Vidyavardhini's College of Engineering and Technology, Vasai

Department of Computer Science & Engineering (Data Science)

AY: 2025-26

Class:	BE-CSE(DS)	Semester:	VII
Course Code:	CSDOL7011	Course Name:	NLP Lab

Name of Student:	Hitesh Shetye
Roll No. :	49
Experiment No.:	1
Title of the Experiment:	Identifying and Critically Reviewing Research Papers on a Selected NLP Application
Date of Performance:	
Date of Submission:	

Evaluation

Performance Indicator	Max. Marks	Marks Obtained
Performance	5	
Understanding	5	
Journal work and timely submission	10	
Total	20	

Performance Indicator	Exceed Expectations (EE)	Meet Expectations (ME)	Below Expectations (BE)
Performance	4-5	2-3	1
Understanding	4-5	2-3	1
Journal work and timely submission	8-10	5-8	1-4

Checked by

Name of Faculty   :

Signature   :

Date   :

CSDOL7011: Natural Language Processing Lab

Aim: To perform a critical literature review of five research papers on a chosen NLP application, focusing on the problem addressed, solution proposed, and limitations identified in each work.
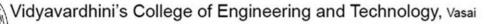
Objective: To critically review five research papers on an NLP application, analyzing problems, solutions, and limitations.

Tools Required:

1. Research databases: Google Scholar, IEEE Xplore, ACM Digital Library, Springer, Elsevier, or arXiv
2. MS Word / Google Docs or LaTeX for documentation
3. Internet access

Procedure:

1. Select a Real-World NLP Application:
   a. Choose any one topic such as:
      i. Machine Translation
      ii. Sentiment Analysis
      iii. Text Summarization
      iv. Question Answering Systems
      v. Chatbots
      vi. Named Entity Recognition
      vii. Information Retrieval, etc.
2. Search and Select 5 Research Papers:
   a. Preferably from peer-reviewed journals or conferences.
   b. Papers must be recent (preferably from the last 5–7 years).
   c. Ensure papers are directly relevant to the chosen application.
3. Read and Analyze Each Paper:
   a. Focus on these three aspects for each paper:
      i. Problem Statement: What problem or challenge does the paper address?
      ii. Proposed Solution: What model, algorithm, or framework is presented?

      iii. Critical Evaluation: What are the limitations, gaps, or areas for improvement?

4. Prepare the Review Document:

    a. Create a structured table or section for each paper with:

      i. Title, Authors, Year, and Source

      ii. Summary of problem

      iii. Summary of solution

      iv. Critical remarks and insights

5. Submit the Review:

    a. The write-up should be minimum 4–5 pages.

    b. Include references in standard citation format (APA/IEEE/MLA).

Description of the Experiment:

This experiment introduces students to academic research and the process of reviewing scientific literature in the field of NLP. It allows them to explore cutting-edge developments, analyze technical approaches, and reflect critically on current limitations. This lays the foundation for their final-year projects or research internships.

Detailed Description of the NLP Technique:

Since this experiment is open-ended, the NLP techniques will vary based on the papers selected. However, students will encounter:

- Deep learning architectures (e.g., LSTM, BERT, GPT, T5)
- Statistical NLP approaches (e.g., N-gram models, HMMs)
- Evaluation metrics (e.g., BLEU, ROUGE, Accuracy, F1-score)
- Datasets commonly used for training and evaluation

Conclusion:

Name of the Paper	Solution Used	Drawbacks / Limitations	Areas of Improvement
1] Fine-tuning BERT with BiLSTM for Fine-grained Movie Reviews Sentiment Analysis — Gibson Nkhata et al.	1) Fine-tuning BERT with BiLSTM 2) SMOTE for oversampling 3) NLP Augmenter (NLPAUG) 4) Heuristic for overall polarity	1) High computational cost due to BiLSTM + BERT 2) Complex integration of heuristic logic 3) Limited to English and specific datasets	1) Reduce training time with lighter models 2) Test generalization on multilingual datasets 3) Explore transformer-only architectures
2] Sentiment Analysis of Movie Reviews Using Machine Learning — Subhashree Darshana et al.	1) ML models: KNN, SVM, Random Forest, Logistic Regression 2) TF-IDF, n-grams 3) Lemmatization, visualization (WordCloud)	1) Accuracy not clearly reported 2) No deep learning methods 3) Limited feature analysis	1) Add deep learning comparisons 2) Provide quantitative benchmarks 3) Test on larger and more diverse datasets
3] Sentiment Analysis for IMDb Movie Review Using SVM Method — Fidya Farasalsabila et al.	1) SVM with TF-IDF and BoW 2) Grid search for parameter tuning 3) Preprocessing with standard NLP pipeline	10 Single algorithm focus (SVM only) 2) Binary classification only 3) Lacks model generalization discussion	1) Incorporate ensemble or hybrid models 2) Test with multi-class classification 3) Evaluate deep learning comparisons

4] Sentiment Analysis Using NLP (Multi-modal) — Om Jagtap et al.	1) NLP + Computer Vision (FER) 2) TextBlob, VADER, Flair, Text2Emotion 3) Real-time analysis with Streamlit	1) Lacks quantitative accuracy metrics 2) No deep learning model integration 3) Limited facial emotion dataset validation	1) Integrate deep learning (CNN, LSTM) 2) Improve FER model with custom datasets 3) Add multi-language sentiment analysis
5] Sentiment Analysis of IMDb Movie Review Using Deep Learning — Shubham Vaibhav Kanse et al.	1) LSTM & RNN 2) Tokenization, Lemmatization, Padding 3) Word Embedding, TF-IDF	1) Exact accuracy not given 2) RNN model limitations with longer dependencies 3) No benchmark comparison	1) Provide clearer accuracy metrics 2) Compare with transformer-based models (e.g., BERT) 3) Include fine-grained analysis

CSDOL7011: Natural Language Processing Lab