Text

Description automatically generated

**Activity based**

**Project Report on**

**AI Business Intelligence**

**Project Phase - I**

**Submitted to Vishwakarma University, Pune**

**Under the Initiative of**

**Contemporary Curriculum, Pedagogy, and Practice (C2P2)**

**By**

**Shravan Sudhir Meshram**

**SRN No : 202101425**

**Roll No : 31**

**Div : E**

**Third Year Engineering**

**Faculty In charge:- Prof. Moumita Pal**

**Date Of Project Phase 1:- 16 February 2024**

**Department of Computer Engineering**

**Faculty of Science and Technology**

**Academic Year**

**2023-2024 Term-II**

**Business Intelligence : Phase I**

**Project Name : Social Media Sentiment Analysis**

### ****Introduction:****

Social Media Sentiment Analysis, as it provides a direct window into the thoughts and feelings of customers, stakeholders, and the general public. By integrating sentiment analysis into BI processes, organizations can enhance their decision-making processes, improve customer relations, and stay ahead of market trends. Social Media Sentiment Analysis within the BI framework involves the systematic extraction and interpretation of sentiments embedded in the vast sea of user-generated content across social media channels. This process empowers BI professionals to gauge public perception, track brand sentiment, and identify emerging trends with unprecedented granularity.

Social Media Sentiment Analysis is a pivotal component in the realm of Business Intelligence, offering a unique lens into the sentiments expressed across digital platforms. By employing advanced natural language processing and machine learning techniques, BI professionals can decipher the positive, negative, or neutral tones in user-generated content.

### Problem Statement

Understanding customer sentiments on social media platforms is crucial for businesses to manage their online reputation, identify areas for improvement, and tailor marketing strategies. This project focuses on developing a Business Intelligence (BI) solution for Social Media Sentiment Analysis, extracting insights from social media data

### Objective

The primary objective is to design and implement a BI system that integrates with social media data sources, performs sentiment analysis, and provides actionable insights into customer opinions, trends, and sentiment shifts.

### ****Motivation:****

Social Media Sentiment Analysis stems from the recognition that social media has become a powerful and influential communication channel, shaping public opinions and influencing consumer behaviour. Several compelling motivations drive the need for this integration:

1. Understand Consumer Sentiments:

* Gain insights into how customers feel about products, services, and brands.

1. Enable Real-Time Responses:

* React swiftly to changing sentiments on social media, staying ahead of trends and issues.

1. Attain Competitive Edge:

* Predict market trends and gain a competitive advantage by adapting strategies proactively.

1. Manage Brand Reputation:

* Monitor and respond to sentiments, protecting and enhancing brand reputation.

1. Inform Data-Driven Marketing:

* Tailor marketing strategies based on customer sentiment for more effective campaigns.

1. Support Strategic Decision-Making:

* Provide decision-makers with valuable insights for formulating informed and strategic plans.

1. Enhance Customer Experience:

Identify areas for improvement in products or services, ultimately enhancing customer satisfaction.

### ****Literature Review:****

* Social Media Sentiment Analysis into Business Intelligence (BI) reveals a growing body of research and practical applications in the intersection of these two domains. Scholars and practitioners alike recognize the significance of leveraging social media data for enhancing BI processes. Here's a concise overview of the key themes found in the literature:
* Real-time Analytics: Many studies underscore the importance of real-time analytics in the context of social media sentiment analysis within BI. The ability to process and analyse vast amounts of social data in real-time enables organizations to make timely, informed decisions and respond promptly to emerging trends.
* Decision Support Systems: Scholars explore how integrating social media sentiment analysis into BI contributes to the development of more effective decision support systems. By providing actionable insights derived from sentiment analysis, BI systems become instrumental in strategic decision-making across various business functions.
* Customer Experience Management: The literature emphasizes the role of social media sentiment analysis in enhancing customer experience management. Businesses can proactively address customer concerns, identify areas for improvement, and personalize services based on sentiments expressed on social media, ultimately fostering stronger customer relationships.
* Brand Reputation and Crisis Management: Several studies delve into the use of social media sentiment analysis for brand reputation management and crisis communication. Organizations can monitor sentiments, identify potential PR issues, and respond swiftly to protect and repair their brand image during crises.
* Challenges and Ethical Considerations: Researchers acknowledge the challenges associated with social media sentiment analysis, including the dynamic nature of language, slang, and evolving expressions. Privacy concerns and ethical considerations related to the use of personal data in sentiment analysis also emerge as critical themes in the literature.

### ****Existing Work :****

In this project, we undertake exploratory data analysis (EDA) on the Social Media Sentiment dataset. The process includes:

1. **Data Collection:** Collect social media data from various platforms, including customer reviews, comments, and mentions.

2. **Data Preprocessing:** Preprocess the social media data to handle noise, irrelevant information, and ensure data consistency. Clean text data, handle emotions, and address any data quality issues that may affect sentiment analysis.

3. **Feature Engineering:** Identify and engineer features that contribute to sentiment analysis. This may include sentiment scores, sentiment trends over time, and the identification of key topics or keywords associated with positive or negative sentiments.

4. **BI Dashboard Development**: Design and implement a user-friendly BI dashboard that visualizes key sentiment analysis metrics. Include components for monitoring overall sentiment trends, identifying sentiment influencers, and assessing the impact of marketing campaigns.

### ****Future Scope:****

The current analysis serves as a foundation for more in-depth investigations and modelling to predict employee performance. Future steps could include:

1. **Deeper Insights:** Nuance beyond basic sentiment, analyzing visuals, and complex emotions.
2. **Personalization:** Tailored experiences, targeted marketing, and relevant content.
3. **Beyond Marketing:** Individual reputation, public opinion tracking, campaign impact assessment.
4. **Algorithm Advancements:** Increased accuracy with context, cultural nuances, and evolving language.
5. **Emerging Applications:** Mental health monitoring, hate speech detection, safer online environment.
6. **Challenges:** Privacy, bias, manipulation, need for responsible and transparent use.

### ****Future Plan:****

The next phase of this project involves:

1**. Sentiment Score Calculation:** Implement features for calculating sentiment scores from social media text data. Utilize natural language processing (NLP) techniques to assess the polarity of customer opinions.

2. **Trend Analysis:** Develop features for trend analysis, tracking sentiment changes over time. Identify patterns, spikes, or dips in sentiment that may coincide with specific events, product launches, or marketing efforts.

3. **Influencer Identification:** Integrate features for identifying social media influencers who impact sentiment. This includes recognizing individuals or accounts whose opinions carry significant weight within the online community.

4. **Brand Mention Analysis:** Analyze brand mentions within the BI system. Understand the context and sentiment associated with brand mentions to assess the overall perception of the brand in the social media landscape

### ****Dataset:****

### 

### ****Data Preprocessing:****

Importing Libraries:

import re

import nltk

import pandas as pd

from nltk.corpus import stopwords

from nltk.tokenize import word\_tokenize

from nltk.sentiment.vader import SentimentIntensityAnalyzer

from sklearn.feature\_extraction.text import TfidfVectorizer

from sklearn.decomposition import LatentDirichletAllocation

Loading Dataset:

# Load your dataset

df = pd.read\_csv('ST.csv')

Finding The Missing Values in the Dataset:

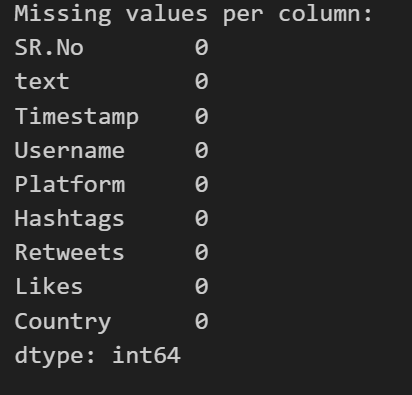
# Count missing values per column

missing\_values = df.isnull().sum()

# Display missing values per column

print("Missing values per column:")

print(missing\_values)



Text Pre-Processing:

# Text preprocessing function

def preprocess\_text(text):

    text = re.sub(r'http\S+', '', str(text))  # Remove URLs

    text = re.sub(r'@\w+|\#', '', str(text))   # Remove mentions and hashtags

    text = text.encode('ascii', 'ignore').decode('ascii')  # Remove emojis

    text = text.lower()  # Convert to lowercase

    tokens = word\_tokenize(text)  # Tokenize

    stop\_words = set(stopwords.words('english'))

    tokens = [token for token in tokens if token not in stop\_words]  # Remove stopwords

    cleaned\_text = ' '.join(tokens)  # Join tokens

    return cleaned\_text

Applying TF-IDF Vectorization

# TF-IDF Vectorization

tfidf\_vectorizer = TfidfVectorizer()

tfidf\_matrix = tfidf\_vectorizer.fit\_transform(df['cleaned\_text'])

# Latent Dirichlet Allocation (LDA) for topic modeling

lda\_model = LatentDirichletAllocation(n\_components=2, random\_state=42)

lda\_matrix = lda\_model.fit\_transform(tfidf\_matrix)

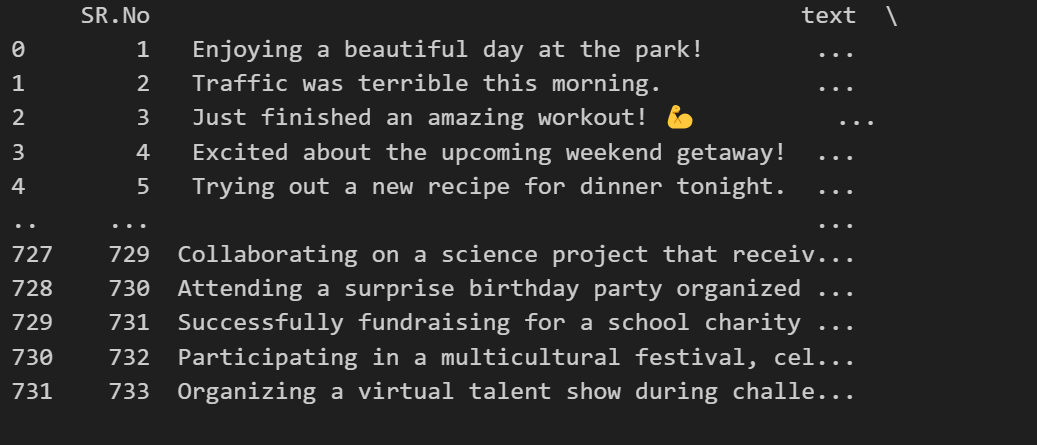
# Add sentiment score and LDA topic features to dataframe

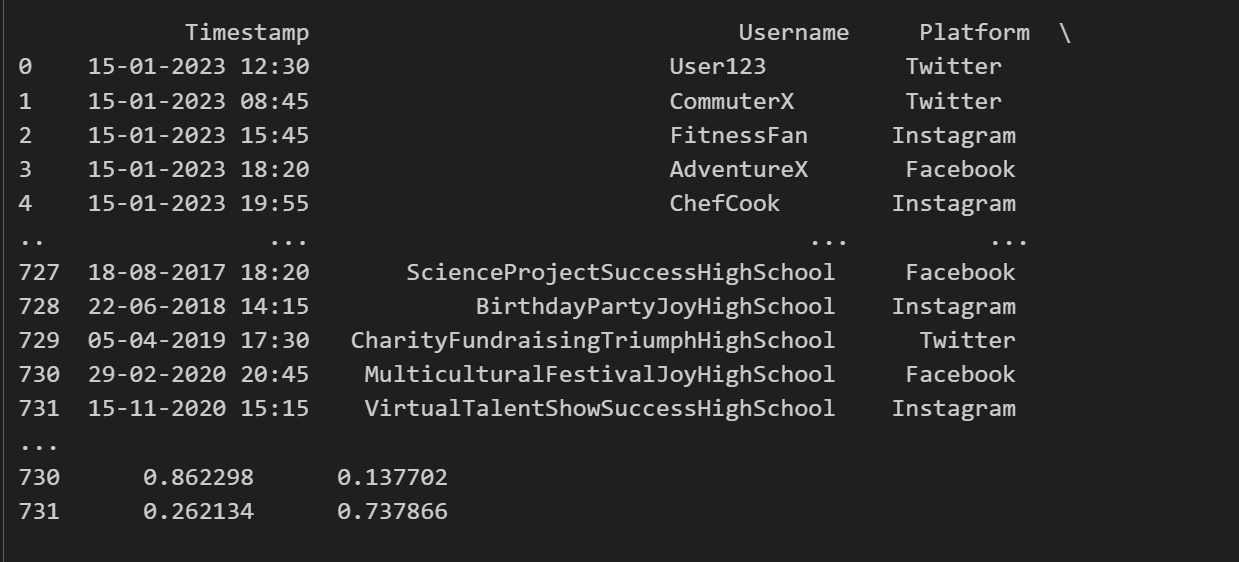
df['topic\_1\_prob'] = lda\_matrix[:, 0]

df['topic\_2\_prob'] = lda\_matrix[:, 1]

# Display the dataframe with engineered features

print(df)





**BI Dashboard Development**:

