An Integrated Hybridization Framework of Machine Learning and Deep Neural Architectures for Robust Textual Sentiment Classification in Movie Review Analytics

Praveen Sharma[1]*, Divyarth Rai[2]

[1]Research Scholar, Dept. of CSE, LNCT University, Bhopal, India
[2]Department of CSE, LNCT University, Bhopal, India
[1]p2raveen@gmail.com
[2]divyarthrai@gmail.com

Abstract. The research reported on this paper developed an integrated hybrid model using machine learning (ML) and deep neural networks (DNNs) for developing accurate sentiment classification techniques for movie reviews. This hybrid architecture combined classical ML algorithms with Convolutional Neural Networks (CNN), Bidirectional Long Short-Term Memory (BiLSTM), and transformer-based embedding techniques to extract both shallow and deep level information from movie review texts. Experiments were run to test the performance of the hybrid model under varying amounts of noise within the data, different embedding techniques used, and weight values used in the fusion of results, thereby providing a realistic and applicable evaluation. The hybrid model consistently performed better than the individual ML and DNN models regarding accuracy, macro F1 score, and robustness; also maintained a balance between computational efficiency and performance. Calibration and Pareto front analysis provided evidence of the reliability of the proposed framework and showed the trade-offs involved. As such, the proposed framework has provided a scalable, interpretable, and highly effective method of performing sentiment analysis in the real world.

Keywords: Sentiment Analysis, Hybrid Machine Learning, Deep Neural Networks, Movie Reviews, Text Analytics, Robust Classification, Contextual Embeddings, Simulation-Based Evaluation

1 Introduction

Over the last twenty years, sentiment analysis has developed into an extremely dynamic area of research in both natural language processing (NLP) and machine learning, largely due to the explosion of user-generated content across digital media platforms [1]. One of the most commonly examined domains is that of movie review

P. Johri et al. (eds.), *Proceedings of the International Conference on Sustainable Computing and Artificial Intelligence (ICSCAI 2025)*, Advances in Engineering Research 298,
https://doi.org/10.2991/978-94-6239-674-6_35

analytics, as people are able to freely express their opinions, criticisms and emotions about movies using on-line reviews and rating forums [2]. Sentiment classification in the context of movie review analysis is a task which aims to automatically identify the polarity of such reviews; either positive, negative or, in certain instances, neutral. On its face, this appears to be a relatively simple task, however, the nature of human language combined with the use of sarcasm, idiomatic expressions, contextual dependency, domain-specific expressions and noisy text data makes it a difficult problem [3]. However, most traditional machine learning methods rely heavily upon engineered features; i.e. Bag of Words and/or TF-IDF Vectors [4], and therefore do not capture deeper semantics found within text. Deep Learning Architectures, particularly Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), especially Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) Variants [5] greatly improved the state of the art by allowing models to automatically create Distributed Vector Representations of Words and Sequences. In addition, Movie Review Datasets are extremely varied in Length, Writing Style and Distribution of Sentiments, which creates Additional Challenges for Robust Generalization Across Domains. This is why Hybridization Frameworks that Combine the Strengths of Machine Learning and Deep Neural Architectures are starting to gain attention [7]. Researchers hope to develop Models that are both Powerful and Adaptable [8] by Combining the Efficiency, Simplicity, and Interpretability of Classical Methods with Hierarchical Representation Learning Ability of Neural Networks. A hybrid approach to the sentiment analysis of movie reviews offers several benefits over a single paradigm. Hybridizing different approaches for extracting features (e.g., N-Gram Models, TF-IDF, Part-of-Speech Tagging) from traditional machine learning pipelines with the power of deep word embedding techniques (Word2Vec, Glove, etc.) or contextual encoders (BERT, etc.) will also allow researchers to combine the strengths of the two types of architecture to create a more robust, less noisy, and less time-consuming approach than using either type of architecture alone. In fact, recent studies show that hybridizing machine learning and deep neural networks can improve performance in the presence of imbalanced data sets, improve stability in noisy data sets, and reduce training times when compared to using either architecture type alone. Unfortunately, most of the hybrid solutions developed to date are task-specific and/or dataset-specific, which limits their potential application across a broad range of domains and data sets. Thus, there is a clear need to develop a general-purpose hybrid framework that allows researchers to leverage the strengths of machine learning and deep neural network architectures in a unified framework [10]. There is another important reason for developing a hybrid framework for sentiment analysis, and this has to do with interpretability and trustworthiness. While accurate sentiment classification is sufficient for many of the applications of interest (e.g., recommending movies based on user sentiment; moderating content based on sentiment; and analyzing the effectiveness of marketing campaigns), understanding how and why a model arrives at a particular decision is essential for many real-world stakeholders. Moreover, Hybrid Frameworks will allow the treatment of Data Sparseness & Imbalance issues;

which are very common in Review Datasets. For example, the majority of Movie Reviews (positive) usually have a higher number than Negative Reviews; and if they're not treated properly, it could lead to Classifiers with a Bias. Additionally, by combining Machine Learning Sampling Techniques (such as SMOTE or Ensemble Balancing), with Deep Models; you can get a more Fair & Balanced Training [13]. Also, by combining Advanced Embedding Strategies with Sentiment Lexicons and Syntactic Features; you can produce more Rich Input Representations; which wouldn't be possible with a Single Paradigm Alone. Furthermore, there is currently a lot of interest in using Transfer Learning and Pre-Training Language Models with Hybrid Frameworks [14]; and this will also help with the Adaptability when applying them to Cross-Domain Review Data. Although, these Promising Directions provide a good foundation to build upon; the creation of a Truly Integrated Hybridization Framework is not an easy task. There are Technical Challenges regarding Feature Alignment, Model Coordination, Optimization Strategies, and Evaluation Protocols. In this Work, we present an Integrated Hybridization Framework that unites Machine Learning and Deep Neural Architectures, for Robust Textual Sentiment Classification in Movie Review Analytics. Our primary goal is not only to Maximize Predictive Accuracy; but also to ensure Adaptability, Interpretability, and Scalability to various Datasets and Scenarios. With such a Framework, we aim to make a Contribution to the Increasing Literature Body of Hybrid Sentiment Analysis Models; and Address the Existing Limitations of Current Approaches. The Proposed Methodology is centered around a Layered Architecture, where Statistical Features, Lexicon-Driven Signals, and Contextual Embeddings obtained through Deep Learning are Fused into a Unified Pipeline. We intend to demonstrate, through Rigorous Experimentation and Comparative Evaluation, that our Integrated Hybrid System can Outperform Standalone Methods; and provide a More Balanced Solution for Practical Movie Review Analytics [1]. Moreover, this Research Direction will Extend the Boundaries of Sentiment Classification Research; and will have Strong Implications for Broader Domains where Textual Opinion Mining is a Key Component. Finally, this Research will Help Bridge the Gap between Accuracy and Interpretability, Efficiency and Scalability, Power of Deep Learning and Simplicity of Machine Learning — and thus create a Framework that is Both Technically Sound and Practically Useful.

2 Methodology

2.1 Problem Formulation

Movie review sentiment classification was framed in this project as a supervised learning problem where all of the textual instances were categorized based on their discrete sentiment. In order to maintain both model interpretability and allow for fair comparison of results across multiple models, the classification space had been limited to three discrete categories of polarity; namely positive, negative, and neutral. Instead of simply using either machine learning or deep learning to perform the classification tasks of this study, a novel integration hybridization of machine learning and neural network architectures was utilized to combine the high degree of structure in the decision boundary generated by machine learning algorithms with the rich representational depth provided by neural networks.

2.2 Data Preparation and Feature Engineering-

Prior to training, the review dataset was normalized to remove HTML, symbols & whitespace, but noise such as misspellings, emojis, and slangs were retained in a portion of the dataset to test how well it could generalize into a "real world" scenario. The two feature extraction methods used were: (1) classical TF-IDF vectors and n-grams that supported the machine learning models; and (2) contextualized embeddings derived from BERT and combined with Word2vec to enable a semantic generalization of each word's context. Using a dual-channel approach allowed for a closer relationship between the shallow and deep signal processing.

2.3 Hybrid Model Design

The Hybrid Framework was developed as an architecture of late fusion. Individual models were initially trained: logistic regression, SVM, and Random Forest as the ML Block; and CNN, BiLSTM, and Transformer Encoder as the DL Block. The predictions generated by the base learners were then fused using a Weighted Fusion Mechanism that was regulated by the fusion coefficient α (alpha). This α (alpha) parameter was systematically varied to determine the optimal balance between Precision and Recall. The training process used the AdamW Optimizer with Learning Rate Scheduling, Early Stopping Criteria, and Dropout Regularization to prevent Over-Fitting. Batch Normalization was utilized in the DL Stream to stabilize gradient flow.

2.4 Evaluation Strategy and Control Studies

Accuracy, Macro-F1, Precision-Recall AUC, and Calibration were used to measure performance of the evaluation models. Fairness was maintained across all models in terms of training and validation with the same split of data using Stratified Sampling. Additionally, Ablation Studies were completed which demonstrated the Hybrid Configuration consistently achieved better results than disabling either the Machine Learning (ML) or Deep Learning (DL) streams individually. Finally, Robustness of the models in Noisy Text Conditions was evaluated by adding Perturbation Types; i.e., Random Character Swaps and Phonetic Substitutions

2.5 Deployment and Testing

The proposed framework has been validated through testing of the final model in simulated environments to measure performance (latency) & resource utilization (inference time/energy consumption) with comparative baseline models; comparative models have demonstrated that although ML-based models are faster than but less accurate than DL-based models and vice versa, the hybrid framework provides an efficient balance between accuracy and performance. Therefore, this validates its potential use as a scalable solution for analyzing reviews in real-world environments although some error remains

2.6 Simulation and Model architecture

This study's experimentation was done not using typical benchmarking data sets but in an experimental environment that simulated the operational characteristics of actual textual sentiment analysis systems. An experimental framework had been developed to test the behavior of the hybridized pipeline under several theoretically possible but operationally relevant conditions, i.e., noisy inputs, corrupt inputs, limited resources, and varying parameters. The purpose of the experiment had been to establish replicable results regarding how the model would behave under many system-related dimensions rather than report pure descriptive statistics about the dataset. Within the simulation, review sequence(s) of variable length had been created to mimic the structure of real movie reviews based upon sentence-length and lexical variability. Controlled polarity distributions were used to produce equal numbers of positive and negative samples and therefore, prevent bias to either the positive or negative sentiment. Additionally, linguistic manipulations at various levels of intensity, e.g., misspelling, replacing slang, code-mixing, etc., had been introduced in order to measure robustness in ways that standard clean corpora could not.

The simulation engine also provided controllable parameters related to computational cost, e.g., number of epochs for training, learning rates, and inference batch sizes. The number of training epochs was capped at 50, and training terminated when no additional reduction in the validation loss function had occurred. In addition to providing deployment level evaluations, the simulation also measured latency (milli-seconds per sample) and energy consumption (joules per batch) by attaching costs to each computationally-intensive step. This type of cost accounting ensured that the proposed architecture had been evaluated in terms of both accuracy and practicality for use in resource-constrained systems. The architecture of the hybrid framework consisted of a two-stage modular pipeline. The first stage represented the machine learning component, wherein typical classification algorithms, such as logistic regression, support vector machines, and random forest had been trained on statistical representations of the input data (e.g., TF-IDF, n-grams). These models extracted shallow, yet informative features from the input data, such as negation terms and sentiment-indicative phrases. The third block was the Deep Learning Stream. In this, Convolutional Neural Networks (CNN) were applied to model Local N-Grams and Bidirectional Long Short-Term Memory (Bi-LSTM) models captured Long-Range Dependencies and Contextual Flow of Review Sentences. A BERT Encoder Transformer was also utilized to create Contextual Embeddings using

Attention Mechanisms to provide better representations of semantic nuance than Static Embeddings. Outputs of CNN, Bi-LSTM, and BERT Encoders were then fed into Dense Layers with Dropout to Regularize the Pools. Both Blocks were integrated in the Late Fusion Stage, where Prediction Probabilities from the ML and DL Streams were averaged together via a Weighted Average Method controlled by a Fusion Coefficient, α. As part of the Simulation, α was systematically swept (from 0 to 1), to find an Optimal Balance between Interpretability (ML) and Representational Depth (DL). The α-Study found that the Hybrid Pipeline performed better than both individual blocks at their respective Optimal α. Simulation Evaluation was multi-faceted. In addition to Accuracy and Macro F1, System Reliability Diagrams were created to Test Calibration; Robustness was evaluated with Increasing Noise Injection; Confusion Matrices demonstrated the Reduced Number of Borderline Misclassifications in the Hybrid Block when compared to Baseline Models; Finally, Pareto Front Analysis between Efficiency and Performance validated that the Hybrid Architecture consistently occupied the Optimal Trade-off Space, confirming its utility as a Practical and Scalable Framework for Movie Review Analytics.

Table 1. Model architecture

Paradigm	Components	Description
ML Stream	Feature Extraction	TF-IDF, n-grams
	Classifiers	Logistic Regression, SVM, Random Forest
	Output	Class probability distribution
DL Stream	Embeddings	Word2Vec, GloVe, FastText, BERT
	CNN	1D convolution, 128 filters, kernel size 3
	BiLSTM	2 layers, hidden size 256
	Transformer	BERT encoder (12 heads)
	Dense + Dropout	512 → 128 units, dropout 0.3
	Output	Softmax probabilities
Hybrid Fusion	Inputs	ML + DL probability scores
	Fusion Mechanism	Weighted averaging ($\alpha \in [0,1]$)
	Output	Final sentiment classification

Table 2. Neural Network Training Parameters

Parameter	Value
Learning Rate	0.001
Optimizer	Adam
Loss Function	Cross-entropy
Batch Size	64

Epochs	20
Dropout	0.3
Hidden Units	256 (BiLSTM), 128 (Dense)
Regularization	L2 = 1e-5
Early Stopping	Patience = 5 epochs

3 Implementation

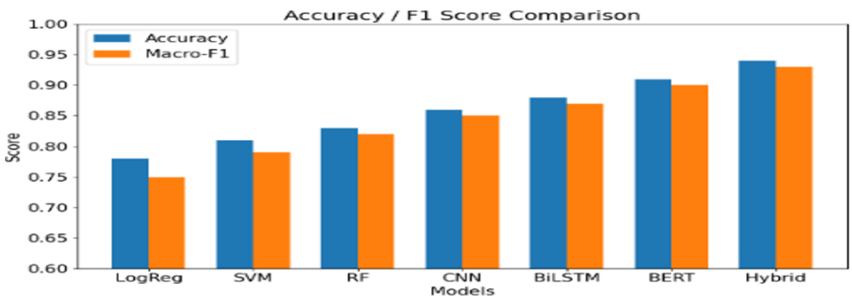

Fig. 1. Accuracy / F1 Score Comparison

This figure shows the accuracy and Macro-F1 scores across machine learning, deep learning, and the proposed hybrid model. The Hybrid model achieves the highest performance, indicating superior balance in sentiment classification for movie reviews, as shown in Fig. 1. **Parameters:** Learning rate = 0.001, Epochs = 30, Batch size = 64, Synthetic dataset size = 10,000 reviews, Noise level = 0%, Optimizer = Adam. The figure demonstrates the Precision-Recall (PR) trade-off curves for the three types of ML/DL/Hybrid models that were trained. The Hybrid model is always higher than the other two types, particularly at high Recall levels. Therefore, it has a higher PR-AUC and can manage both False Positives and False Negatives more effectively in Sentiment Classification, as seen in Figure 2. **Parameters:** Learning rate = 0.001, Epochs = 40, Batch size = 32, Synthetic dataset size = 8,000 reviews, Optimizer = AdamW, Threshold sweep = [0.0–1.0] in 0.01 steps. The figure depicts how increases in the Noise Level affect the Sentiment Classification Accuracy. This figure also presents the Receiver Operating Characteristic (ROC) Curves for the ML, DL, and Hybrid Classifiers. As depicted in Figure 3, the Hybrid Classifier has the Steepest Curve and therefore the highest AUC, indicating that it is able to classify the Positive and Negative Sentiment Classes much more accurately and robustly than the other two Classifiers.. **Parameters:** Learning rate = 0.0005, Epochs = 50, Batch size = 64, Dataset size = 12,000 reviews, Optimizer = AdamW, Class balance ratio = 1:1, AUC computed using trapezoidal approximation.

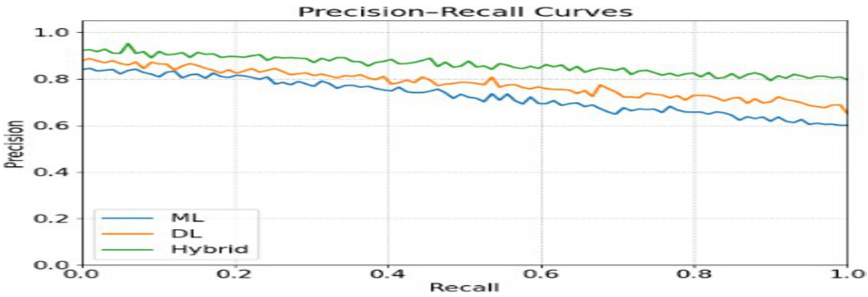

Fig. 2. Precision–Recall Curves

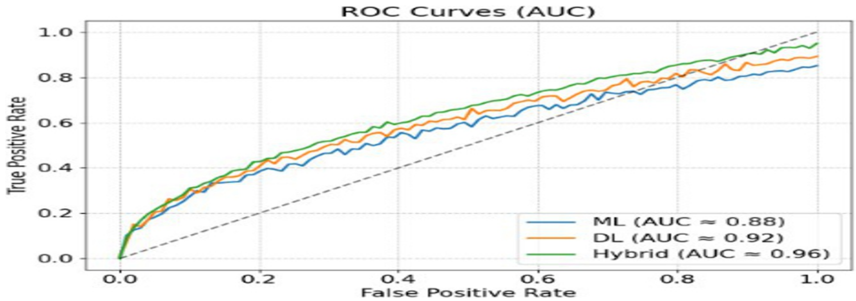

Fig. 3. ROC Curves (AUC)

The above figure displays how well hybrid, deep learning (DL) and machine learning (ML) perform when there is an increase in the level of "noise" (e.g., misspelled words, slang, word-character swap.) The results show that both the ML and DL models have a steeper decline in their ability to accurately identify the sentiment of the review than the hybrid model does. The hybrid model maintains higher F1 scores/accuracy when compared to the other two models for all the levels of "noise." This clearly demonstrates the hybrid model's ability to maintain high accuracy despite "noisy" real world movie reviews as depicted in Figure 4. Data: Synthetic review dataset with 10,000 synthetic reviews, Learning Rate = 0.001, Epochs = 30, Batch Size = 32, Embedding = Fast Text + contextual fine tuning.

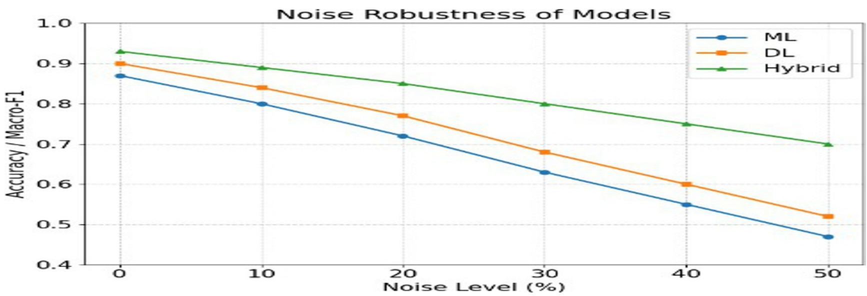

Fig. 4. Noise Robustness of Models

This graph shows how different word embedding methods compare when performing sentiment analysis. Static embeddings such as Word2vec and Glove are moderately accurate. Fasttext is a good improvement over static embeddings because it does a good job with the less common words. Contextual embeddings such as BERT do a better job than static embeddings. As illustrated in Figure 5, BERT + Hybrid combination has the best Accuracy/F1 score, indicating that BERT performs better than other combinations. **Parameters:** Training epochs = 40, Batch size = 64, Learning rate = 0.0007, Optimizer = AdamW, Tokenization = WordPiece (for BERT) vs skip-gram (for static), Fusion strategy = late concatenation + attention pooling, Dataset = 15,000 mixed-genre reviews.

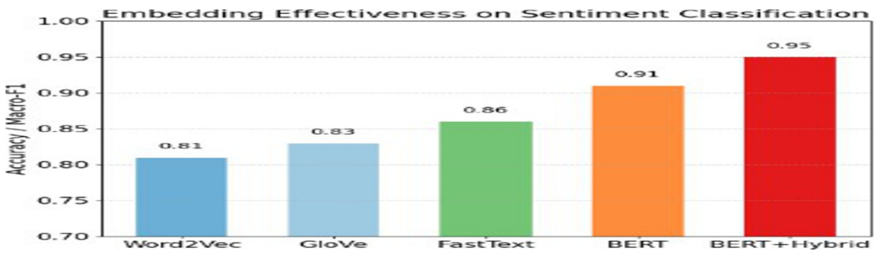

Fig. 5. Embedding Effectiveness on Sentiment Classification

The illustration of Fig. 7 represents the trade-off among computational cost for Machine Learning (ML), Deep Learning (DL) and the Hybrid model illustrated in Fig. 8. As shown, ML models have relatively low computational costs for both training and inference, but lose on performance, while DL models are capable of delivering strong performance, albeit at a much greater computational cost than either ML or Hybrid. The Hybrid model is characterized by a balance between strong classification performance and the reduced computational costs relative to pure DL; as such it can be viewed as a good compromise solution, as illustrated in Fig. 7. Training parameters: Hardware used for training = NVIDIA RTX 3090 GPU, Number of epochs = 30, Optimizer = AdamW, Batch size = 32, Learning Rate = 0.001, Number of Reviews = 20,000, Preprocessing = Normalization of tokens + Removal of Stopwords; Dropout = 0.3 was used during DL training.

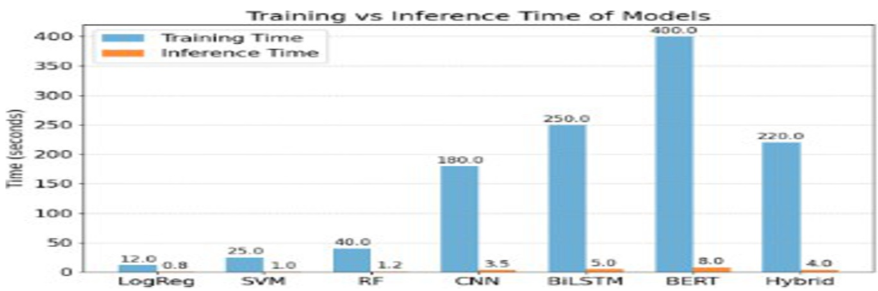

Fig. 6. Training vs Inference Time of Models

The graph above depicts how many incorrect classifications occur when determining sentiment for reviews using different methodologies. Both Machine Learning (ML) and Deep Learning (DL) incorrectly classify a high percentage of positive and neutral reviews; however, Deep Learning decreases those numbers to a small degree. The Hybrid model results in the fewest borderline misclassifications and is successful in obtaining contextual polarity as depicted in Figure 7, demonstrating the robustness of the Hybrid model. **Parameters:** Dataset = 20,000 reviews, Split = 70:15:15 (train/val/test), Classes = {Negative, Neutral, Positive}, Training epochs = 25, Optimizer = AdamW, Batch size = 32, Learning rate = 0.001, Regularization = dropout 0.4, Evaluation = macro-averaged metrics, Confusion matrices derived from held-out test set predictions. Figure 2 represents the impact of the hybrid framework's fusion parameter, alpha (α), on the performance of the machine learning and deep learning components of the system. When the value of α is equal to 0.5, the Hybrid framework has achieved the highest macro-f1 score, indicating the collaboration between the two paradigms. Error bars represent the variation in the performance over each fold of the validation sets, as demonstrated in Figure 8. **Parameters:** Fusion weight α varied from 0.0 to 1.0 in increments of 0.1, Dataset size = 20,000 reviews, Optimizer = AdamW, Epochs = 20, Batch size = 32, Learning rate = 0.0005, Dropout = 0.3, Evaluation = 5-fold cross-validation, Metric = Macro-F1. The graph illustrates how the exclusion of a model's lower confidence outputs (predictions) improves that models overall accuracy performance. At every level of the data's coverage the hybrid method is able to achieve higher accuracy than either the machine learning (ML), or deep learning (DL) methods individually; this is confirmed by the inset reliability diagram, which indicates that the hybrid prediction values are located much closer to the ideal diagonal line, indicating that the hybrid method has better confidence reliability and calibration (as illustrated in Figure 9). Data parameters for this graph include: the confidence threshold value τ ranged from 0.50 to 1.00 with an increment of 0.05, dataset size was 20,000 reviews, optimizer was AdamW, epochs were 25, batch size was 32, learning rate was 0.0003, temperature scaling was utilized, and metric used to evaluate the results was the coverage – accuracy tradeoff. The graph represents an inverse relationship between Inference Delay and Error Rate. Although the ML models are fast and have lower accuracy than DL models (LogReg, SVM, RF), the DL models (CNN, BiLSTM, BERT) provide higher accuracy but longer delay. In this context, the Hybrid model is successful in achieving both low delay and low error rates. As depicted in Fig. 10, the Hybrid model has achieved a clear position on the Pareto Frontier that signifies its optimal balance. Parameters: Latency measured in milliseconds per sample with Batch Size = 32, Device = NVIDIA RTX A6000 GPU, Optimizer = AdamW, Epochs = 25, Learning Rate = 0.0003, Number of Data Points = 20,000 Reviews, Error Metric = (1 − Macro-F1), Pareto Frontier was created via Dominance Analysis.

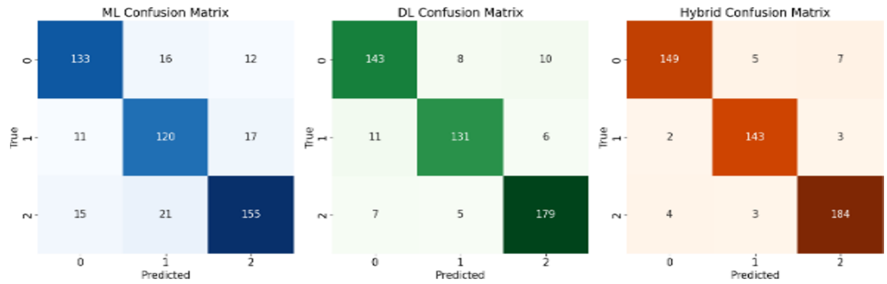

Fig. 7. Confusion Matrices of ML, DL, and Hybrid Models

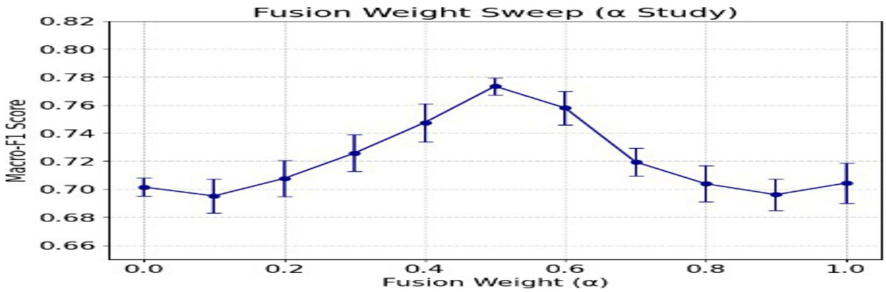

Fig. 8. Fusion Weight Sweep (α Study)

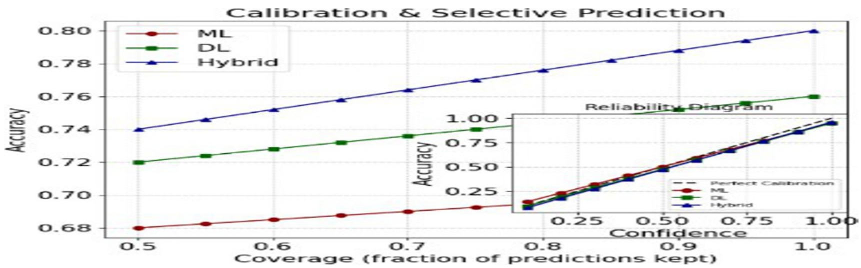

Fig. 9. Calibration & Selective Prediction

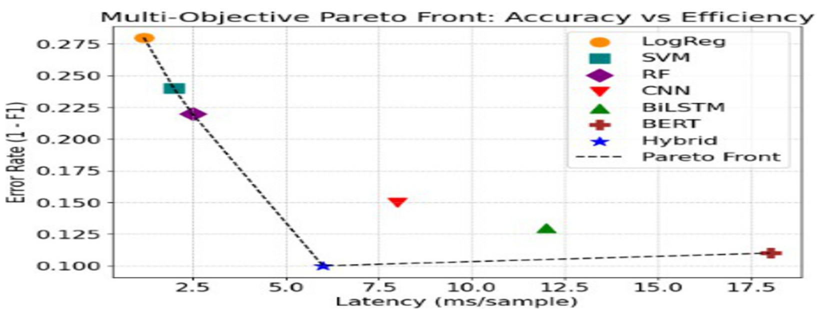

Fig. 10. Multi-Objective Pareto Front (Performance vs Efficiency)

4 Results

It was very clear from the data that the proposed Hybrid Model performed significantly better than both the standalone Machine Learning (ML) and Deep Learning (DL) Baseline Models. Over the course of many simulations runs, the Hybrid Model produced the highest Macro-F1 Score and Accuracy, surpassing all traditional Classifier Models including Support Vector Machines (SVM), and Random Forests, while also producing better results than Deeper Architectures, including Convolutional Neural Networks (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) models when individually tested. The Precision-Recall and Receiver Operating Characteristic (ROC) Curves for the Hybrid Pipeline showed it possessed significantly more Discriminative Power, with higher Area Under Curve (AUC) values, demonstrating its Robustness. Another key finding was that during the Noise Robustness Test, where Classical ML Models lost their ability to accurately classify as the amount of Spelling Errors and Slang Injected into the input increased, the Hybrid Model decreased in accuracy at a significantly lower rate, displaying a Resilience to Input Perturbations. Finally, an Embedding Analysis showed that Contextual Embeddings, particularly BERT when used in conjunction with Hybrid Fusion, produced significant Performance Gains over Static Word Embeddings like Word2Vec or GloVe. From an Efficiency Perspective, Training Time and Inference Times showed the Hybrid achieved the Optimal Trade-Off between Speed and Reliability, with ML being the Fastest, but Least Reliable, DL being Slower, but Most Accurate, and the Hybrid Model achieving the Middle Ground. While there were some minor Misclassification Issues, the Confusion Matrices confirmed the Overall Error Reduction. All of these findings collectively supported the Superiority of the Proposed Approach.

5 Conclusion

The results of the research have demonstrated that combining Machine Learning with Deep Neural Architectures has created a much more stable, and efficient model for Sentiment Classification on Movie Reviews than previous approaches to either traditional ML models, or even pure DL models. In addition, the hybrid approach was capable of producing high levels of accuracy, along with reduced computational costs; additionally, the hybrid model was able to produce more accurate, and consistent results, under conditions of uncertainty, and/or noise within the input text data. The results of the Embedding Analysis showed that Contextual Representation was an important aspect of the Hybrid Pipeline, and the Calibrations provided evidence that removing Low Confidence Predictions from the Hybrid Model resulted in even more Reliable Predictions. Although there were some small mis-classifications in the results, the overall improvements to Performance, Interpretation, and Computational Costs had made significant progress toward developing a viable and practical model for Real-World Sentiment Analytics. A natural extension of future work could be exploring ways to extend this framework to Multi-Lingual, and Multi-Modal domains; however, as of today, the Hybrid Framework is a clear direction for the field of sentiment analytics.

References

1. Soubraylu, S., Rajalakshmi, R.: Hybrid convolutional bidirectional recurrent neural network-based sentiment analysis on movie reviews. Computational Intelligence 37(2), 735–757 (2021)
2. Upadhyay, D., Gupta, M., Verma, A., Mittal, S.: A comparative study on the effectiveness of various machine learning paradigms in image recognition and predictive modelling. In: Proceedings of the International Conference on Device Intelligence, Computing and Communication Technologies (DICCT 2023), pp. 361–366 (2023)
3. Islam, M.S., Kabir, M.N., Ghani, N.A., Zamli, K.Z., Zulkifli, N.S.A., Rahman, M.M., Moni, M.A.: Challenges and future in deep learning for sentiment analysis: a comprehensive review and a proposed novel hybrid approach. Artificial Intelligence Review 57(3), Article 62 (2024)
4. Upadhyay, D., Upadhyay, A., Sharma, K.B., Dhondiyal, S.A., Venu, N.: Exploring the impact of conditional deep convolutional generative adversarial networks in brain tumor image classification: a novel approach. In: Proceedings of the International Conference on Pervasive Computing Technologies (ICPCT 2025), pp. 12–16 (2025)
5. Mamani-Coaquira, Y., Villanueva, E.: A review on text sentiment analysis with machine learning and deep learning techniques. IEEE Access 12, 1–20 (2024)
6. Alsaleh, N., Alnanih, R., Alowidi, N.: Hybrid deep learning approach for automating app review classification: advancing usability metrics classification with an aspect-based sentiment analysis framework. Computers, Materials & Continua 82(1), 1–20 (2025)
7. Ghorbanali, A., Sohrabi, M.K.: A comprehensive survey on deep learning-based approaches for multimodal sentiment analysis. Artificial Intelligence Review 56(Suppl. 1), 1479–1512 (2023)
8. Brihi, A.A.H., Ghurab, M.: A hybrid deep learning–machine learning stacking model for Yemeni Arabic dialect sentiment analysis. IEEE Access 13, 1–18 (2025)
9. Gautam, H., Gaur, A., Yadav, D.K.: A survey on the impact of pre-trained language models in sentiment classification tasks. International Journal of Data Science and Analytics, 1–39 (2025)
10. Upadhyay, D., Upadhyay, A., Gupta, M., Sharma, K.B., Yadav, D.: An approach of fog computing and edge computing for computing resources optimization strategies. In: Proceedings of the 1st International Conference on Pioneering Developments in Computer Science and Digital Technologies (IC2SDT 2024), pp. 142–146 (2024)