Business Requirements Document (BRD) - News Sentiment Analysis Project

**Introduction**

**1.1 Purpose:**

The purpose of this Business Requirements Document (BRD) is to outline the requirements for conducting sentiment analysis on news articles. The project will utilize datasets from the Trending News Classifier and the Video Pattern Categorization Team. The primary objective is to analyze the sentiments expressed in news articles and video content to gain valuable insights into public opinions and emotions.

**1.2 Scope:**

The scope of this project encompasses the following components:

* **Data Collection:** Obtain datasets from the Trending News Classifier and Video Pattern Categorization Team. The datasets should include news articles and video content along with corresponding sentiment labels (e.g., positive, negative, neutral).
* **Data Preprocessing:** Clean and preprocess the textual data from news articles, including the removal of noise, punctuation, and special characters. For video content, extract relevant features like visual content, audio signals, or text subtitles.
* **Sentiment Analysis:** Develop sentiment analysis models using natural language processing (NLP) techniques for news articles and video content. Generate sentiment scores for each item and calculate the overall sentiment for specific topics or categories.
* **Integration:** Integrate the sentiment analysis models for news articles and video content to gain insights from both types of data. Ensure seamless communication between the components to obtain a unified analysis.
* **Performance Evaluation:** Assess the performance of the sentiment analysis models using appropriate evaluation metrics, ensuring accuracy and efficiency.
* **Security:** Implement measures to protect user data, ensure the confidentiality of analyzed content, and adhere to relevant data protection and privacy regulations.
* **Web Interface (Optional):** Develop a user-friendly web interface using Django or Flask, allowing users to interact with the sentiment analysis system conveniently.

**Business Objectives**

**2.1 Data Collection**

* **2.1.1** Obtain datasets from the Trending News Classifier and Video Pattern Categorization Team. The datasets should contain labeled news articles and video content, along with their corresponding sentiment labels.

**2.2 Sentiment Analysis**

* **2.2.1** Develop sentiment analysis models for news articles and video content using appropriate NLP and machine learning techniques.
* **2.2.2** Generate sentiment scores for each news article and video, indicating the polarity of sentiments (positive, negative, neutral).
* **2.2.3** Calculate the overall sentiment for specific topics or categories based on the sentiment scores of individual articles and videos.

**2.3 Integration**

* **2.3.1** Integrate the sentiment analysis models for news articles and video content to gain unified insights from both types of data.
* **2.3.2** Ensure real-time and seamless communication between the sentiment analysis components.

**2.4 Visualization**

* **2.4.1** Create informative and visually appealing visualizations to present sentiment analysis results effectively.
* **2.4.2** Use appropriate charts, graphs, or word clouds to display sentiment distributions for different topics or video categories.

**2.5 Performance Evaluation**

* **2.5.1** Assess the performance of the sentiment analysis models using standard evaluation metrics such as accuracy, precision, recall, and F1 score.
* **2.5.2** Optimize the models for efficient sentiment analysis and response times.

**2.6 Security**

* **2.6.1** Implement measures to protect user data and ensure the confidentiality of analyzed news articles and video content.
* **2.6.2** Adhere to relevant data protection and privacy regulations, such as GDPR or any other applicable laws.

**2.7 Web Interface (Optional)**

* **2.7.1** Develop a user-friendly web interface using Django or Flask to interact with the sentiment analysis system.
* **2.7.2** Allow users to input news articles or video content for sentiment analysis through the web interface.
* **2.7.3** Ensure the web interface is responsive, compatible with different browsers, and provides an intuitive user experience.

**Assumptions and Constraints**

**3.1 Assumptions**

* **3.1.1** Users will provide relevant and valid news articles and video content for sentiment analysis.
* **3.1.2** Sufficient and reliable datasets will be available from the Trending News Classifier and Video Pattern Categorization Team.

**3.2 Constraints**

* **3.2.1** The project must be completed within the specified timeframe and budget constraints.
* **3.2.2** The sentiment analysis system must comply with the limitations and capabilities of the available datasets.

**Risks and Mitigation Strategies**

**4.1 Risk: Insufficient or low-quality data from the datasets may impact the accuracy of sentiment analysis.**

Mitigation: Conduct thorough data quality checks and preprocessing to ensure data integrity. If needed, consider data augmentation techniques to improve the dataset's quality.

**4.2 Risk: Technical challenges in developing the sentiment analysis models for news articles and video content.**

Mitigation: Collaborate with experienced data scientists and NLP experts to overcome technical complexities. Conduct extensive testing and validation of the models to ensure accuracy.

**Sign-off**

By signing below, the stakeholders acknowledge their understanding and agreement with the requirements outlined in this Business Requirements Document for the News Sentiment Analysis Project.