UNITED
Group of Institutions
Allahabad • Greater Noida

Summer Training Project Presentation

On

Text-Emoji Recommender

BACHELOR OF TECHNOLOGY

COMPUTER SCIENCE AND ENFINEERING

Name : Vaishali Pal Roll no. 2402841530084

Under Supervision of

Mrs. Pallavi Shukla (Assistant Professor)

Introduction

The Text Emoji Recommender suggests suitable emojis for any text using Machine Learning and NLP. It makes conversations more expressive, engaging, and easy to understand.

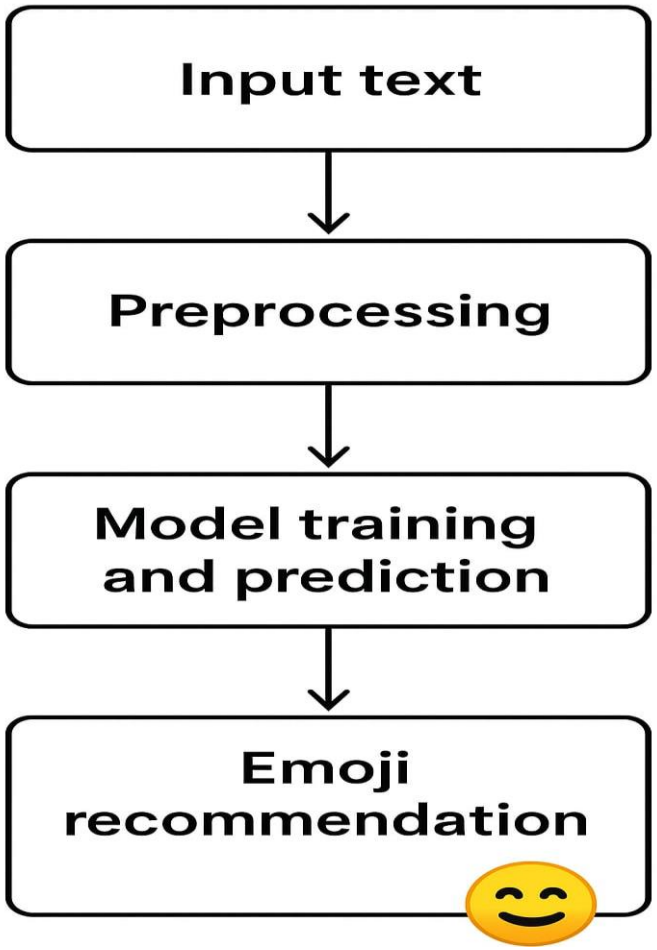
Objectives

To build a machine learning–based system that recommends suitable emojis for text messages, enhancing digital communication by making it more expressive, faster, and intuitive.

Hardware and Software Requirements

Hardware Requirement	Software Requirement
PC - Dell(intel CORE i5)	Programming Language-Python
RAM - 2-4 GB RAM	Libraries-pandas, numpy, scikit-learn(sklearn), matplotlib.pyplot, itertools, joblib, streamlit, ngrok.
Storage- 250mb	ML Framework-scikit-learn
Processor-13th Gen Intel(R) Core(™)i5-1340P,1900 Mhz, 12 Core(s), 16 Logical Proc…	IDE- Google Colab (Python version:3.11), ➡ Google Colab runs on Ubuntu 20.04 LTS as its notebook backend OS.
	Dataset- is self made(manually curated) (no existing dataset was suitable for text emoji recommender)

Data Flow Diagram



Future Work

- Context-Aware Recommendation
- Multi-Emoji Prediction
- Emoji Sense Disambiguation
- Cross-Language Support
- Multi emoji recommendation per sentence
- Real time chat integration

Source of Dataset

The dataset was manually created by pairing short text messages with suitable emojis. It was curated to cover diverse conversational contexts like emotions, greetings, and daily activities, ensuring high-quality and relevant training data for the recommender system.

Methodology

- Manually curated text–emoji dataset
- Text preprocessing & IDF feature extraction
- Training ML models (Logistic Regression, Linear Regression)
- Model evaluation using accuracy, MSE, classification report & confusion matrix
- Emoji prediction for new text inputs

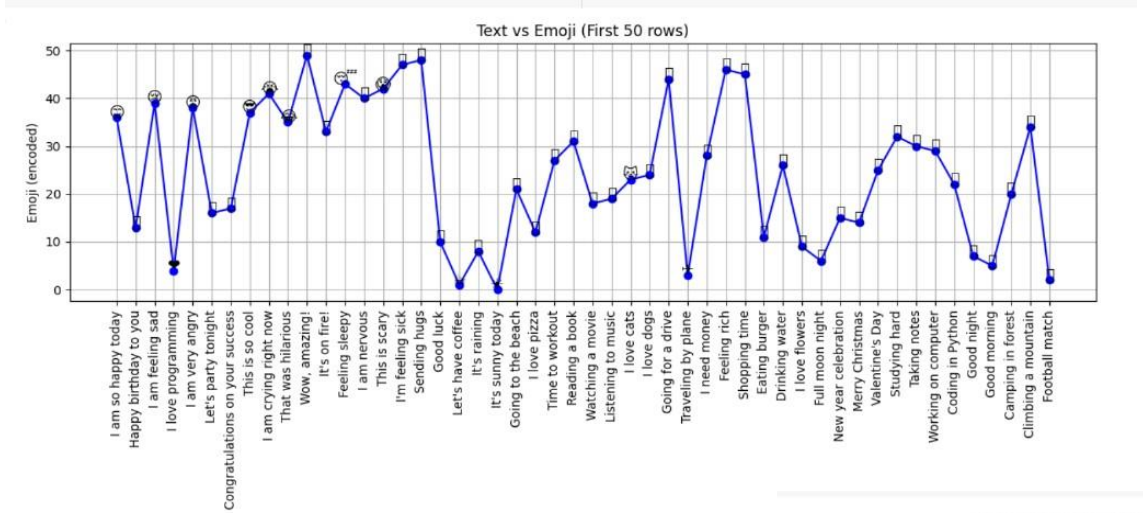
Result

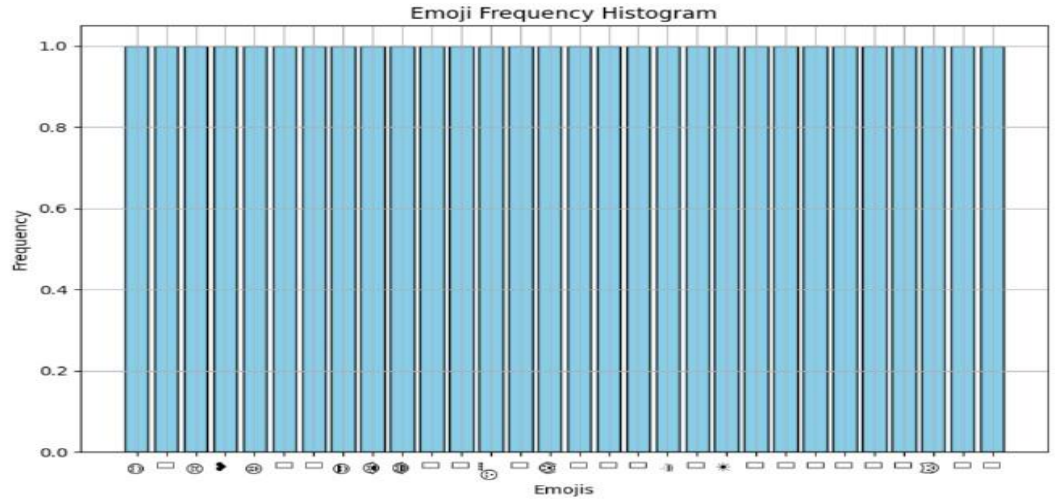
The model is able to recommend relevant emojis for different text inputs.

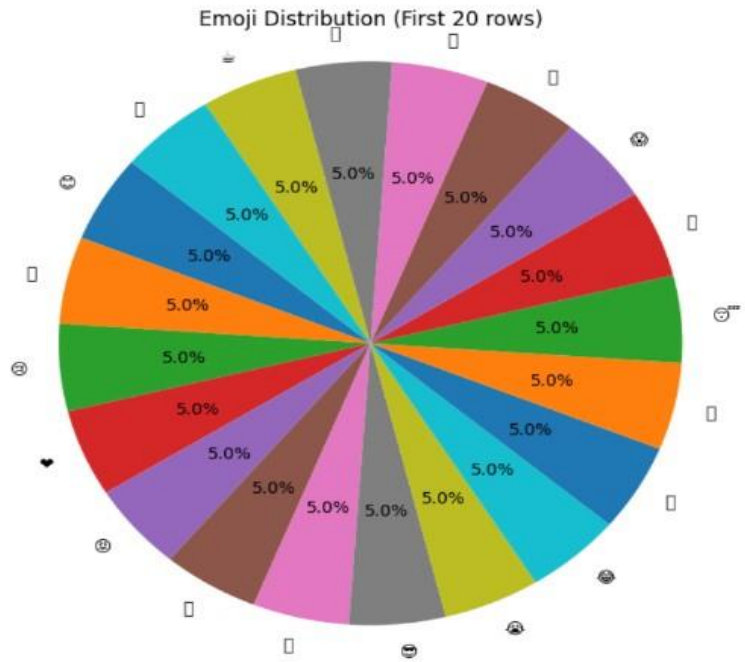
“I am sad” → 😢😭

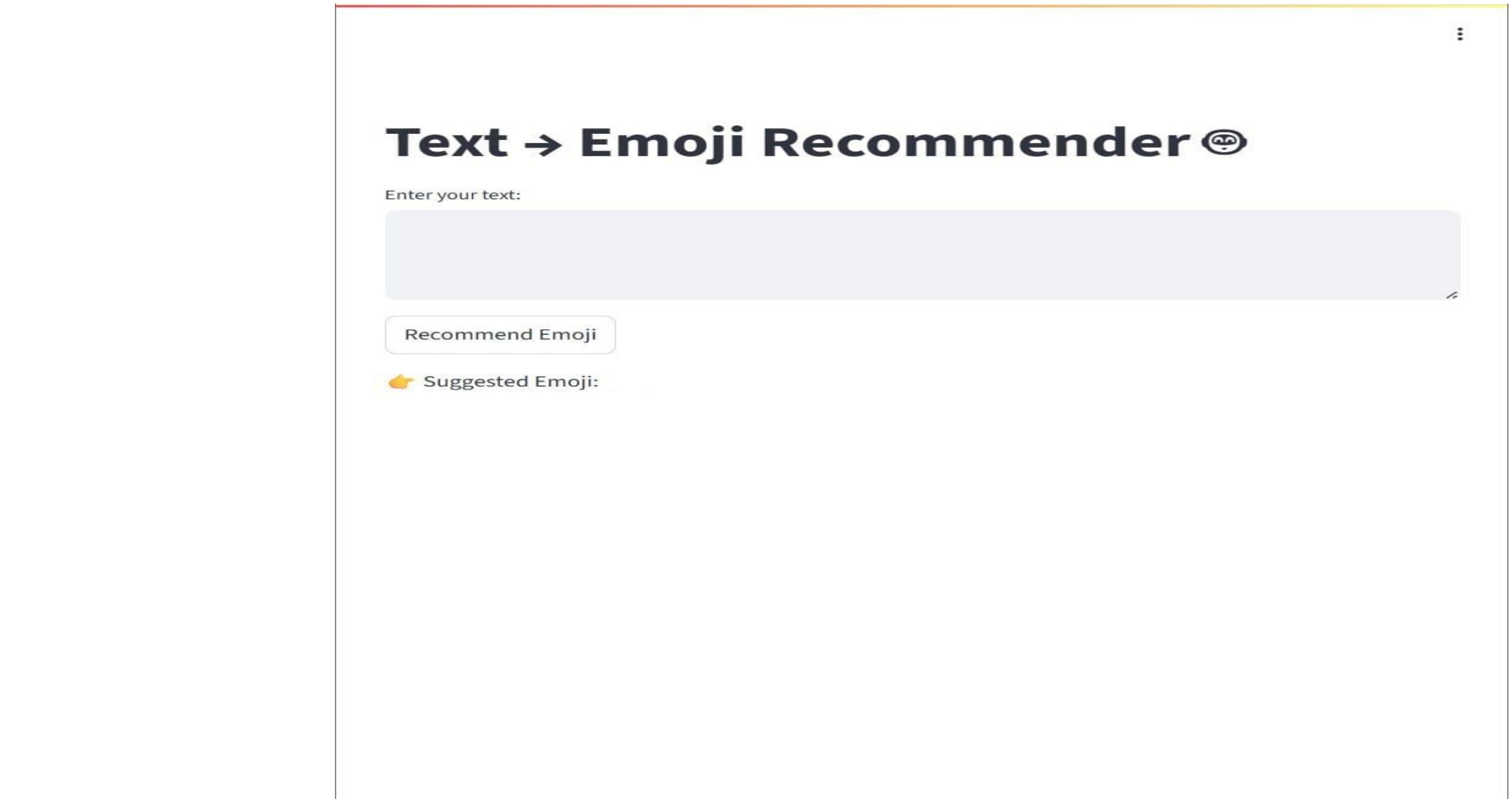
“I bought a new car” → 🚗🚙

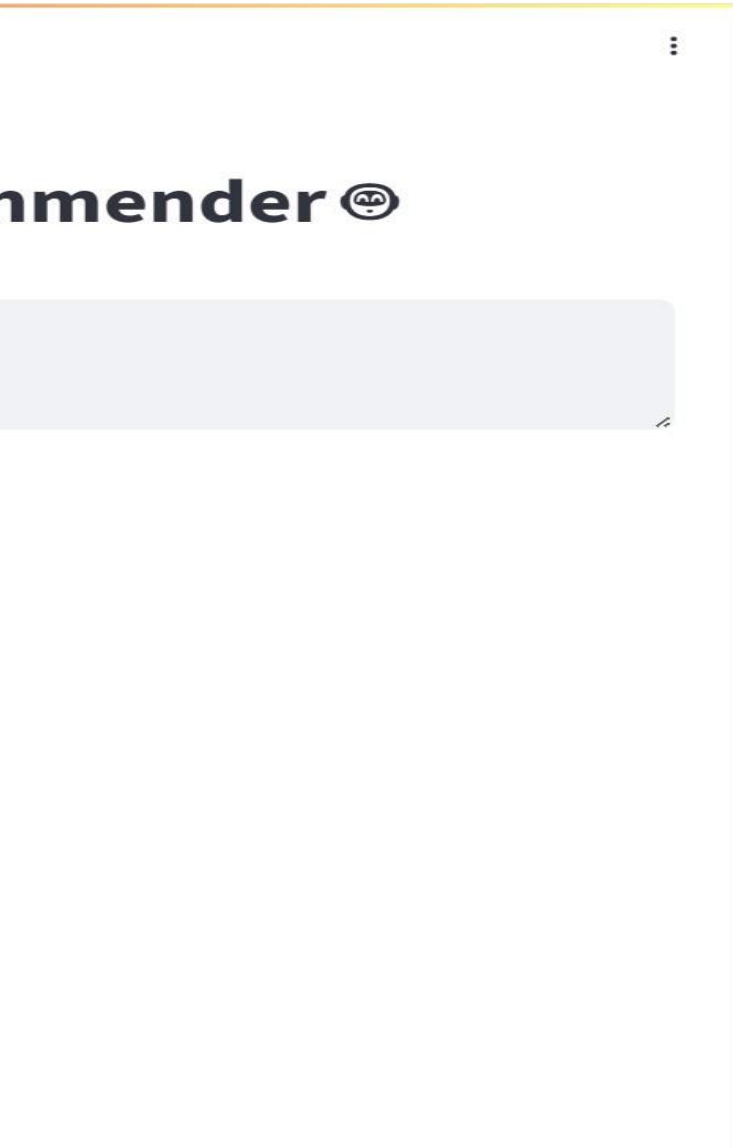
“I am reading a book” → 📚📖


Text vs Emoji (First 50 rows)
Emoji (encoded)
I am so happy today
Happy birthday to you
I am feeling sad
I love programming
I am very angry
Let's party tonight
Congratulations on your success
This is so cool
I am crying right now
That was hilarious
Wow, amazing!
It's on fire!
Feeling sleepy
I am nervous
This is scary
I'm feeling sick
Sending hugs
Good luck
Let's have coffee
It's raining
It's sunny today
Going to the beach
I love pizza
Time to workout
Reading a book
Watching a movie
Listening to music
I love cats
I love dogs
Going for a drive
Traveling by plane
I need money
Feeling rich
Shopping time
Eating burger
Drinking water
I love flowers
Full moon night
New year celebration
Merry Christmas
Valentine's Day
Studying hard
Taking notes
Working on computer
Coding in Python
Good night
Good morning
Camping in forest
Climbing a mountain
Football match


Emoji Frequency Histogram
Frequency
Emojis


Emoji Distribution (First 20 rows)
5.0%


Text → Emoji Recommender
Enter your text:
Recommend Emoji
Suggested Emoji:


Text → Emoji Recommender
Enter your text:
book
Recommend Emoji
Suggested Emoji:

Conclusion

- Built a text–emoji recommender using ML (IDF + classifiers).
- Achieved accurate emoji predictions on manually curated dataset.
- System enhances communication by making text more expressive.

References

- Inspiration from Wine & Salary dataset codes for ML workflow
- Guidance from group PDF (ST25 GEN-AI-NLP 1)
- Used docs/tutorials of scikit-learn, pandas & streamlit