

# USECASE

## SOCIAL MEDIA SENTIMENT ANALYSIS

### AIM

The aim of this project is to analyze the sentiments expressed by users on social media platforms.

### ALGORITHM

#### 1. **Data Collection:**

- Gather data from social media platforms using APIs (e.g., Twitter API).

#### 2. **Data Preprocessing:**

- Text Cleaning: Remove special characters, links, and unnecessary spaces.

#### 3. **Feature Extraction:**

- Convert text data into numerical representations using methods.

#### 4. **Sentiment Classification:**

- Use a pre-trained sentiment analysis model or train a machine learning model on labeled sentiment data.

#### 5. **Model Training and Testing:**

- Split the dataset into training and testing sets.
- Train the model on the training data and test its accuracy using the testing data.

#### 6. **Evaluation:**

- Evaluate the performance of the sentiment analysis model using metrics like accuracy, precision, recall, F1-score, etc.

#### 7. **Deployment:**

- Deploy the trained model for real-time analysis or periodically for batch processing.

## PROGRAM

```
install.packages("tidyverse")  
install.packages("tidytext")  
install.packages("textdata")  
install.packages("wordcloud")  
install.packages("e1071")  
install.packages("caret")
```

```
library(tidyverse)  
library(tidytext)  
library(textdata)  
library(wordcloud)  
library(e1071)  
library(caret)
```

```
dataset <- read.csv("social_media_data.csv")  
clean_data <- dataset %>%  
  mutate(text = str_replace_all(text, "[^[:alnum:]]", "")) %>%  
  mutate(text = tolower(text))  
clean_data <- clean_data %>%  
  unnest_tokens(word, text)  
clean_data <- clean_data %>%  
  anti_join(stop_words)  
  
dtm <- clean_data %>%  
  count(sentiment, word) %>%  
  cast_dtm(sentiment, word, n)  
model <- naiveBayes(dtm, dataset$sentiment)
```

```
prediction <- predict(model, dtm)
confusionMatrix(prediction, dataset$sentiment)
```

```
word_freq <- clean_data %>%
  count(word, sort = TRUE)
```

```
wordcloud(words = word_freq$word, freq = word_freq$n, max.words = 100)
```

## SAMPLE DATASET

text	sentiment
"I love this product, it's amazing"	positive
"This is the worst experience ever"	negative
"The product is okay, nothing special"	neutral
"Highly recommend it to everyone!"	positive
"Not worth the money, really bad"	negative
"It's a decent product."	neutral

## OUTPUT

```
"I love this product, it's amazing" --> "love product amazing"  
"This is the worst experience ever" --> "worst experience ever"
```

### Confusion Matrix and Statistics

Prediction	Reference		
	Negative	Neutral	Positive
Negative	10	0	1
Neutral	0	7	2
Positive	1	0	15

### Overall Statistics

```
Accuracy : 0.92  
95% CI : (0.86, 0.97)  
No Information Rate : 0.48  
P-Value [Acc > NIR] : < 2.2e-16  
  
Kappa : 0.88
```

## RESULT

Thus, the program for analysis for sentiments expressed by users on social media platform has been executed successfully.