# SENTIMENT ANALYSIS – THEORY
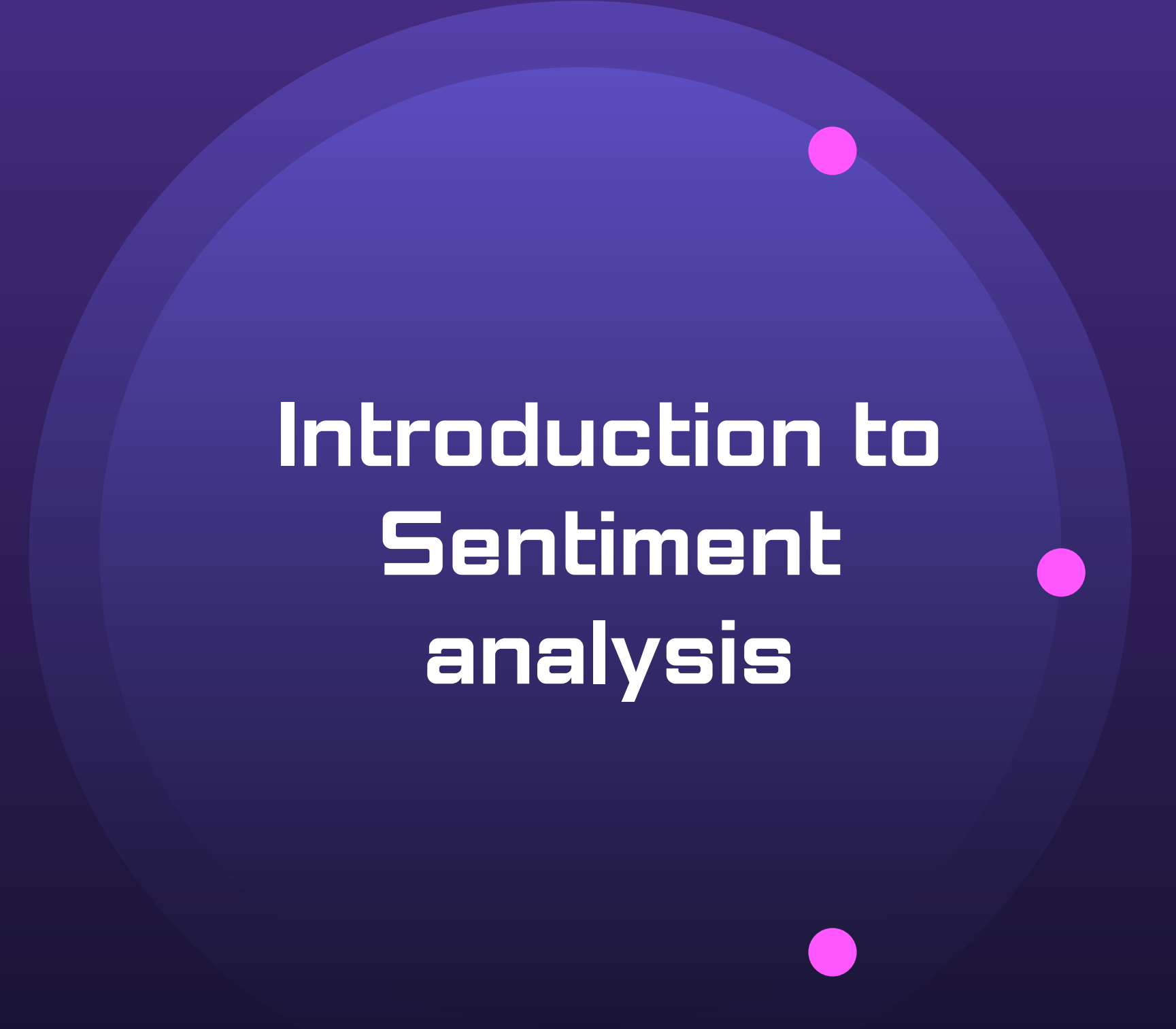
Presented by Tannu Chaudhary & Lekha Mazumder

# Introduction to Sentiment analysis

## What is Sentiment analysis

Sentiment Analysis, also called Opinion Mining, is a technique used to identify and extract emotions, opinions, or attitudes expressed in a piece of text.

## Purpose

To determine whether the sentiment is:
- Positive
- Negative
- Neutral

## Examples

- "I love this phone." → Positive
- "This phone is bad." → Negative
- "This is a smartphone." → Neutral
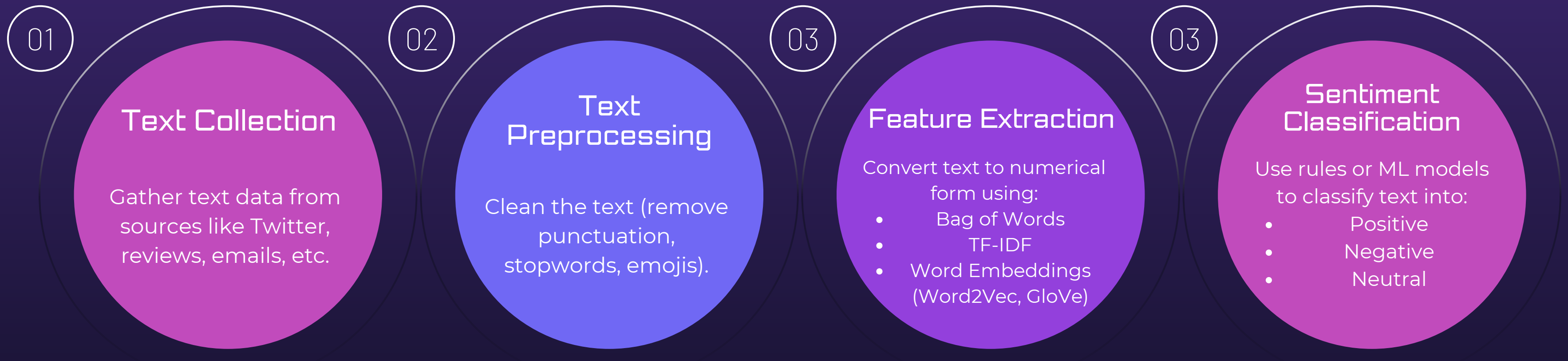
# Why Sentiment analysis is important

## Key Reasons

- Helps businesses understand customer feedback
- Assists in decision-making based on public opinion
- Tracks brand reputation on social media
- Used in politics, finance, healthcare, etc.

## Real life examples

- Analyzing Twitter data to predict election trends
- Evaluating product reviews on Amazon or Flipkart
- Monitoring public reactions to a new policy or event

# How Sentiment Analysis Works (Step-by-step)

**01**

## Text Collection

Gather text data from sources like Twitter, reviews, emails, etc.

**02**

## Text Preprocessing

Clean the text (remove punctuation, stopwords, emojis).

**03**

## Feature Extraction

Convert text to numerical form using:
- Bag of Words
- TF-IDF
- Word Embeddings (Word2Vec, GloVe)

**03**

## Sentiment Classification

Use rules or ML models to classify text into:
- Positive
- Negative
- Neutral

# Approaches to Sentiment Analysis

01 Lexicon based approach

02 Machine Learning-Based Approach

03 Hybrid Approach

# Lexicon based approach

## What is it?

This approach uses a predefined dictionary of words (called a lexicon) where each word has a sentiment score—either positive, negative, or neutral.

## How is it works ?

- Each word is given a score:
- "Happy" = +1
- "Bad" = -1
- The sentiment of the sentence is calculated by adding the scores of all words.

## Example

Sentence: "The camera is good but the battery is bad."
- "Good" = +1
- "Bad" = -1
- → Total Score = 0 → Neutral sentiment

## Pros

- Simple and easy to implement
- Doesn't require training data

## Cons

- Cannot understand context or sarcasm
- Limited by the quality of the dictionary

# Machine Learning-Based Approach

## What is it ?

In this method, the computer is trained using example texts (with known sentiments) to learn patterns and predict the sentiment of new texts.

## How is it works ?

- Provide a large dataset with labeled examples (positive/negative)
- Convert text into features (like word counts or embeddings)
- Train a model (like Naive Bayes, SVM, or deep learning)
- Predict sentiment on new text based on learned patterns

## Example

If the model has seen:
- "I love this phone." → Positive
- "I hate this app." → Negative

Then when it sees:
- "This phone is amazing." → It predicts: Positive

## Pros

- Understands context better
- More accurate with large
-  datasets

## Cons

- Needs a lot of labeled data
- Takes time and computational resources to train

# Hybrid Approach

## What is it ?

This approach combines both Lexicon-based and Machine Learning methods to get the advantages of both.

## How is it works ?

- Use a sentiment dictionary to assign initial scores to words
- Then apply a machine learning model for better interpretation
- Often used in real-world applications for improved accuracy

## Example

- The word "great" may get a positive score from the lexicon
- The ML model checks sentence structure, context, and decides if overall sentiment is positive or mixed

# Pros

- More accurate than using either approach alone
- Works well in complex or real-world data

# Cons

- More complex to design and implement

# Application of Sentiment analysis

01

## Business

Analyzing customer
feedback
Product improvement

02

## Social media
monitoring

Tracking brand mentions
Measuring public mood

03

## Finance

Predicting stock
trends from
news/tweets

04

## Politics

Analyzing public
opinion during
elections

05

## Healthacre

Understanding
patient feedback in
surveys

## Benefits

- **Automates large-scale opinion analysis**
- **Supports data-driven decisions**
- **Improves customer satisfaction**

## Challenges

- **Difficult to detect sarcasm or humor**
- **Words may have multiple meanings**
- **(e.g., "This movie is sick!" = good or bad?)**
- **Slang and multilingual text can reduce accuracy**

# Conclusion

- Sentiment Analysis is a powerful tool to understand opinions from text.
- It combines NLP techniques with smart algorithms.
- Widely used in many industries to improve decisions and services.
- Future trends include:
  - Emotion detection
  - Multilingual sentiment analysis
  - Real-time feedback systems

# Thank you
# for your attention!!!