Development of an Optimized Convolutional Neural Network Architecture for Sentiment Analysis of Movie Reviews with a Large-scale Dataset

N. Dhamayanthi, B. Lavanya*
Department of Computer Science, University of Madras
Chennai, India

Abstract
Sentiment analysis plays a pivotal role in deciphering opinions and emotions expressed in movie reviews, contributing to enhanced understanding and decision-making in the film industry. In this paper, we present the development of an optimized Convolutional Neural Network (CNN) architecture for sentiment analysis of movie reviews using a large-scale benchmark dataset from the IMDB database. The research delves into meticulous experimentation with hyperparameters, including variations in convolutional layers and dimensions, to optimize model performance. Preprocessing techniques, tokenization, and feature vector creation were integral steps, enabling the transformation of textual data into numerical representations for efficient analysis. The proposed methodology showcases promising results, achieving an accuracy rate of 88.92% which favourably compares to the research published using the same benchmark dataset. Notably, the fine-tuned CNN architecture demonstrates robustness and efficacy in sentiment classification, signifying potential advancements in sentiment analysis for large datasets.

Keywords: *Machine Learning; Text Mining; Natural Language Processing; Sentiment Analysis; Deep Learning*.

1. Introduction

Sentiment analysis in movie reviews stands as a pivotal task in natural language processing, offering valuable insights into public opinion and consumer sentiments. With the exponential growth of digital content and user-generated reviews, the need for robust and accurate sentiment analysis methodologies has become increasingly pronounced. Leveraging the advancements in deep learning, particularly Convolutional Neural Networks (CNNs), has presented a promising avenue for effectively gauging sentiment within textual data. This research endeavors to contribute to this burgeoning field by developing an optimized CNN architecture tailored specifically for sentiment analysis in the domain of movie reviews. The primary objective is to design a model capable of discerning sentiment nuances within extensive datasets sourced from repositories like the IMDB database. Through meticulous experimentation and fine-tuning of hyperparameters, our research aims to establish an accurate and adaptable CNN architecture, addressing the challenges posed by the variability and complexity inherent in movie review data. This research aims to enhance the accuracy and dependability of sentiment analysis by developing a model that is both robust and scalable, which could significantly contribute to its broader application and understanding in various real-world scenarios. Key contributions of this research are outlined as follows:

1. We develop an optimized Convolutional Neural Network architecture (CNN) for sentiment analysis of movie reviews using a gold standard large benchmark dataset from IMDB using 50000 reviews.
2. We conducted several experiments with hyperparameter tuning including variations in convolutional layers and dimensions which optimized the model performance.
3. To evaluate the effectiveness of our model, we use key performance indicators including Accuracy, Recall, Precision, and the F1-score. These metrics provide a comprehensive assessment of our model's performance.

4. When benchmarked against the methodologies recently applied to the same dataset, our methodology showcases promising results, achieving an accuracy rate of 89.92%

The paper is organized as follows. Section 2 talks about the related work. Section 3 of this paper outlines the proposed methodology. Moving into Section 4, the results and discussions delve into the empirical findings along with a comparison of the state-of-the-art research in sentiment analysis. Finally, Section 5 encapsulates the paper with a conclusive summary.

2. Related work

In the dynamic field of sentiment analysis, a multitude of techniques and models have been developed to understand and classify emotions and opinions expressed in text. Studies such as [10] have delved into the analysis of movie reviews, exploring the impact of review length and readability on classification accuracy with discourse-aware methods using deep neural networks. Different techniques such as Naïve Bayes [1] [11], SVM [1] [9], LSTM [3] [8], K-Nearest Neighbor [11], Decision Tree [9], and ensemble methods like Random Forest [11] [6], stochastic gradient boosting [6] have been employed, demonstrating varying levels of accuracy and performance in classifying sentiments within different datasets and domains.

The COVID-19 pandemic brought new challenges and focus areas for sentiment analysis. Research like [2] employed various deep learning models and traditional classifiers to gauge sentiments on social media during the pandemic's early phases. Meanwhile, [7] introduced the Hybrid Heterogeneous Support Vector Machine (H-SVM) for sentiment classification, and [18] combined four deep learning and one classical supervised machine learning model to analyze tweets related to COVID-19 from eight countries, highlighting the role of social media in raising awareness.

To address the challenges of dimensionality in datasets, Author [4] introduced a novel deep neural network structure combining a bidirectional convolutional recurrent neural network with a unique group-wise enhancement approach, while [5] demonstrated the efficacy of heterogeneous ensemble techniques in sentiment analysis. The complexity of sentiment analysis is further illustrated in [14], which introduces a unique mathematical approach using Game Theory and the SentiWordNet lexicon for review sentiment analysis.

In the financial sector, [19] presents a hybrid model that combines deep learning with sentiment analysis to predict stock prices, utilizing a Convolutional Neural Network and an LSTM Neural Network. Similarly, [28] introduces Polaris, a system for real-time analysis and prediction of users' sentimental trajectories on social media. A comprehensive survey of sentiment analysis is provided in [20], which details its applications, methodologies, and challenges.

Advancements in aspect-based sentiment analysis are represented in [15] with a new graph convolutional network model and [25]'s ALDONAr, an enhanced version of the ALDONA model. [22] introduces RDLS, a deep-learning approach for summarizing opinions from multiple documents, while [16] presents an unsupervised model for aspect term extraction using a guided LDA approach.

Further innovations are seen in [29]'s CNN-based method for sentiment analysis, [27]'s Gated Alternate Neural Network (GANN), and [23]'s Semantics Perception and Refinement Network (SPRN). Novel feature selection methods for sentiment classification are proposed in [26] with an Iterated Greedy metaheuristic algorithm, and [31] introduces a multiobjective optimization-based weighted voting ensemble method.

Authors [13] present a novel method for aspect-level sentiment classification using a modified Bidirectional LSTM, and [32] assesses the effectiveness of four machine learning algorithms in online reviews sentiment classification. The role of social media in product analysis and development is explored in [17] with a social media-based opinion retrieval system, and [30] introduces a new method using text-based Hidden Markov Models (TextHMMs) for sentiment analysis. Lastly, [21] and [24] propose hybrid filter models and deep learning models, respectively, for enhancing the classification in sentiment analysis tasks.

These studies collectively highlight the importance of text preprocessing, diverse machine learning algorithms, and domain-specific analysis for accurate sentiment classification in different contexts.

3. Methodology

Fig. 1. illustrates our proposed methodology. We utilize a large benchmark dataset of 50000 movie reviews sourced from the IMDB database. The Preprocessing phase involves essential text normalization techniques, encompassing error rectification, lowercase standardization, removal of extraneous punctuation, and

---

handling of HTML tags and special characters. Subsequently, the preprocessed text undergoes tokenization, segmenting it into tokens and facilitating the creation of a word-index map crucial for converting text into numerical representations. This enables the creation of numerical feature vectors conducive to machine learning analyses. As an integral step in model development, the dataset is partitioned into training and testing subsets, adopting an 80-20 split, respectively. The Convolutional Neural Network (CNN) architecture design entails meticulous hyperparameter optimization, involving fine-tuning based on the model's performance in sentiment analysis.

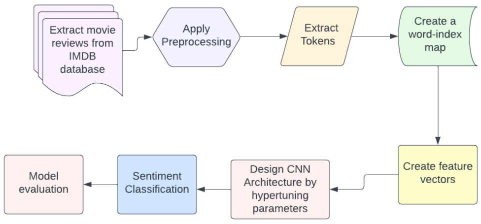

Fig. 1. Proposed Methodology

3.1 DATASET DESCRIPTION

We utilized the benchmark dataset tailored for sentiment classification of movie reviews, as introduced by Andrew et al. [12], comprising a total of 50,000 reviews evenly distributed between 25000 positive and 25000 negative sentiments. To ensure diversity and mitigate potential rating correlations, no more than 30 reviews were included for any single movie. The partitioning of the dataset into distinct training and testing sets safeguards against the model memorizing movie-specific terms and associated labels, preserving its generalizability. Negative reviews, labelled as "0" are characterized by ratings ≤4 out of 10, while positive reviews, labelled as "1" are defined by ratings ≥7 out of 10. Our training set consists of 80% (40,000 reviews), while the remaining 20% (10,000 reviews) constitute our testing set. Fig. 2. showcases a snapshot of the dataset, represented as a Python DataFrame.

3.2 PROPOSED CNN ARCHITECTURE

After rigorous experimentation and fine-tuning of hyperparameters, a refined Convolutional Neural Network (CNN) architecture, represented in Fig. 3. emerged, pivotal for achieving the desired results. Set with an embedding dimension (D) of 25, the architecture comprises a sequential arrangement of layers. The initial input layer is followed by an embedding layer, effectively mapping the vocabulary to a dense vector space. Subsequently, a series of three one-dimensional convolutional layers, each followed by max-pooling layers, serve to capture hierarchical features in the text. Starting with 32 filters and progressing to 64 and 128 filters successively, these convolutional layers are complemented by max-pooling operations to down-sample feature maps. A global max-pooling layer aggregates the maximum values across the temporal sequence, distilling the most salient features. The architecture culminates in a dense layer employing a hyperbolic tangent (tanh) activation function, producing a single output for sentiment classification.

	reviews	sentiment
0	Apart from the fact that this film was made (...	0
1	I loved this movie - the actors were wonderful...	1
2	This show makes me(and many others) hate their...	0
3	I just saw this movie for the second time with...	1
4	When the long running 'Happy Ever After' came ...	1
...	...	...
49995	OK, I would give this a 1, but I'm gonna give ...	0
49996	Joseph L. Mankiewicz is not remembered by most...	1
49997	Anyone who has seen Ali G before, should be we...	1
49998	If you're in the middle of a ferocious war and...	1
49999	this is the worst movie i've ever seen. i'm no...	0

50000 rows × 2 columns

Fig. 2. Screenshot of the dataset from Python DataFrame

3.3 TOOLS LEVERAGED FOR BUILDING THE CNN ARCHITECTURE

In the development of the Convolutional Neural Network (CNN) architecture for sentiment analysis of movie reviews, a suite of essential tools and libraries played pivotal roles. TensorFlow, alongside its user-friendly API Keras, served as the cornerstone for constructing and training the deep learning model. Leveraging

TensorFlow's robust framework and Keras' streamlined interface, the architecture's design and implementation were facilitated, enabling seamless experimentation and fine-tuning of hyperparameters. Complementing this, Pandas, a powerful data manipulation library, aided in handling and preprocessing the dataset, ensuring its readiness for training. NumPy, renowned for its numerical computation capabilities, facilitated efficient matrix operations and mathematical computations within the neural network layers, enhancing computational efficiency.

Additionally, the data serialization module Pickle ensured the storage and retrieval of processed data and trained models, while Matplotlib provided insightful visualizations, offering a comprehensive view of model performance and data trends. Collectively, these tools and libraries were instrumental in the construction, training, and evaluation of the CNN architecture, contributing significantly to the success of sentiment analysis within the domain of movie reviews. We evaluated the performance using key metrics including Accuracy, Precision, Recall, and F1-score.

3.4 EVALUATION METRICS

For comparing the results while hyperparameter tuning, we have used four evaluation metrics Accuracy, precision, Recall & F1-Measure. Accuracy measures the ratio of correctly predicted instances to the total number of instances in a dataset. It provides an overall assessment of the model’s correctness.

$$\text{Accuracy} = \frac{True\ Positives + True\ Negatives}{Total\ Instances}$$

Precision quantifies the ability of a classifier to correctly identify positive instances. It represents the ratio of true positive predictions to all positive predictions made by the model.

$$\text{Precision} = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

Recall measures the ability of a classifier to identify all positive instances. It is the ratio of true positive predictions to the total actual positive instances in the dataset.

$$\text{Recall} = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

F1-Score is a harmonic mean of precision and recall. It considers both false positives and false negatives equally in the evaluation.

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

4. RESULTS AND DISCUSSION

The progression of experiments, which began with a modest collection of 100 reviews and expanded to encompass 50,000 reviews, demonstrated a noticeable pattern in the performance of the model. As the volume of reviews increased, the accuracy of the sentiment analysis model exhibited a notable ascent. Initial evaluations with a smaller dataset showed an accuracy of 43.44% for 100 reviews, gradually improving to 46.97% for 200 reviews and 49.29% for 500 reviews. With the further augmentation in the dataset size, the accuracy continued to rise consistently, reaching 85.58% accuracy when trained and tested on the complete dataset of 50,000 reviews. Notably, a significant boost in accuracy was observed as the dataset surpassed certain thresholds, notably achieving a sharp increase from 69.55% at 5,000 reviews to 78.62% at 10,000 reviews, eventually plateauing around 85% as the dataset scaled beyond 25,000 reviews. Fig. 4 represents the Accuracy of CNN deep learning network across number of reviews.

In an effort to optimize the Convolutional Neural Network (CNN) architecture for sentiment analysis, a comprehensive hyperparameter tuning experiment was conducted by varying the number of convolutional layers from 1 to 9, as illustrated in Fig. 5. While keeping a fixed Dimension of 30 and consistent Relu and tanh activations across all configurations, the evaluation parameters were tracked and are summarized in Table I. The experiment revealed intriguing insights into the model's performance metrics. Notably, configurations with 3 convolutional layers demonstrated a significant improvement in accuracy, precision, recall, and F1-score, showcasing a balanced performance across multiple evaluation metrics. Based on this analysis and to strike a balance between performance and complexity, the decision was made to proceed with a CNN architecture comprising 3 convolutional layers.

Continuing the optimization of our Convolutional Neural Network (CNN) architecture, a focused investigation was undertaken, maintaining a constant configuration of 3 convolutional layers while varying the embedding dimensions from 5 to 45, as illustrated in Fig. 6. The loss values observed across training and validation sets for each dimension over four epochs were recorded, as shown in the figure. Notably, the evaluation metrics such as accuracy, precision, recall, and F1-score were also captured and summarized in Table II. These metrics offered invaluable insights into the model's performance across different embedding dimensions. Particularly, configurations with 25 dimensions exhibited remarkable performance, demonstrating higher accuracy and balanced precision, recall, and F1-score metrics compared to other dimensions. Consequently, after careful evaluation of these results, it was decided to proceed with 25 dimensions as the optimal configuration for our CNN architecture.

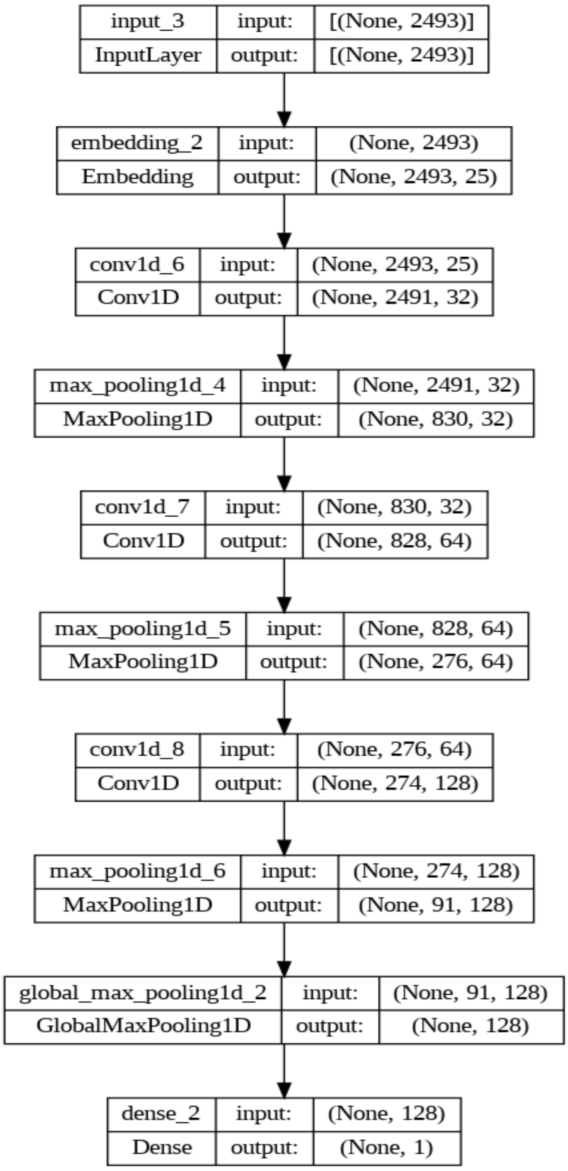

Fig. 3. Optimized CNN Architecture created from Python code

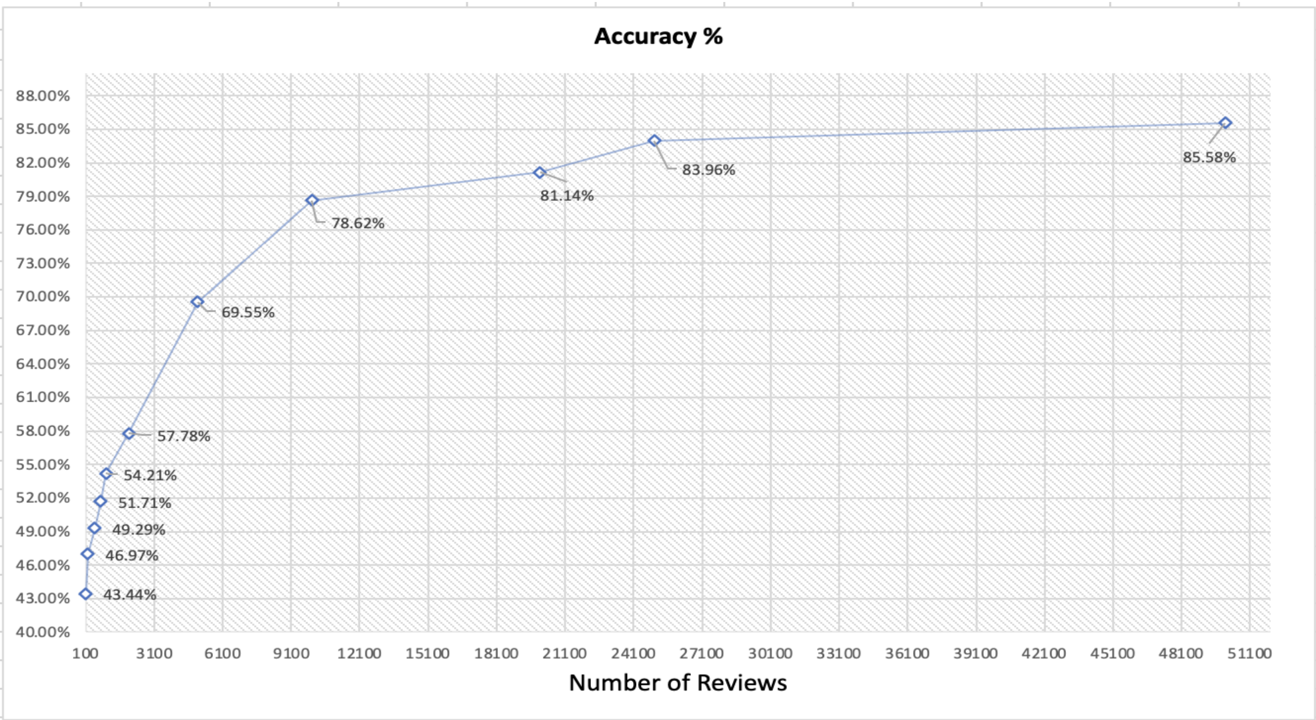

Fig. 4. Accuracy of CNN deep learning network across number of reviews (first Experiment)

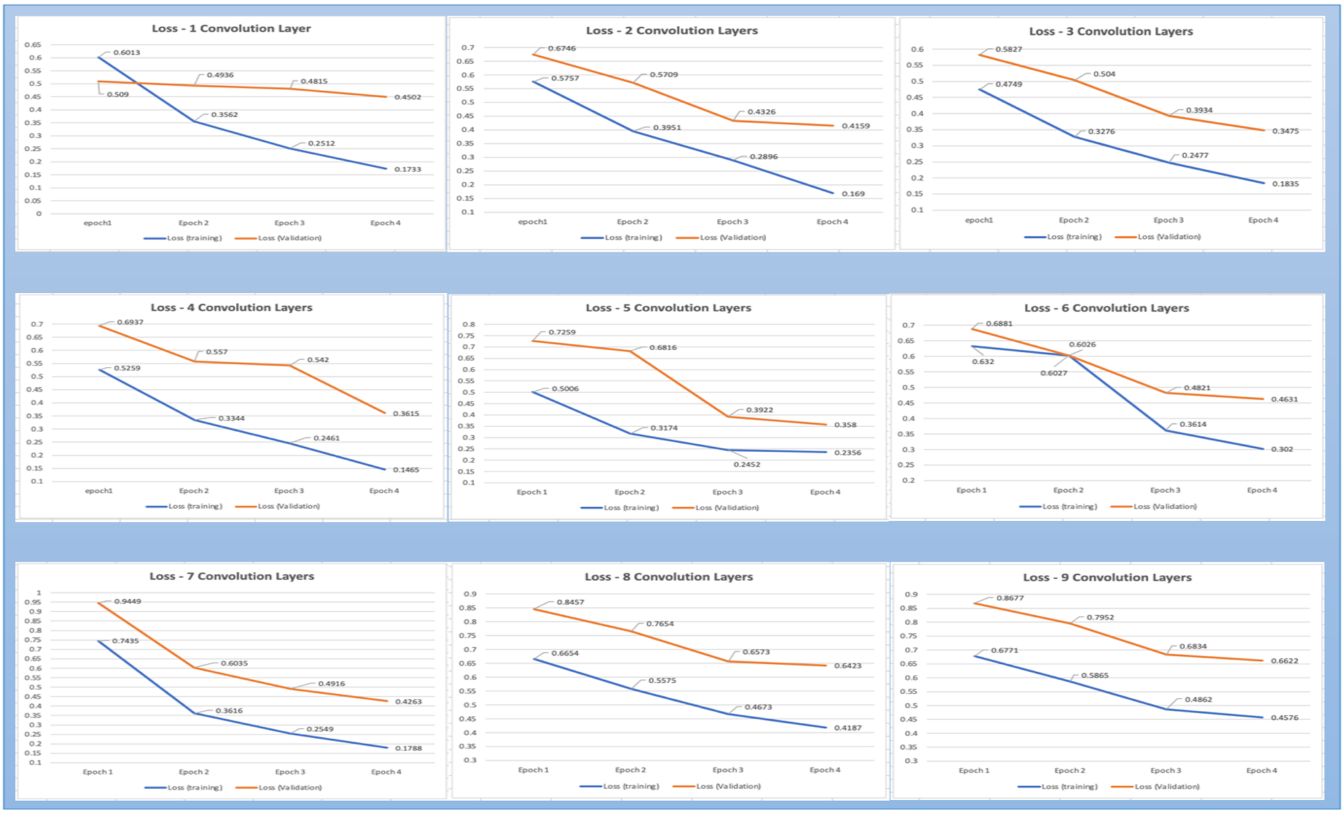

Fig. 5. Hyperparameter tuning for CNN with 50000 movie reviews using 1 to 9 convolution layers

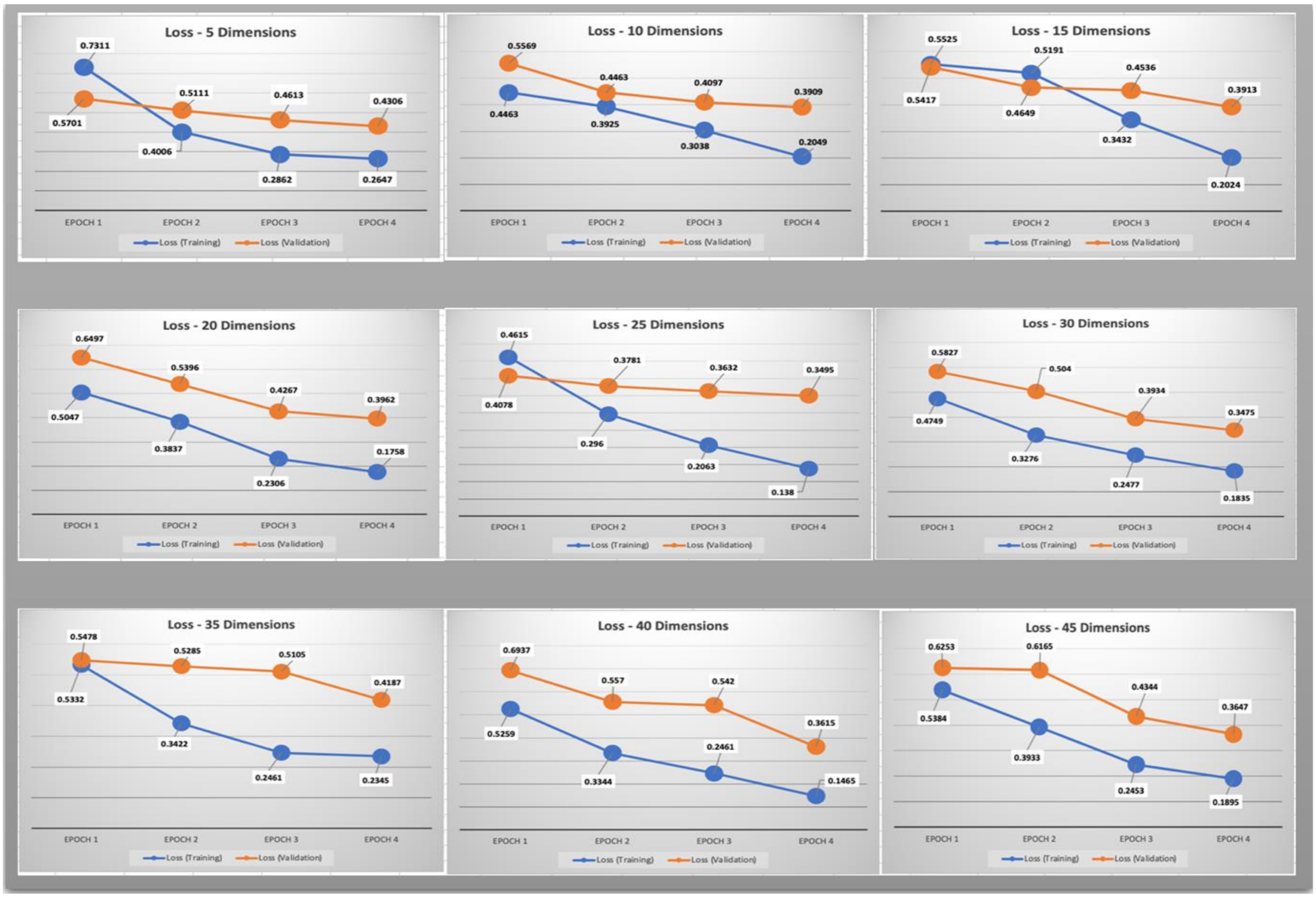

Fig. 6. Hyperparameter tuning for CNN with 50000 movie reviews using 5 to 45 dimensions

After extensive experimentation and analysis, our research culminated in the development of an optimized Convolutional Neural Network (CNN) architecture tailored for sentiment analysis of movie reviews. The final architecture, comprising three convolutional layers with dimensions set at 25, has been rigorously fine-tuned and selected based on meticulous evaluation across various configurations. This architecture demonstrates robust performance with improved accuracy, precision, recall, and F1-score metrics compared to alternate configurations tested throughout the study. The careful selection of the architecture's parameters aims to strike a balance between computational efficiency and predictive efficacy, positioning it as an optimal model for sentiment analysis tasks on movie review datasets.

4.1 COMPARISON WITH THE EXISTING STUDIES USING THE SAME BENCHMARK DATASET

Our proposed methodology provides an improvement in accuracy in sentiment analysis of movie reviews, as evidenced by its comparative performance against existing studies in the field. Across a spectrum of previous research conducted by Mathias et al. [10], Palak et al. [11], Atiqur et al. [9], Lin et al. [8], and Harsh et al. [1], our approach yields promising outcomes, achieving an accuracy rate of 88.92% as illustrated in Fig. 7. While the landscape of sentiment analysis in movie reviews has seen commendable efforts in recent years, our methodology's high accuracy underscores its effectiveness and potential for further enhancements in this domain.

5. CONCLUSION

In conclusion, the development of an optimized Convolutional Neural Network (CNN) architecture for sentiment analysis in movie reviews has demonstrated remarkable performance. The architecture, meticulously designed and fine-tuned through comprehensive experimentation, achieved an accuracy of 88.92%, outperforming existing methodologies in the field. This research has underscored the significance of tailored

CNN architectures for sentiment analysis tasks, especially when handling vast datasets like the one sourced from the IMDB database.

The benefits of this work extend beyond enhanced accuracy. The optimized architecture showcased robustness and resilience, maintaining its efficacy across varying dataset sizes and convolutional layer configurations. This versatility highlights its potential for scalability and adaptability to diverse sentiment analysis tasks and datasets beyond movie reviews.

Looking ahead, further research can explore several promising avenues. Experimentation with different neural network architectures beyond CNNs, such as recurrent neural networks (RNNs) or attention mechanisms, could offer insights into improved sentiment analysis. Additionally, investigating transfer learning techniques or ensemble models might enhance generalization to unseen data and mitigate overfitting concerns. Our research opens doors for future endeavors in sentiment analysis by laying a solid foundation with an optimized CNN architecture.

TABLE I. PERFORMANCE OF CNN WITH DIFFERENT CONVOLUTION LAYERS

Conv Layer	Accuracy	Precision	Recall	F1-Score
1	0.8596	0.8014	0.9268	0.8596
2	0.8558	0.8106	0.9041	0.8548
3	0.8799	0.8651	0.9032	0.879
4	0.8737	0.8591	0.8921	0.8753
5	0.8711	0.86	0.8749	0.8674
6	0.8441	0.8325	0.8638	0.8479
7	0.8588	0.8825	0.8222	0.8513
8	0.6795	0.6383	0.8542	0.7306
9	0.6473	0.6233	0.7597	0.7106

TABLE II. PERFORMANCE OF CNN ACROSS DIFFERENT DIMENSIONS

Dimension	Accuracy	Precision	Recall	F1-score
5	0.8664	0.8273	0.9071	0.8653
10	0.8744	0.839	0.8979	0.8675
15	0.8643	0.8136	0.9186	0.8629
20	0.8652	0.8401	0.8839	0.8614
25	0.8892	0.8748	0.9109	0.8925
30	0.8799	0.8561	0.9032	0.879
35	0.8676	0.8269	0.9031	0.8633
40	0.8737	0.8591	0.8921	0.8753
45	0.8702	0.8492	0.8883	0.8683

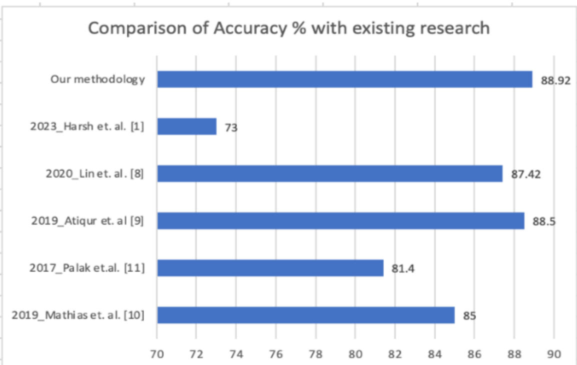

Fig. 7. Accuracy % comparison with existing research

REFERENCES

[1] Harsh Sharma, Satyajit Pangaonkar, Reena Gunjan and Prakash Rokade, "Sentimental Analysis of movie reviews using Machine Learning," *ITM Web of Conferences*, Vol. 53, June 2023.

[2] S. Vernikov, A. Lyras and A. Kanavos, "Multiclass sentiment analysis on COVID-19-related tweets using deep learning models", *Neural Computing and Applications*, Vol. 34, pp. 19615–19627, November 2022.

[3] A. Pimpalkar and J.R. Raj, "MBiLSTMGloVe: Embedding GloVe knowledge into the corpus using multi-layer BiLSTM deep learning model for social media sentiment analysis", *Expert Systems with Applications*, Vol. 203, October 2022.

[4] Aytuğ Onan, "Bidirectional convolutional recurrent neural network architecture with group-wise enhancement mechanism for text sentiment classification", *Journal of King Saud University - Computer and Information Sciences*, Vol. 34, pp. 2098-2117, May 2022.

[5] Jacqueline Kazmaier and Jan H. van Vuuren, "The power of ensemble learning in sentiment analysis", *Expert Systems with Applications*, Vol. 187, January 2022.

[6] N. Dhamayanthi and B. Lavanya, "Sentiment Analysis Framework for E-Commerce Reviews Using Ensemble Machine Learning Algorithms," in *Data Engineering and Intelligent Computing. Advances in Intelligent Systems and Computing*, Vol. 1407, V. Bhateja, S.C. Satapathy, C.M. Travieso-González, V.N.M. Aradhya, Eds. Springer, Singapore, May 2021.

[7] H. Kaur, S.U. Alankar, B. Alankar, and V. Chang, "A Proposed Sentiment Analysis Deep Learning Algorithm for Analyzing COVID-19 Tweets", *Information Systems Frontiers*, Vol. 23, pp. 1417-1429, April 2021.

[8] Lin Li, Tiong-Thye Goh, and Dawei Jin, "How textual quality of online reviews affect classification performance: a case of deep learning sentiment analysis", *Neural Computing and Applications*, vol. 32, pp. 4387-4415, May 2020.

[9] A. Rahman and M. S. Hossen, "Sentiment analysis on movie review data using Machine Learning Approach," *2019 International Conference on Bangla Speech and Language Processing (ICBSLP)*, Sylhet, Bangladesh, pp. 1-4, September 2019.

[10] Mathias Kraus and Stefan Feuerriegel, "Sentiment analysis based on rhetorical structure theory: Learning deep neural networks from discourse trees", *Expert Systems with Applications*, vol. 118, pp. 65-79, March 2019.

[11] Palak Baid, Apporva Gupta and Neelam Chaplot, "Sentiment analysis of movie reviews using Machine Learning Techniques", *International Journal of Computer Applications*, Vol. 179, pp. 45-49, December 2017.

[12] Andrew L. Maas et. al, "Learning Word Vectors for Sentiment Analysis," in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, Association for Computational Linguistics. 2011, pp. 142–15.

[13]Yanping Huang, Qian Liu, Hong Peng, Jun Wang, Qian Yang, and David Orellana-Martín, "Sentiment classification using bidirectional LSTM-SNP model and attention mechanism", *Expert Systems with Applications*, Vol. 221, July 2023.

[14] Neha Punetha and Goonjan Jain, "Bayesian game model based unsupervised sentiment analysis of product reviews", *Expert Systems with Applications*, Vol. 214, March 2023.

15] Qiang Lu, Xia Sun, Richard Sutcliffe, Yaqiong Xing, and Hao Zhang, "Sentiment interaction and multi-graph perception with graph convolutional networks for aspect-based sentiment analysis", *Knowledge-Based Systems*, Vol. 256, November 2022.

[16] Manju Venugopalan and Deepa Gupta, "An enhanced guided LDA model augmented with BERT based semantic strength for aspect term extraction in sentiment analysis", *Knowledge-Based Systems*, vol. 246, June 2022.

[17] Necip Gozuacik, C. Okan Sakar, and Sercan Ozcan, "Social media-based opinion retrieval for product analysis using multi-task deep neural networks", *Expert Systems with Applications*, vol. 183, November 2021.

[18] Mohammad Ehsan Basiri, Shahla Nemati, Moloud Abdar, Somayeh Asadi, U., and Rajendra Acharrya, "A novel fusion-based deep learning model for sentiment analysis of COVID-19 tweets", *Knowledge based systems*, 228, September 2021.

[19] Nan Jing, Zhao Wu, and Hefei Wang, "A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction", *Expert Systems with Applications*, Vol. 178, September 2021.

[20] Marouane Birjali, Mohammed Kasri, and Abderrahim Beni-Hssane, "A comprehensive survey on sentiment analysis: Approaches, challenges and trends", *Knowledge-Based Systems*, Vol. 226, August 2021

[21] Erick Odhiambo Omuya, George Onyango Okeyo, and Michael Waema Kimwele, "Feature Selection for Classification using Principal Component Analysis and Information Gain", *Expert Systems with Applications*, Vol. 174, July 2021.

[22]Asad Abdi, Shafaatunnur Hasan, Siti Mariyam Shamsuddin, Norisma Idris, and Jalil Piran, "A hybrid deep learning architecture for opinion-oriented multi-document summarization based on multi-feature fusion", *Knowledge-Based Systems*, Vol. 213, February 2021.

[23] Wei Song, Zijian Wen, Zhiyong Xiao and Soon Cheol Park, "Semantics perception and refinement network for aspect-based sentiment analysis", *Knowledge-Based Systems*, Vol. 214, February 2021.

[24] Ranjan Kumar Behera, Monalisa Jena, Santanu Kumar Rath, and Sanjay Misra, "Co-LSTM: Convolutional LSTM model for sentiment analysis in social big data", *Information Processing & Management*, Vol. 58, January 2021.

[25] Donatas Meškelė and Flavius Frasincar, "ALDONAr: A hybrid solution for sentence-level aspect-based sentiment analysis using a lexicalized domain ontology and a regularized neural attention model", *Information Processing & Management*, Vol. 57, May 2020.

[26] Osman Gokalp, Erdal Tasci, and Aybars Ugur, "A novel wrapper feature selection algorithm based on iterated greedy metaheuristic for sentiment classification", *Expert Systems with Applications*, Vol. 146, May 2020.

[27] Ning Liu and Bo Shen, "Aspect-based sentiment analysis with gated alternate neural network", *Knowledge-Based Systems*, Vol. 188, January 2020.

[28] SoYeop Yoo, JeIn Song, and OkRan Jeong, "Social media contents based sentiment analysis and prediction system", *Expert Systems with Applications*, Vol. 105, September 2018.

[29]Gichang Lee, Jaeyun Jeong, Seungwan Seo, CzangYeob Kim, and Pilsung Kang, "Sentiment classification with word localization based on weakly supervised learning with a convolutional neural network", *Knowledge-Based Systems*, Vol. 152, July 2018.

[30] Mangi Kang, Jaelim Ahn, and Kichun Lee, "Opinion mining using ensemble text hidden Markov models for text classification", *Expert Systems with Applications*, Vol. 94, March 2018.

[31] Aytuğ Onan, Serdar Korukoğlu, and Hasan Bulut, "A multiobjective weighted voting ensemble classifier based on differential evolution algorithm for text sentiment classification", *Expert Systems with Applications*, Vol. 62, November 2016.

[32] Abinash Tripathy, Ankit Agrawal, and Santanu Kumar Rath, "Classification of sentiment reviews using n-gram machine learning approach", *Expert Systems with Applications*, Vol. 57, September 2016.