# 📚Chatbot

## Project Title

Book Genre Chatbot using TF-IDF and Cosine Similarity

## Objective

To develop a simple chatbot that can recommend a book genre based on user queries using basic Natural Language Processing (NLP) techniques.

## Problem Statement

Users often find it difficult to choose a book genre that aligns with their interests. This chatbot helps users by recommending a genre based on a natural language query.

## Dataset Description

This project uses a small, simulated dataset of book genres and their short descriptions:  
  
Fantasy - A magical world full of dragons, wizards, and epic quests.  
Science Fiction - Explores futuristic technologies, space travel, and alien life.  
Romance - Love stories filled with emotion, passion, and relationships.  
Horror - Scary tales involving supernatural elements, fear, and suspense.  
Mystery - Whodunits and detective stories with twists and turns.

## Technologies Used

• Python 3.x  
• Pandas  
• Regular Expressions (Regex)  
• Scikit-learn (TfidfVectorizer, cosine\_similarity)

## Methodology

1. Text Cleaning:  
- Convert text to lowercase.  
- Remove punctuation and special characters.  
- Strip whitespaces.  
  
2. Vectorization:  
- TfidfVectorizer converts the cleaned descriptions into numerical vectors.  
  
3. Similarity Matching:  
- User input is transformed into a vector and compared with genre vectors using Cosine Similarity.  
- The genre with the highest similarity score is selected.

## How the Chatbot Works

1. User Input: The user enters a sentence (e.g., 'I love stories with magic and wizards').  
2. Preprocessing: The sentence is cleaned and vectorized.  
3. Similarity Check: The chatbot calculates similarity between user query and genre descriptions.  
4. Response: Returns the genre with the highest similarity score above a threshold.

## Sample Query

get\_genre\_response("I love stories about future and space adventures")  
  
Output:  
You might be interested in: Science Fiction

## Code Structure

Book\_Genre\_Chatbot.ipynb  
├── Step 1: Import libraries  
├── Step 2: Create dataset  
├── Step 3: Clean text using regex  
├── Step 4: Vectorize using TF-IDF  
├── Step 5: Define chatbot function  
├── Step 6: Example interaction

## How to Run the Notebook

1. Open Jupyter Notebook or Google Colab.  
2. Create a new notebook.  
3. Paste the code.  
4. Run each cell in order.  
5. Use get\_genre\_response("your query") to chat.

## Possible Enhancements

• Expand the dataset to include more genres.  
• Add a Flask or Streamlit interface.  
• Use BERT or GPT for better responses.  
• Store user preferences for personalized recommendations.

## Conclusion

This project shows how to build a basic genre recommendation chatbot using TF-IDF and cosine similarity. It's lightweight, interpretable, and easy to deploy.