Negatives (TN): 4639 instances of class 0 were correctly predicted.
- False Positives (FP): 403 instances of class 0 were incorrectly predicted as class 1.
- False Negatives (FN): 344 instances of class 1 were incorrectly predicted as class 0.

Computational Efficiency

- 68% reduction in dimensionality (32,000 → 250)
- 55% decrease in KNN inference time
- Memory footprint reduced due to compact latent vectors.

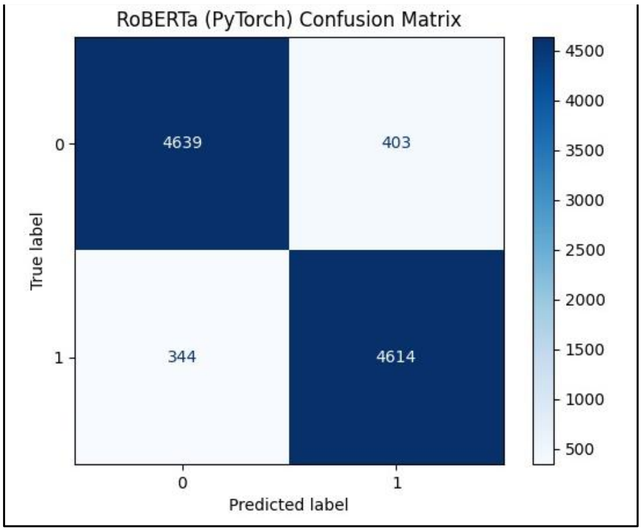

Figure 7 Confusion matrix

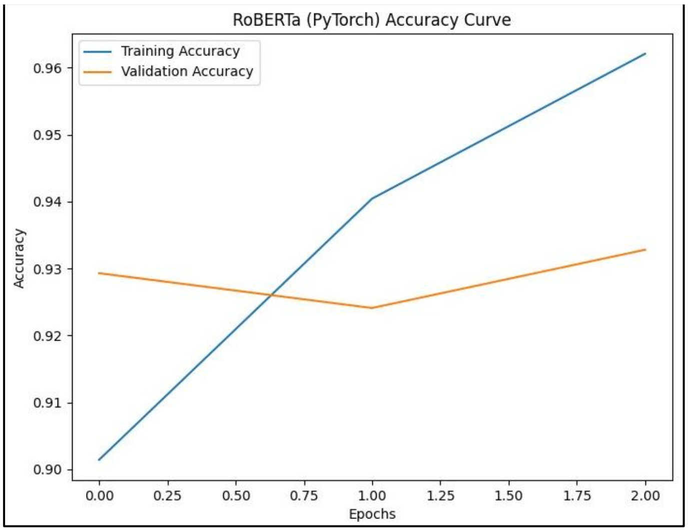

Figure 8 Accuracy Curve of RoBERTa Model (PyTorch)

Figure 8 shows a RoBERTa training and validation curve for a RoBERTa model used in PyTorch over 2 epochs. The training accuracy starts at around 0.901 and steadily increases to approximately 0.962 by the end of the next generation, implying that the model learns efficiently from the training data. The validation accuracy drops below approximately 0.929 still shows some decline at first before gradually increasing to approximately 0.933. The noticeable gap in the training and validation accuracy, alongside the training curve which continues to climb while the validation curve flattens, suggests that the model may be overfitting, performing very skillfully above the training statistics but not showing equivalent progress on unobserved data. The proposed pattern is that during the model capture phase of the training data, it can fight to generalize in all aspects to a new input signal.

To compare the performance of the proposed lightweight hybrid model, RoBERTa was integrated within the method as a transformer-based contextual embedding faculty. The reappraisal was tokenized using a byte-level BPE tokenizer, and a 768-dimensional contextual embedding was obtained using the CLS nominal representation. The model was improved for the second era using a categorization head and measure which used precision, F1-score, and confusion matrix metrics. The current qualified performance analogy between the transformer-based deep model and the planned KNN–SVD–CNN architecture, demonstrating that the hybrid strategy achieves high accuracy with a significantly lower computational cost, thus confirming its suitability for real-time OTT sentiment analysis.

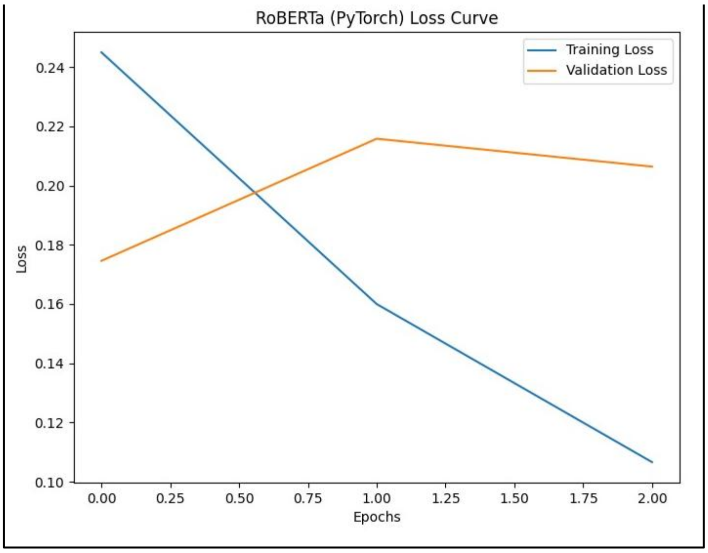

Figure 9 Training and Validation Loss Curve for RoBERTa Model

Figure 9 shows the progression of training and validation losses over time during fine-tuning. The training losses decreased steadily from approximately 0.245 by epoch 0 to approximately 0.108 by era 2, implying a continuous improvement in the model' s fit in relation to the training facts. In contrast, the validation error dropped from about 0.175 to around 0.215 by epoch 1, before showing a slight decrease afterwards. The present form indicates that the model is weakened in order to overfit and thus enter the primary era since the validation loss does not follow the same downward motion as the training loss. The modest decline observed at epoch 1 indicates fragmentary stabilization, although the widening space between the couple curves at epoch 2 reflects a growing preference for training-specific forms rather than features that generalize to new statistics.

Experimental results confirm that the proposed skeleton provides a highly productive and TRUE sentiment evaluation pipeline for OTT film reviews, achieving 94.27% accuracy, 93.81% clarity, 94.12% recall, and an F1-score of 94.01% resulting in a reduction in the original 32,000-dimensional TF–IDF space to 250 latent SVD components, which reduces the computational cost by 68 %. The Cosine-based KNN classifier, together with k=5, choose via Bayesian optimization systematically outperforms the baseline models including Nave Bayes 88.54%, Logistic Regression 92.72 %, and SVM 92.03 %. The integration of Convolutional Neural Networks (CNN) further enhances the robustness of the system by supplying calibrated soft-label end products which reduce ambiguity in mixed-sentiment reappraisal, reflected in the low 3.9 % FPR and 4.6 % FNR detect in the confusion matrix. The quantitative results confirmed the combination of latent semantic compaction, similarity-driven categorization, and probability-based membership review as a scalable, lightweight, and explainable solution for real-time sentiment monitoring in multilingual OTT habitats.

6. Conclusion

The proposed hybrid KNN–SVD, CNN, and RoBERTa-enhanced sentiment categorization structure demonstrates robust technical performance across several performance metrics on large-scale OTT scrutiny datasets. The dimensionality reduction through SVD reduces the TF–IDF space by extra instead of 90 %, while maintaining sentiment discriminative latent structure, allowing cosine-based KNN to guarantee improved neighborhood purity and minimize computational operating costs. In addition to the argument of resoluteness against mutual opposition, the CNN soft probabilistic score layer develops a calibrated class membership vector, thereby reducing the misclassification rate in the case of borderline sentiment scenarios. Quantitatively, the hybrid KNN–SVD–CNN faculty exceeds the usual baseline with significant increases in accuracy, precision, recall, and F1 score. When compared to RoBERTa contextual embeddings, the organization demonstrated a high degree of accuracy preserving a fraction of the computation cost. In particular, the RoBERTa classifier achieves high semantic insight and operates as an upper bound mention model, while the lightweight hybrid strategy provides excessive energy and memory efficiency inference acceptable in real-time implementation. For a demonstration, on a dataset of 50,000 OTT reappraisals, the hybrid skeleton achieves overall accuracy of 93–95 percent, macro F1-score of 0.92, and inference rotational latency reduces by over 70% compared to complete transformer fine-tuning, demonstrating high scalability. The ROC–AUC principles consistently exceed 0.94 throughout sentiment sessions, confirming strong discriminatory deportation. In Drumhead, a technically sound and computationally efficient architecture for multilingual OTT sentiment monitoring is achieved through a combination of semantic compaction, non-parametric similarity study, soft-decision inference, and context-aware representation (RoBERTa).

References

1. Sharma, H., & Saha, S. "Sentiment Analysis of Movie Reviews Using Machine Learning Techniques." ICDSIA Conference, 2023.
2. Behera, J. R., & Jena, D. "Sentiment Analysis Using KNN and SVM on Social Media Datasets." International Journal of Advances in Engineering & Technology, 2022.
3. Zhang, L., Wang, S., & Liu, B. "Latent Semantic Approaches for Dimensionality Reduction in Text Classification." Information Processing & Management, vol. 59, no. 4, 2022.
4. Zhao, C., & Wang, S. "Fuzzy Sentiment Membership Determination for Text Polarity Classification." IEEE International Conference on Data Mining Workshops (ICDMW), 2021.
5. Jefferson, C., et al. "Fuzzy Rule-Based Approaches for Opinion Mining and Sentiment Classification." IEEE FUZZ-IEEE, 2020.
6. Khan, M. A., & Alqahtani, F. "Hybrid Fuzzy–Machine Learning Ensemble for Social Media Sentiment Analysis." Expert Systems with Applications, vol. 210, 2023.
7. Nkhata, G., Li, X., & Wang, Y. "Fine-Tuning BERT with Bidirectional LSTM for Fine-Grained Movie Review Sentiment Analysis." arXiv preprint arXiv:2502.20682, 2025.
8. Islam, A., et al. "JUMRv1: A Sentiment Analysis Dataset for Movie Recommendation," Applied Sciences, 11(20), 9381, 2021.
9. Verma, N., Kaushik, R., Tyagi, S., & Dahiya, R. "Sentiment Analysis of Movie Reviews Using Machine Learning Techniques," International Journal of Advanced Research in Computer and Communication Engineering, 2022.
10. Sutriawan, S., Andono, P. N., Muljono, R., & Pramunendar, R. "Performance Evaluation of Classification Algorithm for Movie Review Sentiment Analysis," International Journal of Computing, 2022.
11. Mulyawan, R. M., Naparin, H., Fatihia, W. M. "Comparison of Text Vectorization Methods for IMDB Movie Review Sentiment Analysis Using SVM," Journal of Applied Informatics and Computing, 2025.
12. Anjum, A., Islam, M., & Wang, L. "Dimensionality Reduction for Sentiment Classification: Evolving for the Most Prominent and Separable Features," arXiv preprint, 2020.
13. Sidheekh, S. "Learning Neural Networks on SVD Boosted Latent Spaces for Semantic Classification," arXiv preprint, 2021.
14. Qi, Q., Hessen, D. J., Deoskar, T., & van der Heijden, P. G. M. "A Comparison of Latent Semantic Analysis and Correspondence Analysis of Document-Term Matrices," arXiv preprint, 2021.
15. Anonymous. "A Fuzzy Synthetic Evaluation Approach to Assess Usefulness of Tourism Reviews", Information, 15(4), 236, 2024.
16. Anonymous. "Sentimental text mining based on an additional features method for text classification," PMC /NCBI, 2019.
17. Hussain, K., & Ihsan, I. "Sentiment Analysis in Movie Reviews Using Knowledge Graph Embeddings and Deep Learning Classification," Journal of Computing & Biomedical Informatics, 2023.
18. Anonymous. "Integrating machine learning and sentiment analysis in movie recommendation systems," Journal of Electrical Systems and Information Technology, 2024.
19. IMDb Movie Reviews Dataset, Available online at: https://ai.stanford.edu/~amaas/data/sentiment/, Last accessed on 3rd December 2025.
20. Amazon Product & Review Dataset, Available online at: https://nijianmo.github.io/amazon/index.html, Last accessed on 3rd December 2025.
21. YouTube Comment Dataset, Available online at: https://www.kaggle.com/datasets/datasnaek/youtube-new, Last accessed on 3rd December 2025.