|  |  |  |
| --- | --- | --- |
| **KONERU LAKSHMAIAH EDUCATION FOUNDATION**  **AZIZ NAGAR, HYDERABAD**  **DEPARTMENT OF ECE**  **Project Proposal** | | |
| **1.0** | **Details of Candidates:** | (i) G Veera Yasaswini (2210040029)  (ii) P Ram Charan Tej (2210040037)  (iii)K Akshitha (2210040067)  (iv)A Manisha (2210040068) |
| **Course of Study:** | B.TECH/ECE |
| **Year:** | IV |
| **Semester:** | I |
| **2.0** | **Course Details:** | 22IE4053A   ENGINEERING CAPSTONE PROJECT - PHASE 1 |
| **3.0** | **Name of Supervisor:** | Dr. Ngangbam Phalguni Singh,  Associate Professor, KLEF/ECE |
| **4.0** | **Proposed Title:** | Enhancing Sentiment Analysis of Movie Reviews |

**August, 2025**

* 1. **Introduction**
  2. **General Introduction**

The rapid rise of web-based reviews has ensured sentiment analysis has become an accepted tool to read the public psyche, especially within the entertainment sector. Reviews for films, specifically, play an important role in audience choices and offer worthwhile feedback regarding viewer tendencies. Examination of these textual data allows stakeholders, such as film makers, producers, and advertising agencies, to assess the feelings of audiences and refine upcoming productions. Classic sentiment analysis methods are mostly lexicon based, and they are best represented by tools like TextBlob and VADER that categorize sentiments with pre-specified words and polarity scores. Although these methods are easy to interpret and use, they have the disadvantage of context sensitive sentiments and sophisticated linguistic patterns. To mitigate such shortcomings, more advanced models like Recurrent Neural Networks (RNNs) and their extensions, i.e., Gated Recurrent Units (GRU), have emerged as top-of-the line solutions because they are capable of learning sequentially complex patterns in text data efficiently.

In this research, we propose a hybrid method that combines word embeddings generated through Word2Vec with traditional and deep learning-driven sentiment analysis techniques. Word2Vec embeds text data in dense vector spaces, thus enriching semantic understanding of words. TextBlob and VADER are used as the basis for sentiment classification, and GRU is incorporated to identify contextual dependencies, thus increasing classification accuracy. The IMDB movie review corpus serves as a baseline to determine the efficacy of the approaches.

The primary goals of this work are presented below:

* For comparison of lexicon-based versus deep learning-based sentiment analysis techniques.
* For analyzing whether Word2Vec embeddings can improve sentiment classification accuracy.
* To compare the performance of GRU in handling compound sentiments with traditional methods.

By combining traditional and deep learning-based sentiment analysis techniques, this research provides an exhaustive comparative study, emphasizing the strengths and weaknesses inherent in each technique in movie review classification.

* 1. **Problem Statement**

The problem addressed in this research revolves around the limitations of traditional lexicon-based sentiment analysis methods in handling complex linguistic patterns, context sensitivity, sarcasm, and domain dependence. Existing approaches such as TextBlob and VADER, while effective for short and simple texts, often struggle with nuanced and compound sentiments. Furthermore, traditional methods fail to leverage contextual word relationships that modern deep learning approaches like Word2Vec and GRU can model effectively. This project aims to overcome these limitations by combining both traditional and deep learning-based techniques for improved sentiment classification accuracy on movie reviews.

* 1. **Objectives of the study**

(i) To compare lexicon-based versus deep learning-based sentiment analysis techniques.  
(ii) To analyze whether Word2Vec embeddings can improve sentiment classification accuracy.  
(iii) To compare the performance of GRU in handling compound sentiments with traditional methods.

* 1. **Scope of the Project**

The scope of this project includes designing, implementing, and evaluating a hybrid sentiment analysis framework that integrates Word2Vec embeddings, lexicon-based approaches (TextBlob, VADER), and deep learning models (GRU). The evaluation will be based on the IMDB movie review dataset, with performance metrics including accuracy, precision, recall, and F1-score. The results will be compared across methods to identify strengths, weaknesses, and best practices for movie review sentiment analysis.

* 1. **Literature Review**

Several research studies have explored sentiment analysis using a range of techniques, from traditional lexicon-based approaches to modern deep learning methods. One hybrid sentiment analysis approach integrated rule-based, supervised, and machine learning models, improving classification effectiveness by balancing precision and recall through a higher F1 score. Another work focused on Twitter sentiment analysis, automatically collecting and processing a corpus for opinion mining, and building a classifier that outperformed earlier strategies in categorizing positive, negative, and neutral sentiments.

Emoji prediction has also been addressed as a text classification problem in multiple studies. Using a dataset of 1.5 million tweets, researchers trained LSTM-RNN and CNN models, finding that CNN achieved higher accuracy and F1-scores than both LSTM and baseline models. Beyond social media, a Street Illumination Mapping algorithm was proposed using smartphone sensors and an IoT-cloud framework, generating high-granular nighttime lighting maps that surpassed traditional methods when tested with real data from Kolkata.

Text classification beyond sentiment analysis has also contributed to methodological advances. WebClassify, a naïve Bayes-based tool with a modified multinomial model, demonstrated improved accuracy with larger vocabularies. Similarly, fuzzy set-based collaborative filtering models combined expert opinions with user feedback to solve cold-start problems in recommendation systems. In multilingual sentiment analysis, a triangulation method was developed to construct more precise sentiment dictionaries compared to standard machine-translated lists.

Other studies focused on opinion mining systems capable of extracting product features and summarizing sentiments from online reviews using IR-based feature extraction. Sentiment classification applied to travel blog reviews compared Naïve Bayes, SVM, and character-based N-gram models, concluding that SVM and N-gram consistently outperformed Naïve Bayes with accuracies above 80%. In probabilistic NLP, improvements to variational EM algorithms using mixture-based posteriors allowed the inclusion of soft constraints, increasing model flexibility.

Comparative studies between multivariate Bernoulli and multinomial naïve Bayes models revealed that the multinomial variant performs better with large vocabularies, achieving accuracy improvements of up to 27%. Sentiment analysis has also been applied to the education domain, where a study analyzing over 45,000 tweets found predominantly negative sentiment worldwide, but some positive trends in low-income countries, along with gender-based differences in subjectivity. For web content, a sentiment-oriented crawling framework was proposed to rapidly extract opinionated text from film and hotel reviews, finding Naïve Bayes superior for films and comparable to K-NN for hotel reviews.

Competitions like SemEval 2014 have driven sentiment analysis advancements, with systems blending word representations and traditional features to achieve notable gains. SO-CAL, a lexicon-based sentiment system using annotated dictionaries for polarity, strength, intensification, and negation, achieved consistent results across domains, validated via crowdsourced dictionary checks. Aspect-level sentiment analysis also progressed through unsupervised hierarchical rule-based methods, achieving high recall on benchmark datasets and outperforming several state-of-the-art models.

Sentiment analysis itself, also called opinion mining, is a key Natural Language Processing (NLP) task that classifies text as positive, negative, or neutral. It is used in industries such as business, social media, entertainment, and finance to monitor public opinion, improve decision-making, and enhance customer experience. In the context of movie reviews, it provides filmmakers with insights into audience reactions and helps viewers make informed choices. Traditional approaches like lexicon-based models are easy to use but often fail to capture sarcasm, negation, and nuanced context, prompting the rise of machine learning and deep learning models.

The importance of sentiment analysis in the film industry is underscored by its ability to reveal audience emotions, optimize promotional strategies, and improve content. Streaming platforms such as Netflix and Amazon Prime leverage sentiment trends to enhance recommendation systems. Machine learning techniques, supported by TF-IDF vectorization, facilitate real-time review monitoring, enabling better interaction with audiences and supporting brand reputation management.

Specific applications in movie review analysis have used datasets of over 20,000 reviews, preprocessed to remove HTML tags, special characters, and stopwords, followed by stemming. These were transformed into numerical features using TF-IDF, and models like Logistic Regression and Decision Tree Classifier were tested. Results showed Decision Trees achieving over 71% accuracy with balanced precision and recall.

Despite these advancements, sentiment analysis faces ongoing challenges. Linguistic ambiguity makes word meaning highly context-dependent, sarcasm and irony are difficult for models to detect, and domain dependence limits model generalization across topics. Handling negations like “not bad” requires sophisticated interpretation, and distinguishing subjective opinions from objective facts remains complex. Aspect-Based Sentiment Analysis (ABSA) demands fine-grained evaluation of sentiment toward specific attributes rather than entire texts, while fake reviews and unbalanced datasets can distort predictions. Finally, there is growing interest in capturing more nuanced emotions beyond simple positive/negative classification, such as joy, anger, or sadness.

Overall, the literature indicates a clear trend toward hybrid approaches that combine the interpretability of lexicon-based models with the contextual depth of deep learning architectures. This project builds on that direction by integrating Word2Vec embeddings, TextBlob, VADER, and GRU to create a robust sentiment analysis framework for movie reviews.

1. **Abstract:**

Sentiment analysis is a key Natural Language Processing (NLP) task used to determine opinions and emotions from text, with significant applications in the entertainment industry where movie reviews influence audience decisions and provide valuable feedback to filmmakers. Traditional lexicon-based approaches such as TextBlob and VADER are easy to interpret but struggle with complex linguistic patterns, sarcasm, negations, and context sensitivity, limiting their effectiveness in nuanced sentiment classification. This project proposes a hybrid sentiment analysis framework that combines Word2Vec embeddings, which capture semantic word relationships, with both lexicon-based methods and a Gated Recurrent Unit (GRU)-based deep learning model to improve classification accuracy. Using the IMDB Movie Review Dataset of 50,000 balanced positive and negative reviews, the system applies preprocessing steps including lowercasing, punctuation removal, stopword elimination, and stemming. The models will be evaluated on accuracy, precision, recall, and F1-score. It is expected that the GRU model with Word2Vec embeddings will outperform traditional methods, demonstrating better handling of contextual dependencies and complex sentiment expressions, ultimately providing a comprehensive comparative analysis of lexicon-based and deep learning approaches for movie review classification.

1. **Methodology**

The present work utilizes the IMDB Movie Review Dataset, containing 50,000 reviews equally split between positive and negative sentiments. Preprocessing involves lowercasing, punctuation and special character removal, stopword removal, and stemming. Word2Vec is used for embedding words into 100-dimensional vectors, while TextBlob and VADER are employed for lexicon-based sentiment scoring. A GRU-based deep learning model processes sequential dependencies for classification, with evaluation metrics including accuracy, precision, recall, and F1-score.

1. **Expected Output**

The expected output is a sentiment analysis model that outperforms traditional lexicon-based approaches in accuracy, precision, recall, and F1-score. Specifically, the GRU with Word2Vec embeddings is expected to demonstrate superior handling of contextual and compound sentiments compared to TextBlob and VADER. The final system will provide a comprehensive comparative analysis of the methods, offering insights into their strengths and weaknesses for movie review classification.

1. **Other Relevant Information**
   1. **Financial Arrangements**

The budget is given below:

|  |  |  |  |
| --- | --- | --- | --- |
| **S/N** | **ITEM** | **DESCRIPTION** | **COST (INR)** |
| 1 | Computer Resources | High-performance computing system for model training and testing | 25,000 |
| 2 | Software & Libraries | Python libraries (TensorFlow, Gensim, Scikit-learn, NLTK, Pandas), open-source | 0 |
| 3 | Dataset Acquisition | IMDB Movie Review Dataset (publicly available) | 0 |
| 4 | Internet & Cloud Services | Cloud storage and computation services for backup and large-scale model training | 5,000 |
| 5 | Miscellaneous | Printing, documentation, and project report preparation | 2,000 |
|  | **Grand Total** |  | **32,000** |

* 1. **Duration (chart required)**

This project will be completed in one year. The proposed schedule is given below:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | **SL.NO.** | **TASK NAME** | **2025 JUL** | **AUG** | **SEP** | **OCT** | **NOV** | **DEC** | | --- | --- | --- | --- | --- | --- | --- | --- | | 1 | Literature review | ✔ | ✔ |  |  |  |  | | 2 | Data collection & system analysis |  | ✔ | ✔ |  |  |  | | 3 | System Design and Development |  |  | ✔ | ✔ |  |  | | 4 | Prototype testing & installation |  |  |  | ✔ | ✔ |  | | 5 | Writing report |  |  |  |  | ✔ | ✔ | | 6 | Submission |  |  |  |  |  | ✔ | |

**10.0 References (MINIMUM OF 10)**

[1] Rudy Prabowo and Mike Thelwall, “Sentiment Analysis: A Combined Approach,” Journal of Informetrics, Vol. 3, Issue 2, pp. 143-157, 2009.  
[2] Alexander Pak, Patrick Paroubek, “Twitter as a Corpus for Sentiment Analysis and Opinion Mining,” International Conference on Language Resources and Evaluation, pp. 1320-1326, 2010.  
[3] Mary Margarat Valentine, Ms. Veena Kulkarni, Dr. R. R. Sedamkar, “A Model for Predicting Movie’s Performance using Online Rating and Revenue,” International Journal of Scientific Engineering Research, Vol. 4, No. 9, pp. 277-282, 2013.  
[4] Luda Zhao, Connie Zeng, “Using Neural Networks to Predict Emoji Usage from Twitter Data,” Semantic Scholar, pp. 1-6, 2017.  
[5] G. S. Tomar, S. Verma, Ashish Jha, “Web Page Classification using Modified Naive Bayesian Approach,” IEEE TENCON-2006, pp. 1-4, Nov 2006.  
[6] H.-A. W. Li-Chen Cheng, “A novel fuzzy recommendation system integrated with experts’ opinion,” IEEE International Conference on Fuzzy Systems, 2011.  
[7] Josef Steinberger, “Creating sentiment dictionaries via triangulation,” Decision Support Systems, Vol. 53, No. 4, pp. 5, 2012.  
[8] Kushal Dave, “Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews,” Proc. 12th Int. Conf. World Wide Web, pp. 10, 2003.  
[9] Z. Z. R. L. Qiang Ye, “Sentiment classification of online reviews to travel destinations by supervised learning,” Expert Systems with Applications, 2009.  
[10] C. P.-R. Mahesh Joshi, “Generalizing Dependency Features for Opinion Mining,” Proceedings of ACL-IJCNLP 2009 Conference Short Papers, 2009.

**CANDIDATES**

Name: **G Veera Yasaswini** Reg. No.**2210040029**

Signature: ……………………… Date: …………

Name:  **P Ram Charan Tej** Reg. No.**2210040037**

Signature: ……………………… Date: …………

Name: **K Akshitha** Reg. No.**2210040067**

Signature: ……………………… Date: …………

Name: **A Manisha** Reg. No.**2210040068**

Signature: ……………………… Date: …………

**SUPERVISOR**

1. Comments by Supervisor:

………………………………………………………………………………………………………

……………………………………………..…………………………..……………………………

………………………………………………………………………………………………………

……………............................................

Date: ……............ Name: ……....……….…………..

Signature: .…………………........