Social Monkey

Project Team

Ahmed Ali	22I-0825
Qusai Mansoor	22I-0935
Ahmad	22I-1288

Session 2022-2026

Supervised by

Dr. Sana Aurangzeb

Department of Computer Science

National University of Computer and Emerging Sciences
Islamabad, Pakistan

Sept, 2025

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Human beings have always been expressive creatures. From the prehistoric hunter-gatherers to the modern civilized man, we have felt the need to be heard and to share our emotions through every channel available [10, 14]. Before the technological revolution, most communication took place in person, with two people or a group exchanging ideas and feelings face to face. But with the rise of the internet and the easy access to smart devices like phones, a new virtual world has emerged where communication between distant people is not only possible but seamless [16, 17]. All of this is facilitated in the modern world through the use of social media. Research shows around 5.4 billion people were social media users in 2025 which amounts to 65 percent of the world population. [9].

There is an ongoing debate about the potential positive and negative side effects of social media usage, especially when it is of extreme nature [1, 21]. Yet no one denies that social media has deeply penetrated the day-to-day activities of people from every walk of life—whether it is a businessperson or an employee, a husband or a wife, the young or the old—social media is now a permanent part of society [6, 18]. Increasingly, businesses and companies use it to push their brand and establish a unique identity [2, 8]. But with this shift to a virtual mode of communication, one fundamental problem arises: how do you judge the sentiment or opinion of others accurately in this medium? [4, 11]

This is where it becomes important to understand the value of sentiment analysis. At its essence, sentiment analysis means judging what a person thinks or feels about a particular subject [4, 11]. In the context of social media, it captures the emotion or sentiment that a user tries to convey through text [3, 12]. Even though other forms of content like video and audio are ubiquitous on social platforms, the real challenge lies in understanding how audiences react to this content—reactions which are often expressed in the form of written comments [3, 11].

That is where Social Monkey comes into play.

Social Monkey is designed to go beyond simple metrics such as likes and views. Our

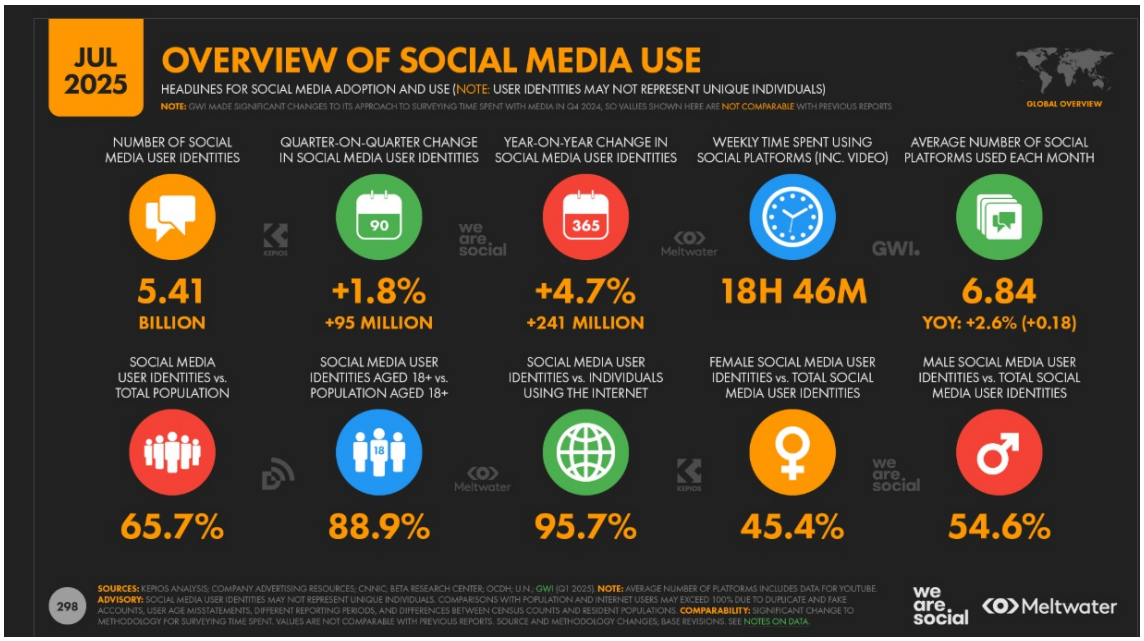

Figure 1.1: Global Social Media Usage - DataReportal

system seeks to uncover the emotional undercurrent of audience reactions and transform it into actionable insights for creators and businesses [12]. By analyzing posts and comments, the platform identifies not only whether an audience response is positive or negative, but also the deeper emotional layers—such as joy, admiration, sarcasm, or frustration—that traditional sentiment tools often miss [3, 4].

The value of this approach lies in its practicality. Instead of overwhelming users with raw data, Social Monkey presents insights in the form of clear dashboards and forecasts [6, 22]. Creators can see which of their posts resonate most, identify emotional triggers that drive engagement, and even predict the likely performance of a draft before publishing it [22]. This feedback loop allows users to optimize their communication strategies, refine their content, and build a stronger connection with their audience [2, 8].

What makes Social Monkey unique is its combination of three elements: emotion-aware sentiment analysis, engagement forecasting, and post optimization [3, 22]. Together, these features provide a holistic solution. Emotion analysis reveals how the audience feels, forecasting predicts how a draft might perform, and optimization helps rewrite or adjust content to maximize engagement [12]. This synergy not only empowers creators to produce more effective posts but also assists brands in maintaining their reputation and avoiding negative backlash before it escalates [6].

In summary, the purpose of this project is to bridge the gap between raw social media interactions and meaningful, actionable understanding. By combining modern sentiment and emotion analysis techniques with predictive and generative capabilities, Social Monkey aims to become an intelligent assistant for creators, marketers, and businesses

who want to truly understand and engage with their audiences in today's digital society [11, 22].

1.1 Problem Statement

Despite having numerous social media analytics tools in market, many competitor platforms such as Hootsuite simplify emotional analysis by classifying mentions into only six basic emotions (anger, disgust, fear, joy, surprise, sadness) using the Parrott method, leaving out many nuanced emotional states that creators care about [13, 19]. Furthermore, modern forms of communication especially among Gen-Z use slang, emojis, internet-speak, code-switching, and cultural memes that evolve rapidly; standard sentiment tools often struggle to keep up with these changes and misinterpret emotional tone [7, 15]. Another often-overlooked capability in many tools is engagement forecasting: while most platforms provide retrospective analytics (likes, comments, sentiment trends), few predict how a draft post will likely perform or suggest optimizations ahead of publication [5, 20]. Social Monkey aims to fill these gaps: it will provide detailed, emotion-rich analysis, handle informal and evolving language, and offer predictive engagement insights so users can understand how a draft might perform before posting.

1.2 Motivation

In recent years, social media has changed a lot. People no longer use it just to share life updates. Now, feelings and identity play a very important role [2]. Creators and brands are judged not only by what they say but also by how their audience feels. A single post can make people happy, angry, caring, or upset. Simple tools that only show whether a post is positive, negative, or neutral are not enough anymore because they cannot capture the full range of emotions people express online [11].

One key motivation for Social Monkey is that many popular tools, like Hootsuite, only detect six basic emotions: anger, disgust, fear, joy, surprise, and sadness [19]. Because of this limit, they miss more subtle feelings such as admiration, gratitude, sarcasm, or shame [3]. These emotions may seem small, but they have a strong hidden effect on how people engage with posts [19].

Today's digital communication—especially among Generation Z—depends a lot on slang, memes, emojis, abbreviations, and local cultural expressions. These ways of speaking change very quickly. Studies show that Gen Z slang can often be unclear without cultural or situational context, and this confusion frequently causes mistakes in sentiment analysis [7, 15].

Another problem is forecasting. Many existing tools only give reports after a post is published, showing how people reacted [20]. Very few can predict how a post will perform before it goes live. This leaves creators guessing if their caption, tone, or hashtag will work well. Social Monkey's forecasting feature is designed to solve this problem [5].

Finally, Social Monkey brings together advanced emotion detection, an understanding of slang and emojis, and the ability to forecast engagement. It also provides clear suggestions to improve posts [22]. The main goal is simple: creators, marketers, and brands need more than just raw data. They need deeper insights that help them improve their style, avoid confusion, and build stronger connections with their audience [8].

1.3 Problem Solution

Social Monkey is a web application designed to solve the problems outlined in Section 1.1. It helps creators, marketers, and brands understand emotions in social media posts and comments. Unlike basic tools that only detect simple emotions, Social Monkey provides detailed emotional insights. It also predicts how posts might perform and suggests ways to improve them. The goal is to help users create better content and connect with their audience.

The application focuses on analyzing user-authorized posts and comments from platforms like Twitter and Instagram. It uses advanced techniques to detect complex emotions like admiration, sarcasm, or gratitude [3]. It also understands slang, emojis, and cultural expressions that are common among Generation Z [7]. This helps avoid mistakes in analyzing emotions. Additionally, Social Monkey forecasts engagement for draft posts, so users know how their content might perform [5]. It provides clear suggestions to optimize posts for better results [22].

The main objectives of Social Monkey are:

- To provide detailed emotion analysis beyond basic sentiments like positive or negative.
- To accurately interpret slang, emojis, and cultural expressions in social media text [15].
- To predict how a draft post will perform before it is published [5].
- To offer simple suggestions to improve posts and increase audience engagement [22].
- To create an easy-to-use dashboard that shows emotional insights and trends [20].

By combining these features, Social Monkey helps users understand their audience better. It reduces the guesswork in creating posts and helps avoid negative reactions. The application is limited to text analysis and does not include video or image generation. It uses secure connections to access data and ensures user privacy with basic encryption. This solution makes social media management easier and more effective for creators and businesses [8].

1.4 Stake Holders

Social Monkey is designed to help various groups who use social media to connect with their audience. Below is a list of key stakeholders who will benefit from or contribute to the project.

- **Social Media Content Creators**: These are individuals like influencers or bloggers who post on Twitter and Instagram. They need to understand how their audience feels about their posts. Social Monkey helps them see emotions like joy or sarcasm in comments and predict how new posts might perform. This helps them create better content.

- **Marketers and Brand Managers**: Companies and marketing teams use social media to promote their brands. They want to know what emotions their posts trigger and how to improve engagement. Social Monkey provides detailed emotion analysis and suggestions to make their posts more effective.

- **Social Media Platforms (Twitter and Instagram)**: These platforms provide the data that Social Monkey analyzes. They benefit indirectly because tools like Social Monkey encourage users to create better content, which keeps people active on their platforms.

- **Project Team**: The developers, including Ahmed Ali (22I-0825), Qusai Mansoor (22I-0935), and Ahmed Ali (22I-1288), build and test Social Monkey. They are responsible for creating the application and ensuring it works well.

- **Supervisor (Dr. Sana Aurangzeb)**: As the project supervisor, Dr. Sana guides the team. She ensures the project meets academic standards and provides technical advice.

- **National University of Computer and Emerging Sciences (NUCES)**: The university supports the project by providing resources and a platform for development. It benefits by showcasing innovative student work.

- **End Users (General Audience)**: These are the people who follow creators or brands on social media. They benefit indirectly because Social Monkey helps create content that matches their emotions and interests.

Each stakeholder plays a role in making Social Monkey successful. Creators and marketers gain tools to improve their posts. Platforms like Twitter and Instagram get better content. The project team and supervisor ensure the application is built well, while the university supports its development. Together, they help Social Monkey achieve its goal of making social media communication clearer and more engaging.

Chapter 2

Project Description

To address the lack of in-depth sentiment and emotional analysis on social media, Social Monkey offers a solution.

2.1 Scope

Social Monkey is a web application developed as an **emotion-aware** social media helper. It analyses user-authorised comments and social media posts to provide fine-grained emotional insights, engagement forecasting, and emotion-driven post optimisation. The initial step would be to establish the **Data Pipeline** and preprocess the Data, which will then be used to achieve **Emotional Analysis**. This will establish the backbone to implement the **Visualisation, Emotional Forecasting and Optimisation** features.

The scope for implementation is limited to platforms such as Twitter and Instagram. Additionally, it excludes Automated Post Scheduling, Image/Video generation, and Multilingual support for emotional analysis. With that, the Project will implement basic OAuth and token encryption.

By focusing on these core capabilities within the defined boundaries, the Project will deliver a proof of concept. This will showcase the feasibility and impact of an emotion-aware social media assistant. Additionally, it will clearly identify future extensions such as full platform integrations, automated publishing, and advanced content generation.

2.2 Modules

The Project is divided into six (6) modules with respect to technical complexity and project understandability; however, these modules would be implemented in four iterations.

2.2.1 Module 1 - Data Ingestion and Preprocessing

Description:

This module securely connects to user-authorised social media accounts or accepts uploaded exports, fetches posts and comments, and normalises the data for downstream analysis. For which Open source and official, paid and free APIs would be used

1. Connect accounts via OAuth 2.0 (Twitter/Instagram) or accept CSV exports for demo.
2. Preprocessing: emoji normalisation, slang lookup, language detection.
3. Store raw JSON + normalised data (Postgres / Mongo).
4. Provide REST endpoints for post retrieval and status.

Acceptance criteria:

Successful OAuth-based ingestion for at least one platform; raw + cleaned records stored; demo dataset of $\geq$ 1000 posts.

2.2.2 Module 2 - Deep Sentiment and Emotional Analysis

Description:

This module builds and operates the emotion analysis engine. It contains two linked workflows: The training pipeline that builds and validates the model using public datasets and in-domain labelled data.

1. Collect and prepare training data from public emotion/sentiment corpora (e.g., GoEmotions, TweetEval) and from a curated set of in-domain examples annotated by the team to capture slang, emoji and platform style.
2. Fine-tune transformer-based models on this combined dataset, selecting model checkpoints according to validation metrics (macro-F1, per-class F1).

3. Evaluate and calibrate the model. Record model metadata (dataset versions, hyperparameters, validation scores) for traceability.

4. Store the chosen model artefact and register a version for deployment.

Acceptance criteria:

Training: model achieves target validation performance, and a versioned model artifact is saved.

2.2.3 Module 3 - Visualisation and Emotion-Based Views

Description:

An interactive dashboard where users view emotion breakdowns, find top posts by emotion, and inspect trend charts and heatmaps.

1. **Emotion-Aware Performance Dashboard:** Shows detailed audience mood distribution for each post (e.g., 30% joy, 20% admiration, 15% sarcasm, 10% anger) rather than a simple positive/negative split.

2. **Negative Emotion Trigger Detection:** Flags posts or comment threads generating anger, sadness or disappointment.

3. **Emotion Heatmaps by Post Type:** Breaks down which emotions dominate per content format (e.g., reels evoke more joy; static posts get more admiration; memes trigger sarcasm).

4. **Gen-Z Slang and Emoji Insight Report:** Detects and interprets slang/emojis (“rizz”, “no cap”) in comments, translating them into clear emotional intent.

Reads the emotion data produced by Module 2 and aggregates it for visual display.

Acceptance criteria:

The dashboard displays aggregates and allows users to filter and export reports. Heatmap and “posts-by-emotion” views load example datasets correctly within interactive UI.

2.2.4 Module 4 - Engagement Forecasting

Description:

Predict expected engagement for a draft post and monitor live posts for sudden negative trends (early-warning alerts).

1. Draft simulator: user pastes a caption (and optionally image metadata) and receives a predicted engagement figure and likely emotional response.
2. Variant comparison: rank alternative captions by predicted performance.

Forecasting utilises historical performance signals and emotional outputs to estimate likely likes/comments/shares.

Acceptance criteria:

Forecasts are returned for drafts in the UI and compared to a simple baseline (show improvement in reported metrics).

Alerts appear for simulated cases and provide clear actionable suggestions.

2.2.5 Module 5 - Emotion-Driven Post Optimisation and Assistant

Description:

A helper that rewrites or suggests caption variants tailored to target emotions (e.g., "inspire joy" vs "encourage gratitude") and formats them per platform.

1. Generate 2–4 caption variants for a given draft tuned to desired emotional outcomes.
2. Provide platform-specific variants (Instagram concise + hashtags, X hook-style).
3. Show predicted uplift for each variant (using Module 4 forecasts).
4. Tone safety checks: warn about possible sarcasm or harmful wording.

Uses prompt-driven text rewriting (local or API-based LLM) combined with the forecast model to score variants before showing them to the user.

Acceptance criteria:

For given test drafts, the assistant returns multiple variants and shows forecasted scores. Variants include platform-specific formatting and safety warnings where applicable.

2.2.6 Module 6 - Admin, Export and Settings

Description:

Workspace and account settings, logging, export scheduling, and data privacy controls.

1. Manage connected accounts and tokens (refresh/revoke).

2. Job monitor for ingestion and model inferences.
3. Export scheduling for regular reports.
4. Simple privacy controls: data retention window and workspace deletion.

Acceptance criteria:

Admin can view ingestion status, revoke a connected account, and perform a workspace data deletion operation.

2.3 Tools and Technologies

Provided are the tools and Technologies which will be used for the Social Monkey Project.

Front-End (User Interface)

- **Framework:** React.js (or Next.js) for building the web interface.
- **Styling:** CSS for a modern, responsive layout consistent with the SocialMonkey brand colors.
- **Charts and Visualisation :** Implementation for dashboards, emotion heatmaps and post-performance graphs.
- **Prototyping:** Figma for wireframes, logo, mock-ups and flow diagrams before development.

Back-End (Core Services and APIs)

- **Language and Framework:** Python with FastAPI to expose RESTful APIs.
- **Authentication:** OAuth 2.0 to fetch posts and comments securely under user consent.
- **Data Storage:** PostgreSQL for structured data (users, posts, analytics) and MongoDb.

Machine Learning Stack

- **Datasets:** GoEmotions (multi-label emotions), TweetEval (Twitter sentiment/emotion tasks), Sentiment140, SemEval, sarcasm and toxicity corpora and many others.

- **Models:** BERTweet for Twitter data, XLM-RoBERTa for posts/comments, DistilBERT or RoBERTa for toxicity/sarcasm detection and many more.
- **Libraries:** HuggingFace Transformers and PyTorch for training and inference; Scikit-learn for metrics and many more.

Version Control, Testing and CI/CD

- **Versioning:** Git/GitHub for code; DVC or MLflow for dataset/model version control.
- **Testing:** Pytest (backend) and Jest (frontend).
- **CI/CD:** GitHub Actions to automate build, test and deploy.

2.4 Work Division

Table 2.1: Table 1

Name	Registration	Responsibility/ Module / Feature
Mr. Ahmed Ali	22i-0825	(Module 1- Feat 1-4) Data Ingestion and Preprocessing
Mr. Ahmad	22i-1288	(Module 1- Feat 1-4) Data Ingestion and Preprocessing
Mr. Qusai	22i-0935	(Module 1- Feat 1-4) Data Ingestion and Preprocessing
Mr. Ahmed Ali	22i-0825	(Module 2- Feat 1-4) Deep Sentiment and Emotional Analysis
Mr. Ahmad	22i-1288	(Module 2- Feat 1-2) Deep Sentiment and Emotional Analysis
Mr. Qusai	22i-0935	(Module 2- Feat 3-4) Deep Sentiment and Emotional Analysis
Mr. Ahmad	22i-1288	(Module 3- Feat 1-4) Visualisation and Emotion-Based Views
Mr. Qusai	22i-0935	(Module 3- Feat 1-4) Visualisation and Emotion-Based Views
Mr. Ahmad	22i-1288	(Module 4- Feat 1-2) Engagement Forecasting
Mr. Ahmed Ali	22i-0825	(Module 4- Feat 1-2) Engagement Forecasting
Mr. Ahmed Ali	22i-0825	(Module 5- Feat 1-4) Emotion-Driven Post Optimisation and Assistant
Mr. Ahmad	22i-1288	(Module 5- Feat 1-2) Emotion-Driven Post Optimisation and Assistant
Mr. Qusai	22i-0935	(Module 5- Feat 1-4) Emotion-Driven Post Optimisation and Assistant
Mr. Ahmed Ali	22i-0825	(Module 6- Feat 1-4) Admin, Export and Settings
Mr. Ahmad	22i-1288	(Module 6- Feat 1-2) Admin, Export and Settings
Mr. Qusai	22i-0935	(Module 6- Feat 1-4) Admin, Export and Settings

2.5 TimeLine

The Social Monkey timeline is divided into four iterations.

- Iteration 1 will perform Data Collection and preprocessing, and deliver an initial implementation of the Visualisation Feature.
- Iteration 2 will perform Deep Sentiment and Emotional Analysis, and will deliver the complete Visualization Feature.
- Iteration 3 will utilize the sentiment analysis performed in the previous iteration to complete the related features and deliver the Engagement Forecasting Feature.
- Iteration 4 will perform the Post Optimisation Module as well as work on the final touches for the Project.

Each iteration will complete a module and will provide a deliverable, with which the Project is expected to be completed till the end of April.

Table 2.2: Table 2

Iteration#	Time frame	Tasks/Modules
01	Sept-Oct	Core Infrastructure and Visualization
02	Nov-Dec	Sentiment and Emotional Analysis
03	Jan-Feb	Engagement Forecasting
04	Mar-Apr	Optimization and Final Integration

Chapter 3

Conclusions and Future Work

Social Monkey offers a new way to analyze social media by providing detailed emotion detection, engagement forecasting, and post optimization. It solves the problem of basic tools that miss complex emotions like sarcasm or admiration and struggle with slang and emojis. The application helps creators and brands understand their audience better on platforms like Twitter and Instagram. It provides clear insights and suggestions to make posts more engaging. This makes social media communication easier and more effective.

In the future, Social Monkey could include more platforms like Facebook or TikTok. Adding features like:

- Automated Post Scheduling
- Multilingual Emotion Analysis
- Video Analysis
- Real-time comment tracking

These updates will help Social Monkey reach more users and offer deeper insights for social media success.

Bibliography

[1] Cecilie Schou Andreassen. The relationship between social media use and psychological outcomes. *Computers in Human Behavior*, 2017.

[2] G. Appel et al. The future of social media in marketing. *Journal of the Academy of Marketing Science*, 2020.

[3] Dorottya Demszky et al. Goemotions: A dataset of fine-grained emotions. *ACL Anthology*, 2020.

[4] M. Elsalam. Sentiment analysis of text incorporating emojis. *International Journal of Computer Applications*, 2022.

[5] R. Xu et al. Predicting participant engagement in a social media-delivered lifestyle intervention. *PMC*, 2022.

[6] Smart Insights. Global social media trends. *Smart Insights*, 2025.

[7] Eliza M. Jeresano and Marigrace D. Carretero. Digital culture and social media slang of gen z. *Educational Linguistics Journal*, 2022.

[8] Andreas Kaplan and Michael Haenlein. Users of the world, unite! the challenges and opportunities of social media. *Business Horizons*, 2010.

[9] DataReportal / Kepios. Global social media statistics 2025. *DataReportal*, 2025.

[10] Stephen C. Levinson. The origin of human communication: evolution and expression. *Philosophical Transactions of the Royal Society B*, 2014.

[11] Y. Mao et al. Sentiment analysis methods, applications, and challenges: A survey. *Pattern Recognition Letters*, 2024.

[12] Binny Mathew et al. Hatexplain: A benchmark dataset for explainable hate speech detection. *AAAI Workshop*, 2020.

[13] Abhijit Mondal and Swapna S. Gokhale. Multi-label classification of parrott's emotions. *KE SEKE Conference Proceedings*, 2021.

[14] M. Morissan. The history of human communication: How did humans build language and become world leaders. *ResearchGate*, 2023.

[15] Vinka Ganita Puspita and Ardik Ardianto. Code-switching and slang: An analysis of language dynamics in the everyday lives of generation z. *Universitas Terbuka Linguistic Journal*, 2024.

[16] Lee Rainie. The future of digital communication. *Pew Research Center*, 2019.

[17] Internet Society. The internet and human communication: A global shift. *Internet Society Research*, 2021.

[18] Statista. Global daily social media usage. *Statista Reports*, 2023.

[19] Hootsuite Support. About insights mention classification. *Hootsuite Help Center*, 2025.

[20] M. Trunfio. Conceptualizing and measuring social media engagement. *PMC*, 2021.

[21] Jean M. Twenge. Social media use and its link to mental health. *Journal of Adolescent Research*, 2020.

[22] J. Wang et al. Social media popularity prediction with multimodal fusion. *Information Processing & Management*, 2023.