Analytica: Novelty Report

Aditya Srivastava (2201013)
Deepanika Gupta (2201062)
Kartik Saini (2201103)

1 Introduction

Analytica is a novel full-stack web-based platform for real-time Twitter sentiment analysis, emotion detection, and toxicity identification. It leverages transformer-based models from CardiffNLP and offers multilingual tweet analysis, dynamic visualization, and user engagement features. This report presents the novelty of Analytica in comparison to existing systems like the VADER-based sentiment analysis tool [1], TweetEval baselines [2] & TweetNLP [3].

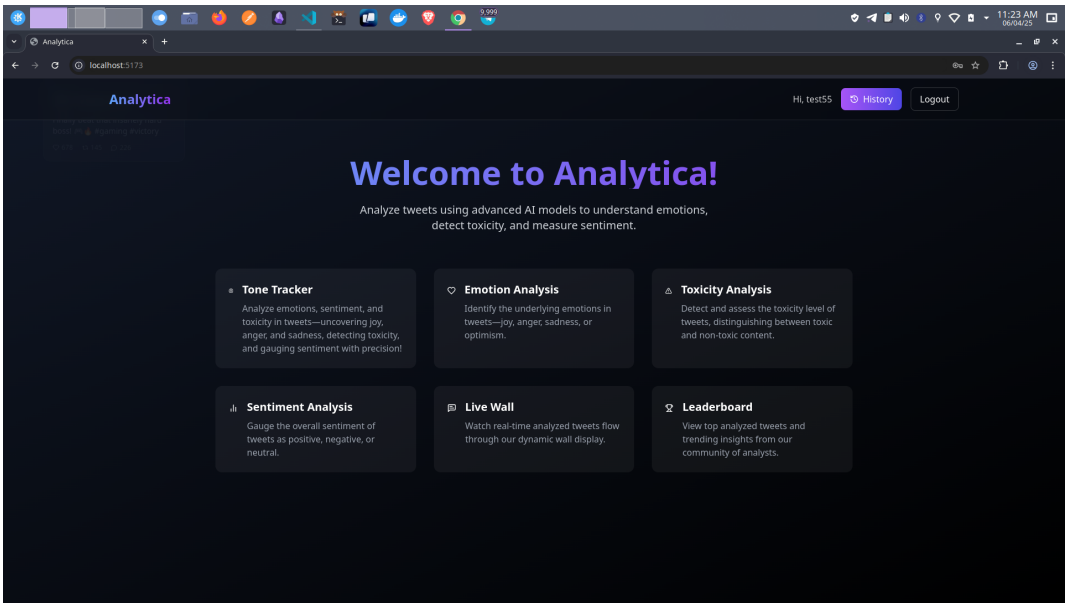

Figure 1: User dashboard showing access to all analysis modules and personalized features.

2 Qualitative Improvement

1. Integrated Multi-Dimensional Analysis

- Analytica can perform both simultaneous sentiment, toxicity, and emotion classification as well as individual classification on each tweet using fine-tuned cardiffnlp/twitter-

roberta-base models.

- In contrast, tools like VADER support only sentiment analysis, and TweetEval/TweetNLP are limited to individual task pipelines and doesn't give combined analysis.

2. Real-Time and Multilingual Analysis

- Analytica scrapes tweets live via Selenium, and auto-translates non-English tweets for analysis—enabling a truly global and real-time social media monitoring system.
- Competing systems are either static (TweetEval) or English-only (VADER, TweetNLP).

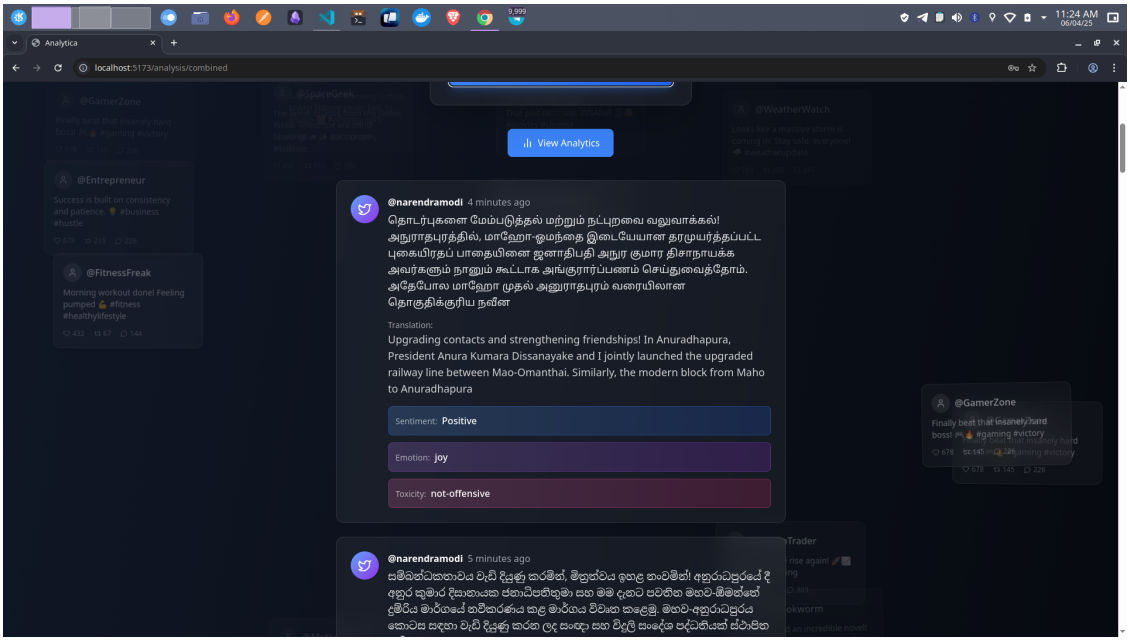

Figure 2: Realtime, combined analysis of a non-english tweet

3. User History + Leaderboard

- Analytica allows registered users to track analysis history and see the statistics on a leaderboard.
- These are not present in other tweet analytics tools like TweetNLP.

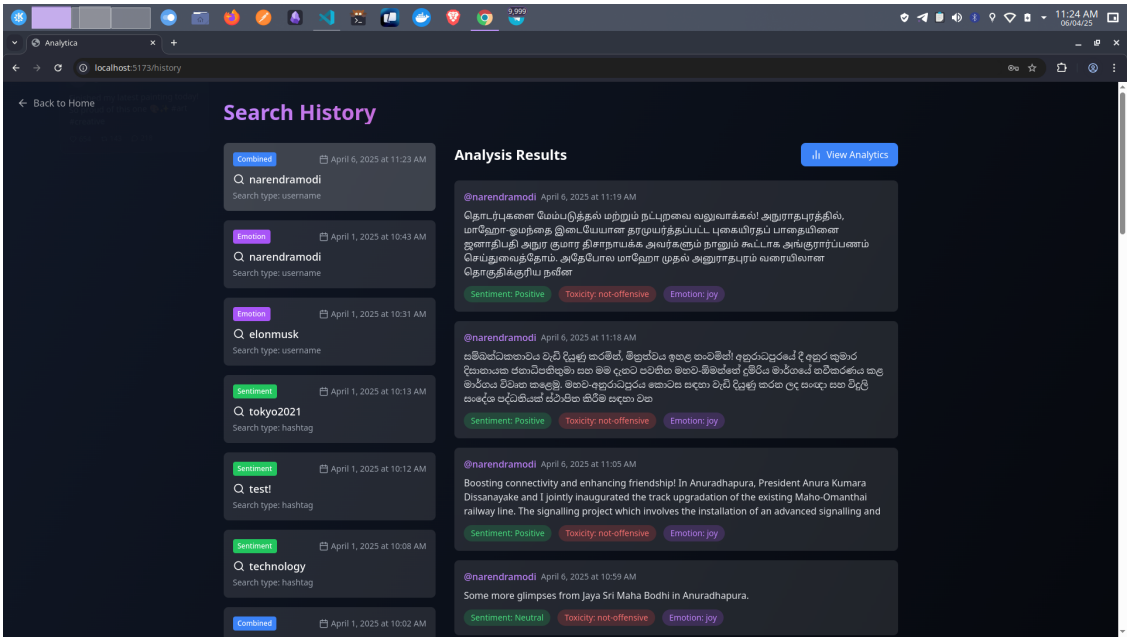

Figure 3: User search history showing past analyses with timestamp and analysis badges.

Figure 4: Leaderboard ranking top analyzed users and hashtags across analysis categories.

4. Interactive Web-Based Interface

- Analytica offers a fully interactive, responsive web interface built using React and Django, allowing users to input queries, visualize results in real time, and track their activity via login and history features.

- In contrast, VADER and TweetEval offer no UI, and TweetNLP offers limited demo-based interfaces.

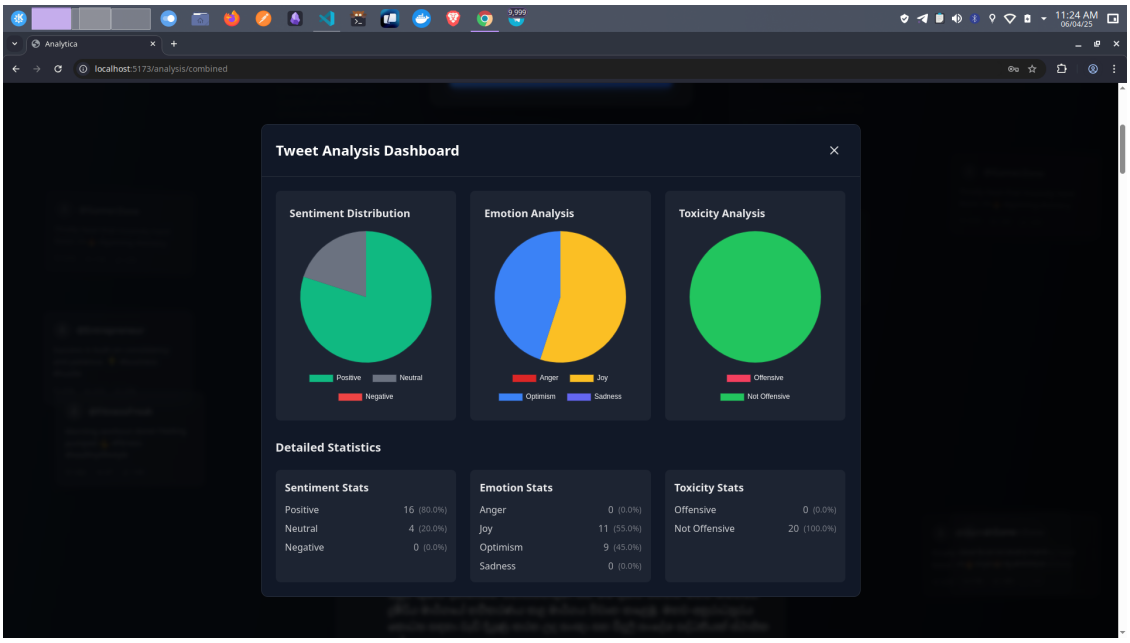

Figure 5: Dashboard displaying sentiment, emotion, and toxicity analysis in pie-chart and statistical formats.

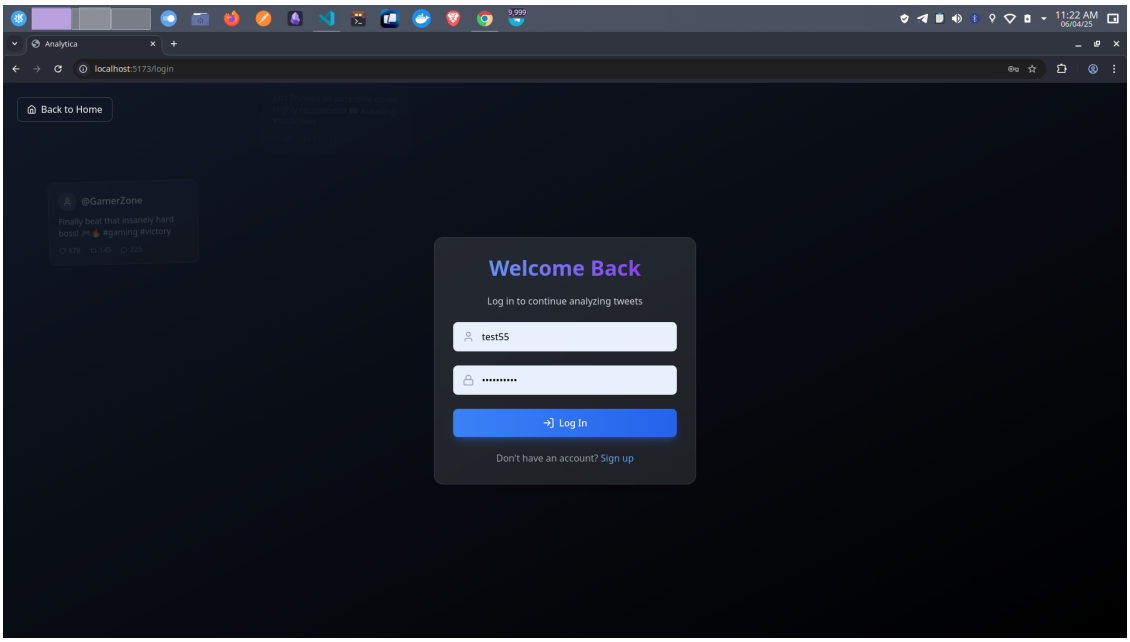

Figure 6: Login interface with secure access for registered users.

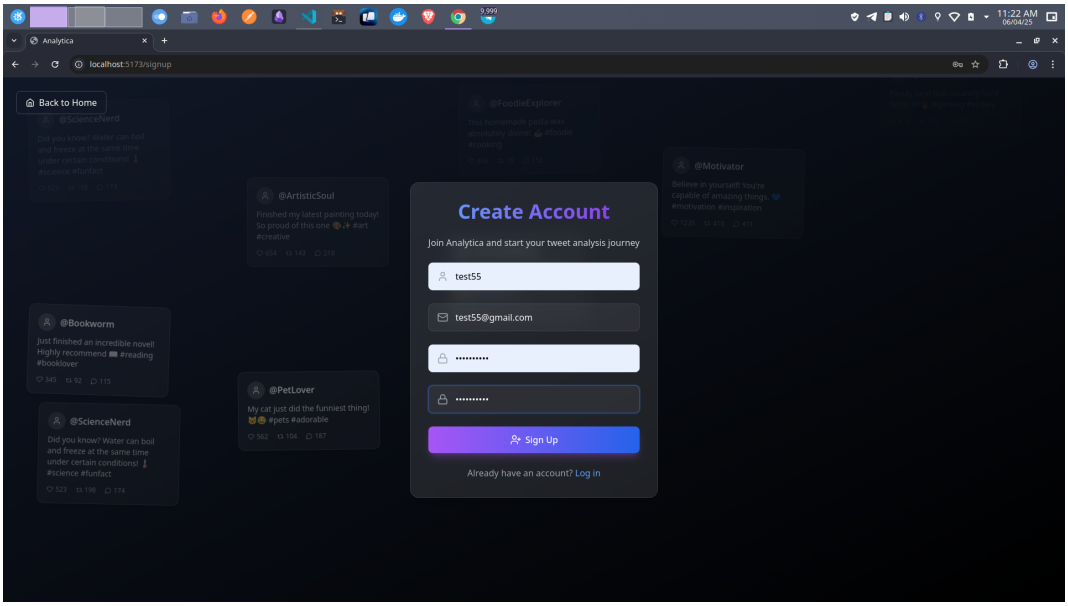

Figure 7: Signup interface to create a new user account for accessing history and leaderboard.

5. Analytica Special - Live Wall

- The LiveWall is a novel feature that displays tweets with animated, color-coded badges for each analysis type, enhancing UX and usability for non-technical users like journalists or researchers or brands looking for their market analytics.
- It dynamically updates and visualizes the results, offering a deeper, more interactive understanding of public discourse.

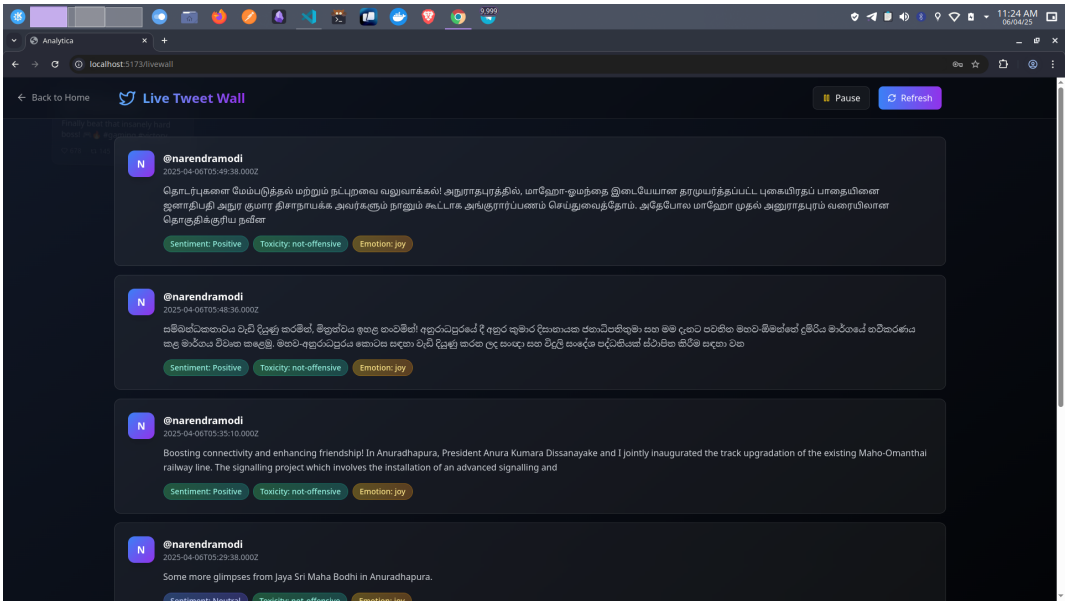

Figure 8: Live Tweet Wall with real-time sentiment, emotion, and toxicity badges.

3 Quantitative Improvement

3.1 Comparison with VADER-Based System

Table 1: Quantitative Comparison: Analytica vs VADER-Based System

Feature	**VADER-Based System**	**Analytica**
Emotion Detection	Not supported	**80.2%** accuracy
Toxicity Detection	Not supported	**82.2%** accuracy
Multilingual Support	English only	**Yes (via translation)**
Visualization	Basic static charts	**LiveWall with badges**
User History	Not available	**Yes**
Leaderboard	Not available	**Yes**

3.2 Comparison with TweetEval Benchmarks (2020)

Table 2: Model Accuracy Comparison with TweetEval Baselines

Task	**TweetEval Baseline (RoB-Base)**	**Analytica (CardiffNLP)**
Sentiment Classification	71.3%	**73.7%**
Emotion Recognition	76.1%	**80.2%**
Offensive Language Detection	79.5%	**82.2%**

3.3 Comparison with TweetNLP

Table 3: Quantitative Comparison: Analytica vs TweetNLP

Feature	**TweetNLP**	**Analytica**
Multilingual Support	English-only (except sentiment)	**Yes (for all 3 tasks)**
Personalization	None	**Login system, search history**
Visualization	Only pie-chart of analysis	**Additional LiveWall + Leaderboard**

4 Real-World Applications of Analytica

Analytica is designed for real-time, multilingual, and multi-dimensional tweet analysis, making it suitable for a wide range of applications:

- **Political Campaign Monitoring:** Track public sentiment, emotional tone, and toxicity of discourse during elections.

- **Brand Reputation Management:** Companies can monitor customer feedback and detect crises in real time.
- **Hate Speech Detection:** Identify and analyze offensive content spreading on social media platforms.
- **Social Movements Analysis:** Monitor global public opinion and emotion during protests or advocacy campaigns.
- **Academic Research:** Use tweet data for linguistic, psychological, or social trend studies with contextual NLP.

5 Conclusion

Analytica extends the capabilities of standard tweet sentiment tools by integrating state-of-the-art models with real-time data handling, multilingual translation, dynamic visualization, and user-focused features. Compared to the research systems like TweetEval, TweetNLP and application-level tools using VADER, Analytica delivers superior performance, broader applicability, and greater utility.

References

[1] A. R. Pai, M. Prince, and P. C. V, "Real-time twitter sentiment analytics and visualization using vader," in *2022 2nd International Conference on Intelligent Technologies (CONIT)*, IEEE, Karnataka, India: IEEE, 2022.

[2] F. Barbieri, J. Camacho-Collados, L. Neves, and L. Espinosa-Anke, "Tweeteval: Unified benchmark and comparative evaluation for tweet classification," in *Findings of the Association for Computational Linguistics: EMNLP 2020*, Online: Association for Computational Linguistics, 2020. [Online]. Available: `https://arxiv.org/abs/2010.12421`.

[3] J. Camacho-Collados, K. Rezaee, T. Riahi, *et al.*, "Tweetnlp: Cutting-edge natural language processing for social media," *arXiv preprint arXiv:2206.14774*, 2022. [Online]. Available: `https://arxiv.org/abs/2206.14774`.