Anshu Verma, 2026, 14:2
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science,
Engineering and Technology
An Open Access Journal

Sentiment Analyzer: A Multi-Method Sentiment Analysis System

Anshu Verma, Anup Kumar Choudhary, Jyotiraditya Kathua, Mohit Sharma
Department of Computer Science & Engineering, MGM's College of Engineering and Technology, Noida, India
Supervisor: Mr. Shashikant Maurya, Asst. Professor

Abstract-Sentiment analysis has emerged as a critical application of natural language processing (NLP) in the digita l age. This paper presents Sentiment Analyzer, a comprehensive multi-method sentiment analysis system that combines lexicon-based methods (VADER, TextBlob), machine learning (ML) classifiers, and ensemble techniques to provide accurate and robust sentiment detection. The system implements a modular architecture with components for text preprocessing, sentiment analysis, emotion detection, emoji analysis, and result visualization. A FastAPI-based REST API enables programmatic access, while an interactive Streamlit dashboard provides a user-friendly interface. The ML pipeline employs TF-IDF vectorization with Logistic Regression, Naive Bayes, and Support Vector Machine classifiers. Experimental evaluation on the SST-2 benchmark demonstrates ensemble classification accuracy of 91.3%, outperforming standalone VADER (71.3%) and basic Logistic Regression (81.2%). API endpoints respond in under 50 ms for single-text analysis, and batch processing of 100 texts completes in under 450 ms.

Index Terms—Sentiment analysis, NLP, VADER, TextBlob, machine learning, SVM, TF-IDF, REST API, ensemble methods, emotion detection.

I. INTRODUCTION

The rapid proliferation of social media platforms, review websites, and online forums has generated an unprecedented volume of user-generated content expressing opinions, attitudes, and emotions. Sentiment analysis, also known as opinion mining, is a subfield of NLP that identifies subjective information from text, typically classifying it as positive, negative, or neutral [1]. Applications span business intelligence, social media monitoring, political analysis, and healthcare [2].

Traditional lexicon-based approaches use predefined dictionaries associating words with sentiment polarity. While interpretable, they struggle with context-dependent language, sarcasm, and domain-specific vocabulary. Machine learning (ML) techniques have improved accuracy substantially, though they require labeled training data and can be opaque in their decision-making [5].

This paper presents Sentiment Analyzer, a system that addresses these limitations through a mult i-method approach. The contributions of this work are: (1) integration of VADER, TextBlob, and ML classifiers into a unified analysis engine; (2) an ensemble method achieving 91.3% accuracy on SST-2; (3) an emotion detection module covering eight affective categories; (4) a production-grade REST API and interactive dashboard; and (5) a fully reproducible ML training pipeline with cross-validated model selection.

The remainder of this paper is organized as follows. Section II reviews related work. Section III describes the system architecture. Section IV presents implementation details. Section V reports experimental results. Section VI discusses limitations and future work. Section VII concludes the paper.

II. RELATED WORK

Hutto and Gilbert [1] introduced VADER, a lexicon and rule-based sentiment analyzer optimized for social media text. VADER encodes grammatical rules for negation, boosters, and punctuation intens ity, achieving competitive performance on Twitter data without requiring training data.

The advent of deep bidirectional transformers via BERT [2] established new state-of-the-art benchmarks for sentiment classification through deep contextual understanding of language. Vaswani et al. [19] introduced the attention mechanism underlying these transformer architectures.

Zhang et al. [3] and Abubakar and Farik [4] surveyed aspect-based sentiment analysis (ABSA), demonstrating that fine-grained analysis provides richer insights than document-level polarity. Medhi [5] and Rai and Singh [6] provided comprehensive surveys of deep learning and ML techniques for sentiment classification, respectively.

Kumar and Sharma [18] demonstrated that ensemble methods combining multiple classifiers consistently outperform individual models. Cambria et al. [10] identified key challenges in opinion mining including subjectivity, sarcasm, and domain adaptation. Despite this progress, existing tools typically provide single-method analysis, lack REST API integration, and omit emotion and emoji analysis— gaps this work addresses.

III. SYSTEM ARCHITECTURE

Overview

Sentiment Analyzer adopts a layered modular architecture comprising five principal components: (1) Core Analysis Module, (2) Machine Learning Module, (3) Database Module, (4) API Module, and (5) Dashboard Module. Each module exposes a well-defined interface, enabling independent development, testing, and deployment.

Core Analysis Module

The core module implements four analyzers: VADER, TextBlob, Emotion Detector, and Emoji Analyzer. A SentimentEngine class orchestrates these analyzers and returns an aggregated result. VADER computes a compound score in [−1, 1] applying booster, negation, and punctuation rules. TextBlob returns polarity in [−1, 1] and subjectivity in [0, 1] using the Pattern library. The Emotion Detector uses a 80 - word weighted lexicon covering eight categories: joy, sadness, anger, fear, surprise, disgust, trust, and anticipation.

1. Machine Learning Module

The ML module implements a complete training pipeline: (1) text preprocessing (Unicode normalization, lowercasing, URL removal, stop-word elimination via NLTK's 179-word list, Porter Stemming); (2) TF-IDF vectorization with max_features=5000, ngram_range=(1,2), sublinear_tf=Tr ue, and L2 normalization; and (3) classifier training supporting Logistic Regression, Naive Bayes, and SVM (linear kernel). Models are persisted via joblib serialization.

API and Dashboard Modules

A FastAPI-based REST API exposes six endpoints: POST /analyze, POST /analyze/batch, GET /history, GET /statistics, and supporting documentation endpoints. Pydantic schema validation ensures robust input handling. The Streamlit dashboard provides single-text analysis, CSV batch processing, and interactive Plotly visualizations including sentiment distribution pie charts, confidence bar graphs, and word clouds.

Database Module

A SQLAlchemy ORM layer backed by SQLite persists analysis results. The schema records input text, sentiment label, confidence score, analysis method, emotion classification, and timestamp. Paginated history queries and aggregate statistics queries support the dashboard's trend analysis features.

IV. IMPLEMENTATION

ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science,
Engineering and Technology
An Open Access Journal

VADER Algorithm

VADER's scoring pipeline: (1) tokenize input; (2) look up each token's base score in the sentiment lexicon; (3) apply booster multiplier (1.293) for intensifiers; (4) apply negation scaling (0.74) within a three-token window; (5) apply ALL-CAPS shift (±0.733) and exclamation augmentation (+0.292 per mark, max 3); (6) normalize compound = raw_sum / $\sqrt{(\text{raw\_sum}^2 + 15)}$. Classification thresholds: compound ≥ 0.05 → Positive; ≤ −0.05 → Negative; else Neutral.

TF-IDF Vectorization

Term frequency is computed as TF(t,d) = count(t,d) / |d|. Inverse document frequency uses scikit - learn's smooth formulation: IDF(t,D) = log[(1+|D|)/(1+df(t,D))]+1 to prevent division-by- zero. The TF-IDF weight is the product TF(t,d)×IDF(t,D). With sublinear_tf=True, raw TF is replaced by 1+log(TF), reducing the dominance of high- frequency terms. Bigrams (ngram_range=(1,2)) capture phrase-level sentiment cues such as "not good".

Technology Stack

Layer	Te chnology	Purpose
Interface	Streamlit 1.28+	Interactive dashboard & visualizations
API	FastAPI 0.104+	REST endpoints with OpenAPI docs
ML	scikit-learn 1.3+	Classifiers, TF-IDF, evaluation
NLP	VADER, TextBlob, NLTK	Lexicon-based analysis & preprocessing
Database	SQLAlchemy 2.0 / SQLite	Persistent analysis history
Visualization	Plotly, WordCloud	Interactive charts

Table I
Technology Stack Overview

V. EXPERIMENTAL RESULTS

Unit Testing: Lexicon-Based Methods

VADER and TextBlob were tested against labeled sentiment examples covering positive, negative, and neutral categories. Both achieved 100% accuracy on the standard eight-sample test suite (Table II). Extended edge-case testing (14 additional cases including ALL-CAPS, negation, and mixed sentime nt) confirmed robustness.

Input Te xt	Expe cted	Compound	Result
I love this product, it's amazing!	Positive	0.865	PASS
This is terrible, worst purchase ever	Negative	−0.807	PASS
The meeting is at 3pm	Neutral	0.000	PASS
What a wonderful experience!	Positive	0.782	PASS
I hate this so much	Negative	−0.571	PASS

TableII
Vader Unit Test Results (Selected)

ML Classifier Comparison

Three classifiers were trained on the SST-2 binary sentiment dataset using the project's preprocessing pipeline. SVM achieved the highest accuracy (89.1%) with precision and recall both at 0.89, followed by Logistic Regression (87.5%) and Naive Bayes (84.2%).

Training times are well under 1 second for all models (Table III).

Model	Accuracy	Precision	Recall	F1	Train (s)
Logistic Regression	87.5%	0.88	0.87	0.87	0.12
Naive Bayes	84.2%	0.85	0.84	0.84	0.08
SVM (Linear)	89.1%	0.89	0.89	0.89	0.15

TABLE III
ML Classifier Comparison on SST-2

Comparative Evaluation

Table IV compares Sentiment Analyzer against standalone tools on SST-2. The Ensemble method (91.3% accuracy, F1=0.90) outperforms standalone VADER (71.3%) by approximately 20 percentage points and basic Logistic Regression without the project's pipeline (81.2%) by 10 points. The additiona l latency of the Ensemble (48 ms vs. 8 ms for standalone VADER) remains within interactive response-time thresholds.

Tool	Accuracy	F1	Latency (ms)	Emotion	API
NLTK VADER (standalone)	71.3%	0.70	8	No	No
TextBlob (standalone)	67.8%	0.66	22	No	No
Basic LR (no pipeline)	81.2%	0.80	30	No	No
SA: VADER enhanced	74.1%	0.73	12	Yes	Yes
SA: ML-SVM	89.1%	0.88	35	Yes	Yes
SA: Ensemble	91.3%	0.90	48	Yes	Yes

TABLE IV
Comparative Analysis with Existing Tools (SST-2 Benchmark)

API and Performance

All six REST API endpoints responded with correct HTTP status codes. The POST /analyze endpoint averaged 45 ms response time; POST /analyze/batch processed 100 texts in 230 ms. Memory footprint ranged from 45 MB for VADER to 220 MB for the dashboard. The SVM confusion matrix (Table V) shows balanced error distribution with 85 true negatives and 91 true positives out of 200 test samples.

	Predicted Negative	Predicted Positive
Actual Negative	85	15
Actual Positive	9	91

TABLE V
SVM Confusion Matrix (200 Test Samples)

VI. LIMITATIONS AND FUTURE WORK

Current Limitations

The system has several methodological constraints. First, VADER and TextBlob are optimized for English and may produce unreliable results for non-English, code-mixed, or highly domain-specific text (medical, legal). Second, the emotion detection lexicon covers only approximately 80 terms across 8 categories, limiting coverage. Third, ML models are static and require retraining to adapt to concept drift. Fourth, TF-IDF does not capture semantic similarity, limiting generalization to unseen vocabulary. Fifth, the Streamlit dashboard is not optimized for concurrent multi-user access.

Future Work

Several enhancements are planned. Integration of transformer-based models (BERT, RoBERTa) [2],[17] will improve contextual understanding and sarcasm detection. Multi- language support via multilingual embeddings will extend coverage beyond English. Aspect-based sentiment analysis (ABSA) [3],[4] will enable fine-grained product-attribute- level insights. Real-time stream processing for live Twitter data and a production containerized deployment (Docker, Kubernetes) are also targeted. Expansion of

the emotion lexicon through crowdsourcing and alignment with the NRC Emotion Lexico n
[11] will improve affective coverage.

VII. CONCLUSION

This paper presented Sentiment Analyzer, a comprehensive multi- method sentiment analysis system addressing critical limitations in existing single-method tools. The system integrates VADER, TextBlob, ML classifiers, and ensemble techniques within a modular architecture, achieving 91.3% accuracy on the SST-2 benchmark—a 20-point improvement over standalone VADER. The addition of emotion detection across eight categories, emoji analysis, a production REST API with six endpoints, and a Streamlit dashboard with interactive visualizations provides functional advantages absent from comparable tools. A fully reproducible ML training pipeline with cross-validated model selection ensures that reported accuracy figures reflect genuine generalization. These results demonstrate that combining mult ip le sentiment analysis paradigms within a well-engineered system yields both accuracy gains and practical utility for deployment in real-world applications.

REFERENCES

1. C. J. Hutto and E. Gilbert, “VADER: A parsimonious rule-based model for sentiment analysis of social media text,” in Proc. 8th Int. Conf. Weblogs and Social Media (ICWSM), Ann Arbor, MI, 2014, pp. 216–225.
2. J. Devlin, M. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectio nal transformers for language understanding,” in Proc. NAACL-HLT, Minneapolis, MN, 2019, pp. 4171–4186.
3. W. Zhang, Y. Tang, and J. Wang, “Aspect-level sentiment analysis: A survey,” IEEE Access, vol. 7, pp. 167450–167470, 2019.
4. L. Abubakar and S. Farik, “Aspect-based sentiment analysis methods and applications: A systematic review,” J. King Saud Univ. Comput. Inf. Sci., vol. 35, no. 2, pp. 435–451, 2023.
5. T. Medhi, “A comprehensive survey on deep learning approaches for sentiment analysis,” IEEE Access, vol. 11, pp. 116405–116430, 2023.
6. S. Rai and A. K. Singh, “A systematic review: Sentiment analysis using machine learning techniques, ” in Proc. Int. Conf. Emerg. Comput. Tech. (ICECT), 2020, pp. 1–6.
7. A. Poudel, S. K. Gautam, and S. K. Singh, “Sentiment analysis of Nepali COVID-19 tweets using machine learning and deep learning,” in Proc. Int. Conf. Artif. Inte l. Mach. Learn., 2022, pp. 1–8.
8. Y. Liu, J. Chen, and H. Peng, “Multi- interactive memory network for aspect based sentiment analysis, ” in Proc. AAAI Conf. Artif. Intell., vol. 34, no. 05, pp. 8202–8209, 2020.
9. Y. Gao, T. Wang, and X. Zhang, “Transformer-based hierarchical attention network for aspect level sentiment classification,” Inf. Process. Manag., vol. 59, no. 3, p. 102872, 2022.
10. E. Cambria, B. Schu ler, Y. Xia, and C. Havasi, “New avenues in opinion mining and sentiment analysis,” IEEE Inte l. Syst., vol. 28, no. 2, pp. 15–21, 2020.
11. M. Hazarika, “Emotion recognition in text: Cha lenges and opportunities,” IEEE Trans. Affect. Comput., vol. 14, no. 2, pp. 980–995, 2023.