

Career Guidance Chatbot – Documentation

Project Overview

Career Guidance Chatbot is an intelligent Streamlit-based web application designed to recommend career roles based on a user's interests or skills. The chatbot uses a machine learning model trained on a dataset of career-related questions and roles. It recommends the best-matching career and shows related information like salary, education, and courses.

Features

- Predict career based on user input (free text)
 - Highlight best-matching career and show top 3 probable matches
 - Show description, salary range, required education, and courses
 - Display model performance metrics (Accuracy, Precision, Recall, F1-score)
 - Sidebar with popular careers and user tips
 - Beautiful and responsive UI with session-state persistence
-

Project Structure

```
career_chatbot/
├── app.py          # Main Streamlit application
├── train_model.py  # train model
├── assets/
│   └── style.css   # Custom CSS for UI styling
├── intent_model.pkl # Trained SVM model
├── vectorizer.pkl  # TF-IDF vectorizer
├── updated_with_all.csv # Dataset with roles, question, answer, salary, all
├── career_guidance_dataset.csv # Dataset used for model training
└── README.md      # Project documentation
```

Machine Learning Model

- **Model:** Support Vector Machine (SVM)
- **Vectorizer:** TF-IDF
- **Input:** User-entered interests or skills
- **Output:** Predicted career role + Top 3 role suggestions with confidence
- **Dataset:** career_guidance_dataset.csv

Dependencies

Install the required libraries using:

```
pip install streamlit pandas joblib scikit-learn plotly
```

How It Works

1. **Load Assets:** Model, vectorizer, and data are cached for fast access.
 2. **User Input:** Text describing skills/interests is collected.
 3. **Prediction:** Input is cleaned, vectorized, and passed to the model.
 4. **Result:** Best matching career is displayed with details.
 5. **Extras:** Optional sections show salary, education, and courses.
 6. **Evaluation:** Metrics and confusion matrix are plotted.
-

UI/UX Highlights

- Custom header and card layout with CSS (`style.css`)
 - Sidebar: Popular roles as buttons + user tips
 - Progress bars for top 3 career matches
 - Toggle buttons for extra information
 - Session state is used to remember user inputs and button states
-

Evaluation Metrics

Model performance (calculated using test set):

- **Accuracy:** Measures overall correct predictions
 - **Precision:** Relevance of predicted results
 - **Recall:** Coverage of actual correct answers
 - **F1-Score:** Balance between precision and recall
 - **Confusion Matrix:** Heatmap showing actual vs. predicted classes
-

Summary Stats

- ☐ **Unique Careers:** Number of distinct roles in the dataset
 - ☐ **Total Questions:** Number of training questions
 - ☐ **Algorithm:** Support Vector Machine (SVM)
-

Example Prompts

You can try these queries:

- What is the job description for an AI Researcher?
 - A typical day for a Data Analyst?
 - What does a QA Engineer do?
-

Future Improvements

- Add feedback loop to improve prediction accuracy
 - Integrate external APIs (e.g., Coursera or LinkedIn Learning)
 - Include videos/interviews for each career
 - Enable voice input and multilingual support
 - Add user account system and personalization
-

Developer Info

- **Developed By:** Shoaib
- **Last Updated:** 2025
- **Framework:** Streamlit
- **License:** MIT