VU VISHWAKARMA UNIVERSITY
Maximising Human Potential

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.
2. Enable Real-Time Responses:
- React swiftly to changing sentiments on social media, staying ahead of trends and issues.
3. Attain Competitive Edge:
- Predict market trends and gain a competitive advantage by adapting strategies proactively.
4. Manage Brand Reputation:
- Monitor and respond to sentiments, protecting and enhancing brand reputation.
5. Inform Data-Driven Marketing:
- Tailor marketing strategies based on customer sentiment for more effective campaigns.
6. Support Strategic Decision-Making:
- Provide decision-makers with valuable insights for formulating informed and strategic plans.
7. 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:

	A	B	C	D	E	F	G	H	I	J
3	Just finished an amazing workout! ðŸ'ª			15-01-2023 15:45	FitnessFa	Instagrar	#Fitness :	20	40	USA
4	Excited about the upcoming weekend getaway!			15-01-2023 18:20	Adventur	Facebook	#Travel #	8	15	UK
5	Trying out a new recipe for dinner tonight.			15-01-2023 19:55	ChefCook	Instagrar	#Cooking	12	25	Australia
6	Feeling grateful for the little things in life.			16-01-2023 09:10	Gratitude	Twitter	#Gratitud	25	50	India
7	Rainy days call for cozy blankets and hot cocoa.			16-01-2023 14:45	RainyDay	Facebook	#RainyDa	10	20	Canada
8	The new movie release is a must-watch!			16-01-2023 19:30	MovieBuf	Instagrar	#MovieNi	15	30	USA
9	Political discussions heating up on the timeline.			17-01-2023 08:00	DebateTa	Twitter	#Politics	30	60	USA
10	Missing summer vibes and beach days.			17-01-2023 12:20	BeachLov	Facebook	#Summer	18	35	Australia
11	Just published a new blog post. Check it out!			17-01-2023 15:15	BloggerX	Instagrar	#Blogging	22	45	USA
12	Feeling a bit under the weather today.			18-01-2023 10:30	Wellness	Twitter	#SickDay	7	15	Canada
13	Exploring the city's hidden gems.			18-01-2023 14:50	UrbanExp	Facebook	#CityExpl	12	25	UK
14	New year, new fitness goals! ðŸ'ª			18-01-2023 18:00	FitJourne	Instagrar	#NewYeal	28	55	USA
15	Technology is changing the way we live.			19-01-2023 09:45	TechEnth	Twitter	#Tech #In	15	30	India
16	Reflecting on the past and looking ahead.			19-01-2023 13:20	Reflectio	Facebook	#Reflectic	20	40	USA
17	Just adopted a cute furry friend! ðŸ¾			19-01-2023 17:10	PetAdopt	Instagrar	#PetAdop	15	30	Canada
18	Late-night gaming session with friends.			20-01-2023 00:05	GamerX	Twitter	#Gaming	18	35	UK
19	Attending a virtual conference on AI.			20-01-2023 11:30	TechConf	Facebook	#AI #Tech	25	50	USA
20	Winter blues got me feeling low.			20-01-2023 15:15	WinterBl	Instagrar	#WinterB	8	15	USA
21	Sipping coffee and enjoying a good book.			21-01-2023 08:40	Bookwor	Twitter	#Reading	22	45	India
22	Exploring the world of virtual reality.			21-01-2023 13:20	VRExplor	Facebook	#VR #Virt	15	30	USA
23	Productive day ticking off my to-do list.			21-01-2023 16:45	Productiv	Instagrar	#Product	30	60	USA
24	Just finished a challenging workout routine.			22-01-2023 09:15	FitnessW	Twitter	#Fitness :	20	40	UK
25	Celebrating a milestone at work! ðŸŽ‰			22-01-2023 14:30	CareerMi	Facebook	#Career #	12	25	Canada
27	Sunday brunch with friends.			22-01-2023 12:00	BrunchBu	Instagrar	#Brunch :	15	30	UK
28	Learning a new language for personal growth.			23-01-2023 16:20	Language	Facebook	#Languag	25	50	India
29	Quiet evening with a good book.			23-01-2023 19:45	BookLove	Instagrar	#Reading	15	30	Australia
30	Reflecting on the importance of mental health.			24-01-2023 11:30	MentalHe	Twitter	#MentalH	22	45	USA
31	New painting in progress! ðŸŽ¨			24-01-2023 15:00	ArtistInA	Facebook	#Art #Pai	12	25	UK
32	Weekend road trip to explore scenic views.			24-01-2023 17:30	RoadTrip	Instagrar	#RoadTri	18	35	UK
33	Enjoying a cup of tea and watching the sunset.			25-01-2023 18:00	SunsetWi	Twitter	#TeaTime	15	30	India
34	Coding a new project with enthusiasm.			25-01-2023 13:15	CodeEnth	Facebook	#Coding #	30	60	USA
35	Feeling inspired after attending a workshop.			26-01-2023 09:45	Worksho	Instagrar	#Inspirat	25	50	USA
36	Winter sports day at the local park.			26-01-2023 14:20	WinterSp	Twitter	#WinterS	15	30	Canada
37	Quality time with family this weekend.			26-01-2023 17:40	FamilyTir	Facebook	#FamilyT	22	45	UK
38	Attending a live music concert tonight.			27-01-2023 20:00	MusicLov	Instagrar	#Music #	18	35	USA
39	Practicing mindfulness with meditation.			27-01-2023 12:30	MindfulN	Twitter	#Mindful	15	30	India
40	Trying out a new dessert recipe.			27-01-2023 16:10	DessertE>	Facebook	#Dessert	12	25	Canada
41	Excited about the upcoming gaming tournament.			28-01-2023 09:00	GamingEi	Instagrar	#Gaming	30	60	USA
42	Planning a garden makeover for spring.			28-01-2023 14:45	GardenPl	Twitter	#Gardeni	25	50	USA
43	Celebrating a friend's birthday tonight! ðŸŽ‚			28-01-2023 19:30	Birthday	Facebook	#Birthday	15	30	UK
44	Feeling accomplished after a productive day.			29-01-2023 17:00	Productiv	Instagrar	#Accomp	22	45	India

Data Preprocessing:

Importing Libraries:

```
import re
```

```python
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:

```python
# Load your dataset
df = pd.read_csv('ST.csv')
```

Finding The Missing Values in the Dataset:

```python
# Count missing values per column
missing_values = df.isnull().sum()

# Display missing values per column
print("Missing values per column:")
print(missing_values)
```

```
Missing values per column:
SR.No          0
text           0
Timestamp      0
Username       0
Platform       0
Hashtags       0
Retweets       0
Likes          0
Country        0
dtype: int64
```

Text Pre-Processing:

```python
# 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

```python
# 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)
```

```
     SR.No                                         text  \
0        1   Enjoying a beautiful day at the park!      ...
1        2   Traffic was terrible this morning.         ...
2        3   Just finished an amazing workout! 💪         ...
3        4   Excited about the upcoming weekend getaway!  ...
4        5   Trying out a new recipe for dinner tonight.  ...
..     ...                                             ...
727    729   Collaborating on a science project that receiv...
728    730   Attending a surprise birthday party organized ...
729    731   Successfully fundraising for a school charity ...
730    732   Participating in a multicultural festival, cel...
731    733   Organizing a virtual talent show during challe...
```

```
           Timestamp                                    Username     Platform  \
0      15-01-2023 12:30                                  User123      Twitter
1      15-01-2023 08:45                                CommuterX      Twitter
2      15-01-2023 15:45                                FitnessFan    Instagram
3      15-01-2023 18:20                                AdventureX     Facebook
4      15-01-2023 19:55                                 ChefCook     Instagram
..          ...                                            ...           ...
727    18-08-2017 18:20      ScienceProjectSuccessHighSchool         Facebook
728    22-06-2018 14:15         BirthdayPartyJoyHighSchool           Instagram
729    05-04-2019 17:30   CharityFundraisingTriumphHighSchool         Twitter
730    29-02-2020 20:45     MulticulturalFestivalJoyHighSchool       Facebook
731    15-11-2020 15:15     VirtualTalentShowSuccessHighSchool      Instagram

...
730      0.862298          0.137702
731      0.262134          0.737866
```

BI Dashboard Development: