**A COURSE KNOWLEDGE BASED RECOMMENDER SYSTEMS USING CONSTRAINT BASED METHOD**

**MONICAH ATIENO OMONDI**

**SCT211-0616/2018**

**SUPERVISOR**

**DR KENNEDY OGADA**

A research project submitted to the Department of Computing in the School of Computing and Information Technology in partial fulfilment of the requirement for the award of a Degree in Computer Science by Jomo Kenyatta University of Agriculture and Technology.

**2022**

# ABSTRACT

Recommender systems have been used in a variety of industries to deliver individualized recommendations to users, including entertainment, e-commerce, agriculture, healthcare, and education. When using online systems, recommender systems can help to overcome the problem of a user being inundated with information. Courses are now made available online on portals for students to apply, thanks to the digitization of the course application procedure at higher education institutions. There are far too many courses for a student to explore thoroughly before making a decision. As a result, students are choosing and being assigned to courses in which they are uninterested, necessitating the development of a course recommender system that proposes a short selection of courses that are most relevant to the student. This project aims on developing a knowledge base recommender system for providing personalized course recommendation to students based on their personality and performance. In contrast to other recommender systems, the knowledge-based recommender system will be more accurate in this study since it gets rid of common restrictions such the cold-start, new item, and grey sheep problems. It also leverages domain information that is free of noise, making its recommendations more reliable. Reviewing recommender system methodologies and creating a course knowledge-based recommender system are the main goals of this work.

Table of Contents

[ABSTRACT 2](#_Toc108178073)

[CHAPTER 1: INTRODUCTION 4](#_Toc108178074)

[1.1 Introduction . 4](#_Toc108178075)

[1.2 Background 4](#_Toc108178076)

[1.3 Problem statement 6](#_Toc108178077)

[1.4 Proposed solution 7](#_Toc108178078)

[1.5 Main objectives 7](#_Toc108178079)

[1.5.1 Specific objectives 7](#_Toc108178080)

[1.6 Justification 7](#_Toc108178081)

[CHAPTER 2: LITERATURE REVIEW 8](#_Toc108178082)

[2.1 Introduction 8](#_Toc108178083)

[2.2 Recommender systems 8](#_Toc108178084)

[2.2.1 Content Based recommender system 8](#_Toc108178085)

[2.2.2 Collaborative Filtering 9](#_Toc108178086)

[2.2.3 Hybrid recommender system 11](#_Toc108178087)

[2.2.4 Knowledge-Based recommender system 12](#_Toc108178088)

[2.3 Course Selection 13](#_Toc108178089)

[2.3.1 Personalities 14](#_Toc108178090)

[CHAPTER 3: RESEARCH METHODOLOGY 16](#_Toc108178091)

[3.1 DATA COLLECTION 16](#_Toc108178092)

[3.2 SYSTEM METHODOLOGY 16](#_Toc108178093)

[3.3 REQUIREMENTS 17](#_Toc108178094)

[3.4 BUDGET 18](#_Toc108178095)

[3.5 TIMELINES 18](#_Toc108178096)

[REFERENCES 20](#_Toc108178097)

# CHAPTER 1: INTRODUCTION

## Introduction

This chapter starts with a description of the background where relevant concepts are defined and also reviews relevant research done in this area. The background is followed by problem statement, proposed solution, main and specific objectives and justification.

## 1.2 Background

Recommender systems are information filtering systems that deal with the problem of information overload (Osadchiy et al, 2019) by filtering vital information fragment out of large amount of dynamically generated information according to user’s preferences, interest, or observed behavior about item (Alhijawi et al, 2020). Recommender system has the ability to predict whether a particular user would prefer an item or not based on the user’s profile. There are several types of recommendation techniques that exist today. The most commonly used techniques can be classified in three categories: collaborative filtering techniques, content-based filtering techniques and hybrid systems (Chang et al, 2022). These recommender systems identify trends among a large number of users. The trends, identified on the basis of the user’s behaviors, are then used to classify new users. The resulting classification allows the generation of a recommendation under the hypothesis that users belonging to the same class will have and prefer a similar behavior. Companies like Amazon, Google and Facebook apply these types of recommendation algorithms in order to provide their users with books and movies, information sources and ads, and potential friends suggestions, respectively.

Collaborative recommender system builds a model from a user’s past behavior, activities, or preferences and makes recommendations to the user based upon similarities to other users (Ghasemi et al, 2021). Typically, the intuition behind collaborative filtering is that if users have similar preferences in the past, they tend to have the same interest in the future (Chen et al, 2021). Collaborative filtering system tries to find other like-minded users and then recommends the items that are most liked by them. Content-based recommender systems use item features to recommend other items similar to what the user likes, based on their previous actions or explicit feedback (Javed et al, 2021). It makes recommendations by using keywords and attributes assigned to objects in a database (e.g., items in an online marketplace) and matching them to a user profile. The user profile is created based on data derived from a user’s actions, such as purchases, ratings (likes and dislikes), downloads, items searched for on a website and/or placed in a cart, and clicks on product links.

The hybrid technique combines more than one type of recommendation technique and ensures that they complement each other by replacing weakness of one type with strength of the other type (Ndung'u et al, 2021). Hybrid recommender systems can produce outputs which outperforms single component systems by combining these multiple techniques. A knowledge-based recommender system on the other hand generates recommendations on the basis of the domain knowledge. It recommends items based on specific domain knowledge about how certain item features meet users’ needs and preferences and, ultimately, how the item is useful for the user. Constraint-based systems and case-based are the type of knowledge-based recommender systems (Ricci et al., 2020). Case-based recommenders determine recommendations on the basis of similarity metrics whereas constraint-based recommenders predominantly exploit predefined knowledge bases that contain explicit rules about how to relate customer requirements with item features

One of the most important stages in a young adult is the transition from high school to higher level learning institutions. Trying to choose a career at this stage is one of the most challenging moments in this stage of life. At this point the stakes are high and whatever one chooses will be stuck with them for a long time. Students need guidance when selecting their course choices for higher education. In Kenya, students lack well planned and organized career guidance in school (Njogu et al, 2019). Unfortunately, a major percentage of people at this stage are not equipped with the right knowledge to make the right decision. Many people around them could be giving contradictory advises or they could be disturbingly oblivious of the critical junction the are at. Further, ( Newa et al, 2021) found out that 63.3% of students admitted in public universities in Kenya were dissatisfied with the degree course because they were placed in degree courses they didn’t have a passion for.( Kanyingi-Maina & M.W, 2020) recommends that students should be encouraged to make career choice decisions in areas that they have or can acquire knowledge easily, skills and have interests as it is likely to promote productivity when the student is doing what they are interested in.

In the Kenyan job market, there have been many discussions regarding the lack of skill from graduates. On the supply-side, there is evidence that Kenyans leave school without suitable skills for the workplace (Mumbe & K., 2020). If we look at this problem from the root, we realize that most graduates end up taking courses not compatible to them. When a person partakes a course, they have interest in, they are likely to go an extra mile and do more research and teach themselves skills at a personal level thus by the time they are done with the course they not only have the theoretical knowledge but also have the practical skills. An ideal course recommender system that uses the students' interests and performance to recommend a list of courses that match their interest and performance will ensure students are assigned to courses they are interested in and may easily gain knowledge.

## 1.3 Problem statement

A Knowledge-Based Recommender System (KBRS) distinguishes itself among the various types of recommender systems by applying another technique to produce a recommendation. While the collaborative filtering technique, the content-based technique and the hybrid technique are popular means for the generation of recommendation, they have a number of limitations: the new user problem, the new item problem, the grey sheep problem, limited content analysis, over-specialization, and data sparsity (Guo et al, 2020). The knowledge-based recommender system addresses these limitations. When using the knowledge-based approach, no large data set is necessary and the cold-start, new item and the grey sheep problem are thus avoided. Also, because the domain knowledge, on which are based the recommendations, is noise-free the recommendations are more reliable (Kovaliuk et al, 2021). The only limitation faced by the KBRS is the construction of the knowledge base, which usually is a complicated task that demands considerable domain knowledge, and expertise in knowledge representation. The knowledge-based system is thus the most suitable approach to use.

## 1.4 Proposed solution

This project aims at developing a knowledge-based recommender system using the constraint-based method that recommends courses to students based on their personality and performance. This way the student is pointed to direction most suitable for them and the information overload is cut down drastically. A student will be required to enter his/her high school grades. A questionnaire will then be used to test the student’s personality using the Holland’s Code developed by John Holland. The Holland code known as the Holland’s Theory of Vocational Personalities in Work Environment provides a framework on understanding career interests that are therefore used in career guidance. The personalities will then be mapped to the various courses available and the system will then generate a list of courses most suitable to the student.

## 1.5 Main objectives

Review recommender systems techniques and develop a course knowledge-based Recommender System based on interest and performance.

### 1.5.1 Specific objectives

* Analyze recommender systems techniques
* Develop a course knowledge-based recommender system
* Test the developed system

## 1.6 Justification

Education is currently one of the most important long-term investments. By applying this knowledge-based course recommender system, a higher number of students will end up taking the appropriate courses and the application procedure to higher education institutions will be streamlined. Students' productivity will also rise, making instruction more fun for lecturers while also providing motivated and skilled graduates for the labor market. As a result, this system benefits not only students, but also instructors and companies.

# CHAPTER 2: LITERATURE REVIEW

## 2.1 Introduction

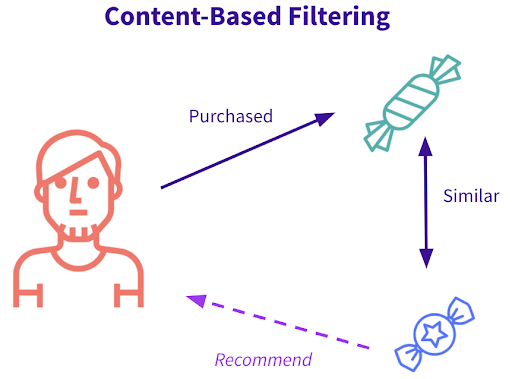
Recommender systems are discussed at the beginning of this chapter. In-depth descriptions of the different types of recommender systems are provided, along with discussions of pertinent research on those particular types of recommender systems. In addition, the various personalities as classified by Dr. John Holland are examined.

## 2.2 Recommender systems

Recommender system is an area of research in machine learning. The recommender system's main idea is to build relationship between the products, users and make the decision to select the most appropriate product to a specific user (Mohamed et al., 2019). Product in this case can represents many things such as courses, movies, amount of loan a financial institution can offer, houses, restaurants, books, articles and news among others. These systems help to improve the quality and the decision-making process. These systems have several advantages such as benefiting users in finding items of their interest, helping item providers in delivering their items to the right user, identifying products that are most relevant to users, personalizing content and helping websites improve user engagements (Kulkarni et al., 2020). The main types of recommender system include content-based recommenders, collaborative filtering, hybrid recommenders and knowledge-based recommenders.

### 2.2.1 Content Based recommender system

Content-based filtering uses item features to recommend other items similar to what the user likes, based on their previous actions or explicit feedback (Javed et al., 2021). Building a content-based recommender system involves recommending items similar to those that have been preferred by the user in the past. These systems are scalable, work well independent of the number of users in the system and does not experience cold start issues since it takes considerations of historical preferences of a user and property of an item. However, this system requires enough details about the item to be provided so as to differentiate products precisely without which poor accuracy is experienced (Amara et al., 2020).



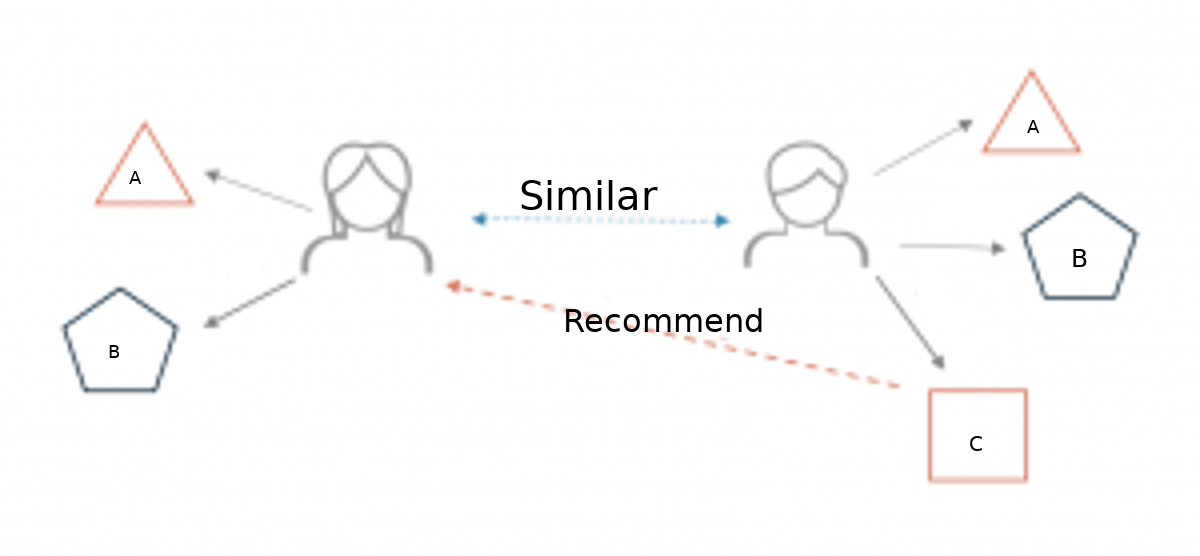
(Mokarrama et al., 2020) proposed a recommender system that will help the prospective students of Bangladesh in choosing the most suitable private universities for getting admission. In this proposed recommender system, a list of top-K private universities is recommended to the students who are willing to get admitted to the private universities using content-based filtering technique. To attain this goal they considered six parameters, namely grade point average (GPA) of secondary school certificate (SSC) examination, GPA of higher secondary certificate (HSC) examination, total GPA, tuition fees, university ratings, and university rankings. They used cosine similarity algorithm that calculates the similarities between a prospective student and a university using the six parameters. They showed that applying content-based recommender filtering technique by taking into consideration the university information and user preferences, a prospective student can be provided recommendation of the most suitable universities. The researchers noted that the system can be extended by integrating user feedback and making use of collaborative filtering techniques to constitute a hybrid system. They also noted that the more universities are considered, system performance can be evaluated more precisely and it will be more interactive to the user.

(Morsomme et al., 2019) developed a content-based course recommender system for the Liberal Arts bachelor of the University College Maastricht, the Netherlands. The system aimed at complementing academic advising and

helping students make better-informed course selections. The system recommended courses whose content best matched the student’s academic interests, issued warnings for courses that are too advanced given the student’s academic background and suggested suitable preparatory courses. They based the course recommendations on a topic model fitted on course descriptions, and the warnings on a sparse predictive model for grade based on students past academic performance and level of academic expertise. They used preparatory courses which consisted of courses whose content had the best preparatory value according to the predictive model. They found that course recommendations were relevant for a wide range of academic interests present in the student population and that students found recommendations for courses at other departments especially helpful. They also discovered that preparatory courses often lacked coherence with the target course and needed to be improved. For future work they saw the need to differentiate structure-related and content-related topics which seemed particularly difficult to do. They also saw the need to expand the course data to course manuals which contained detailed description of the courses’ content. They also noted that Kullback Leibler distance could be used to increase the coherence between preparatory courses and target course.

### 2.2.2 Collaborative Filtering

Collaborative filtering recommends items to users based on their historical rating information. It is based on the assumption that users with similar taste in the past will have similar preferences in the future. The items that other users with similar tastes liked in the past are recommended to the target user (Koren et al., 2022). The likeness in taste of two users is computed with regards to the likeness in the past ratings of the users. Currently there are two approaches for collaborative filtering, memory-based (user based) and model-based (item-based) algorithms. Model-based collaborative filtering algorithms use data mining techniques in order to develop a model of user ratings, which is used to predict user preferences while the memory-based algorithms, also known as nearest-neighbour methods, treats all user items by means of statistical techniques in order to find users with similar preferences (Yu et al., 2022). Collaborative filtering boasts being simple to implement and accurate. However, they suffer from cold start problem where they fail to recommend for first time users whose information is not on the system or offer recommendations about an item that was just incorporated into the system and, therefore, has few, if any, evaluations from users(Srifi et al., 2020).



(Fan et al, 2021) K-Nearest Neighbours (KNN) is a machine learning algorithm to find clusters of similar users based on certain metrics such as ratings. It is an effortless but productive machine learning algorithm. It is effective for classification as well as regression. However, it is more widely used for classification prediction. KNN groups the data into coherent clusters or subsets and classifies the new inputted data based on its similarity with previously trained data. The input is assigned to the class with which it shares the most nearest neighbours. Most collaborative filtering techniques are built by implementing this algorithm.

(Bhumicitr et al., 2019) proposed a recommender system that recommends university elective courses based on the similarities between the courses and the courses taken by the student. Collaborative based recommendation using Pearson Correlation Coefficient and Alternating Least Square (ALS) were subjected on dataset of academic records of university students. ALS was found to be the best performed and deployed in the recommender system. The researcher further proposes the use of other information apart from the student enrolment data in order to incorporate the behavior of the student for further recommendation.

(Gupta et al., 2020) proposed a movie recommendation system using collaborative filtering. The main objective was to recommend movies using the item-based technique. The proposed approach used the KNN algorithm to find the distance between the target movies with every other movie in the dataset and then it ranks the top k nearest similar movies using cosine angle similarity. This started by the extraction of the dataset to gather information about the target movie and the user’s rating. The collaborative filtering then followed with the formatting of the rating dataset so that it can be consumed by the KNN model, to remove the huge dataset handling problems. The dataset is reduced according to the popularity removing the noisy error pattern to get the sparse matrix. Next cosine similarity was used to find the distance between the target movie and other movies, which gave the top k nearest neighbor. And finally, the required recommended list of movies was displayed with descending order of distance. For future work they recommended combining content-based filtering with the collaborative filtering in order to minimize the errors and improve the performance as a hybrid approach.

### 2.2.3 Hybrid recommender system

A hybrid recommender system is a unique subtype of recommender system that combines different techniques of different types, for example, mixing content based filtering and collaborative filtering. It is also possible to mix different techniques of the same type, like naïve Bayes based content-based filtering plus kNN based content-based filtering Also, mixing same type of techniques with different datasets can be possible (Zhang et al, 2019). The most common hybridizing methodology is combining different techniques of different types, for example, mixing Content-based and collaborative filtering. (Esteban et al., 2020). The shortcomings of both content-based and collaborative filtering techniques are overcome by this strategy. The combination can be done in various ways like including employing content and collaborative-based methods to generate predictions separately, then combining the predictions, or simply enhancing a content-based approach with the capabilities of collaborative-based methodsand vice versa (Khan et al., 2020).

(Alper et al., 2021) proposed a hybrid recommender system that uses student and course information with collaborative filtering and content-based filtering models. The proposed system provided consistent recommendations by using explicit and implicit data, without predefined association rules. The collaborative filtering algorithms used grades as rating values. The content-based filtering algorithms utilized text-based information about students and courses by converting them into feature vectors using natural language processing methods. In the combination phase of the hybrid recommender system, only one of the collaborative filtering and one of the content-based filtering models are used with different ensembling methods. The suggested hybrid recommender system achieved outperforming results for all evaluation metrics. The results showed the values of the rank-aware for the individual models and the hybrid models with different combinations. In particular, for content-based filtering with Bayesian personalized ranking, the hybrid model performed better than any algorithm in practice. They noted that for future work, more detailed models using course evaluation questionnaires, various grade scaling methods, and course success thresholds can be used.

### 2.2.4 Knowledge-Based recommender system

Knowledge-based systems recommend items based on specific domain knowledge about how certain item features meet users’ needs and preferences and, ultimately, how the item is useful for the user. Knowledge based systems tend to work better than others at the beginning of their deployment but if they are not equipped with learning components, they may be surpassed by other shallow methods that can exploit the logs of the human/computer interaction (El Bouhissi et al., 2021). The knowledge-based system can further be grouped into case based or constraint based. In terms of used knowledge, both systems are similar: user requirements are collected, repairs for inconsistent requirements are automatically proposed in situations where no solutions could be found, and recommendation results are explained. The major difference lies in the way solutions are calculated. Case-based recommenders determine recommendations on the basis of similarity metrics whereas constraint-based recommenders predominantly exploit predefined knowledge bases that contain explicit rules about how to relate customer requirements with item features (Ricci et al., 2022). Constraint-based methods are particularly well suited for recommending complex products such as financial services or electronic consumer goods.

(Santamaria et al., 2020) proposed developing knowledge-based recommender system for tourist attraction area selection in Ethiopia using a case-based reasoning approach. They aimed to design a recommender system for tourist attraction areas and visiting time selection that can assist experts and tourists to make timely decisions that helps them to get fast and consistent advisory service so that visitors can identify tourist attraction areas that have the highest potential of success/satisfaction and that match their personal characteristics. For the development of the recommender system, essential knowledge was acquired through semi-structured interview and document analysis. Domain experts and visitors were interviewed to elicit the required knowledge about the selection process of attraction area. The acquired knowledge was modeled using hierarchical tree structure and it was represented using feature value case representation. The main data source used to develop the system for tourist attraction area selection was previous tourist cases collected from national tour operation and ministry of culture and tourism. As a retrieval algorithm, nearest neighbor retrieval algorithm was used to measure the similarity of new case (query) with cases in the case base. Accordingly, if there was a similarity between the new case and the existing case, the system assigned the solution (recommended attraction area and visiting time) of previous case as a solution to new case. To decide the applicability of the prototype system in the domain area, the system was evaluated by involving domain experts and visitors through visual interaction using the criteria of easiness to use, time efficiency, applicability in the domain area and providing correct recommendation. The researchers further recommended expanding this work by adding other important attributes such as housing preference, level of education, marital status, and purchasing habits by making a direct survey of successful visitors.

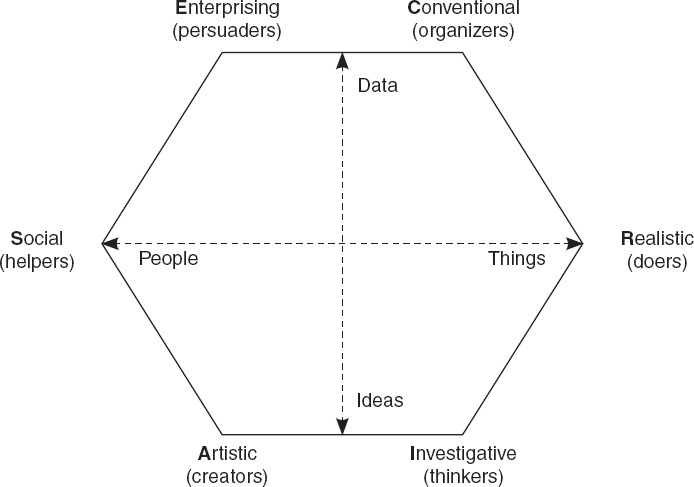
## 2.3 Course Selection

One of the many crucial decisions students must make that will have long-term consequences is their course of study (Njogu & S.W., 2019). This is due to the fact that having a career that aligns with one's interests, abilities, and values greatly boosts the likelihood of one's social and economic success as well as their sense of fulfillment and pleasure. Students encounter a dilemma throughout the course selection process since it can be challenging to align their career choices with their skills, interests, and performance.

### 2.3.1 Personalities

The Holland Theory developed by John Holland enables people to determine their career personality, vocation interest, and work personality. It was first proposed in 1959 and was initially conceived of as a trait and factor theory (Zainudin et al., 2020). Today, Holland's theory is regarded as one of the most useful and extensively studied career theories. This is due to how straightforward and easy the theory is. Holland's theory aids in career planning and describes how changes to the workplace environment and job characteristics can result from interactions between people and their surroundings (Harahap et al., 2020). Holland distinguishes six personality types—realistic, intellectual, artistic, social, entrepreneurial, and conventional—that correspond with various work environments. Prior research has demonstrated that personality types and environmental factors can aid students in making career choices (Ding et al., 2020).

The six personality types Realistic(R), Intellectual(I), Artistic(A), Social(S), Entrepreneur(E) and Conventional(C) are assessed to determine the degree of resemblance to the personalities and a three-letter code is generated to summarize the career interest of a person. The three most notable personalities will be identified when a student completes a questionnaire. The personality with the highest score from the three prominent personalities will then be the person’s most dominant personality and it will greatly influence the choice of career and the satisfaction of that career. The other two personalities from the three-letter code have a lesser role but they are still significant in career selection (Coenen et al., 2021).



According to Holland's RIASEC theory of personality types, Realistic individuals (the “doers”) have a strong preference for hands-on work that is manual in nature. These individuals prefer to work with tangible and physical objects over working with people or theoretical ideas (Ja-Hwung et al., 2021). Investigative individuals (the “thinkers”) enjoy thinking analytically and scientifically. These individuals prefer to use observational and logical skills to solve highly complex and abstract problems. Artistic individuals (the “creators”) are creative and imaginative and think outside the box. These individuals prefer ambiguity and freedom in solving problems. Social individuals (the “helpers”) prefer human interactions over mechanical work. These individuals have a strong preference for group settings and derive satisfaction from helping others. Enterprising individuals (the “persuaders”) are often skilled public speakers and are highly sociable. These individuals tend to have strong leadership skills and to navigate toward positions of power (Ja-Hwung et al., 2021). Finally, Conventional individuals (the “organizers”) have a strong preference for activities that are explicit and structured in nature. These individuals thrive in environments that are consistent and predictable (Ja-Hwung et al., 2021).

## 2.4 Gap

Recommender systems have been applied in various domains as a means to avoid information overload among users of online systems. Course recommendation is an issue that research is continuously being conducted. Prototypes have been developed to recommend courses using knowledge- based systems, collaborative filtering and clustering techniques. Student performance and previous enrolment data have been used to recommend these courses. .(Kanyingi-Maina & M.W, 2020) There is however need to incorporate the student’s interest in recommending these courses. This literature review brings out the need to incorporate the interest of the students and their performance in developing a course recommender system.

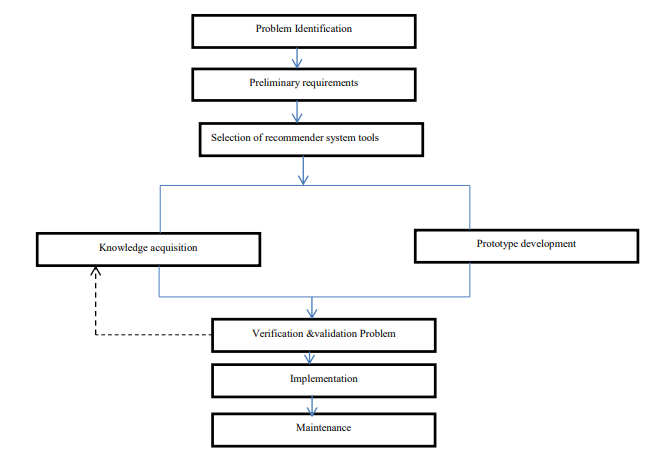
# CHAPTER 3: RESEARCH METHODOLOGY

## 3.1 Introduction

This chapter focuses on the system methodology process and the research process for this project.

### 3.2 System methodology

For this project the knowledge-based system methodology will be used to guide the steps for developing the course recommender system.



* **Problem Identification**

The process of problem identification involves the development of clear, straightforward problem statements that can be linked directly with the specific goals and objectives. In this case the problem to be solved by the knowledge-based system was identified.

* **Preliminary requirements**

This stage involves the acquisition of the initial requirement to set the stage for the design of the system. Knowledge acquisition also starts here and continues through later stages.

* **Selection of recommender system tools**

This stage involves the selection of the various tools to be used in developing the system. Django will be used to develop this system. The SQLite database will be used to store the data for this project which includes the course details, grade details, the interest questions and the user details.

* **Knowledge acquisition and system development**

This stage involves performing two activities in a continual process until the desired result is achieved. Knowledge acquisition is a permanent and crucial activity throughout all stages of designing, implementing and maintaining the system. It consists of the elicitation and interpretation of data on the functioning of expertise in some domain, in order to design, build, extend, adapt or modify a knowledge-based system. The knowledge acquisition will be in the form of interviews. The knowledge acquired will be continually updated to the knowledge base until the system works as expected.

* **Verification and Validation Problem**

This stage involves verifying that the knowledge acquitted is valid and not ambiguous to ensure it does not violate the expert’s expectation. Further reasoning is verified to eliminate invalid reasoning that may have arisen from incorrect translation of the knowledge.

Validation is done on the system to ensure that hidden errors are captured and rectified before proceeding to release the system. The knowledge acquisition and system development stage are invoked when an issue arises at this stage

* **Implementation**

This stage involves deploying the system. The website will be deployed on Heroku which is a container-based cloud Platform as a Service (PaaS) that enables developers to deploy, manage, and scale modern apps for free.

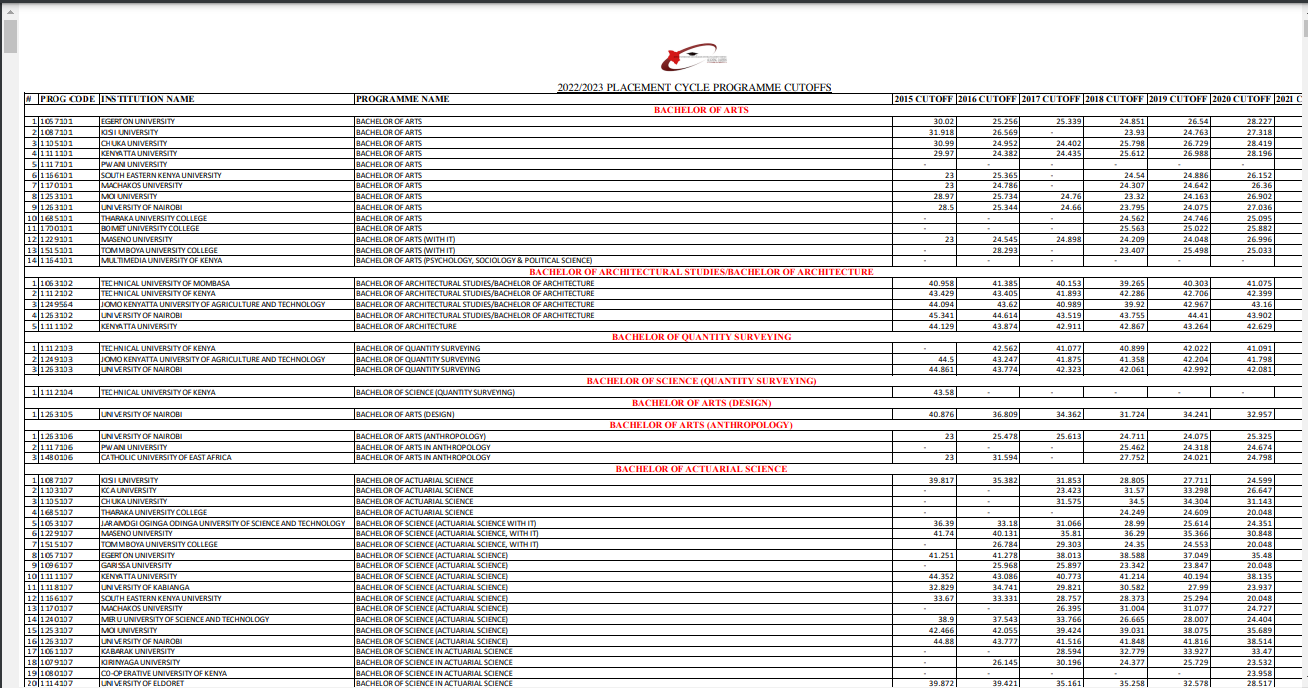
* **Maintenance**

This stage involves implementing any changes that the software might need after deployment. This can include handling residual bugs that were not able to be patched before launch or resolving new issues that crop up due to user reports.

## 3.3 Research Methods

### 3.3.1 Data Collection

The course data to be used for this project will be collected from the KUCCPS website. The course data will contain the courses together with the minimum grade requirement for the different courses. This course data will be updated with relevant course personality as directed by the career guidance expert who will map the courses available locally to the RIASEC personalities where they correspond with occupations. The universities offering the specific courses will also be included in the course data.



The holland self-assessment questions to be used for generating the RIASEC personalities will be obtained from Columbia university website. This document contains a list of general questions the student is required to answer in order generate the student course personality. The document contains the list of questions and also the formula used to generate the RIASEC personality.



### 3.3.2 Data Preparation

The available courses collected from the KUCCPS student’s portal were mapped to the courses that led to the occupations. The courses were then mapped to the personality types that best fit the courses based on the interests of the student. The mapping of the occupations to the courses and thereafter to the RIASEC personality was done with the help and guidance of a Career Guidance Professional.

### 3.3.3 System Development

The system will be developed in a continual process of knowledge acquisition and updates will be made on the knowledge base with the new domain knowledge obtained from the domain expert.

### 3.4.4 Evaluation

Evaluation will be taken to verify the effectiveness of the system in recommending courses to students based on their interest and performance at their secondary school. Precision, true positivity rate and false positivity rate will be used to measure the performance of the system in recommending course to secondary school graduate student.

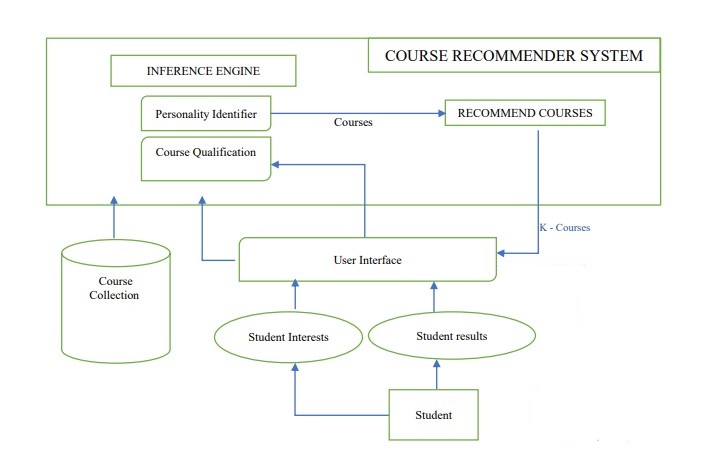
# CHAPTER 4: SYSTEM ANALYSIS AND DESIGN

## 4.1 Introduction

This chapter outlines the process of analysis and design of the course recommender system.

## 4.2 Design model

The design shows how the system is constructed and describes the critical components of the system. It shows the relationship between the different components.



### 4.2.1 Personality Identifier

The users are subjected to a RIASEC personality test of sixty questions. Each question is mapped to a personality type. Each positive response to a question is weighed as 1. The weights of each personality are determined by aggregating all responses. The next step is to get the top personality with the highest weight and that represents the student’s dominant personality.

### 4.2.2 Course Qualification

This module verifies whether the student’s grades has met the minimum requirement in order to qualify for that particular course. The grading system used for this system is as below.

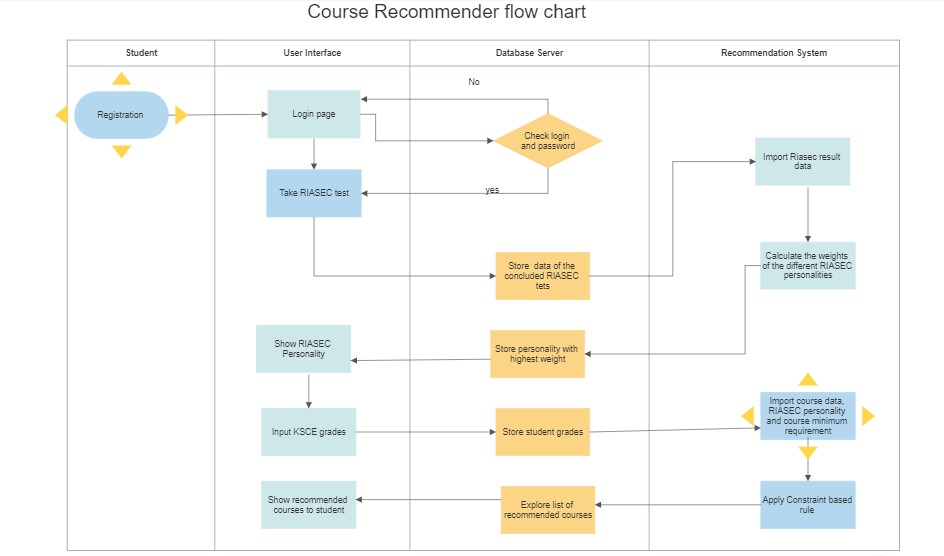
|  |  |
| --- | --- |
| Grade | Percentage range |
| A | 81-100 |
| A- | 74\_80 |
| B+ | 68-73 |
| B | 63-67 |
| B- | 53-62 |
| C+ | 55-59 |
| C | 50-54 |
| C- | 45-49 |
| D+ | 40-44 |
| D | 35-39 |
| D- | 30-34 |
| E | 0-28 |

### 4.2.3 Course recommendation

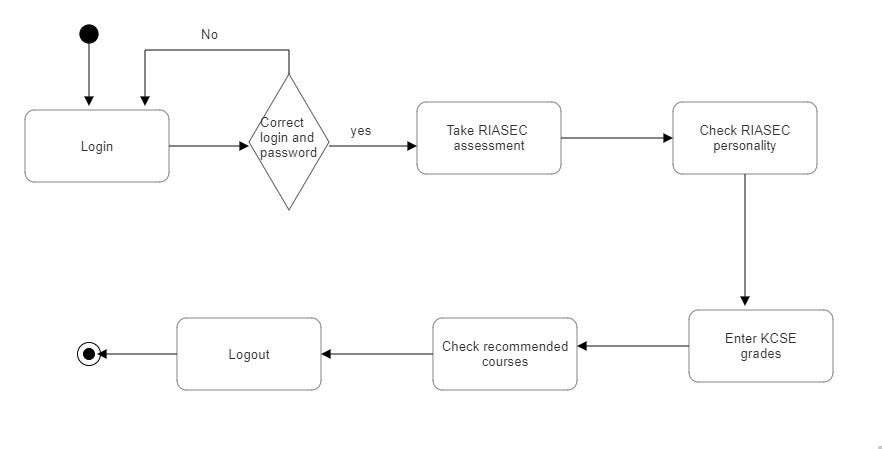
This module is used to filter courses based on inputs received from the course qualification and the personality identifier module. The available courses are thus reduced to a list of courses that the student matches based on their identified dominant personality and their performance. The courses will include degree, diploma and certificate level. The number of courses to be used in this will highly depend on the number of courses the career guidance expert was able to map to the different RIASEC personalities. A breakdown of the distribution of the course’s levels is as below:

|  |  |  |  |
| --- | --- | --- | --- |
| PERSONALITY | CERTIFICATE | DIPLOMA | DEGREE |
|  |  |  |  |
| REALISTIC | 14 | 16 | 18 |
| INVESTIGATIVE | 14 | 15 | 19 |
| ARTISTIC | 14 | 16 | 10 |
| SOCIAL | 15 | 16 | 18 |
| ENTERPRISING | 16 | 17 | 18 |
| CONVENTIONAL | 15 | 17 | 17 |

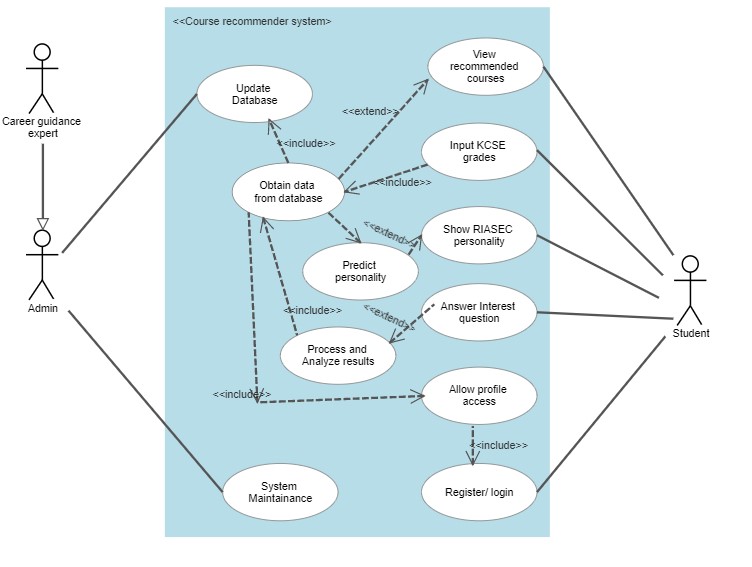
## 4.3 Flow chart



## 4.4 Activity Diagram



## 4.4 Use Case Diagram



# CHAPTER 5: SYSTEM IMPLEMENTATION AND TESTING

## 5.1 Introduction

## The implementation of the suggested project and the system's testing are the main topics of this chapter.

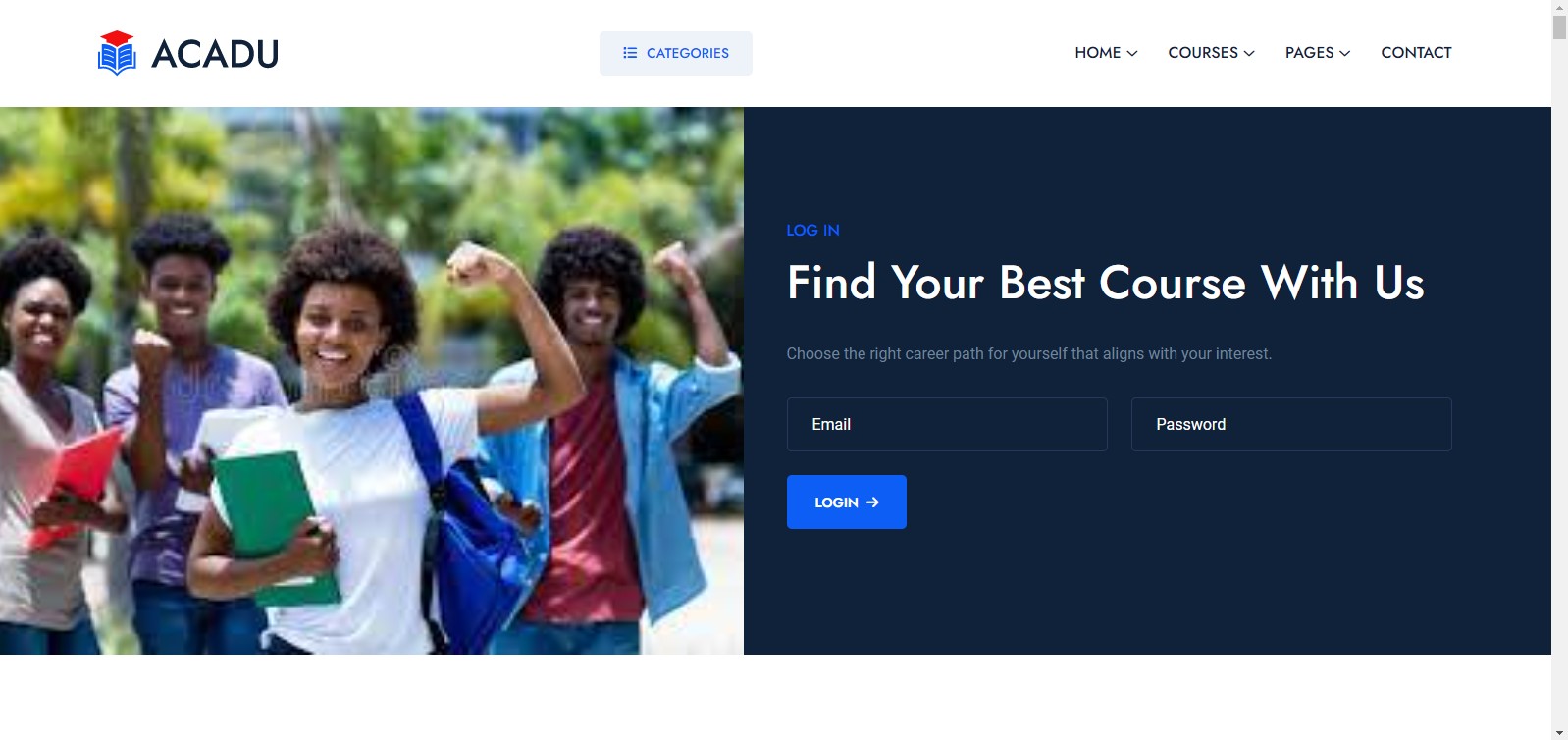
## 5.2 Application development

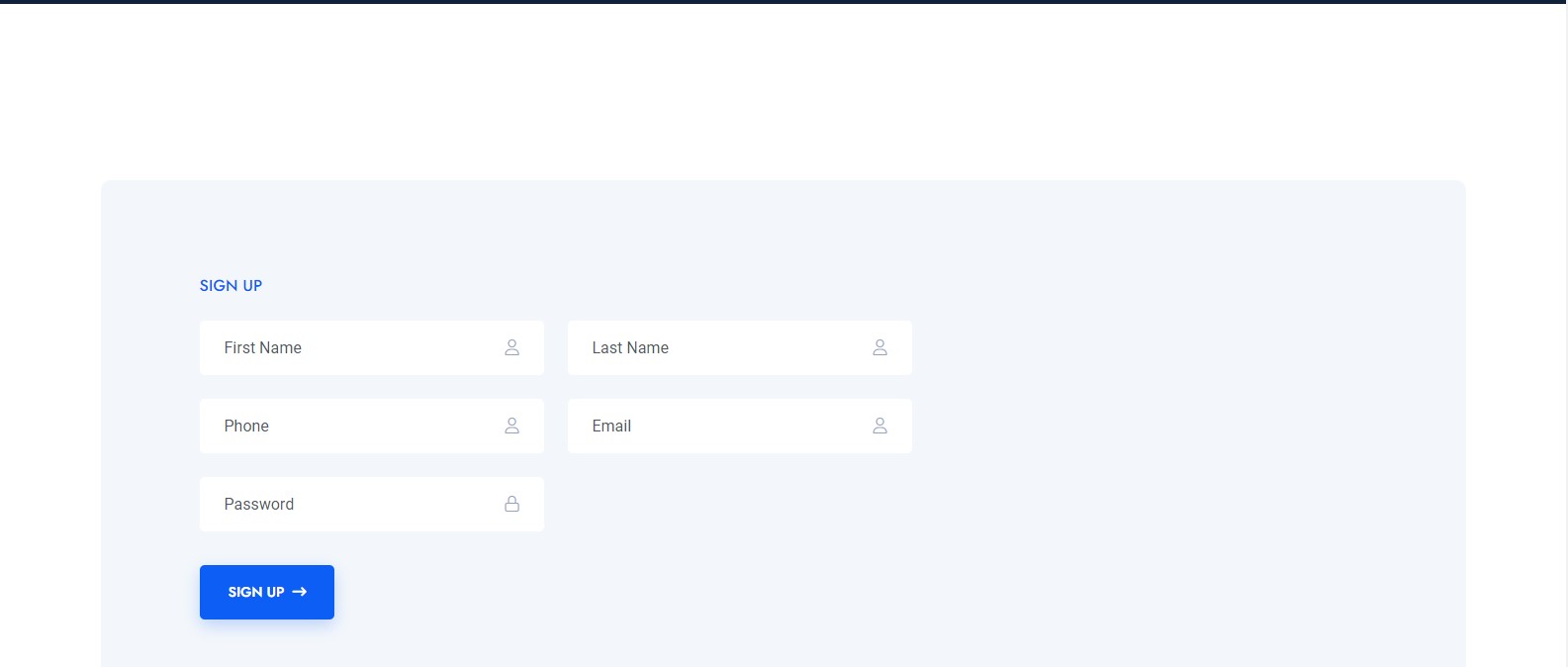
The tools used in the design and implementation of the project are:

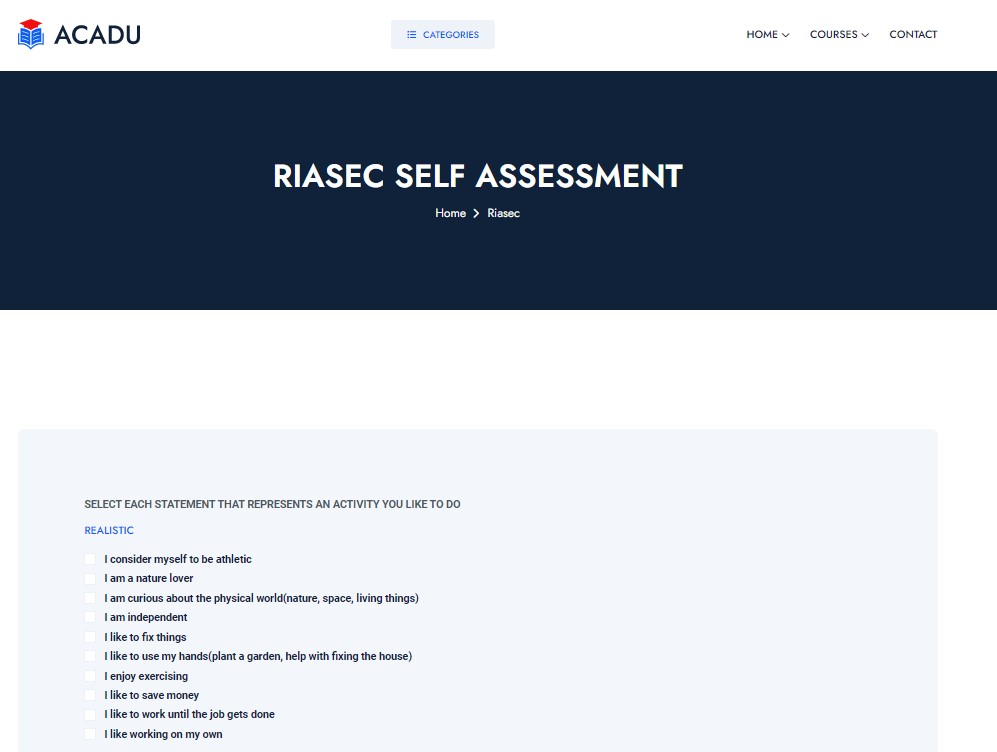
Django: this is a high-level python web-based framework

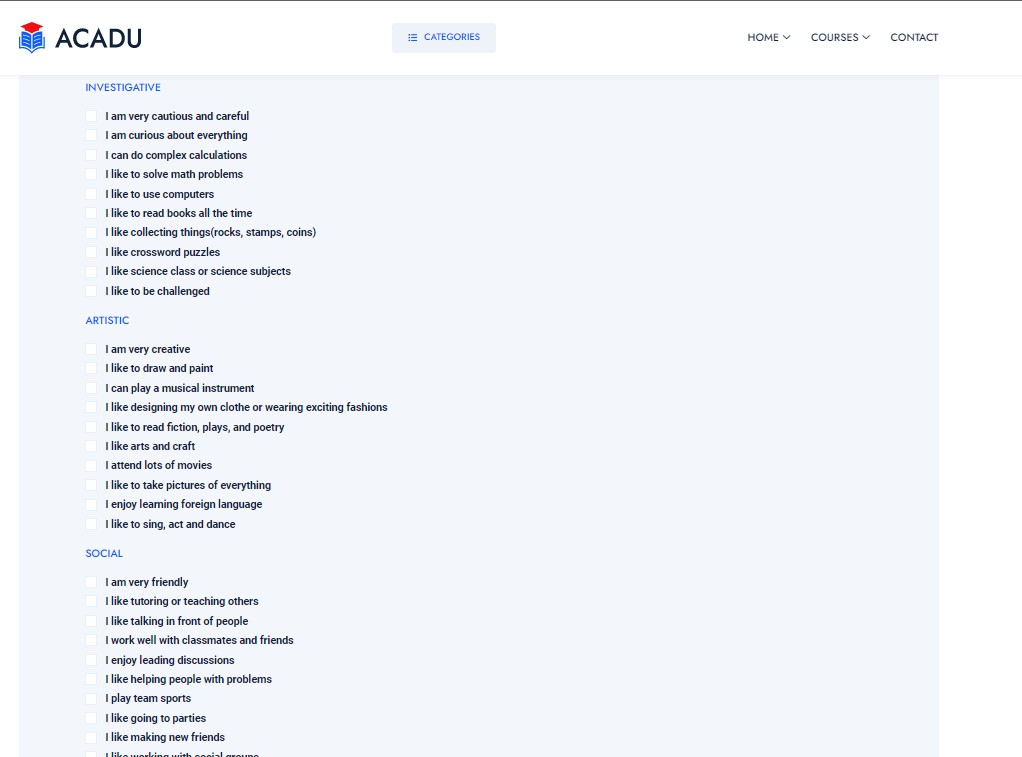
Sqlite3: this is the default database engine that comes by default in Django

### 5.2.1 System pages

The user first sees an onboarding screen when they first access the website, where they may obtain a quick description of what the website is about and further instructions. In order to proceed with the course recommendation process, a new user must first sign up.



The RIASEC self-assessment page opens after registering. The user must carefully study all of the statements and check off any activities that match their interests.

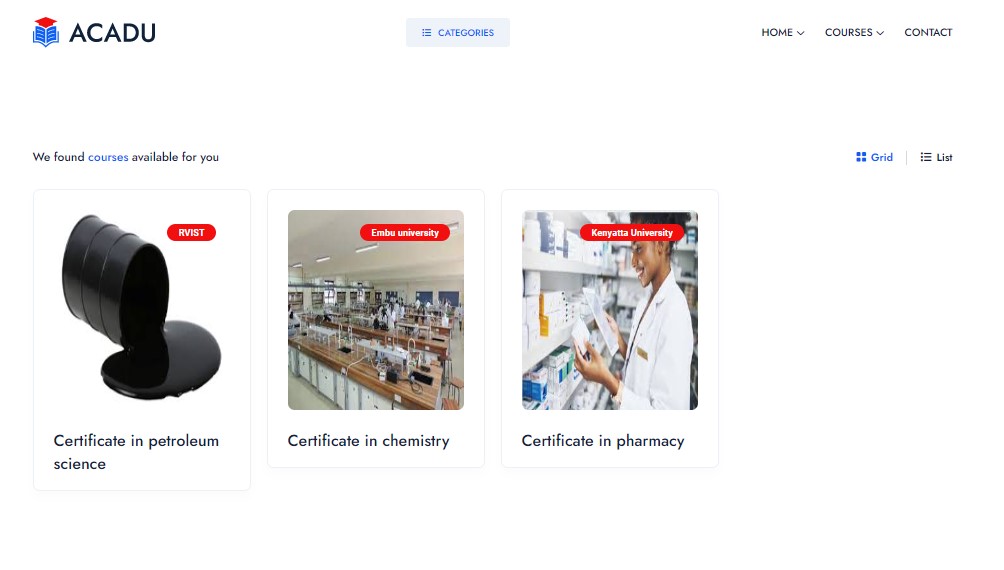


After completing the assessment, the user's RIASEC personality will be determined using the results. The user is then forwarded to another page where their riasec personality is displayed and the user is then asked to enter his/her Kenya Certificate of Secondary Education grades so that the suggested courses will also match his/her results.





The user will be directed to the recommended courses page after picking the grades, where all of the courses that fit his or her riasec personality and KCSE grades will be listed.



## 5.3 System Testing and evaluation

The project was evaluated by testing the satisfaction of students to the courses recommended. 35 random students were requested to use the system and their satisfaction with the recommended courses recorded. The students were instructed to determine whether they liked the suggested courses and whether there were any other courses in the database that weren't suggested but were on the list of options that matched their personality and performance and that they would want to pursue.

The result of the evaluation was recorded in a confusion matrix as shown below:

|  |  |  |
| --- | --- | --- |
|  | Liked Courses | Not Liked Courses |
| Recommended Courses | 45 | 11 |
| Not Recommended courses | 4 | 55 |

The accuracy of the prototype was then calculated from the above results

|  |  |  |  |
| --- | --- | --- | --- |
| Accuracy | Precision | Recall | F\_Measure |
| 86.95 | 80.35 | 91.84 | 85.71 |

## The accuracy of the system, which is 86.95% according to the table below, demonstrates the level of acceptable performance. The f measure, which is 85.71%, represents the system's effectiveness as a whole. Both findings demonstrate that the system performs pretty well.

## 5.4 Discussion of results

An accuracy of 86.95% is considered to be a respectable result by those evaluating the system's performance. The accuracy suggests that it is possible to efficiently recommend courses to students based on their performance and interests. It is clear that a knowledge-based course recommender system can be employed to suggest courses to students given that it has practical understanding of how a course satisfies a student's demands and the capacity to deduce the connection between a student's interest and performance and potential course recommendations.

## 5.5 Appraisal of the project

### 5.5.1 Strengths

* The knowledge-based system is able to recommend courses to students with a high accuracy.
* The system helps to reduce the time and cost associated with making the important decision of choosing a course based on performance and interests.

### 5.5.2 Weaknesses

* The system knowledge based is limited in quality and completeness as there are very many courses offered in universities but the system has limited number of courses.
* The system doesn’t recommend courses to students with grades lower than D in their Kenya Certificate of secondary Education
* The RIASEC theory used in this project is not always perfect. The personalities associated are approximate but not precise hence it doesn’t work for everyone (Usslepp et al, 2020).

# CHAPTER 6: CONCLUSION AND RECOMMENDATION

## 6.1 Conclusion

Recommender systems have been employed in entertainment, e-commerce, agriculture, healthcare and education among other industries to provide personalized suggestions to users based on information available. In education, recommender systems have been employed to suggest courses to students based on their performance and ratings. The goal of this study was to recommend courses to students by taking into account their performance and interests. A knowledge-based course recommender system was developed and its effectiveness was evaluated to measure its effectiveness in recommending courses.

This knowledge-based course recommender system offers several advantages over prior course recommender systems. It provides expert level knowledge that is based on John Holland RIASEC theory and also on the expertise of the career guidance counsellor who helped map the different courses to the different RIASEC personalities. This recommender system was also built to be relevant to the Kenya locality as the courses picked are all found within the local universities. This system also saves on the cost associated with having to source for career guidance counsellor services when making course decisions during the transition from high school to tertiary level of education. This approach turns a simple course recommender into a discovery tool for both relevant RIASEC personalities and courses. It empowers students to make informed decisions about their academic plans and to discover exactly works best for them based on their interests.

## 6.2 Recommendations

This study primarily focuses on the RIASEC personalities by John Holland, which has some drawbacks, such as its inability to account for all different types of people. Future research could incorporate additional theories, such as the Myers-Briggs personality types, which distinguish more personality types based on careers. A small number of courses are available for this research. Future iterations of the project could be enhanced to incorporate all locally offered courses. Future research must also look into how to rank the different courses so that students can be given recommendations for the best courses.

### 

# REFERENCES

Osadchiy, T., Poliakov, I., Olivier, P., Rowland, M., & Foster, E. (2019). Recommender system based on pairwise association rules. *Expert Systems with Applications*, *115*, 535-542..

Alhijawi, B., & Kilani, Y. (2020). The recommender system: a survey. *Int. J. Adv. Intell. Paradigms*, *15*(3), 229-251..

Chang, Y., Shu, L., Du, E., Chen, C., Zhang, Z., Zheng, Z., ... & Xing, X. (2022). GraphRR: A multiplex Graph based Reciprocal friend Recommender system with applications on online gaming service. *Knowledge-Based Systems*, 109187.

Ghasemi, N., & Momtazi, S. (2021). Neural text similarity of user reviews for improving collaborative filtering recommender systems. *Electronic Commerce Research and Applications*, *45*, 101019..

Chen, Y. C., Hui, L., & Thaipisutikul, T. (2021). A collaborative filtering recommendation system with dynamic time decay. *The Journal of Supercomputing*, *77*(1), 244-262.

Javed, U., Shaukat, K., Hameed, I. A., Iqbal, F., Alam, T. M., & Luo, S. (2021). A review of content-based and context-based recommendation systems. *International Journal of Emerging Technologies in Learning (iJET)*, *16*(3), 274-306..

Ndung'u, R. N., Kamau, G. N., & Wambugu, G. M. (2021). Developing Hybrid-Based Recommender System with Naïve Bayes Optimization to Increase Prediction Efficiency..

Ricci, F., Rokach, L., & Shapira, B. (2022). Recommender Systems: Techniques, Applications, and Challenges. *Recommender Systems Handbook*, 1-35.

Njogu, S. W. (2019). *Analysis of factors influencing career choice among public secondary schools students in Meru County Kenya* (Doctoral dissertation, KeMU).

Newa, E. O., Kathungu, B., & Wasanga, C. (2021). Factors influencing uptake of career counselling services among university students in Nairobi City County, Kenya. *African Social Science and Humanities Journal*, *2*(1), 47-61.

Kanyingi-Maina, M. W. (2020). Influence of Career Guidance Programmes on Change of Programme of Study among First-Year Undergraduate Students in Kenyan Universities. *Journal of Education*, *3*(7).

Mumbe, K. (2020). *Institutional Factors And The Acquisition Of Employable Skills Among Students In Technical, Vocational Education And Training Institutions In Kilifi County, Kenya* (Doctoral dissertation, University of Nairobi).

Guo, Q., Zhuang, F., Qin, C., Zhu, H., Xie, X., Xiong, H., & He, Q. (2020). A survey on knowledge graph-based recommender systems. *IEEE Transactions on Knowledge and Data Engineering*.

Kovaliuk, T., Kobets, N., Pasichnyk, V., & Kunanets, N. (2021, September). Knowledge-Based Recommender Systems for the Construction of Individual Educational and Career Trajectories. In *2021 IEEE 16th International Conference on Computer Sciences and Information Technologies (CSIT)* (Vol. 2, pp. 266-272). IEEE.

Mohamed, M. H., Khafagy, M. H., & Ibrahim, M. H. (2019, February). Recommender systems challenges and solutions survey. In *2019 International Conference on Innovative Trends in Computer Engineering (ITCE)* (pp. 149-155). IEEE.

Javed, U., Shaukat, K., Hameed, I. A., Iqbal, F., Alam, T. M., & Luo, S. (2021). A review of content-based and context-based recommendation systems. *International Journal of Emerging Technologies in Learning (iJET)*, *16*(3), 274-306.

Amara, S., & Subramanian, R. R. (2020, March). Collaborating personalized recommender system and content-based recommender system using TextCorpus. In *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)* (pp. 105-109). IEEE.

Mokarrama, M. J., Khatun, S., & Arefin, M. S. (2020). A content-based recommender system for choosing universities. *Turkish Journal of Electrical Engineering and Computer Sciences*, *28*(4), 2128-2142.

Koren, Y., Rendle, S., & Bell, R. (2022). Advances in collaborative filtering. *Recommender systems handbook*, 91-142.

Yu, M., Quan, T., Peng, Q., Yu, X., & Liu, L. (2022). A model-based collaborate filtering algorithm based on stacked AutoEncoder. *Neural Computing and Applications*, *34*(4), 2503-2511.

Srifi, M., Oussous, A., Ait Lahcen, A., & Mouline, S. (2020). Recommender systems based on collaborative filtering using review texts—a survey. *Information*, *11*(6), 317.

Bhumichitr, K., Channarukul, S., Saejiem, N., Jiamthapthaksin, R., & Nongpong, K. (2019, July). Recommender Systems for university elective course recommendation. In *2019 14th international joint conference on computer science and software engineering (JCSSE)* (pp. 1-5). IEEE.

Esteban, A., Zafra, A., & Romero, C. (2020). Helping university students to choose elective courses by using a hybrid multi-criteria recommendation system with genetic optimization. *Knowledge-Based Systems*, *194*, 105385.

Khan, Z., Iltaf, N., Afzal, H., & Abbas, H. (2020). DST-HRS: A topic driven hybrid recommender system based on deep semantics. *Computer Communications*, *156*, 183-191.

Alper, A. R. I. K., Okyay, S., & Nihat, A. D. A. R. (2021). Hybrid Course Recommendation System Design for a Real-Time Student Automation Application. *Avrupa Bilim ve Teknoloji Dergisi*, (26), 85-90.

El Bouhissi, H., Adel, M., Ketam, A., & Salem, A. B. M. (2021). Towards an Efficient Knowledge-based Recommendation System. In *IntelITSIS* (pp. 38-49).

Ricci, F., Rokach, L., & Shapira, B. (2022). Recommender Systems: Techniques, Applications, and Challenges. *Recommender Systems Handbook*, 1-35.

Santamaria-Granados, L., Mendoza-Moreno, J. F., & Ramirez-Gonzalez, G. (2020). Tourist recommender systems based on emotion recognition—a scientometric review. *Future Internet*, *13*(1), 2.

Njogu, W., Kibaara, T., & Gichohi, P. (2019). How career guidance services affect career choice among public secondary school students in Meru County, Kenya. *African Journal of Emerging Issues*, *1*(8), 1-13.

Zainudin, Z. N., Rong, L. W., Nor, A. M., Yusop, Y. M., & Othman, W. N. W. (2020). The relationship of holland theory in career decision making: A systematic review of literature. *Journal of Critical Reviews*, *7*(9), 884-892.

Harahap, N. A., Amalianingsih, R., & Hidayat, D. R. (2020). Type of Personality in Making Career Decisions Based on the John L Holland Theory. *Jurnal Ilmiah Bimbingan Konseling Undiksha*, *11*(1).

Ding, Y., Wang, Q., Hourieh, N., & Yu, Q. (2020). vocational personality types in college engineering students in relation to academic achievement. *journal of employment counseling*, *57*(1), 27-47.

Coenen, J., Borghans, L., & Diris, R. (2021). Personality traits, preferences and educational choices: A focus on STEM. *Journal of Economic Psychology*, *84*, 102361.

Ja-Hwung Su, Yi-Wen Liao, Jia-Zhen Xu, Yu-Wei Zhao, A Personality-Driven Recommender System for Cross-Domain Learning Based on Holland Code Assessments, Sustainability, 10.3390/su13073936, **13**, 7, (3936), (2021).

Usslepp, N., Hübner, N., Stoll, G., Spengler, M., Trautwein, U., & Nagengast, B. (2020). RIASEC interests and the Big Five personality traits matter for life success—But do they already matter for educational track choices?. Journal of personality, 88(5), 1007-1024.

# APPENDICES

### 1. REQUIREMENTS

Hardware Requirements

1. Laptop
2. Hard disk

Software requirements

1. Vs Code environment
2. Django
3. SQLite
4. Window 10

Other requirements

1. Internet
2. Printing

### 2 BUDGET

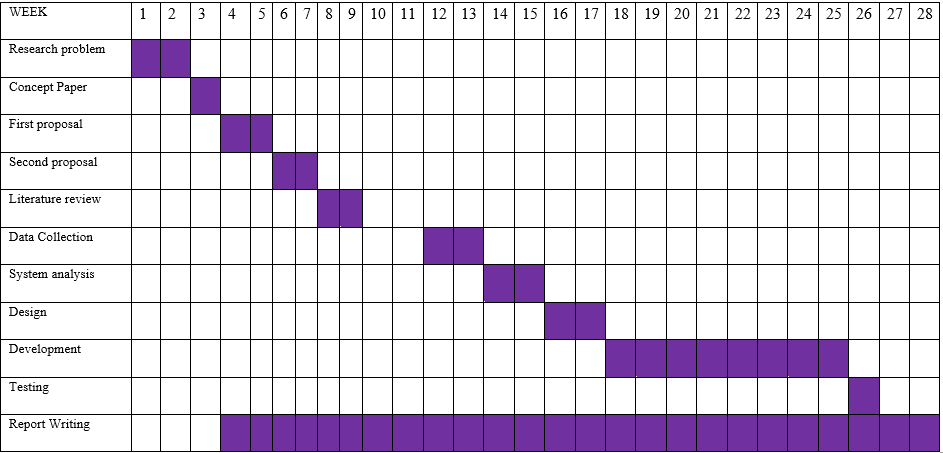
|  |  |  |  |
| --- | --- | --- | --- |
| ITEM | UNIT PRICE | QUANTITY | TOTAL |
| 1. Internet | Kes 5,000 | \_\_ | Kes 5,000 |
| 1. Printing | Kes 500 | \_\_ | Kes 500 |
| 1. Career guidance expert | Kes 7,500 | \_\_ | Kes 7,500 |
| TOTAL |  |  | Kes 13,000 |

### 3 TIMELINES

#### 3.1 Project schedule

|  |  |  |  |
| --- | --- | --- | --- |
| Task | Start Date | Completion date | Deliverable |
| Defining research problem | 23-05-2022 | 03-06-2022 | Problem statement  Feedback from supervisor |
| Concept paper | 04-06-2022 | 11-06-2022 | Notes from supervisor |
| First proposal | 12-06-2022 | 24-06-2022 | Proposal  Notes from supervisors |
| Second proposal | 25-06-2022 | 06-07-2022 | Updated proposal  Notes from supervisors |
| Literature Review | 07-07-2022 | 25-07-2022 | Literature review Document  Notes from supervisors |
| Data Collection and Analysis | 08-08-2022 | 15-08-2022 | Data |
| System Analysis | 16-08-2022 | 26-08-2022 | System requirements |
| Design | 27-08-2022 | 09-09-2022 | Use case |
| Development | 10-09-2022 | 11-11-2022 | Prototype |
| Testing | 12-11-2022 | 25-11-2022 | Working Model |
| Documentation | 12-06-2022 | 16-12-2022 | Project Documentation |

#### 3.2 Gantt Chart



### Project code

This section provides the code used to develop the system