AY: 2025-26

Class:	BE	**Semester:**	VII
Course Code:	CSDOL7011	**Course Name:**	Natural Language Processing

Name of Student:	Konisha Jayesh Thakare
Roll No. :	71
Experiment No.:	1
Title of the Experiment:	Identifying and Critically Reviewing Research Papers on a Selected NLP Application
Date of Performance:	15.07.2025
Date of Submission:	22.07.2025

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 : Dr. Tatwadarshi P. Nagarhalli

Signature :

Date :

Aim: To identify and critically review research papers on the application of Sentiment Analysis in Natural Language Processing, in order to understand methodologies, results, and research gaps.

Theory:

Natural Language Processing (NLP) enables machines to understand and process human language. One of its most widely used applications is **Sentiment Analysis (SA)**, which determines the emotional tone of text — positive, negative, or neutral.

Applications of Sentiment Analysis include:

- Product review mining (e.g., Amazon reviews).
- Social media opinion analysis (e.g., Twitter, Facebook).
- Market prediction and brand monitoring.
- Customer feedback analysis in businesses.

Research in SA has evolved from lexicon-based approaches (dictionary lookups) to machine learning models (SVM, Naïve Bayes), and now to deep learning and transformers (CNN, LSTM, BERT).

Critically reviewing research papers helps identify strengths, limitations, and opportunities in these evolving approaches.

Procedure:

1. Select NLP application → Sentiment Analysis.
2. Identify relevant research papers.
3. Extract details: problem, methodology, datasets, evaluation, results.
4. Compare the works critically.
5. Conclude with insights and future directions.

Critical Review of Research Papers:

Paper 1:
Title: *Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts*
Author: Yoon Kim
Conference: EMNLP 2014
Summary: Proposed CNN-based models for sentiment classification on sentence-level tasks. Outperformed traditional ML baselines.

Dataset: Stanford Sentiment Treebank (SST).
Strengths:

- Captures local word patterns effectively.
- Requires minimal pre-processing.

Limitations:

- Cannot handle long-term dependencies well.
- Needs large labeled data for good performance.

Paper 2

- **Title**: *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*
- **Authors**: Devlin et al.
- **Conference**: NAACL 2019
- **Summary**: Introduced BERT, a transformer-based model trained on large corpora. Achieved state-of-the-art results across multiple NLP tasks, including sentiment analysis.
- **Dataset**: GLUE Benchmark, SST-2, IMDB reviews.
- **Strengths**:
  - Context-aware embeddings (bidirectional).
  - Transfer learning reduces task-specific training effort.
- **Limitations**:
  - Computationally expensive (large model size).
  - Requires fine-tuning for domain-specific tasks.

Paper 3

- **Title**: *Sentiment Analysis of Twitter Data Using LSTM Networks*
- **Authors**: Saif Hassan et al.
- **Journal**: Procedia Computer Science, 2018
- **Summary**: Used Long Short-Term Memory (LSTM) networks to classify sentiment from Twitter data, addressing sequential dependencies in short messages.
- **Dataset**: Twitter Sentiment140 dataset.
- **Strengths**:
  - Handles sequential data better than CNNs.
  - Performs well on short, noisy texts like tweets.
- **Limitations**:
  - Training is slow due to sequential nature.
  - Sensitive to preprocessing quality.

Comparative Review

- CNN (Paper 1): Good for feature extraction, but weak on long-term dependencies.
- LSTM (Paper 3): Handles sequential context better but is slower to train.
- BERT (Paper 2): Outperforms both CNN and LSTM by using bidirectional attention, but requires heavy computation.

Observation:

- The trend in research shows a shift from traditional ML → Deep Learning → Transformer-based architectures.

- While performance has improved, computational cost and explainability remain open challenges.

Conclusion:

From this review, it is evident that Sentiment Analysis has significantly benefited from advancements in deep learning and transformer-based models. Early CNN and LSTM approaches improved upon traditional machine learning methods but faced limitations in capturing long-term dependencies and contextual meaning. Transformer-based models such as BERT have demonstrated superior performance by leveraging pre-training on large corpora and generating context-aware embeddings, thereby achieving state-of-the-art results across sentiment analysis benchmarks. However, challenges such as high computational cost, resource requirements, and lack of interpretability remain. Future research should therefore focus on developing lightweight, domain-adaptive, and explainable models that balance accuracy with efficiency to enable wider and more practical applications of sentiment analysis in real-world scenarios.