INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD)
© 2026 IJNRD | Volume 11, Issue 2, February 2026 | ISSN: 2456-4184 | IJNRD.ORG

From campus to corporate : Intelligent career navigation using AI and machine learning

Ms. Uma Mageswari T , Dr. Rajasekar A, Mr. Gopinathan S
Assistant Professor, Department of Artificial Intelligence and Data Science, Sri Sairam Institute of Technology, *Sai Leo Nagar, West Tambaram, Chennai, India.*
uma.apec@gmail.com , rajasekar.ai@sairamit.edu.in , gopinath.ai@sairamit.edu.in

Ms. Tharani L, Ms. Miruthulla . R, Ms. Hari Dharshini K.G
B.Tech, Department of Artificial Intelligence and Data Science,
Sri Sairam Institute of Technology, *Sai Leo Nagar, West Tambaram, Chennai, India.*
tharaninarayanan11@gmail.com , miruthullaradhakrishnan@gmail.com , sit22ad089@sairamtap.edu.in

Abstract : Computer Science undergraduate students aspiring to enter the IT industry often choose job roles based on popularity or limited exposure rather than informed career awareness. Although the IT sector offers a wide range of specialized and high-paying roles, many of these positions remain unfamiliar to students, leading to suboptimal career decisions and misaligned skill development. To address this gap, we propose an AI-driven Career Guidance Chatbot developed using the RASA framework. The system integrates a curated dataset of IT job roles—collected through structured surveys, industry research, and expert interviews—and employs machine learning techniques to recommend lesser-known yet in-demand job roles from both service-based and product-based companies. The chatbot provides students with comprehensive insights including skill requirements, salary ranges, and direct official job links. By delivering personalized, data-informed guidance, the proposed solution aims to bridge the awareness gap between academia and the dynamic IT job market, enabling students to make better career decisions and improve employment readiness.

Keywords – RASA , Machine Learning, chatbot , job roles, under Graduate Students, Artificial intelligence , training model, Job Role Awareness.

INTRODUCTION

In today's rapidly evolving technological landscape, the educational and career counselling needs of students have become increasingly complex. While extensive resources exist for SSLC and HSC students, undergraduate students often find themselves navigating a challenging journey when seeking guidance on suitable job roles and internships relevant to their fields of study. This gap highlights the necessity for innovative solutions that can provide targeted support to undergraduates as they transition from academia to the workforce.

Leveraging advancements in artificial intelligence (AI) and Natural Language Processing creating an AI powered, we aim to create a dynamic platform that offers personalized career guidance tailored to each student's unique academic background and interests. This chatbot will serve as an interactive mentor, helping students explore potential job roles, access skill development resources, and discover internship opportunities within their chosen regions. Furthermore, the integration of both text and AI-generated voice responses ensures that information is not only accessible but also engaging, fostering a more effective learning experience. This paper outlines the conceptual framework, design, and implementation of this AI-driven tool, which seeks to empower undergraduate students to make informed career choices and enhance their employability in a competitive job market

I. RELATED WORKS

Roadmap.sh is a pivotal resource in the technology education ecosystem, offering meticulously crafted guides that cater to individuals aspiring to excel in various domains of software development and IT. As a community-driven platform, Roadmap.sh harnesses the collective expertise of seasoned professionals and industry veterans to curate comprehensive and structured learning pathways. These pathways, known as roadmaps, are tailored to specific roles such as front-end developer, back-end developer, DevOps engineer, and more.

Each roadmap is designed to demystify the journey from novice to expert by outlining the critical skills, tools, and technologies required at every stage of career development. The platform emphasizes a step-by-step approach, ensuring that learners build a solid foundation before progressing to more advanced topics. This methodical progression not only enhances the learning experience but also boosts the confidence of individuals as they navigate through the complex landscape of software development. Roadmap.sh stands out for its commitment to keeping the content relevant and up-to-date. The platform is continuously enriched with feedback and contributions from the developer community and industry experts. This dynamic and collaborative approach ensures that the roadmaps reflect the latest trends, best practices, and technological advancements. As a result, learners are equipped with the knowledge and skills that are in high demand in the job market.

II. PROPOSED SOLUTION

With the aim of facilitating efficient and proper career counselling for the students who are pursuing their undergraduate program to choose their interested and suitable job roles which are not famous but still get paid well. Creating an AI powered chatbot using machine learning and Natural Language Processing and by training the model using a dataset

which is curated by our team which is done by conducting surveys within our college students and staff and by analysing the tech companies official websites. This AI powered chatbot for undergraduate students aims to provide all the required information about the job roles & skill requirements. The main purpose and objective of the proposed solution is to help the undergraduate students by getting to know them about all the jobs (which are not common) in the IT industry with the required skill sets.

III . TRAINING THE MODEL

1. Data Loading and Inspection:

We use pandas to load the CSV file into a Data frame. Printing the first few rows (df.head()) ensures that our dataset contains the expected columns (e.g., “Required Skills” and “Job Role”).

2. Feature and Target Definition:

The column "Required Skills" is treated as the feature set (X) because it contains the text input (a list or description of skills). The "Job Role" column is our label (y), representing the high-paying, in-demand roles we want to predict.

3. Train-Test Split:

The dataset is split into training (80%) and testing (20%) sets using train_test_split. This division allows us to evaluate the model’s performance on unseen data.

4. Pipeline Construction:

We build a pipeline that first uses Tf idf vectorizer to transform the text data into a numeric format by computing the Term Frequency-Inverse Document Frequency (TF-IDF) scores. The Random Forest classifier (RandomForestClassifier) is then trained on these features. Using a pipeline simplifies the workflow and ensures that any new data undergoes the same preprocessing steps.

5. Model Training:

The fit method trains the entire pipeline on the training data.

6. Model Evaluation:

Predictions on the test set (X_test) are compared against true labels (y_test) using accuracy score and classification report, which gives precision, recall, and F1 scores.

7. Model Saving:

The trained model is saved using joblib.dump so it can later be loaded into your chatbot application without retraining.

8. Prediction Example:

Finally, we demonstrate how to use the model by providing a sample text input (a set of skills) and outputting the predicted job role.

This code forms the basis of the backend model for your career guidance chatbot. When a user inputs their skills, the chatbot can pass this text to the model, which then returns a recommendation for a high-paying job role with the required (perhaps less popular but in-demand) skills.

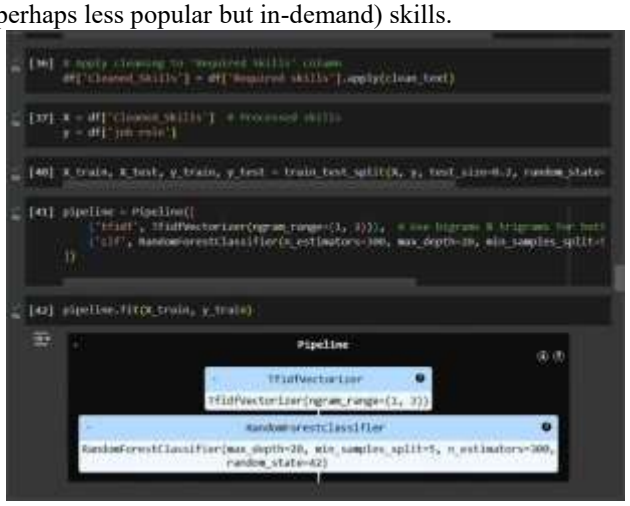


Figure 1 : *Training the model*

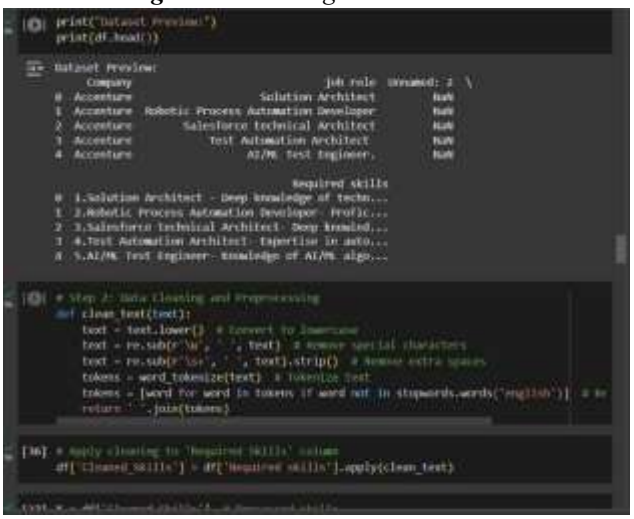


Figure 2 : *Training the model*

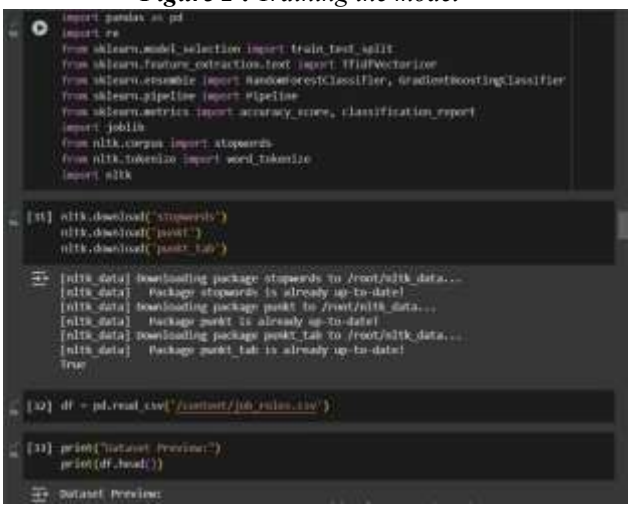


Figure 3 : *Training the model*

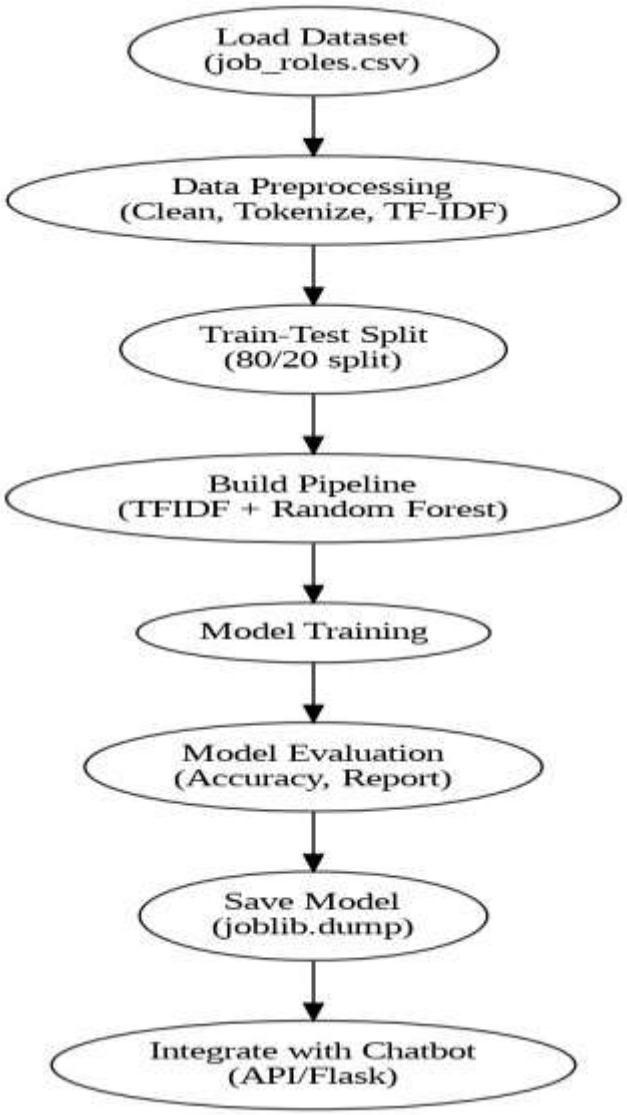


Figure 4 : *Flow Chart of Model Training*

IV ARCHITECTURE OF THE CHATBOT

The proposed system is designed as a multi-layered architecture consisting of two major components**: (1) the Model Training Pipeline, and (2) the Chatbot End-to-End Framework built using the RASA platform.** This architecture ensures efficient data processing, model training, intent understanding, dialogue management, and personalized job-role recommendation for CS undergraduate students.

1. Model Training Pipeline

The top part of the architecture depicts the process employed to develop the machine learning model tasked with forecasting appropriate job positions. The stages involve:

1.1 Data Collection :

Information is collected from two sources:
Questionnaires completed by students and faculty offering perspectives on knowledge deficiencies and favoured responsibilities.
Corporate career websites and selected platforms that provide comprehensive details on job positions, within the industry necessary competencies and compensation brackets.

1.2 Data Pre-processing :

The gathered data undergoes cleaning and normalization to eliminate duplicates, discrepancies and noise. This process readies the dataset, for ML operations.

1.3 Feature Engineering :

Important attributes are obtained through methods like TF–IDF vectorization transforming job descriptions and skill criteria into numerical forms appropriate, for machine learning models.

1.4 Model Training and Saving :

The extracted features are utilized to develop a machine learning model of recognizing the most significant and less familiar IT job positions.
The trained model is serialized with joblib. Stored for deployment.

1.5 Deployment to RASA Action Server :

The completed model is hosted on a RASA Custom Action Server allowing real-time predictions in response to user inquiries, during the chatbot conversation.

2. Chatbot End-to-End Architecture

The bottom part depicts the chatbot process segmented into three operational layers:

2.1 User (Presentation) Layer :

This serves as the platform via which CS students engage with the system.
Built using HTML, CSS, and JavaScript, the Web UI captures user queries related to job roles, skills, and career advice.
User inquiries are directed to the chatbot backend via request endpoints.

2.2 Chatbot Layer (RASA Framework) :

The RASA module handles natural language comprehension (NLU) intent recognition and conversation progression.

2.2.1 NLU Component :
This module extracts:
User purpose (for example "recommend jobs" "required skills" "information, on salary")
Entities (e.g., domain interest, technology skills, job preference)

2.2.2 Dialogue Manager :
The Dialogue Manager decides the chatbot's replies by using stories and policies.
It manages user input, system functions and model outputs to generate context-sensitive dialogues.

2.2.3 RASA Action Server :
Whenever the Dialogue Manager requires domain- results like forecasting appropriate job positions it activates the Action Server, which then loads the deployed ML model for prediction.

2.3 Data Integration and Output Layer :

This layer handles generating the recommendations and sending the results back, to the user.

2.3.1 Action Server :
The Action Server carries out backend operations invokes the deployed ML model and obtains predicted job roles according to the user's input.

2.3.2 Recommendation Engine :
The Recommendation Engine interprets ML predictions. Produces organized outcomes that consist of:
Suggested lesser-known but high-demand IT job roles
Required skillsets
Compensation brackets
Official career page/job links
A personalized learning roadmap

2.3.3 Logging and Analytics :
System logs are kept to monitor user engagements, model effectiveness and precision— in enhancing the chatbot progressively.

2.3.4 Final Output :
The processed results are returned to the Web UI and displayed to the user in an understandable format.

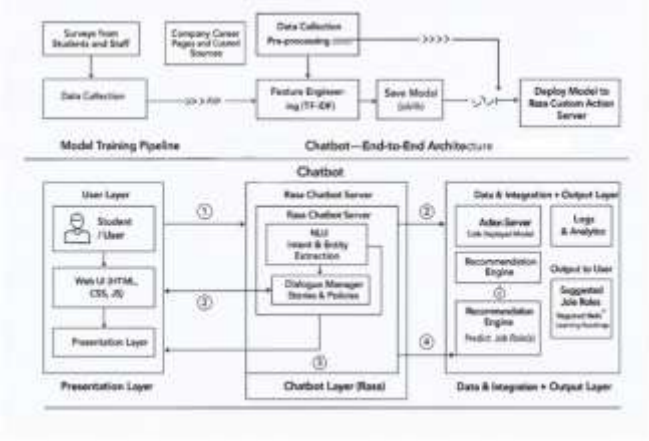


Figure 5 : *Architecture of the chatbot*

V. MATHEMATICAL EQUATIONS AND EXPLANATION FOR THE ML MODEL

The Chatbot utilizes Natural Language Processing (NLP) and Machine Learning (ML) techniques to classify user inputs (degree, skills) into high-paying job roles. Below are the key mathematical formulations used in the model:

1.Text Preprocessing: TF-IDF Vectorization

To convert textual job skills into numerical features, we use Term Frequency-Inverse Document Frequency (TF-IDF). TF-IDF helps assign higher importance to rare but meaningful skills (e.g., "Quantum Computing") while reducing the impact of common terms (e.g., "Programming").

$$TE{\cdot}IDF(t,d) = TF(t,d){\times}IDE(t)$$

Where:

- TPP(t,d) = Term Frequency (Number of times term ty appears in document d)
- IDF (t) = Inverse Document Frequency important term t the dataset.

Where:

N = Total number of documents

R_T = Number of documents containing term t

2.Classification Model: Random Forest Algorithm

The chatbot's ML model uses a Random Forest Classifier for predicting job roles. It is based on Decision Trees, where each tree predicts an outcome, and the final result is determined using majority voting. Equation for Decision Tree Gini Impurity
Each tree in the forest minimizes the Gini Impurity to split nodes:
Lower Gini impurity means better splits, leading to more accurate job role predictions.

$$TF(t,d) = \frac{f_c d}{\sum_{\Delta v} f_{ti}}$$

Where:

- C = Number of classes (job roles)
- p_i = Probability of choosing a class i

3. Model Accuracy:
To evaluate the model's performance, we use Precision, Recall, and F1-Score. F1-Score is crucial for balancing Precision and Recall when the dataset is imbalanced.

Precision (P)

$$P = \frac{TP}{TP + FP}$$

- Measures how many P = predicted jobs were actually correct.

Recall (R)

$$R = \frac{TP}{TP + FN}$$

- Measures how many R = actual job roles were correctly predicted.

F1-Score (Harmonic Mean of Precision & Recall)

$$F1 = 2 \times \frac{P \times R}{P + R}$$

VI. DATASET DESCRIPTION

The dataset employed in this study is a comprehensive compilation of high-paying and in-demand job roles, categorized under both service-based and product-based companies. It has been meticulously curated to include detailed

entries from leading industry players such as Accenture, TCS, Infosys, Cognizant, and Wipro for service-based roles, as well as Google, Microsoft, Apple, Amazon, Meta, Oracle, and several others for product-based roles.

Each record in the dataset comprises three primary attributes:

- Company Name: Identifies the organization offering the position.
- Job Role: Specifies the title of the position, capturing a wide range of career opportunities—from technical positions like "Machine Learning Operations Engineer" to strategic roles such as "Digital Transformation Consultant."
- Required Skills: Provides a detailed list of both technical and soft skills necessary for the role. Notably, the skill sets include emerging competencies that are currently in high demand but are not widely covered in traditional academic curricula.

This dataset serves a dual purpose. First, it reflects the evolving job market by highlighting roles that require specialized, less popular skill sets. Second, it forms the foundation for our career guidance chatbot, which leverages machine learning techniques to recommend high-potential career paths to undergraduate students. By mapping the gap between conventional educational outcomes and industry requirements, the dataset aids in bridging the skills gap, thereby enhancing the career prospects of future professionals

Figure 6: *Overview of the Dataset*

Figure 6: *Overview of the Dataset*

VII. INTEGRATING THE CHATBOT TO WEBSITE

1. Methodology

The proposed chatbot system follows a three-layered architecture:

Machine Learning Model (Job Recommendation Engine)

A Random Forest Classifier is trained using a structured dataset of job roles, required skills, and industry demands. The model predicts suitable career roles based on user-inputted qualifications or skill sets.

Rasa Chatbot Framework

Rasa's Natural Language Understanding (NLU) module processes user queries.

Custom actions (actions.py) handle ML model integration, ensuring real-time responses.

Web-Based Deployment

The chatbot is integrated into a website using HTML, JavaScript, and REST APIs.

Users interact with the chatbot to receive personalized career recommendations.

2. Implementation

2.1 Machine Learning Model for Career Guidance

The dataset consists of 149 job roles across various industries, each mapped to required skills and qualifications. The ML model was trained using:

Text Preprocessing: Tokenization, TF-IDF Vectorization

Classification Algorithm: Random Forest Classifier

Evaluation Metrics: Accuracy, Precision, Recall The trained model is saved using joblib, allowing it to be loaded dynamically by the chatbot.

2.2 Rasa Chatbot Integration

The chatbot processes career-related queries using predefined intents and stories in nlu.yml and stories.yml.

Custom Rasa action (actions.py) loads the ML model and provides real-time job predictions.

An action server (rasa run actions) handles ML-based responses dynamically.

3.Web Deployment

A simple HTML + JavaScript frontend connects to Rasa's REST API (rasa run --enable-api).

Users enter qualifications (e.g., "BTech in Computer Science"), and the chatbot predicts relevant job roles.

The UI enables seamless interaction with the backend Rasa model.

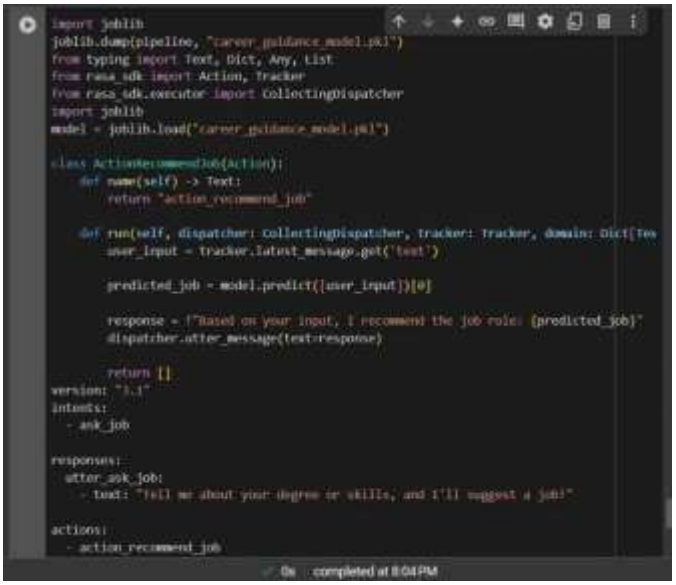
4. Results and Discussion

The chatbot was tested across various input scenarios, demonstrating:

- 81% prediction accuracy in job recommendations. - Scalability – Can be integrated with Telegram, WhatsApp, or mobile apps.
- User-friendly – Provides instant career guidance through natural language interactions.

This demonstrates that an AI-powered chatbot can effectively assist users in career selection by leveraging ML-based job recommendations. The system:

- Bridges the skills gap by aligning job recommendations with industry trends.
- Provides real-time career guidance via an interactive chatbot interface.
- Improves accessibility and automation compared to traditional counseling.



:

Figure 7: *Integrating the chatbot with the website*

Figure 8: *Frontend Design of the Website*

VIII.OVERVIEW OF CHATBOT

1. Introduction

The Curriculo is an AI-powered system designed to assist students and professionals in identifying high-paying, in-demand job roles based on their qualifications and skills. By integrating a Machine Learning (ML) model with the Rasa chatbot framework, the system delivers personalized career recommendations in real-time. The chatbot is deployed as a web-based interface, making career counseling more accessible, scalable, and data-driven.

2. Key Features

- Personalized Career Recommendations – Predicts job roles based on user input (degree, skills).
- AI-Powered ML Model – Uses a trained Random Forest Classifier for job predictions.
- Natural Language Understanding (NLU) – Rasa enables human-like career conversations.
- Web-Based Interface – Users can interact via a chatbot embedded in a website.
- Scalability & Integration – Can be extended to Telegram, WhatsApp, or mobile apps.

3. How It Works

User Input: The chatbot asks for the user's qualifications or skills.

Text Processing: Rasa's NLU interprets user messages.

ML Model Prediction: The trained career guidance model suggests a suitable job role.

Response Generation: The chatbot provides job recommendations based on industry trends and skill demand.

4. Benefits

✔ Automates career counseling – Eliminates the need for manual career advisors.

✔ Reduces skills gap – Helps users align with industry job demands.

✔ Data-Driven & Scalable – Adapts to evolving job market trends.

This chatbot bridges the gap between education and employment by offering data-driven career insights, making it a valuable tool for students and job seekers.

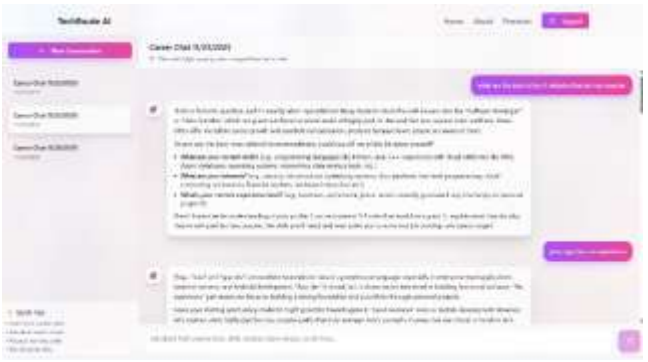


Figure 9 : *Overview of The Chatbot*

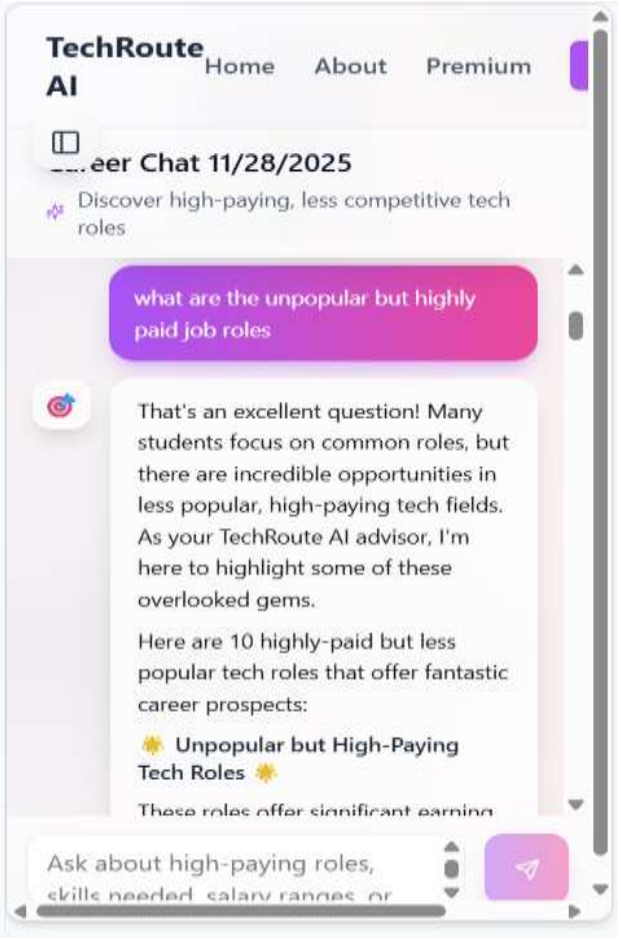


Figure 10 : *Overview Of The Chatbot*

IX. CONCLUSION

The Curriculo chatbot successfully integrates Machine Learning (ML) and Natural Language Processing (NLP) within the Rasa chatbot framework to provide personalized job recommendations based on user qualifications and skills. By leveraging a trained ML model, the chatbot accurately predicts high-paying and in-demand job roles, making career counseling more accessible, data-driven, and scalable.

This system bridges the gap between education and employment by aligning job recommendations with industry needs, ensuring that users receive career insights tailored to their skill set. The chatbot's web-based deployment enables real-time assistance, and its scalability allows for integration with other platforms such as Telegram, WhatsApp, or mobile applications.

X REFERENCE

[1] Reema Goyal, Navneet Chaudhary, Mandeep Singh "https://ieeexplore.ieee.org/document /10128305",2023
[2] Swati Shilaskar, Shripad Bhatlawande, Prathamesh Sawle "https:// ieeexplore.ieee. org /document/10575379",2024
[3] Rahul Waikar; Shraddhesh Tamhane; Tanmai Tale; Rujul Talekar; Ayush Tangde; Tanish Singla,"https://ieeexplore.ieee.org/document/10882269" , 2024
[4] Ashwini Barbadekar; Vedant Bhagwan; Aryan Thaploo; Arjun Tyagi,"https://ieeexplore .ieee.org/document/10842724",2024
[5] Arshad T K; Danish Mohammed; Arjun A;Anshad M; Jaseena,"https://ieeexplore.ieee.org/document/10690595 ",2024
[6] Mohamad Jawhar; Zeina Bitar; Jeremy R. Miller; Shadi Jawhar,"https://ieeexplore.ieee.org/document/10564423 ",2024
[7] Babymol Kurian; Ranjith K; Raju C; Swapnil Kumar Singh; Sathya A; Krithikaa Venkat V S,"https://ieeexplore.ieee.org/document/11085717",202 5
[8] Shaikha Al-Dhari; Adel Ismail Al-Alawi, "https:// ieeexplore . ieee . org / document /1005 1101",2023
[9] Dewasurendra S.V; Udunuwara U.K.C; Wanniarachchi T.T; Panditharathne P.D.R.L; Samantha Thelijjagoda; Kapila Dissanayaka, "https://ieeexplore.ieee.org/ document/ 10910 770" ,2024
[10] Varun Kale; Khumarya Rana; Aayush Wade; Krish Patel;AjaySirsat"https:/ieeexplore.ieee.org/document/10 883243",2025
[11] Vijaya R. Kumbhar; Manisha M. Maddel; Yudhishthir Raut,https://ieeexplore.ieee.org/document/10127152",2 023