Conversational recommended systems for educational paths

Sahil Singh

MSC DATA SCIENCE & ANALYTICS THESIS

Department of Computer Science
Maynooth University, Co. Kildare, Ireland.

A thesis submitted in partial fulfilment of the requirements for the Masters in Data Science and Analytics

Supervisor: Fabiano Pallonetto

Date: August 07, 2023

Declaration

I hereby certify that this material, which I now submit for assessment on the program of study as part of MSc – Data Science and Analytics qualification, is entirely my own work and has not been taken from the work of others - save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: Sahil Singh Date: 07-08-2022

Acknowledgements

Completing this project has been an enriching journey, made possible through the support, guidance, and contributions of numerous individuals and resources. I would like to thank my supervisor Professor Fabiano Pallonetto for being my constant guidance throughout the entire duration of this thesis. I would also like to thank my father for being my constant support system. I would like to acknowledge the contributions of the broader research community, whose published work has provided the theoretical basis for the thesis. Their ideas and innovations have guided our approach and enriched our understanding.

Contents

Abstract

MSc in Data Science and Analytics

Conversational recommended systems for educational paths

by Sahil Singh

In the context of rapidly changing educational and career trajectories, the need for individual support becomes increasingly paramount. This project introduces a conversational recommendation system for educational pathways, a pioneering solution that seamlessly blends machine learning, natural language processing, and data analytics to generate public recommendations. personalized work with tailored courses. This initiative aims to help users effectively align their career aspirations by aligning their skills and interests with labour market needs.

The focus of this project is its interactive chatbot interface. Users engage in dynamic conversations with the chatbot, communicating information such as their job title preference, location, field of study, and skillset. Using Python and advanced machine learning techniques, the system meticulously processes and analyses input data to create an individual career path for each user.

At the heart of this system is a powerful machine learning model that deciphers user preferences and performs job searches accordingly. The chatbot uses web scraping techniques to extract job postings that match the user's criteria, ensuring accurate, real-time recommendations. Additionally, this template suggests potential job titles based on the user's area of research and skills, broadening their horizons and driving in-depth exploration.

Additionally, this project emphasizes integrating education into the equation. In addition to job recommendations, the system assesses users' skills and offers courses that can bridge the gap between their current skills and desired job prerequisites. The course recommendation aspect uses the most advanced algorithm to look at a set of course data and skill match, providing users with curated lists of courses that match their goals their profession.

Conversational recommendation system for educational pathways that excel in personalization and adaptability. By understanding users' current aspirations and capabilities, this project offers a roadmap that seamlessly connects their ambitions with concrete steps. At a time when career transitions are increasingly common and the need for lifelong learning is paramount, this system emerges as an invaluable companion, deftly guiding users to the point of intersection. between professional success and personal satisfaction. Through an innovative combination of machine

learning and human interaction, this project represents a significant step forward in educational and career guidance, promoting lifelong learning, and forging a career trajectory that has meaning.

Chapter 1

Introduction and Motivation

1.1 Introduction

In an era characterized by rapid technological advancements and dynamic changes in the labour market, traditional career planning methods are facing unprecedented challenges. The convergence of globalization, automation, and changing skill requirements has created a complex and ever-changing career landscape. Navigating this terrain requires more than conventional strategies; it requires a tailored, data-driven approach to guide individuals towards meaningful and fulfilling careers. "Conversational recommendation system for educational pathways:
The project "Connecting Work and Courses" emerges as an innovative solution to the intersection of technology and career development, offering a new way to connect individuals with job opportunities and appropriate educational trajectory.

Despite the multitude of resources available for career exploration, individuals often have difficulty identifying the most appropriate career roles that match their unique skills, interests, and aspirations. . Traditional methods, such as job search platforms and standardized career guidance, tend to give generic advice that lacks the necessary personalization. As a result, individuals can make ill-informed career choices, leading to dissatisfaction, unfulfilled potential, and missed opportunities for career growth.

The Conversational Recommender System for Educational Paths project aims to address these challenges by introducing an intelligent and interactive chatbot-based platform. The platform engages users in natural language conversations to gather important information about their job interests, geographic considerations, education, and skills. Using advanced natural language processing techniques and machine learning algorithms, the project aims to decode user input and produce personalized career recommendations.

What is a chatbot?

A chatbot is a computer program or artificial intelligence (AI) application designed to simulate human conversation through text or voice interaction. Chatbots are used to facilitate human-computer communication in a way that feels natural and conversational. They can be integrated into other messaging platforms, websites, applications or interfaces to provide automated responses, information and support to users.

Chatbots are typically designed to understand and respond to user queries, requests, or commands. They use natural language processing (NLP) techniques to understand and interpret user comments. Depending on their complexity, chatbots can perform a variety of tasks, from providing basic information to performing more advanced actions, such as making reservations, responding to customer requests. or even guide the user through complex processes.

There are two main types of chatbots:

Rule-based chatbots:

These chatbots work according to predefined rules and patterns. They follow a decision tree structure, where they associate user input with predefined responses or actions. Rule-based chatbots are best suited for situations with limited and predictable interactions.

AI-powered chatbots:

These chatbots use artificial intelligence and machine learning techniques, especially natural language processing, to understand and respond to user input in a more human-like way. They can learn from interactions and improve over time, allowing them to process more queries and provide more personalized feedback.

Chatbots are used in many different industries and sectors, including customer service, e-commerce, healthcare, finance, and more. They offer benefits such as 24/7 availability, fast response times, scalability, and cost-effectiveness in handling frequent and repetitive tasks. However, their effectiveness depends on the quality of the underlying technology, the accuracy of their understanding of natural language, and the relevance of their answers to user queries. .

In addition, the project recognizes the essential role of education in shaping successful careers. In addition to job recommendations, the system also incorporates an educational aspect by recommending courses that can fill skill gaps and enhance users' qualifications for their desired roles. By seamlessly combining career opportunities with tailored learning pathways, the project offers a holistic approach to career development.

About web scraping:

Web scraping, an integral part of data collection in the digital age, involves the automatic extraction of information from web pages. It serves as a conversion tool that allows individuals and organizations to collect data on a variety of topics for different purposes. Using scripts or automated tools, web scrapers navigate through web pages, mimicking human interaction to access, retrieve, and organize data. This process is especially useful for aggregating information from multiple sources quickly and efficiently.

At the heart of web scraping is the interaction between web servers and the scraping engine. The process begins with the tool sending a request to a specific URL, similar to how a browser accesses a website. In response, the server transmits the requested HTML content, encapsulating the visual and structural elements of the web page. The next step is to parse this HTML content, from which structured data is extracted, providing the basis for data extraction.

Through various techniques, web scraping can extract targeted data elements from the analyzed HTML code. These elements cover a wide range of shapes, including text, images, links, tables, and more. The versatility of web scraping is marked by its applicability in all fields. It can extract product details from e-commerce websites, aggregate articles, collect financial data, and even curate job postings from career portals. The extracted data is then stored in an optional format, from conventional formats like CSV and Excel to a more structured database.

Ethical considerations are an essential aspect of web scraping. Compliance with the terms of use and site guidelines is required to ensure that the scraping process meets legal and ethical standards. In addition, responsible data collection methods prioritize the prevention of server overload, which can affect the site's performance and accessibility to other users. Adhering to these principles not only maintains ethical integrity, but also helps foster positive relationships within the digital ecosystem.

In a world full of data, web scraping presents unprecedented opportunities. It enables researchers, businesses, journalists, and data enthusiasts to collect and act on information that supports

analysis, decision making, and insights. However, its potential also underscores the need for responsible and respectful use. By understanding the nuances of web scraping, individuals can navigate the digital landscape with integrity, opening up a wide range of possibilities while maintaining ethical boundaries.
Here the Project Conversational Recommendation Systems for Educational Pathways offers an innovative solution to the challenges of modern career planning. Through intelligent chatbot interface and data-driven algorithms, the project strives to connect individuals with the right job opportunities and educational resources, guiding them towards a fulfilling and fulfilling future. Career Development. By understanding and adapting to each user's unique needs, the project aims to redefine how individuals navigate their careers in the dynamic and ever-changing job market.

1.2 Problem introduction

In the modern educational and career exploration landscape, individuals often face the daunting challenge of aligning their aspirations, skills, and opportunities. Traditional approaches to career guidance, characterized by standardized assessments and generic advice, often fail to capture many aspects of individual preferences and changing labour market needs. The divergence between conventional methods and the dynamic realities of the professional world has raised a pressing issue – the need for a personalized, technology-enabled solution that helps bridge the gap between career aspirations and possible journeys.

The current problem revolves around the lack of a comprehensive system that integrates job offers with skills upskilling opportunities. As the labour market changes rapidly and industries seek candidates with specific skills, individuals need guidance that not only directs them to the right job title, but empowers them to thrive. develop the skills necessary to obtain these positions. Current career planning tools often focus only on job offers, ignoring the important link between defined job titles and the skills needed to excel in roles. there.

In addition, individuals often have difficulty accessing relevant and up-to-date job postings from a variety of sources. The proliferation of online job platforms has created a fragmented landscape in which manually browsing multiple websites for job openings can be time consuming and overwhelming. This problem is further complicated by the difficulty of matching individual skills with those required by the vacancy. Traditional keyword-based searches can yield inaccurate results and users may miss opportunities due to lack of skill alignment. Therefore, raising the issue around the inadequacy of current career guidance methods aims to provide a holistic and individualized approach. It's clear that an innovative solution that leverages technology, data analytics, and machine learning is needed to seamlessly integrate job offers with upskilling opportunities. This solution will ease the difficulty of accessing relevant job postings, accurately match the user's skills to the job requirement, and provide possible steps to skill development.

Solving this problem has the potential to revolutionize the way individuals navigate their career trajectory. By creating a unified platform that not only recommends suitable job titles, but also guides users to acquire the necessary skills, this project aims to bridge the gap between aspirations and achievements. . Later sections of the project delve deeper into the proposed solution to this problem and describe the methods used to develop the "Conversational Recommendation System for Educational Pathways".

1.3 Solution

The Conversational Recommendation System for Education Pathway project aims to fill this gap by providing an intelligent and interactive platform that engages users in meaningful conversations about their career aspirations. At the heart of this system is a sophisticated chatbot interface that converses with the user, gathering essential information such as job title preferences, geographic considerations, field of study, and available skills. . Through natural language processing and machine learning, this input is analyzed to make suitable recommendations.

The project uses web scraping techniques to track and extract real-time job postings that match user profiles. By identifying opportunities that match an individual's interests and qualifications, the system allows users to explore and pursue new careers with more confidence. Furthermore, the project harmonizes the pedagogical aspect of career development by recommending suitable courses aimed at improving the user's skills, thus bridging the gap between current and current skills. prerequisites of the positions being sought.

1.4 Overview of this dissertation

The thesis is structured to provide a comprehensive understanding of the "Conversational Recommendation System for Educational Pathways". It begins with an exploration of current challenges in modern career planning, highlighting the limitations of traditional approaches. The following sections dive into the technical background of the project, shedding light on the architecture of the chatbot interface, the combination of machine learning algorithms, and the intricacies of web scraping methods.

Furthermore, the thesis delves into data analysis techniques used to match user profiles with relevant job openings. The integration of skill gap-based course recommendations is discussed in detail, with a focus on algorithmic approaches that support this important aspect of the system. In addition, the thesis elaborates on ethical considerations and user security measures integrated into the project.

In summary, the Conversational Recommendation Systems for Educational Pathway project represents a paradigm shift in career counseling and education. By leveraging the synergy between machine learning and human interaction, the project provides a holistic approach to career planning. This thesis attempts to comprehensively elucidate the mechanisms of innovation and impact of this project, highlighting its potential to revolutionize the way individuals chart their career paths in the industry in an increasingly complex and interconnected world.

Chapter 2

Background Literature Review

2.1 Literature Review

A. Gupta and D. Garg, "Applying data mining techniques in job recommender system for considering candidate job preferences," 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Delhi, India, 2014, pp. 1458-1465, doi: 10.1109/ICACCI.2014.6968361.

A. Gupta and D. Garg in their study showed the process of suggesting job recommendations which involve analyzing a candidate's profile to find matches and considering their individual job preferences. Initially, they extracted rules that predict the typical preferences of various user groups. These rules guided the job recommendations for the target candidate, combining content-based matching with the candidate's preferences. These preferences were gathered from mined rules or the candidate's job history. Their approach focused on both candidate preferences and content-based profile matching. This not only enhanced prediction accuracy but also assisted candidates unsure of their job goals by utilizing group preferences. However, for candidates with unique career paths, the system adjusts by emphasizing their recent job preferences and tailoring recommendations accordingly. By tracking the candidate's current job preferences, they were able to prioritize pertinent jobs over irrelevant ones identified through content-based matching. This strategy ensured that only relevant job options are presented to the candidate.

L. G. Rodriguez and E. P. Chavez, "Feature Selection for Job Matching Application using Profile Matching Model," 2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS), Singapore, 2019, pp. 263-266, doi: 10.1109/CCOMS.2019.8821682.

The research by L.G. Rodriguez and E.P. Chavez introduced an approach designed for an agency to extract pertinent details from resumes and assess them based on various characteristics. Once these attributes were identified, the system employed a clustering algorithm to align job seekers' profiles with the prerequisites outlined in job postings from potential employers. This facilitates the agency in identifying suitable candidates for specific positions across different companies, leading to more wellinformed decisions. The attributes were derived from the resumes of 2,283 job seekers, as well as job requirement data sourced from the Public Employment Service Office (PESO) in Pangasinan, Philippines. The collected data were subjected to data mining using an analytical software called WEKA, and the results are statistically presented through rankings. Once the attributes were pinpointed, the proposed system proceeded to utilize a clustering algorithm. This algorithm was employed to match the profiles of job seekers with the criteria laid out in job advertisements by prospective employers.

I. Paparrizos, B. B. Cambazoglu and A. Gionis, "Machine learned Job Recommendation", ***Proceedings of the fifth ACM conference on Recommender systems RecSys '11 ACM,*** **pp. 325-328, 23-oct. 2011.**

A model was developed by I. Paparrizos, B. B. Cambazoglu, and A. Gionis to tackle the challenge of suggesting appropriate employment opportunities for individuals in search of new jobs. The approach framed this recommendation issue as a supervised machine-learning task. The methodology leveraged historical job transitions and associated data from both employees and institutions to anticipate the subsequent job move of an employee. A machine learning model was trained using a substantial volume of job transition records gathered from publicly accessible employee profiles on the internet. In total, four distinct configurations were established. For each configuration, a basic baseline was taken into account, which consistently predicts the most prevalent class within the training data.

Carțiș, AI., Suciu, D.M. (2020). Chatbots as a Job Candidate Evaluation Tool. In: Debruyne, C., ***et al.*** **On the Move to Meaningful Internet Systems: OTM 2019 Workshops. OTM 2019. Lecture Notes in Computer Science(), vol 11878. Springer, Cham. https://doi.org/10.1007/978-3-030-40907-4_19**

Carțiș and Suciu introduced an innovative approach involving an intelligent chatbot to conduct screening interviews. This system offers a more interactive experience for users (job candidates) compared to mundane web forms, simulating a conversation with a real individual. Concurrently, the chatbot assesses user-provided information and assigns scores using a sentiment analysis algorithm integrated with the IBM Watson Personality Insights service. An enhancement proposal entails the customization of result weights. In job applications, there exist general prerequisites, but certain soft skills could be deemed essential skills, such as effective communication for a managerial role. Enabling the adjustability of result weights empowers companies to align outcomes with their specific requirements, ensuring they receive the most pertinent results.

2.2 Tools and other systems

Tool	Description
import requests	The Requests Library is a Python module that allows HTTP requests to be sent to a specified URL. It allows you to retrieve data from web pages, interact with APIs, and perform various web-related tasks. In the context of code, it can be used to send HTTP GET requests to web pages, which can then retrieve or process information.
BeautifulSoup	BeautifulSoup is a Python library that makes web scraping easy. It parses HTML and XML documents, making it easy to extract specific information from web pages. With BeautifulSoup, you can navigate and manipulate HTML elements, extracting text, attributes, and other data. This library is often used in conjunction with the Query library to extract structured data from web pages.
pandas	pandas is a popular data manipulation and analysis library in Python. It provides data structures like DataFrames that allow you to store and manipulate tabular data. By importing it as pd, you can use its functions to easily read, process, and manipulate datasets. In the context of code, it can be used to manage and manipulate data related to courses and their attributes.
TfidfVectorizer	TfidfVectorizer is a component of the sklearn (scikit-learning) library, which is a comprehensive machine learning library for Python. TfidfVectorizer is specifically used to preprocess text data in

	natural language processing (NLP) tasks. It converts a set of text documents into a TF-IDF feature matrix. TF-IDF stands for Term Inverse Document Frequency, a technique that assigns weights to words based on their importance in the document relative to the corpus. This matrix is often used for machine learning tasks involving textual data.
cosine_similarity	cosine_similarity is another function in sklearn library. It is a metric used to measure the cosine of the angle between two non-zero vectors. In the context of the code, it can be used to calculate the cosine similarity between the TF-IDF vectors of user skills and related skills of the courses. This same measure helps identify courses that are closely related to the user's skills and interests.

Chapter 3

Methodologies

3.1 Procedure

The advancement of the "Conversational Recommender Framework for Instructive Ways" venture included a precise handle including different stages, counting information collection, preprocessing, demonstrate advancement, assessment, and integration into a user-friendly chatbot interface. This area traces the techniques utilized in each stage, explaining the complexities of the project's usage.

3.2 Implementation

3.2.1 Data Collection

Coursera data from Kaggle:

For this project, the data collection phase included gathering information from the Coursera platform through a publicly available dataset on Kaggle. The dataset provides a wealth of information about the different courses offered on Coursera, including the course name, the university or institution offering the course, the course URL, and most importantly, the skills involved. to each course. . This dataset acts as a fundamental resource for mapping skills for specific courses, an essential part of the project's recommendation system.

The dataset serves as a starting point, providing insights into the types of skills the different courses aim to develop. This information forms the basis for the project's pedagogical recommendations, as the user's skill profile can match the skills highlighted in different courses. While datasets provide a solid foundation, it is essential to ensure accuracy and alignment with project objectives.

Web Scraping LinkedIn for Job Data:

LinkedIn web scraping is an essential aspect of the project, allowing real-time and relevant work data to be collected to feed the recommendation engine. LinkedIn, as a leading professional networking platform, hosts a wide variety of job postings from different industries and locations. The web scraping process involves extracting job titles, locations, and other relevant details from LinkedIn job listings.

By leveraging web scraping techniques, the project exploited the dynamic labor market landscape, ensuring users received up-to-date and contextual job offers. This process requires a meticulous approach, including managing potential challenges such as changing site structure, potential access restrictions, and ethical considerations associated with web scraping.

Work data collected from LinkedIn forms an important part of the project's recommendation algorithm. It makes it easy to identify job titles that match a user's interests, locations, and fields of study. This data, combined with the Coursera dataset, allowed the project to bridge the gap between career aspirations and possible educational pathways.

In short, the combination of Coursera data collection from Kaggle and web job listings collection from LinkedIn provided a comprehensive foundation for the project's referral system. The Coursera dataset enriches the understanding of skills related to different courses, while LinkedIn job data enables real-time and relevant job recommendations. The aggregation of these data sources allows the project to guide users through their unique educational and career journey.

3.2.2 Performance Metrics

To calculate the effectiveness of the recommender system, performance measures were used. Cosine similarity, which is a standard measure in text analysis, is used to quantify the match between user skills and course descriptions. The performance of the recommender system is assessed by accuracy, recall, and F1 scores, providing insights into the system's ability to accurately predict relevant courses based on information. user input.

3.2.3 Model Evaluation

The model review process involves a series of steps to ensure the quality and relevance of the proposed courses. Initially, user input skills are compared with those associated with each course using a cosine similarity measurement. Cosine likeness may be a concept commonly utilized in machine learning and normal dialect preparing to degree the closeness between two non-zero vectors. It's especially valuable when managing with high-dimensional information, such as content information, where conventional Euclidean remove might not give exact measures of closeness.

In machine learning, cosine closeness evaluates the cosine of the point between two vectors. The esteem of cosine closeness ranges from -1 to 1, where:

- 1 indicates that the two vectors have the same direction **(absolutely similar).**
- 0 indicates that the vectors are orthogonal **(not similar).**
- -1 indicates that the vectors are in opposite directions **(maximum difference).**

cosine_similarity(A, B) = (A . B) / (||A|| * ||B||)

The cosine similarity is 0.4 for this project as the course dataset is limited.

```
def recommend_courses_for_job(job_title, user_skills):
    relevant_skills = ' '.join(job_title)
    user_skills += ' ' + relevant_skills

    tfidf_vectorizer = TfidfVectorizer()
    tfidf_matrix = tfidf_vectorizer.fit_transform(dataset["Skills"])
    user_tfidf = tfidf_vectorizer.transform([user_skills])

    similarities = cosine_similarity(user_tfidf, tfidf_matrix)

    matching_courses = []

    for index, similarity in enumerate(similarities[0]):
        if similarity > 0.4:
            matching_courses.append((dataset.loc[index, "Course Name"],
                                     dataset.loc[index, "University"],
                                     dataset.loc[index, "Course URL"],
                                     similarity))

    matching_courses.sort(key=lambda x: x[3], reverse=True)

    if matching_courses:
        print("Matching courses:")
        for course_name, university, course_url, similarity in matching_courses:
            print(f"Course Name: {course_name}")
            print(f"University: {university}")
            print(f"Course URL: {course_url}")
            print(f"Cosine Similarity: {similarity:.4f}")
            print("=" * 30)
    else:
        print("No matching courses found.")
```

The **recommend_courses_for_skills()** function takes the user's skills as input and recommends relevant courses based on those skills.

- **user_tfidf = tfidf_vectorizer.transform([user_skills])**: This line transforms the user's skills into a TF-IDF vector using the same TF-IDF vectorizer that was used to process the course dataset. This vector represents the user's skills in the same format as the course skill vectors.
- **similarities = cosine_similarity(user_tfidf, tfidf_matrix)**: This line calculates the cosine similarity between the user's skill vector and the skill vectors of all the courses in the dataset.

The result is a similarity matrix where each entry represents the similarity between the user's skills and a course's skills.

- **matching_courses = []**: This list will store the courses that match the user's skills.
- The **for** loop iterates through the similarity matrix. If the similarity between the user's skills and a course's skills is above a certain threshold (0.4 in this case), the course is considered a matching course. The course's name, university, URL, and similarity score are stored in the **matching_courses** list.
- **matching_courses.sort(key=lambda x: x[3], reverse=True)**: The matching courses are sorted based on their similarity scores in descending order, so that the most relevant courses appear first.
- Finally, the recommended courses are printed, including the course name, university, URL, and cosine similarity score. If no matching courses are found, a message is displayed indicating so.

This function calculates the similarity between the user's skills and the skills associated with each course in the dataset. Courses with higher similarity scores are considered more relevant to the user's skills.

An accurate measure that measures the ratio of suggested relevant courses to the total number of suggested courses. High accuracy indicates that a significant portion of the recommended courses actually match the user's skills. Recall, on the other hand, quantified the ratio of suggested courses to the total number of suitable courses available. A high recall value indicates that the system has effectively identified and recommended the most suitable courses. The F1 score, which balances accuracy and recall, provides a balanced assessment of system performance. A higher F1 score indicates a better trade-off between accuracy and recall, demonstrating the system's ability to provide accurate and comprehensive course recommendations.

3.2.4 Integration in Chatbot interface

The highlight of this project is the seamless integration of the recommender system into the user-friendly chatbot interface. Chatbots leverage natural language processing (NLP) to engage users in conversational interactions. Using the nltk library, the chatbot processed user input, identified key terms, and seamlessly integrated them into the recommendation algorithm. Chatbots are designed to ask users for information such as job interests, location, field of study, and skills.

```python
def main():
    print("Welcome to the recommendation system!")
    interests = input("Enter your job interests: ")
    location = input("Where are you located? ")
    field_of_study = input("What is your field of study? ")
    skills = input("Please enter your skills (comma-separated): ")

    recommended_jobs = perform_job_search(interests, location, field_of_study, skills)
    job_titles = [job[0] for job in recommended_jobs]
```

The **main()** function serves as the entry point of the recommendation system. It begins with a welcoming message to provide a user-friendly interface for interacting with the system.

- **interests = input("Enter your job interests: ")**: This line prompts the user to input their job interests. The entered text is stored in the **interests** variable, which will be used as a search keyword.
- **location = input("Where are you located? ")**: Similarly, this line prompts the user to input their location. The provided location will help narrow down job search results.
- **field_of_study = input("What is your field of study? ")**: Here, the user is prompted to input their field of study, which assists in tailoring the job search to their educational background.
- **skills = input("Please enter your skills (comma-separated): ")**: The user is asked to input their skills, separated by commas. These skills are crucial for generating relevant job recommendations.
- **recommended_jobs = perform_job_search(interests, location, field_of_study, skills)**: This line calls the **perform_job_search()** function with the collected user inputs. The function searches for jobs based on the provided criteria and returns a list of recommended job titles along with their locations.
- **job_titles = [job[0] for job in recommended_jobs]**: This line extracts the job titles from the list of recommended jobs. These job titles will be used to present options to the user for further interaction.
- **if __name__ == "__main__":** block ensures that the **main()** function is executed only if the script is run directly (not imported as a module).

This structure encapsulates the user interaction and job search process, providing a user-friendly experience when using the recommendation system.

These entries are then matched to relevant job titles using LinkedIn job listing web scraping techniques. Suggested job titles have been presented to the user, the user can select a specific job that interests them.

```
Welcome to the recommendation system!
Enter your job interests: Data Scientist
Where are you located? London
What is your field of study? Computer Science
Please enter your skills (comma-separated): Python
Recommended jobs:
1. Data Scientist / Machine learning engineer
   Location: London, England, United Kingdom

2. Data Scientist (w/ ML Engineering)
   Location: London, England, United Kingdom

3. Data Scientist
   Location: London, England, United Kingdom

4. Machine Learning Engineer - London
   Location: London, England, United Kingdom

5. Software Developer (Graduate Role)
   Location: London, England, United Kingdom

6. Data Scientist
   Location: London, England, United Kingdom

7. Health Data Scientist
   Location: London, England, United Kingdom
```

Once the user's career interests are established, the system recommends based on the selected job title and the user's skills to recommend suitable courses. Chatbots provide users with a curated list of courses to fill skill gaps and improve proficiency.

```
23. Data Scientist
   Location: London, England, United Kingdom

24. Data Scientist
   Location: London, England, United Kingdom

25. AI Engineer
   Location: London, England, United Kingdom

Enter the number of the job you are interested in: 24
Matching courses:
Course Name: Pandas Python Library for Beginners in Data Science
University: Coursera Project Network
Course URL: https://www.coursera.org/learn/pandas-python-library-beginners-data-science
Cosine Similarity: 0.2476
==============================
Course Name: Understanding and Visualizing Data with Python
University: University of Michigan
Course URL: https://www.coursera.org/learn/understanding-visualization-data
Cosine Similarity: 0.2385
==============================
Course Name: Getting Your Film off the Ground
University: Michigan State University
Course URL: https://www.coursera.org/learn/film-off-ground
Cosine Similarity: 0.2373
==============================
Course Name: Basic Data Processing and Visualization
University: University of California San Diego
Course URL: https://www.coursera.org/learn/basic-data-processing-visualization-python
Cosine Similarity: 0.2362
==============================
```

In conclusion, the methodologies encompassed data collection from a Coursera dataset, implementation of performance metrics for evaluation, and the integration of the recommendation system into an interactive chatbot interface. This holistic approach ensured the accuracy and relevance of recommended courses while enhancing user engagement through conversational interaction. The project's methodologies serve as a foundation for creating a user-centric and effective career guidance tool that seamlessly aligns education with emerging job opportunities.

3.2.5 Overall working of the chatbot

The Course and Job Referral Chatbot is a sophisticated system designed to provide personalized career guidance by recommending relevant courses and job opportunities based on a person's interests and skills. user. The overall functionality of this chatbot involves the seamless integration of multiple components including data collection, natural language processing, recommendation algorithms, and user interaction. Here is an overview of how chatbots work:

User interaction and information collection:

The chatbot begins to interact with the user by greeting them and prompting them to enter their career interests, location interests, field of study, and skills. This dialogue-based approach creates a conversational experience, allowing users to communicate naturally with the chatbot.

Collect data:

Course data:

Chatbots collect course-related data from datasets from platforms like Coursera. This dataset contains information about various courses, including name, university, URL, and related skills.

Job data:

Chatbots leverage web scraping techniques to pull real-time job postings from platforms like LinkedIn. It retrieves job title, location and other related details.

Natural Language Processing (NLP):

Understanding user input:

Chatbots use NLP to understand and interpret user input. It identifies keywords, context and user intent to extract relevant information like career interests, skills, etc.

User Profile:

Chatbots create user profiles by mapping the skills and interests provided by the user. This profile forms the basis for personalized recommendations.

Recommended tools:

Course recommendation:

The chatbot uses the user's field of study and skills to match them to courses in the dataset. It calculates the relevance of each course based on overlapping skills and makes key recommendations.

Recommended use:

Chatbot matches user skills and interests with real-time job postings. It calculates job relevance using algorithms such as cosine similarity and recommends jobs that match the user's profile. Interactions and Outputs:

Presenting recommendations:

The chatbot presents the user with a list of recommended courses and job openings. Users can select a specific job to explore further.

Course details:

If the user selects a course, the chatbot will provide details like the course name, university, and URL for more information.

Job details:

If the user selects a job, the chatbot will display details such as the job title, location of the job.

Chapter 4

Conclusions and Future work

4.1 Conclusions

The Conversational Recommender System for Educational Paths project represents a major step forward in modern career counseling and educational recommendations. In an ever-changing career landscape where the alignment of skills, aspirations and opportunities is paramount, this project offers a transformational solution that leverages cutting-edge technology to empower individuals pursue meaningful and fulfilling careers.

Through a complex combination of natural language processing, machine learning, and data analysis, this project has successfully realized a user-centric, intelligent recommendation system. This system, encapsulated in a user-friendly chatbot interface, is revolutionizing the way individuals navigate complex career exploration and skill enhancement.

From the very beginning of the project, the goal was to reduce the challenges inherent in conventional approaches to career planning. General advice and standardized assessments often fail to summarize many aspects of an individual's career path. This project fills this gap by facilitating personalized conversations with users, engaging them in a dynamic dialogue to capture career interests, geographical considerations, their level of education and skills. The merging of these inputs provides the foundation on which the recommendation system works, ensuring the accuracy and relevance of its recommendations.

The implementation of web scraping techniques to pull job postings in real time from platforms like LinkedIn introduces a dynamic element that helps tailor recommendations to the latest job trends. By providing users with a curated list of job titles that match their interests, the system allows them to discover roles that match their ambitions and qualifications. However, the ingenuity of this project goes beyond work recommendations. Recognizing the symbiotic relationship between education and career advancement, the system expands its ability to recommend suitable courses to fill skills gaps. Integrating career guidance into educational pathways means a holistic approach to personal and professional development. Users are provided not only with job opportunities but also with specific steps to improve their skills and qualifications, effectively equipping them in the competitive job market.

Basically, the project “Conversational recommendation system for educational pathways” embodies the symbiosis between human aspirations and technological capabilities. It turns the traditional static, one-way career planning process into a dynamic, interactive, and user-centered journey. By enabling individuals to harness the power of technology to align their skills, aspirations and opportunities, this project represents a paradigm shift in career exploration and education.

As we are at the forefront of the fourth industrial revolution, where flexible career paths and ever-evolving skillset, this project embodies the essence of adaptability and innovation. It exemplifies the fusion of data-driven insights, artificial intelligence, and human intuition to create a personalized, ever-evolving road to success. "Conversational recommendation system for educational pathways" is not just a project; it heralds a new era in career guidance, where individuals are empowered to shape their destiny with knowledge, insight and confidence.

4.2 Future work

In the future, there are some interesting ways to improve the project "Conversational Referral System for Educational Pathways". The success of the project provides a solid foundation for future improvements and expansions. One promising direction is to enrich the dataset with more job titles, skills, and courses. By combining more diverse data, the accuracy and relevance of the recommender system can be significantly increased. In addition, fine-tuning the model through hyperparameter tuning and advanced natural language processing techniques can lead to more accurately matched skill and lesson recommendations. Another avenue for future work is to explore advanced similarity metrics beyond cosine similarity. By testing metrics like Jaccard similarity, Pearson correlation, or specialized text similarity algorithms, the project was able to uncover more nuanced ways to compare user skills with key content. learning, potentially improving the accuracy of the recommendations.

Due to the dynamic nature of the job market, it is important to incorporate real-time updates for skills and courses. This could involve mining APIs or using web search to gather current information from platforms like Indeed and other job sites, ensuring that users receive up-to-date recommendations. best and more suitable.

To improve personalization, it can be helpful to set up a user feedback mechanism. Allowing users to rate the relevance of suggested courses creates a feedback loop, refines understanding of the personal preference system, and improves recommendations over time.

Exploring the integration of deep learning models, such as neural networks, is another interesting avenue. Models such as the Siamese grid or those incorporating the attention mechanism can uncover complex patterns in the data, potentially leading to more precise skill matching.

Incorporating user-specific factors, such as career aspirations, preferred form of study, and education level, into the recommendation process can further personalize the experience. This can lead to many personalized lesson suggestions tailored to each individual's unique goals. Partnering with online learning platforms like Coursera, edX, or Udemy to directly link recommended courses to subscription options creates a seamless experience for users, making it easy to switch from topic to topic. Export to improve skills.

Finally, expanding the project's global reach by adapting the interface of the chatbot and the recommendation system to support multiple languages and accommodate diverse cultural contexts can help a wide range of audiences. gain access to the benefits of the project.

In short, the project "Conversational Suggestion System for Educational Pathways" is the starting point for a lot of future improvements. By continuously refining the model, capturing new data sources and exploring emerging technologies, the project has the potential to become an indispensable tool, allowing individuals to follow a unique journey, to complete their careers and develop skills in an ever-changing professional landscape.

4.3 Critique of algorithm and reasons for error

An important aspect that requires careful consideration is the use of cosine similarity as the primary measure of similarity. While cosine similarity is a commonly used metric and serves well for initial skill matching, it also has its downsides. Cosine similarity does not take into account the semantic meaning of words; it treats words as independent traits, which can lead to inaccuracies when comparing skills that are contextually related but not syntactically similar. This limitation can lead to false negatives or false positives when recommending skill-based courses. Exploring more advanced similarity metrics that take into account semantic relationships can improve the accuracy of recommendations. Another potential source of error is in the process of removing work from platforms like LinkedIn. The success of the project depends on the availability and consistency of data from these platforms. A change in the format of a job posting, a change in the platform structure, or potential site search limitations may result in incomplete or inaccurate job data retrieval. These errors can lead to suboptimal job recommendations or, in some cases, missed opportunities for users. To overcome this, regular maintenance and adjustment of the scraper mechanism is required to ensure the durability of the system.

In addition, competency-based course design recommendations may ignore the context in which competencies are applied. Skills often come with different skill levels, and the project's current approach treats all skills equally. Combining user skill levels or skill weights can provide more nuanced recommendations that are more relevant to the user's skill development needs.

An inherent challenge of the project is the dynamic nature of the labor market and the rapidly changing skills required. The project's reliance on historical data may not fully capture emerging skills or roles. As a result, the system can recommend courses that were relevant in the past but are no longer relevant to the current needs of the industry. To address this, a continuous feedback loop incorporating real-time labor market information can improve the system's ability to adapt to changing trends.

Additionally, the project's reliance on user-supplied data leads to biases and potential inaccuracies. Users may not accurately assess their skills or provide complete information, resulting in recommendations that do not accurately reflect their abilities. Incorporating methods for verifying and validating user-supplied data, such as skills assessment questions or integration with professional profiles, can improve the credibility of recommendations.

4.4 Contribution to the State of The Art

The Conversational Recommendation System for Educational Pathways project has made significant contributions to modern technology in several key areas:

Personalized career and education recommendations:
The project advances the field of personalized recommender systems by integrating user interests, skills and aspirations. By engaging users in conversation and considering various aspects such as career interests, location, field of study, and skills, the project creates a tailored recommendation experience that goes beyond conventional methods. traditional approach.

Seamless user interaction with Chatbot interface:
Using a chatbot interface improves user engagement and accessibility. This approach fits the growing trend of conversational AI, allowing users to interact naturally and comfortably while receiving valuable recommendations. This contributes to the user experience and accessibility of recommender systems. Competency-based course recommendations:
The project's ability to recommend courses based on user skills meets an important need in career planning. This addition goes beyond job recommendations to provide users with concrete steps to improve their skills, aligning their educational activities with their career goals. This innovative integration bridges the gap between education and employment.

Real-time workflow data integration:
Incorporating web scraping techniques to extract job postings from platforms like LinkedIn gives a dynamic and relevancy to the recommendations. By integrating real-time employment data, the project ensures that users are exposed to the latest trends and opportunities, helping to ensure the timeliness and accuracy of recommendations. Comprehensive approach to skill matching:
The use of cosine similarity to match user skills to course content is a notable contribution. Although cosine similarity is a widely used metric, its application to skill matching in the context of conversational recommendation demonstrates its effectiveness in personalizing educational pathways. sex. This approach emphasizes the importance of skills in shaping careers.

User-driven feedback loop:
By suggesting user feedback mechanisms on the relevance of the course, the project realizes the value

of user feedback. This iterative feedback loop not only improves personalization, but also helps refine the recommendation algorithm over time. This approach is consistent with the user-centered design philosophy of recommender systems.

Integrate with multiple work platforms:
Scalability to integrate data from various employment platforms, such as Indeed, will further enhance project relevance and relevance. This contribution demonstrates the ability to adapt to different data sources and underscores the project's commitment to providing comprehensive recommendations.

Explore future directions:
Project insights into future enhancements, such as advanced similarity discovery, multimodal data consolidation, and internationalization, offer a thought-provoking approach improvement. These leads contribute to the vision of continuous improvement and innovation in the referral field.

Bibliography

[1] *A Survey on Conversational Recommender Systems*
DIETMAR JANNACH(University of Klagenfurt), AHTSHAM MANZOOR (University of Klagenfurt), WANLING CAI (Hong Kong Baptist University), LI CHEN(Hong Kong Baptist University) 2021.

[2] *Unifying Recommender Systems and Conversational User Interfaces*
Alain D. Starke, Marketing and Consumer Behaviour Group, Wageningen University & Research, Netherlands and Department of Information Science and Media Studies, University of Bergen
Minha Lee, Department of Industrial Design, Eindhoven University of Technology, Netherlands 2022.

[3] *Towards Topic-Guided Conversational Recommender System*
Kun Zhou, Yuanhang Zhou, Wayne Xin Zhao, Xiaoke Wang, Ji-Rong Wen 2020.

[4] *Applying data mining techniques in job recommender system for considering candidate job preferences*
Anika Gupta; Deepak Garg IEEE 01 December 2014.

[5] *Web Scraping: State-of-the-Art and Areas of Application*
Rabiyatou Diouf, Edouard Ngor Sarr, Ousmane Sall, Babiga Birregah, Mamadou Bousso, Sény Ndiaye Mbaye 2019 IEEE International Conference on Big Data (Big Data)

[6] *Collaborative job prediction based on Naïve Bayes Classifier using python platform*
Savita Choudhary, Siddanth Koul, Shridhar Mishra, Anunay Thakur, Rishabh Jain
International Conference on Computation System and Information Technology for Sustainable Solutions
Year: 2016 | Conference Paper | Publisher: IEEE

[7] *Course Recommendation System using Content-based Filtering*
Pranshu Mishra, Vishal Jain
7th International Conference on Trends in Electronics and Informatics (ICOEI)
Year: 2023 | Conference Paper | Publisher: IEEE

[8] *A Nobel Approach to Recommend Online Courses by using Hybrid Recommendation System*
Sumita Gupta, Rana Majumdar, Sapna Gambhir
10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)
Year: 2022 | Conference Paper | Publisher: IEEE

[9] *Machine learning driven course recommendation system*
Sara Lazarevic, Tamara Zuvela, Sofija Djordjevic, Srdjan Sladojevic, Marko Arsenovic
21st International Symposium INFOTEH-JAHORINA (INFOTEH)
Year: 2022 | Conference Paper | Publisher: IEEE