Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 06, ACSCON - 2026

Social Media Text Analysis for Mental Health

Mrs. Shruthi B L
Computer Science and Engineering
ACS College of Engineering
Bangalore, India
sruthibl89@gmail.com

Suraj B
Computer Science and Engineering
ACS College of Engineering
Bangalore, India
sk8016833@gmail.com

Sanjeev Hegde
Computer Science and Engineering
ACS College of Engineering
Bangalore, India
sanjeev.hegdev@gmail.com

Sanjay Marathi
Computer Science and Engineering
ACS College of Engineering
Bangalore, India
sanjaymarathi014@gmail.com

***Abstract*—The increasing incidence of mental health issues necessitates innovative diagnostic methods beyond traditional, subjective assessments, particularly due to the sheer volume and nuanced language found on digital communication platforms[cite: 1130]. This project, the Advanced Social Analyzer, presents a unified, scalable, and automated screening system designed to quantify psychological risk from public textual content acquired from Reddit, Instagram, and Twitter (X)[cite: 1131]. The system utilizes a dual-engine analytical pipeline: a primary Machine Learning classifier (Naive Bayes) calculates an Abnormal Probability Score quantifying the statistical likelihood of high-risk language patterns while a secondary lexicon-based sentiment analyzer (VADER) provides granular emotional profiling (Positive, Negative, Neutral scores) for critical contextual validation[cite: 1132]. Data acquisition is achieved through robust multi-channel connectors, employing Asynchronous APIs, external scraping services, and browser automation[cite: 1133]. The analysis is synthesized into a decisive Overall Verdict and visualized via a dynamic, timeseries chart[cite: 1134]. This synergistic approach provides a temporally aware assessment, enabling the detection of behavioral escalations and serving as a tool for content moderation[cite: 1135].**

***Index Terms*—Artificial Intelligence, Machine Learning, Mental Health Prediction, Naive Bayes, Natural Language Processing, Sentiment Analysis[cite: 2052, 2053, 2054, 2055].**

I. Introduction

The rapid growth of digital communication platforms such as Reddit, Instagram, and Twitter has changed how individuals express their emotions, thoughts, and behavioral patterns[cite: 1151]. These platforms provide a real-time glimpse into the psychological state of users because people tend to express personal struggles, stress, frustration, and emotional instability through posts, comments, and captions[cite: 1152]. Unlike traditional communication methods, social media captures spontaneous and unfiltered expressions, making it a powerful source for analyzing mental health indicators[cite: 1153].

Historically, mental health diagnosis relied heavily on scheduled clinical consultations, therapist-patient conversations, and standard psychological questionnaires[cite: 1154]. These methods are accurate but not continuous, not scalable, and often too late to detect early warning signs of emotional decline[cite: 1155]. Moreover, mental health conditions like depression, anxiety, or self-harm ideation often manifest subtly in language patterns long before an individual seeks help[cite: 1156]. With advancements in Artificial Intelligence (AI), Machine Learning (ML), and Natural Language Processing (NLP), it is now possible to analyze patterns in text and identify psychological risks with significant precision[cite: 1157]. These technologies can uncover emotional sentiment, identify abnormal linguistic cues, and build behavior trends across time all of which are impossible to achieve manually[cite: 1158].

A. Problem Definition

Although social media contains rich psychological indicators, there is currently no unified, automated, and real-time system that can analyze text across platforms and determine mental health vulnerability[cite: 1161]. Current mental health monitoring systems rely largely on human-centered processes such as clinical consultations, standard mental health questionnaires, manual text review by moderators, and basic sentiment analysis tools provided by platforms[cite: 1207, 1208, 1209, 1210, 1211].

The limitations of existing systems are significant:

- **Subjective and Periodic:** Traditional evaluations are performed occasionally and depend on user honesty[cite: 1213].
- **Not Scalable:** Humans cannot review millions of posts across platforms[cite: 1214].
- **Lack of Real-Time Monitoring:** Behavioral changes can occur rapidly, but existing tools cannot detect them instantly[cite: 1215].
- **Limited Emotional Depth:** Most available sentiment tools identify only surface-level emotions[cite: 1216].
- **Poor Context Understanding:** Sarcasm, metaphors, and slang make manual interpretation difficult[cite: 1217].

Different platforms have different data formats, content styles, and API restrictions, meaning no universal tool can extract, clean, and standardize text for analysis[cite: 1164, 1165]. Furthermore, negative posts do not always indicate high mental health risk; expressing sadness due to an event is different from expressing signs of hopelessness or self-harm[cite: 1171, 1172]. Most existing tools offer static results but do not show how a person's emotional state changes over days, weeks, or months, which is crucial for detecting escalation[cite: 1178].

B. Objectives

The major objectives are to:

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 06, ACSCON - 2026

- Design a system that automatically collects public posts from Reddit, Instagram, and Twitter[cite: 1242].
- Analyze collected data using a machine learning classifier for psychological risk[cite: 1243].
- Identify emotional sentiment and measure negative, positive, and neutral tones[cite: 1244].
- Combine these findings into a clear final mental health verdict[cite: 1245].
- Provide time-based visualization for tracking emotional changes[cite: 1246].
- Build the system in a simple, efficient, and scalable manner[cite: 1247].

II. Literature Review

The field of computational mental health analysis has grown significantly due to the availability of social media data and advancements in machine learning and natural language processing[cite: 1271].

A. Machine and Deep Learning Approaches for Psychological Risk Detection

Early research on mental health detection relied heavily on traditional machine learning models such as Naive Bayes, Support Vector Machines, Logistic Regression, and Random Forest[cite: 1276]. These models were used mainly because they are computationally efficient, easy to train, and perform well on structured datasets[cite: 1277]. Research shows that Naive Bayes is one of the fastest classifiers, making it suitable for real-time applications[cite: 1279]. These models function by converting text into numeric features using techniques such as Bag of Words or TF-IDF, producing interpretable outputs such as probability scores[cite: 1280, 1281]. However, these models face limitations when analyzing complex, context-dependent emotional language[cite: 1283]. They cannot understand deeper semantic meaning, detect sarcasm, or interpret sequential patterns in sentences[cite: 1284]. Despite these limitations, Naive Bayes is still widely used for initial screening tasks because of its speed and reliability, and this project adopts it for generating the abnormal probability score[cite: 1291].

B. Sequential Deep Learning Models

To address the shortcomings of traditional models, researchers explored deep learning models designed to understand the order and context of words[cite: 1297]. Models such as Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU) can analyze long sequences of text and learn how emotional patterns develop across sentences[cite: 1298]. LSTM networks can remember information over longer text sequences, while GRU is a simplified version of LSTM and performs faster with similar accuracy[cite: 1300]. Bidirectional GRU reads text from both directions, improving contextual understanding[cite: 1300]. Sequential models outperform traditional models when dealing with emotional language, personal experiences, and psychological narratives[cite: 1301]. However, these models require large datasets and extensive training, which increases computational cost[cite: 1302].

C. Transformer-Based Models

The introduction of transformer architectures marked a major advancement in natural language processing[cite: 1305]. Models like BERT and RoBERTa provide contextual embeddings, which means the meaning of each word is understood based on surrounding words[cite: 1306]. Transformers achieve state-of-the-art accuracy in sentiment analysis and mental health prediction, capable of understanding complex expressions, sarcasm, deeper context, and informal language[cite: 1308, 1309]. Many studies report more than 95 percent accuracy in psychological risk classification using transformer models[cite: 1310]. Transformers significantly outperform all previous approaches. However, their computational cost is high, which makes them challenging to deploy for real-time web applications[cite: 1312].

D. Sentiment and Emotional Context Profiling

VADER is one of the most widely used sentiment analysis tools for social media data[cite: 1320]. It is designed to handle informal language, emojis, slang, and abbreviations commonly found on platforms such as Twitter[cite: 1321]. VADER provides positive, negative, neutral, and compound emotion scores[cite: 1323]. It works well on short messages such as tweets and comments, is extremely fast and suitable for real-time analysis, and does not require training data[cite: 1324, 1325, 1326]. However, it cannot detect deeper psychological meanings or understand new slang unless manually added to its lexicon[cite: 1328].

E. Context and Community-Based Risk Prediction

Platforms like Reddit contain threaded conversations where context is built through replies[cite: 1342]. Graph-based models such as Graph Transformer Networks analyze these structures to understand how discussions develop[cite: 1343]. Graph models can detect harmful conversation patterns, predict the escalation of negative or abusive discussions, and capture community influence and group behavior[cite: 1345, 1346, 1347]. Accuracy increases significantly when conversation structure is included[cite: 1348]. Additionally, topic modeling tools like Latent Dirichlet Allocation (LDA) are widely used to identify dominant themes in large datasets[cite: 1351].

III. Software Requirement Specification (SRS)

The Software Requirement Specification provides a detailed description of the functionalities, constraints, and resources required to build the mental health analysis system[cite: 1368].

A. Functional Requirements

These are the essential functions the system must perform:

- The system must allow the user to enter a public username from Reddit, Instagram, or Twitter[cite: 1373].
- The system must strictly validate the format of the entered username (e.g., ensure no special characters are used that are invalid for the platform) and confirm that the specified platform is supported[cite: 1389].

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 06, ACSCON - 2026

- For a successful analysis, the system must attempt to extract a minimum threshold of posts (e.g., the 50 latest posts or comments) from the public profile to ensure statistical significance[cite: 1390].
- The system must record the exact time and date for every extracted text sample to facilitate chronological trend analysis[cite: 1391].
- Text Preprocessing Steps must include normalization (lowercasing), tokenization, stop word removal, and URL/Mention/Hashtag removal to eliminate elements that bias the sentiment or classification models[cite: 1393, 1394, 1395, 1396, 1397].
- The Naive Bayes Classifier must output a raw probability score, $P(\text{Abnormal}|\text{Text})$, which represents the likelihood of the text sample indicating an abnormal psychological state[cite: 1398].
- The VADER analysis must provide the standard four scores: Positive (pos), Negative (neg), Neutral (neu), and Compound score for each text sample[cite: 1399].
- The system must calculate the Mean Abnormal Probability and the Median Compound Sentiment Score across all collected text samples, alongside the percentage of posts falling into the High Risk category (where $P(\text{Abnormal}) > 0.7$)[cite: 1401, 1402].
- The generation of the overall psychological risk verdict (e.g., "Low," "Moderate," "High") must be based on a weighted combination of the Mean Abnormal Probability and the Mean Negative Sentiment Score, using predefined and documented thresholds[cite: 1408].

B. Non-Functional Requirements

1) Performance: The total time from submitting a username to displaying the results dashboard must not exceed 15 seconds for a standard dataset of 50 posts[cite: 1432]. The time required to load all necessary machine learning and NLP models (Naive Bayes, VADER) must be less than 3 seconds upon system initialization[cite: 1433]. The system must be capable of handling simultaneous analysis requests from at least 5 concurrent users without degradation in the 15-second latency target[cite: 1434].

2) Reliability and Security: The Naive Bayes Classifier model must maintain a minimum documented F1-Score of 0.75 on the system's test dataset[cite: 1436]. All extracted text data must be retained exactly as sourced for analysis, and any preprocessing step must be reversible or clearly logged[cite: 1437, 1438]. The system must not store any text or analysis results after the user session has ended, except for necessary system logs that do not contain personal identifiers[cite: 1441]. All social media API keys and scraping credentials must be securely stored as environment variables and never exposed in the source code or client-side application[cite: 1442].

C. Hardware and Software Requirements

The software components require a Python 3.9+ environment for access to modern language features and efficient dependency management[cite: 1447]. Key Flask extensions, such as Flask-WTF for form handling and requests for secure external API communication, are specified[cite: 1452]. The 'scikit-learn' library handles the Naive Bayes model implementation and management, while the 'nltk' (Natural Language Toolkit) library provides VADER sentiment analysis[cite: 1454, 1455]. The 'pandas' library is required for efficient text data manipulation, aggregation, and filtering prior to visualization[cite: 1456].

Hardware requirements include a minimum of 8 GB RAM for a server-based deployment to comfortably handle simultaneous model loading, concurrent user requests, and OS overhead[cite: 1467, 1468]. A Quad-core processor (or better) with a minimum clock speed of 2.0 GHz is required to manage the parallel processing inherent in asynchronous data fetching[cite: 1469]. A minimum symmetrical 50 Mbps internet connection ensures fast and reliable access to social media APIs, alongside 10 GB of SSD storage for the operating system, Python environment, pre-trained ML models, and logs[cite: 1471, 1476].

IV. System Tools

The tools used in the development and execution of the Advanced Social Analyzer system ensure accuracy, speed, and compatibility with social media platforms and machine learning operations[cite: 1560, 1562].

A. Python and Flask Framework

Python is the primary programming language used for developing the system, supporting machine learning, natural language processing, and web-based operations through its rich set of libraries[cite: 1565, 1566]. Python integrates easily with APIs and scraping tools and provides a simple framework for writing the analysis pipeline[cite: 1569, 1570]. Flask is the backend framework used to create the web application, handling user requests, processing usernames, managing routing, and displaying output results[cite: 1573, 1574]. Flask connects the user interface with analytical functions, processing data received from APIs and models, and sending results to the frontend for visualization[cite: 1582, 1583, 1584].

B. Analytical Models

The Naive Bayes classifier provides fast classification suitable for real-time analysis, analyzing text and generating an abnormal probability score[cite: 1588, 1589, 1590]. It is efficient for short social media posts and loads quickly during system startup[cite: 1592, 1593]. VADER determines emotional polarity, identifying negative, neutral, and positive sentiment values, and handling emojis, slang, and informal expressions[cite: 1596, 1599]. VADER validates the abnormal probability by adding emotional context without requiring training data[cite: 1600, 1601].

C. Data Scraping Tools

The Reddit asynchronous client fetches multiple comments quickly, preventing the blocking of the backend and ensuring smooth data collection[cite: 1604, 1605, 1607, 1608]. The

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 06, ACSCON - 2026

Apify scraper collects Instagram captions from public profiles since Instagram does not allow direct text access through regular APIs[cite: 1617, 1618]. Apify provides structured data containing captions and timestamps and handles changes in Instagram layout through automated scraping[cite: 1620, 1623]. Browser automation tools such as Playwright extract tweets dynamically, bypassing API restrictions by mimicking normal browser actions and scrolling through the page to collect older posts[cite: 1626, 1627, 1630, 1632].

V. Methodology

This section defines the precise techniques for data acquisition and standardization across all supported social media platforms[cite: 1489].

A. Data Processing Method

All extracted data, regardless of the source, must be transformed into a unified data structure (e.g., a Pandas Data Frame row) containing exactly two fields: the cleaned and unprocessed raw text of the post/comment, and the UNIX timestamp normalized to the standard Unix epoch format[cite: 1491, 1493, 1494]. A preliminary cleaning step is applied at this stage to remove platform-specific artifacts before the core NLP preprocessing stage[cite: 1495].

The text must be lowercased, and punctuation or special characters removed, except those crucial for VADER sentiment analysis (e.g., emojis and exclamation points)[cite: 1509]. Tokenization occurs before stop word removal[cite: 1509]. For the Naive Bayes model to generalize better, a process of lemmatization (reducing words to their dictionary root) is applied to the tokens[cite: 1511].

B. Analytical Method

The raw text must be converted into numerical features using the Term Frequency-Inverse Document Frequency (TF-IDF) vectorization technique, which weighs words based on their importance in the overall training corpus[cite: 1513]. The system loads the pre-trained and serialized Naive Bayes model and applies the log-likelihood calculation to the input vector to output $P(\text{Abnormal}|\text{Text})$ for each post[cite: 1514, 1515]. The system must verify the integrity and version of the loaded model before every run to ensure consistency[cite: 1516]. The VADER analyzer is applied directly to the pre-processed text to obtain the four standard scores efficiently[cite: 1518]. The resulting Compound score represents the overall emotional valence of the text for the time-series graph[cite: 1519].

C. Aggregation Method

The unified data structure must be sorted descending by the UNIX timestamp to ensure the most recent posts are prioritized in analysis presentation[cite: 1528]. Instead of a simple arithmetic mean, a time-decaying weighted average is applied to both the Abnormal Probability and Negative Sentiment, ensuring more recent posts have a slightly greater influence on the overall verdict than older posts[cite: 1529, 1530]. The system must identify and flag the Top 5 posts with the highest Abnormal Probability for inclusion in the dashboard summary[cite: 1531].

The verdict uses a three-tiered classification instead of binary:

- **Low Risk:** Weighted Average Abnormal Probability $P_A < 0.40$[cite: 1535].
- **Moderate Risk:** $0.40 \leq P_A < 0.65$[cite: 1535].
- **High Risk:** $P_A \geq 0.65$[cite: 1536].

The final verdict is also slightly adjusted upward (e.g., from Low to Moderate) if the Weighted Average Negative Sentiment exceeds a secondary threshold (e.g., 0.50)[cite: 1537].

VI. System Design

System design explains the structural and functional organization of the Advanced Social Analyzer[cite: 1647]. The design follows a modular architecture so that each module performs a specific task without affecting the others[cite: 1649].

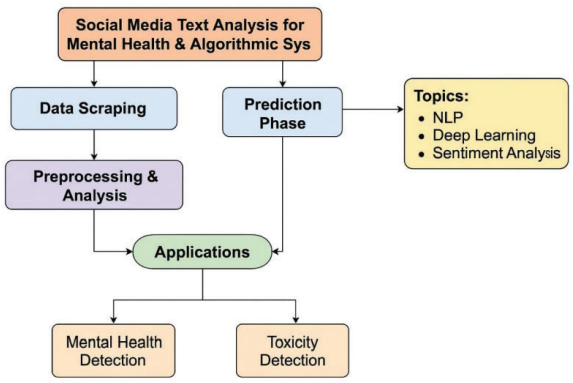

Fig. 1. System Flow Diagram detailing data scraping, preprocessing, prediction phase, and applications including Mental Health Detection and Toxicity Detection[cite: 1650, 1651, 1652, 1653, 1658, 1659].

The system follows a three-tier architecture consisting of the presentation tier, the application tier, and the data tier[cite: 1662]. The presentation tier handles user interaction, receiving a username entered by the user and displaying the results including sentiment scores, abnormal probability, and time-series graphs using HTML, CSS, and Chart.js[cite: 1663, 1664]. The application tier processes the core functions of the system, acting as the central controller to manage requests, execute analysis, and prepare results via the Flask backend, the data extraction modules, the Naive Bayes classification engine, and the VADER sentiment analyzer[cite: 1665, 1666, 1667]. The data tier contains all raw inputs and trained model files, including collected posts from Reddit, Instagram, and Twitter as well as the saved Naive Bayes model and the VADER lexicon[cite: 1672, 1673].

The flow of the Advanced Social Analyzer is divided into structured levels[cite: 1686]. Level 0 shows the simplest overview of the system[cite: 1689]. Level 1 explains how the system gathers posts from each platform[cite: 1692]. Level 2 describes how each post is analyzed by passing the text

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 06, ACSCON - 2026

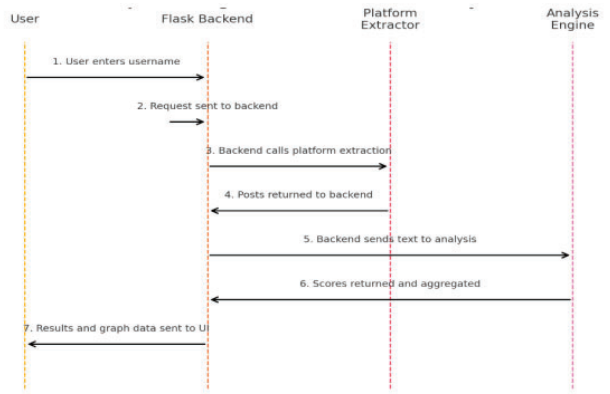

Fig. 2. Sequence Diagram illustrating user interaction with the Flask Backend, Platform Extractor, and Analysis Engine[cite: 1713, 1714, 1715, 1716, 1717].

to the Naive Bayes classifier and VADER to obtain negative, neutral, and positive sentiment scores[cite: 1702, 1703, 1704]. In Level 3, the system sorts all posts chronologically, calculates averages, compares abnormal probability with the threshold, generates the final verdict, and prepares values for Chart.js[cite: 1706, 1707, 1708].

Fig. 3. Activity Diagram outlining the major steps: Validating username, collecting posts, analyzing sentiment, aggregating values, generating verdict, and displaying results[cite: 1742, 1743, 1744, 1745, 1746, 1747, 1748].

VII. System Implementation

System implementation describes how the Advanced Social Analyzer is developed, installed, and executed. The first step must be the creation of a Python Virtual Environment (`venv` or `conda`) to isolate project dependencies, initiated via `python -m venv analyzer_env`. All required Python packages and their specific versions must be listed in a `requirements.txt` file and installed via `pip install -r requirements.txt`. The Chart.js library must be included via a Content Delivery Network (CDN) link within the HTML template files.

A dedicated configuration file (`config.py`) stores application settings, including API keys, the location of the trained model file, and the risk threshold value. The Flask routes must be logically divided into `/` (Main dashboard), `/analyse` (Accepts username and platform via POST), and `/results` (Displays the final analysis via GET).

The data fetching logic is wrapped in asynchronous functions (`async/await`) to ensure the Flask server remains responsive and can handle multiple client requests. The system discards any post/comment where the text content is empty or below a minimum character length (e.g., 5 characters) and attempts a basic language identification check (using libraries like `langdetect`) to skip posts not written in English. The core analysis is implemented as a loop that iterates through the standardized list of posts, calling `classify_nb()` and `analyze_vader()` sequentially.

VIII. System Testing

System testing ensures that the Advanced Social Analyzer functions correctly, produces accurate results, and handles different user inputs reliably[cite: 1821]. The initial stage includes an Environment Sanity Check to verify the virtual environment and Configuration Loading for environment variables[cite: 1825, 1827, 1828]. The Core Module Initialization Test verifies ML Model Integrity by attempting to predict a dummy input vector and tests the VADER Lexicon with a known-value test (e.g., analyzing "This is the best day ever!")[cite: 1829, 1830, 1831, 1832]. An API Connectivity Check performs a lightweight ping to Reddit, Apify, and Twitter to ensure services are reachable without firewall issues[cite: 1833, 1834].

Boundary Value Analysis (BVA) designs tests specifically for boundary conditions, such as inputting usernames exactly at the minimum/maximum length allowed, or testing accounts with exactly the minimum required number of posts for analysis (e.g., 50 posts)[cite: 1842, 1843, 1844]. Equivalence Partitioning groups test data into highly emotional concise posts (Group 1), long descriptive posts (Group 2), and neutral transactional posts (Group 3)[cite: 1845, 1846, 1847, 1848].

White Box Testing uses Path Testing to execute every possible path within critical functions, particularly the if/else logic in data cleaning and final verdict generation[cite: 1851, 1852]. Negative Input Testing is designed to fail by checking invalid characters in the username, a legitimate username on the wrong platform, and a completely private account[cite: 1856, 1857, 1858, 1859]. Non-Functional Load Testing simulates a rapid increase in users (from 1 to 20 concurrent users) to monitor the Flask server response time, ensuring it does not exceed the 15-second latency target[cite: 1869, 1870, 1871].

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 06, ACSCON - 2026

IX. Results

The Advanced Social Text Analyzer system produces a complete psychological analysis based on user-generated social media posts[cite: 1887]. The results include numerical scores, sentiment distribution, behavioral patterns, and an overall verdict[cite: 1888].

A. Summary of Analytical Results

The summary section provides the most important values in a clear and easy-to-understand form[cite: 1891]. The verdict will be displayed using a clear, color-coded, categorical label: "Low Concern" (Green), "Moderate Concern" (Yellow/Amber), or "High Concern" (Red)[cite: 1904]. Alongside the verdict, the system displays the raw Weighted Average Abnormal Probability (WAAP) value (e.g., 0.62), to provide the user with the numerical basis for the category[cite: 1905]. The display explicitly states the range is 0.00 to 1.00[cite: 1909].

Fig. 4. Summary dashboard showing the Overall Analysis Verdict (Low Concern), Avg. Abnormal Probability, Avg. Negative Sentiment, and Avg. Positive Sentiment based on the posts analyzed[cite: 1893, 1895, 1896, 1897, 1898, 1899, 1900].

The sentiment distribution presents the three average sentiment scores (Negative, Neutral, Positive) as a percentage breakdown that totals 100%[cite: 1914]. The Total Number of Posts Analyzed is displayed directly beneath the average scores to validate the statistical significance of the results[cite: 1915].

B. Detailed Post Analysis

Each post is analyzed individually and displayed with its corresponding sentiment values and abnormal probability[cite: 1917]. This provides transparency and allows users or researchers to see how each piece of text contributes to the result[cite: 1918].

The posts must be displayed in reverse chronological order (newest first) to prioritize the most recent behavior[cite: 1928]. The table uses conditional formatting (e.g., coloring the Abnormal Probability score red if it exceeds 0.75) to draw the user's attention to specific high-risk posts immediately[cite: 1933]. A separate, brief section must be dedicated to displaying the Top 5 Posts with the highest calculated Abnormal Probability scores, serving as the immediate psychological evidence supporting the overall verdict[cite: 1936, 1937].

C. Time-Series Trend Visualization

The graphical output is one of the most important features of the system, showing how the user's emotional and psychological patterns change over time[cite: 1939, 1940].

Fig. 5. Detailed post analysis with individual sentiment metrics (Neg, Neu, Pos) and Abnormal Probability per post[cite: 1920, 1921, 1922, 1923].

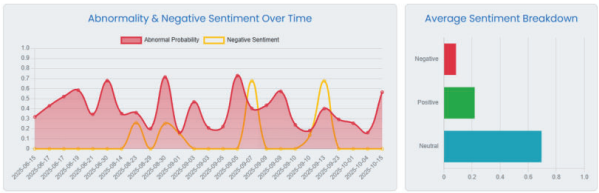

Fig. 6. Time-Series visualization tracking Abnormality and Negative Sentiment Over Time, including an Average Sentiment Breakdown[cite: 1947, 1968, 1969, 1970, 1971].

The graph utilizes two separate Y-axes: one on the left for Abnormal Probability (0.0 to 1.0) and one on the right for the Compound Sentiment Score (-1.0 to +1.0) to prevent scale distortion[cite: 1977, 1978]. Event tooltips appear when hovering over any point on the trend line, displaying the exact timestamp, the score value, and a snippet of the original post text for contextual reference[cite: 1979].

D. Interpretation of Results

The interpretation highlights the importance of convergence (both risk and negative sentiment rising together) as the strongest indicator of escalating concern, and divergence (risk rising while sentiment remains stable) as a signal for subtle, non-emotional psychological shifts[cite: 1986].

If the average abnormal probability is low and negative sentiment is stable, the user shows low concern[cite: 1989]. If abnormal probability is moderate but rising, the user may need observation[cite: 1990]. If abnormal probability and negative sentiment are both high, the user is marked as high concern[cite: 1991]. A sharp, isolated increase in risk followed by a drop suggests an acute emotional event (a temporary crisis or external stressor), rather than a chronic state[cite: 2001]. If the WAAP is sustained above the moderate threshold for a long duration, the system interprets this as a chronic pattern independent of specific recent events[cite: 2006].

The final dashboard follows a top-down information hierarchy: Overall Verdict (High Priority), Summary Scores and Distribution (Medium Priority), Time-Series Visualization

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 06, ACSCON - 2026

(High Priority), and Detailed Post-by-Post Analysis (Low Priority/Expandable)[cite: 2014, 2015, 2016, 2017, 2018].

X. Conclusion

The Advanced Social Analyzer successfully provides an automated and structured method for analyzing psychological risk from social media text[cite: 2024]. By combining a machine learning classifier with a sentiment analysis tool, the system can extract meaningful insights from online communication[cite: 2025]. The classifier generates abnormal probability scores that reflect risk patterns, while the sentiment analyzer identifies emotional tones that further validate the results[cite: 2026]. The system's ability to produce time-series visualizations adds significant value by highlighting emotional and behavioral changes over time[cite: 2027]. This allows researchers, mental health observers, and educators to study long-term patterns rather than relying on isolated posts[cite: 2028].

Although the system does not replace professional diagnosis, it serves as a helpful screening tool for identifying high-risk behavior on social platforms[cite: 2030]. It forms a foundation for further advancements such as deeper emotion classification, transformer-based models, and multimodal analysis that can include images or additional data types[cite: 2031]. Overall, the project demonstrates that artificial intelligence can play a meaningful role in early detection, online safety, and mental health research[cite: 2032].

References

[1] Y. Kasanneni, A. Duggal, R. Sathyaraj, and S. P. Raja, "Effective Analysis of Machine and Deep Learning Methods for Diagnosing Mental Health Using Social Media Conversations," *IEEE Transactions on Computational Social Systems*, 2025[cite: 2034, 2035].

[2] M. Nouman, H. Sara, S. Y. Khoo, and M. A. P. Mahmud, "Mental Health Prediction through Text Chat Conversations," *International Joint Conference on Neural Networks*, 2023[cite: 2036, 2037].

[3] W. Yustanti, A. W. Utami, and P. S. Nautika, "Probabilistic-Based Text Clustering for Optimizing Mental Health Issues Extraction on Social Media," *International Conference on Vocational Education and Electrical Engineering*, 2024[cite: 2038, 2039].

[4] N. Muhamed and A. S. Pillai, "Subclassifying Sadness: Enhancing Sentiment Analysis in Social Media with Large Language Models," *International Conference on Intelligent Computing and Control Systems*, 2025[cite: 2040, 2041].

[5] L. Hebert, L. Golab, and R. Cohen, "Predicting Hateful Discussions on Reddit Using Graph Transformer Networks," *IEEE/WIC/ACM Joint Conference on Web Intelligence*, 2022[cite: 2042, 2043].

[6] R. Amaliyah, "Multi-Stage Multimodal Sentiment Analysis Using Attention Mechanism and BIGRU on Mental Health Topics in Social Media," *IEEE Symposium on Future Telecommunication Technologies*, 2025[cite: 2043, 2044].

[7] M. V. V. Perera and K. M. Piyumal, "Detection of Cyberbullying to Reduce Mental Health Problems Using Machine Learning Algorithms," *International Research Conference on Smart Computing and Systems Engineering*, 2025[cite: 2045, 2046].

[8] N. Prudhvish *et al.*, "DeTox: A Web Application for Toxic Comment Detection and Moderation," *International Conference on Trends in Quantum Computing and Emerging Business Technologies*, 2024[cite: 2047, 2048].