Project Requirements Document

Title: Social Media Sentiment Analysis System

Date: June 10 2024

Version: 1.0

Authors: *Sumayya Sana Syeda*, *Suleman Ali Saif Mir*

Table of Contents

Introduction

Project Overview
The Social Media Sentiment Analysis System aims to analyze social media data to determine public sentiment on various topics, such as products, events, or political issues. This system provides insights into public attitudes and trends, helping businesses, researchers, and policymakers.

Objectives
- Develop a system to gather and analyze social media data.
- Provide real-time sentiment analysis.
- Visualize sentiment trends and user engagement.
- Create a user-friendly interface for searching and viewing sentiment results.

Scope
- Platforms: Twitter and potentially other social media platforms.
- Analysis: Sentiment analysis using natural language processing techniques.
- Visualization: Interactive dashboards and reports.
- Deployment: Web application accessible via cloud services.

Definitions, Acronyms, and Abbreviations
- API: Application Programming Interface
- EDA: Exploratory Data Analysis
- VADER: Valence Aware Dictionary and sEntiment Reasoner
- LSTM: Long Short-Term Memory
- UI: User Interface
- UX: User Experience

General Description

Product Perspective
The system will integrate with social media platforms through APIs and use web scraping tools to collect data. It will preprocess, analyze, and visualize this data, providing insights into public sentiment.

Product Functions
- Data Collection: Gather data using APIs and web scraping tools.
- Data Cleaning: Remove noise and normalize text data.
- Sentiment Analysis: Classify sentiments using NLP techniques.
- Visualization: Present data through interactive dashboards.
- Real-Time Analysis: Provide up-to-date sentiment insights.
- User Interface: Allow users to search and view sentiment analysis.

User Characteristics
- End-Users: Individuals or organizations interested in public sentiment on various topics.
- Researchers: Analysts conducting sentiment analysis studies.
- Businesses: Companies analyzing market trends and customer feedback.
- Policymakers: Government officials and agencies monitoring public opinion.

Constraints
- Data Privacy: Ensure compliance with data protection regulations.
- Performance: System must handle large volumes of data efficiently.
- Scalability: Ability to scale as data volume increases.

Assumptions and Dependencies
- Reliable internet connection for data collection and system access.
- Availability of APIs for data gathering.
- Sufficient computational resources for data processing and analysis.

Specific Requirements

Functional Requirements
1. Data Collection
   - Use Tweepy to gather data from Twitter.
   - Implement web scraping with BeautifulSoup for additional platforms.

2. Data Cleaning
   - Remove special characters, URLs, and stop words.
   - Tokenize and lemmatize/stem text data.

3. Sentiment Analysis
   - Use VADER and TextBlob for initial sentiment classification.
   - Train machine learning models (e.g., logistic regression, SVM, LSTM).

4. Visualization
   - Create bar charts, word clouds, and sentiment timelines.
   - Develop interactive dashboards with Tableau.

5. User Interface
   - Design a web application using Flask or Django.
   - Deploy the application on Heroku or AWS.

User Stories/Use Cases

1. Market Analysis
   - As a business owner, I want to know the public sentiment about my product to improve my marketing strategy.

2. Political Analysis
   - As a policymaker, I want to gauge public opinion on political issues to inform my decisions.

3. Customer Experience
   - As a customer service manager, I want to analyze customer feedback to enhance our services.

Non-Functional Requirements

Performance
- Response Time: Sentiment analysis results should be available within 5 seconds.
- Throughput: The system should handle up to 10,000 requests per hour.
- Scalability: Must support scaling to handle increased data and user load.

Usability
- User Interface: The UI should be intuitive and easy to navigate.
- Accessibility: Ensure the application is accessible to users with disabilities.

Reliability
- Uptime: System should be available 99.9% of the time.
- Failure Recovery: Implement failover mechanisms to ensure continuity.

Security
- Data Protection: Ensure all user data is securely stored and transmitted.
- Authentication: Implement user authentication and authorization.

Maintainability
- Code Structure: Follow best practices for code organization and documentation.
- Ease of Updates: Ensure the system can be easily updated with new features and improvements.

Stakeholder List

- Internal Stakeholders
  - Team Members: Sumayya Sana Syeda, Suleman Ali Saif Mir
  - Management: Sunmaya Sana Syeda
  - IT Department: Suleman Ali Saif Mir

- External Stakeholders
  - Clients
  - End-Users
  - Business Partners
  - Policymakers

Glossary

- API: A set of routines, protocols, and tools for building software and applications.
- EDA: The process of analyzing datasets to summarize their main characteristics.
- VADER: A lexicon and rule-based sentiment analysis tool.
- LSTM: A type of recurrent neural network used in deep learning.
- UI: The space where interactions between humans and machines occur.
- UX: The overall experience of a person using a product.

Appendices

- Appendix A: Data Flow Diagrams
- Appendix B: Sample Data
- Appendix C: User Interface Mockups
- Appendix D: Bibliography and References

Validation

Review Sessions
- Hold meetings with stakeholders to review the requirements document.
- Ensure all requirements are clearly understood and agreed upon.

Prototyping
- Develop prototypes to validate requirements with stakeholders.
- Collect feedback and make necessary adjustments.

Feedback and Approval
- Gather feedback from stakeholders.
- Incorporate changes and obtain formal approval.

Requirement Management

Traceability
- Track requirements throughout the project lifecycle to ensure all are met.

Change Management
- Establish a process for handling changes to requirements.
- Document all changes and their impacts.

Version Control
- Maintain different versions of the requirements document.
- Use version control systems to track changes.

This requirements document serves as a foundational guide for developing the Social Media Sentiment Analysis System. It outlines the project's goals, scope, functionalities, and constraints, ensuring all stakeholders have a clear understanding of the project requirements and expectations.