# Software Requirements Specification (SRS) Analytica

Submitted by:
Aditya Srivastava (2201013)
Deepanika Gupta (2201062)
Kartik Saini (2201103)

## Contents

# 1 Introduction

## 1.1 Purpose

The purpose of this project is to develop a sentiment and toxicity analysis system for Twitter data, enabling users to monitor brand perception, analyze public opinion, and ensure child-safe content filtering. The system will provide real-time sentiment and emotion classification, detect toxic content, and present insights through an intuitive web interface.

## 1.2 Scope

The project aims to serve diverse user groups such as marketing teams, political analysts, and parents/educators by offering tools for real-time sentiment and emotion analysis, content filtering, and data visualization. The system will handle the data efficiently and maintain high accuracy in classification, including for ambiguous tweets.

# 2 Overall Description

## 2.1 Product Features

- Real-time sentiment and emotion analysis.
- Toxicity detection and content filtering.
- Trending topic analysis and exportable sentiment reports.
- Emotion-tuning suggestions for creating positive content.
- User-friendly dashboard for non-technical users.

## 2.2 User Classes and Characteristics

- **Marketing Teams:** Require real-time sentiment analysis and exportable reports.
- **Political Analysts:** Need tools for trend monitoring and sentiment polarization.
- **Parents and Educators:** Focus on content filtering and emotion-tuning suggestions.

## 2.3 Operating Environment

- **Frontend:** HTML, CSS, JavaScript, React.js.
- **Backend:** Python (Django framework).
- **NLP Libraries:** NLTK, spaCy, TextBlob, Transformers (Hugging Face).
- **APIs:** Twitter API for real-time data.
- **Backup and Recovery:** Github.

# 3 System Features and Requirements

## 3.1 Functional Requirements

1. **Real-Time Sentiment Analysis:**
    - **Description:** Analyze tweets to classify them as positive, negative, or neutral.
    - **Inputs:** Text data fetched via Twitter API.
    - **Processing:** Tokenization, model inference using fine-tuned Transformers.
    - **Output:** Sentiment label.
2. **Emotion Analysis:**
    - **Description:** Detect emotions like joy, anger, sadness, and happiness in tweets.
    - **Inputs:** Text data fetched via Twitter API.
    - **Processing:** Pretrained emotion classification model.
    - **Output:** Emotion label.
3. **Toxicity Detection:**
    - **Description:** Identify and flag harmful content based on toxicity scores.
    - **Inputs:** Text data.
    - **Processing:** Pretrained toxicity classification model.
    - **Output:** Toxicity label.
4. **Exportable Reports:**
    - **Description:** Generate detailed sentiment and trend analysis reports.
    - **Inputs:** Aggregated analysis data.
    - **Processing:** Data formatting and visualization.
    - **Output:** Exportable CSV, PDF, or Excel reports.

## 3.2 Non-Functional Requirements

- **User-Friendly Interface:** The application should feature a simple, intuitive interface accessible to both technical and non-technical users.
- **Accuracy:** The system must accurately provide correct sentiment, emotion, and toxicity classification to meet diverse user needs, especially in the case of ambiguous tweets.
- **Maintainability:** Modular codebase to facilitate updates and integration of new features.

# 4 System Attributes

- **Availability:** Operate 24/7 through feasible deployment.
- **Scalability:** Ensure efficient handling of users and workloads with load balancing, modular architecture for new tasks, lazy loading and pagination for frontend optimization, and real-time data handling for adaptability.
- **Backup and Recovery:** Regular backups of trained model and each modularity.
- **Extensibility:** Design the system as independent, self-contained modules, allowing seamless integration of new features or updates without disrupting existing components.
- **Usability:** Provide a simple and intuitive interface accessible to both technical and non-technical users, ensuring ease of interaction.