International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 04 Apr. 2024, pp: 52-62 www.ijaem.net ISSN: 2395-5252

A Counselling System on Area of Specialisation for post-Graduate Studies in Computer Science

S. C. Laz-Eke[1,*], E. C. Ukekwe[2]
[1,2]*Department of Computer Science, University of Nigeria, Nigeria,*

Date of Submission: 25-03-2024 Date of Acceptance: 05-04-2024

Highlights:

- Using sentiment analysis todevelop a unique classification for undergraduate computer science courses under the areas of specialisation in the department.
- Creating a web application that accepts the transcript course scores of a student.
- Developing a simple computational equation to classify students into area of specialisation based on course scores.
- Predicting the most suitable area of specialisation for a student based on the result of the equation.

ABSTRACT:
The importance of academic counselling cannot be over-emphasised especially as it concerns Computer Science. This work identifies the suitable area of specialisation for Post-graduate students of the Computer Science Department, University of Nigeria, Nsukka by using the results obtained from their undergraduate transcript results. Python Natural Language Processing (NLP) was carried out on the undergraduate courses to classify the courses under the respective areas of specialisation. Subsequently, the percentage scores obtained on the respective courses were used to determine the most suitable area of specialisation for the postgraduate students. Hence, a computer system was designed for this purpose. The system makes use of the Structured System Analysis and Design Methodology (SSADM) and is implemented using Visual Studio Code 1.58.2 IDE on a Windows 7 system using HTML, CSS and JavaScript for the front-end, PHP for the back-end, and MySQL database technologies.

I. INTRODUCTION

Post-graduate studies imply further education after graduating with a first degree from a university and entails studying and/or doing research at a more advanced level. It can be a Master's or Doctorate degree. The same is applicable in the department of Computer Science, University of Nigeria, Nsukka.

The department runs a course-based Master's postgraduate studies in five major areas of specialisation which are: systems engineering, software engineering, artificial intelligence, computer networks, and theoretical computer science.

In the undergraduate level, there are 45 computer science courses over the four years of study with two semesters each. In year one, there are 4 major courses offered; In year two, they are 7; In the third year, there are 10 courses; and in the fourth, there are 10 major courses and 12 elective courses. The remaining 2 are service courses to other faculties in the university.

These courses do not have a definite area of specialisation as they cut across different fields. Their contents are in line with NUC (National Universities Commission) and popular/world-wide standards.

In addition, the traditional approach of counselling has been to just browse the internet for ideas (some educational websites and some universities' websites give tips), ask a lecturer or professional (if possible), or look at the area that gives the most money.

These may not look in-depth into the strengths of the student, based on his or her undergraduate result, and may affect the student in future as he or she may not excel best in the chosen field, and/or may not find it fulfilling.

This problem implies that, the society may have square pegs in round holes among so-called masters of computer science. The standard of education/higher institutions may drop overall, as lecturer to student interactions occur.

In the words of the authors in [1], the early classification of university students according to their potential academic performance can be a useful strategy to mitigate failure, to promote the

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 04 Apr. 2024, pp: 52-62 www.ijaem.net ISSN: 2395-5252

achievement of better results and to better manage resources in higher education institutions.

Therefore, seeing the need, it will be solved as follows:

This work classifies all the computer science undergraduate courses into these categories, and uses it to find the top three most suitable based on the results of the student by calculating the average percentage of each and getting the highest three.

In section 1, the concept and the objectives of the project are being introduced. In section 2, the literature of related/similar works is reviewed. In section 3, the system's architecture is analysed and designed. In section 4, the system proper is implemented and then the summary and recommendations are given in section 5.

This system will be best used by the department of computer science, to counsel its prospective post-graduate students.

II. SYSTEM ANALYSIS AND DESIGN

2.1Introduction

Here, the structured system analysis and design methodology (SSADM) is used. It is a set of standards that uses a formal methodical approach to the analysis and application design of information systems. SSADM is based on the waterfall model and it is open to interpretation[2].

2.2 Description of the Existing System

An exact existing system truly does not exist at the time of writing, but a similar one is:

- Browsing the internet for ideas.
- Contacting a lecturer/industry professional for advice.
- Contacting past master's students of the department for advice.

After careful analysis, the following problems were identified in the system:

- No personalised result/output/advice from the system.
- Relative unavailability of past master's students.
- Inconvenience to lecturers/industry professionals.

Not enough will to contact these entities for advice, which in turn may lead to following the crowd.

2.3 Analysis of the Proposed System

2.3.1 Course Classification

The department of computer science in the University of Nigeria offers five areas of specialisation at master's level:

- Systems Engineering
- Software Engineering
- Artificial Intelligence
- Computer Networks, and
- Theoretical Computer Science[3].

The forty-one (41) undergraduate courses offered by the department that are being used are:

COS101 Introduction to Computer Science for Other Faculties (2 units)
COS 102 Computing Practice (2 units)
COS 103 Computer Hardware Organization (2 units)
COS 104 Introduction to Database System (2 units)
COS 105 Introduction to Computer Science for Physical Sciences (2 units)
COS 201 Computer Programming (2 units)
COS 202 Software Engineering I (2 units)
COS 203 Introduction to Microcomputer Systems (2 units)
COS 204 Introduction to Digital System Design (2 units)
COS 231 Assembly Language Programming (2 units)
COS 232 Data Structure (2 units)
COS 242 Data and Computer Communications (2 units)
COS 304 Computer Applications (2 units)
COS 311 Switching Algebra and Discrete Structures (3 units)
COS 331 Operating Systems (3 units)
COS 333 Software Engineering II (2 units)
COS 335 Automata Theory and Formal Languages (2 units)
COS 337 Artificial Intelligence I (2 units)
COS 341 Computer Architecture (3 units)
COS 351 Laboratory for Digital System Design (3 units)
COS 411 Numerical Methods (3 units)
COS 413 System Modeling and Simulation (2 units)
COS 415 Computer Process Control (2 units)
COS 417 Computer System Performance Evaluation (2 units)
COS 419 Operation Research (3 units)
COS 421 Database Design and Management (3 units)
COS 431 Algorithms (2 units)
COS 434 Compiler Construction (3 units)
COS 435 Computer Graphics and Animation (2 units)
COS 436 Expert Systems (2 units)
COS 437 Project Management (2 units)

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 04 Apr. 2024, pp: 52-62 www.ijaem.net ISSN: 2395-5252

COS 438 Artificial Intelligence II (2 units)
COS 439 Games Programming (2 units)
COS 441 Advanced Computer Networks (3 units)
COS 442 Mobile Communication (2 units)
COS 444 Computer Network Security (2 units)
COS 452 Advanced Digital Laboratory (3 units)
COS 461 Organization of Programming Languages (2 units)
COS 463 Structured Programming (2 units)
COS 464 Concurrent Programming (2 units)
COS 471 Web Application Development (2 units)

Excluded are:
COS 490 Project (6 units)
COS 382 Students Industrial Work Experience (6 units)
COS 384 Technical SIWES Report (5 units)
COS 386 SIWES Seminar (4 units)
[4]

2.3.1.1 Methodology

A sentiment analysis model was used to carry out the classification.

A file, named course_contents.txt, contained all the courses and their contents from the departmental handbook, [4], like thus:

<course title><course unit><course content> = <course code> $

A snippet of the file is shown in Figure 3.1 below.

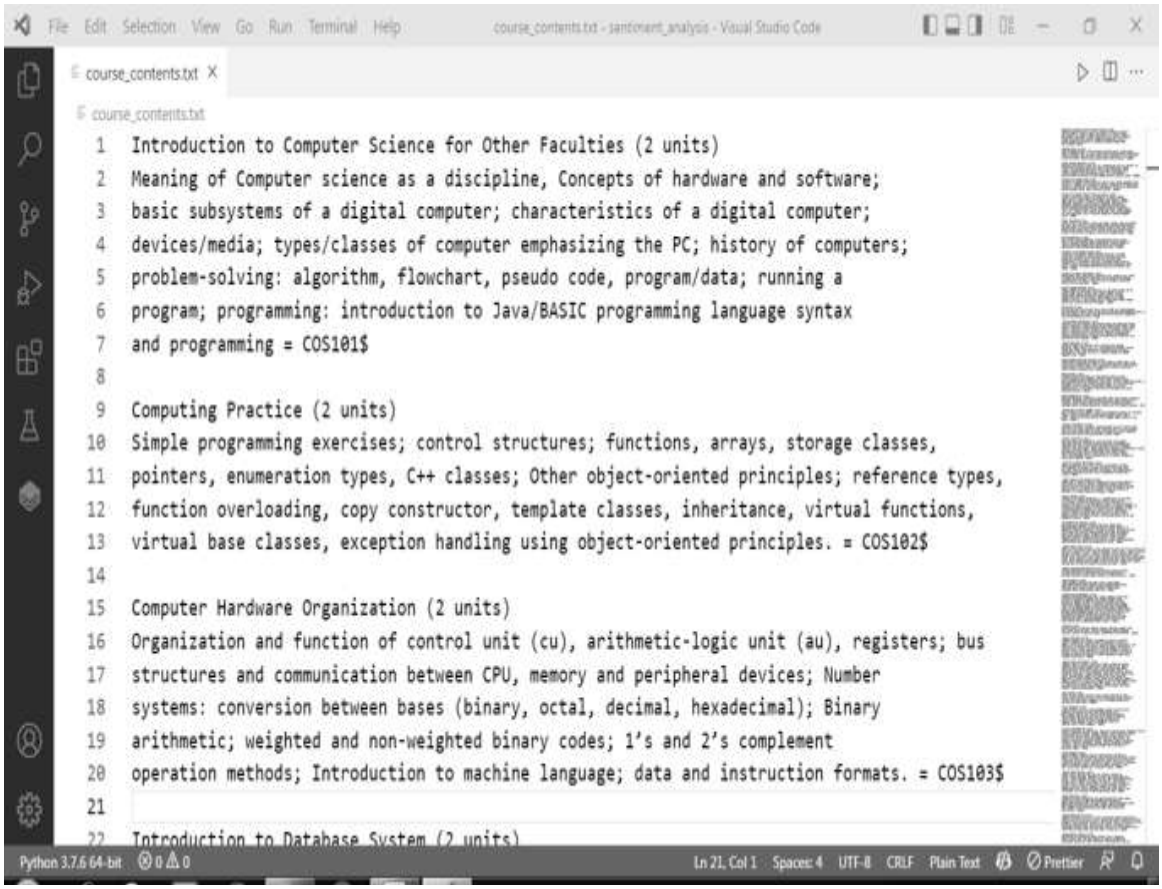

Figure 3.1: course_contents.txt

Next, the contents [3]of each area of specialisation are put in separate files. This is shown in Figure 3.2 below.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 04 Apr. 2024, pp: 52-62 www.ijaem.net ISSN: 2395-5252

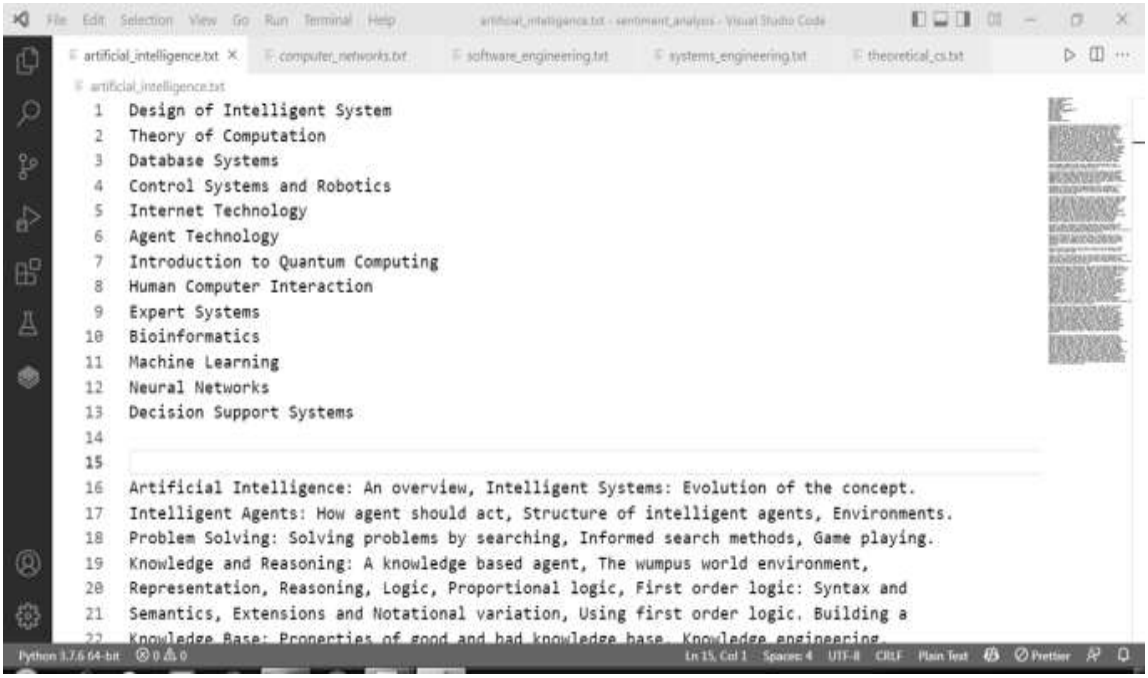

Figure 3.2: Areas of Specialisation Contents

Next, using Python programming language, the simple natural language processing for sentiment analysis procedures are followed for all the files:

- The file is opened/read into the program in utf-8 format.
- All its texts are changed to lowercase.
- Punctuation marks (except '=', '-', and '$') are removed. The excluded ones are for manipulating the texts.
- The words in the text are tokenised.
- Custom stop words are removed. These included too common/vague words that will affect the course similarity count, like 'introduction', 'computer', 'eg', 'etc', 'units'.
- Then the final words are stored in a list.

Now, for each area of specialisation, the code compares the number of similar words with each course, creates a dictionary of the similarity count, and plots its graph, as shown in Figure 3.3, 3.4, 3.5, 3.6, and 3.7.

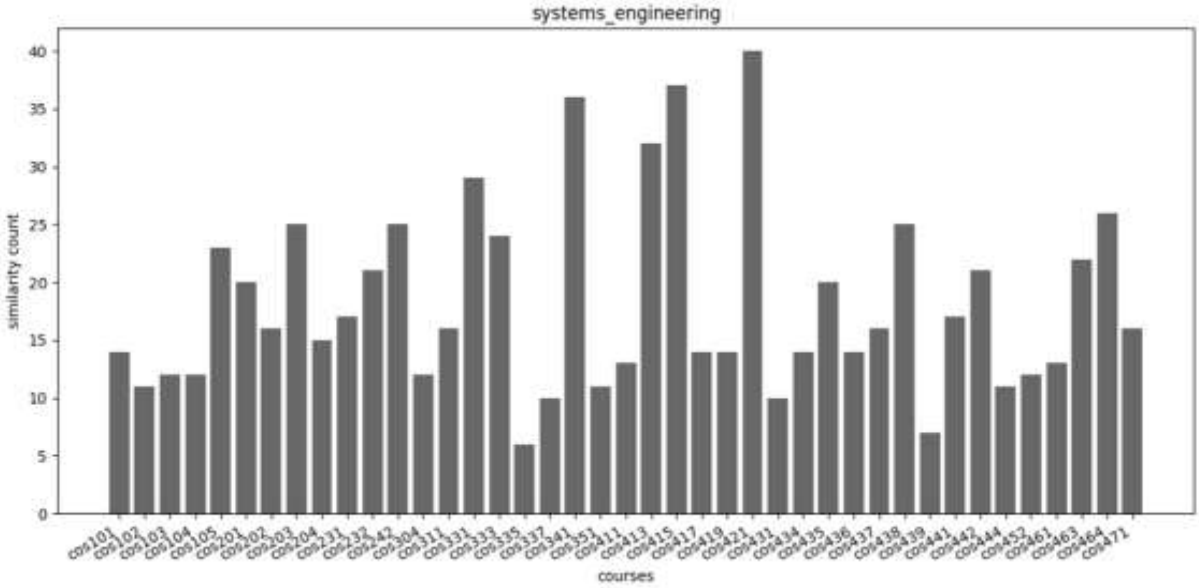

Figure 3.3: Systems Engineering Graph

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 04 Apr. 2024, pp: 52-62 www.ijaem.net ISSN: 2395-5252

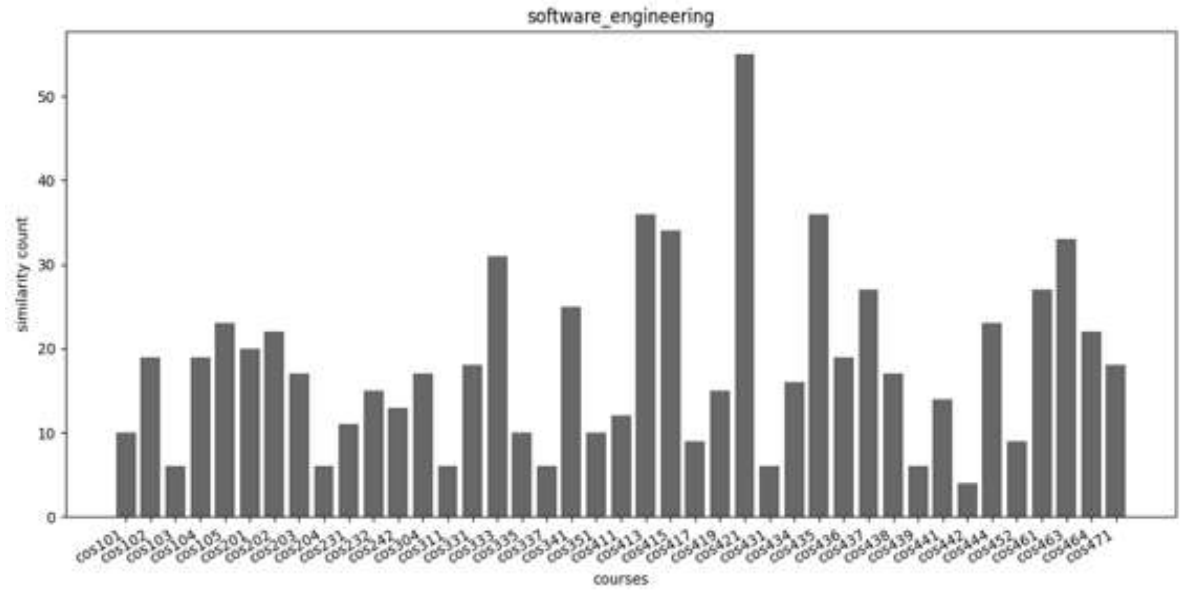

Figure 3.4: Software Engineering Graph

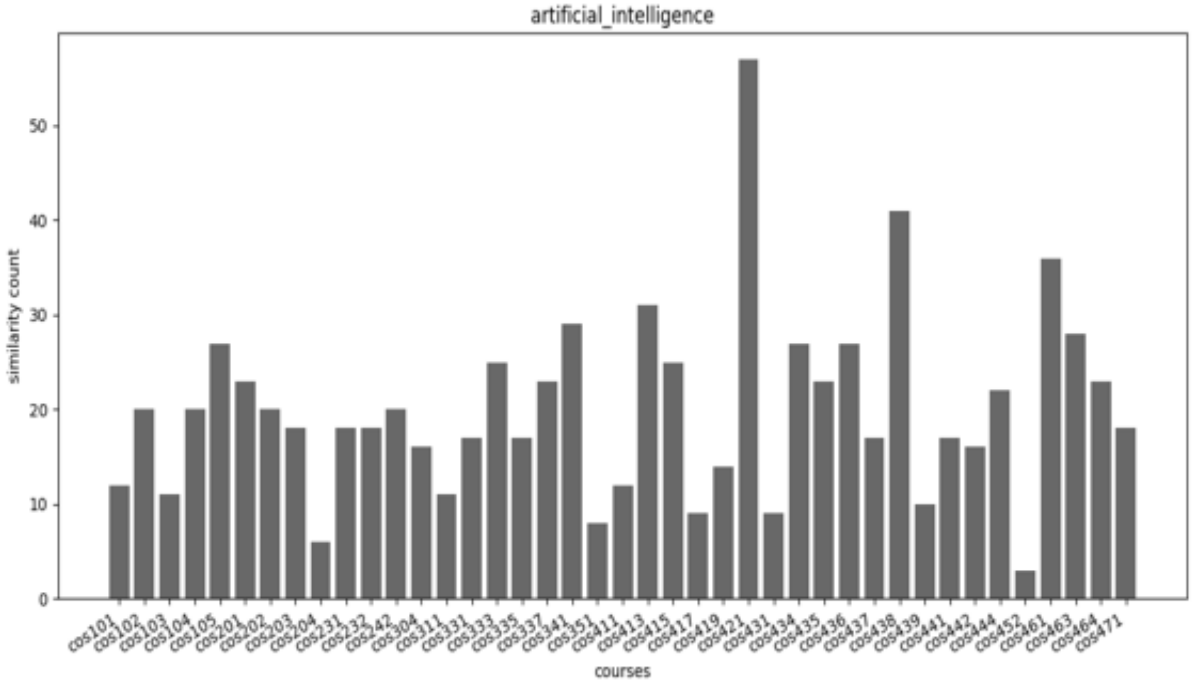

Figure 3.5: Artificial Intelligence Graph

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 04 Apr. 2024, pp: 52-62 www.ijaem.net ISSN: 2395-5252

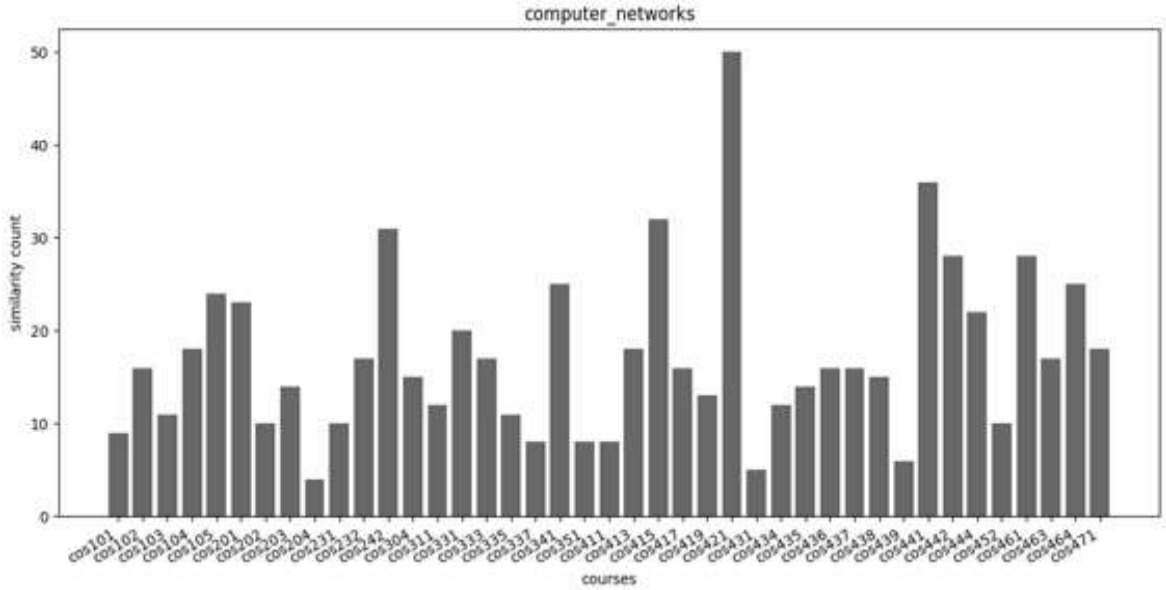

Figure 3.6: Computer Networks Graph

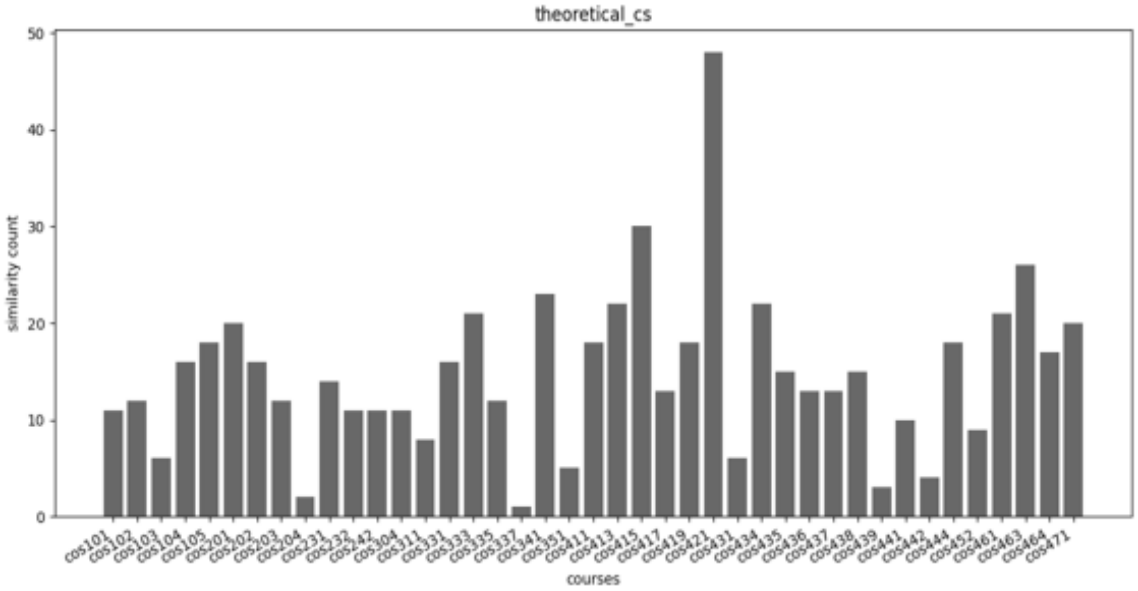

Figure 3.7: Theoretical Computer Science Graph

Then, the courses that pass the 70th percentile (grade A in the University's scoring) of each graph are chosen for that area of specialisation.

The arithmeticformula is: percentile (index) = 70% × number of course similarity counts

Every course that then has a similarity count greater than or equal to the count in the percentile (index) position, when arranged in ascending order, is recorded.

The result of these classifications is as shown in Table 3.1, 3.2 and 3.3 below.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 04 Apr. 2024, pp: 52-62 www.ijaem.net ISSN: 2395-5252

Table 3.1: Course Classification for Systems and Software Engineering

Systems Engineering [14 courses]	**Software Engineering [13 courses]**
COS 421: Database Design and Management(40)	COS 421: Database Design and Management (55)
COS 415: Computer Process Control (37)	COS 413: System Modeling and Simulation (36)
COS 341: Computer Architecture(36)	COS 435: Computer Graphics and Animation (36)
COS 413: System Modeling and Simulation (32)	COS 415: Computer Process Control (34)
COS 331: Operating Systems (29)	COS 463: Structured Programming (33)
COS 464: Concurrent Programming (26)	COS 333: Software Engineering II (31)
COS 203: Introduction to Microcomputer Systems (25)	COS 437: Project Management (27)
COS 242: Data and Computer Communications (25)	COS 461: Organization of Programming Languages (27)
COS 438: Artificial Intelligence II (25)	COS 341: Computer Architecture (25)
COS 333: Software Engineering II (24)	COS 105: Introduction to Computer Science for Physical Sciences (23)
COS 105: Introduction to Computer Science for Physical Sciences (23)	COS 444: Computer Network Security (23)
COS 463: Structured Programming (22)	COS 202: Software Engineering I (22)
COS 232: Data Structure (21)	COS 464: Concurrent Programming (22)
COS 442: Mobile Communication (21)	

Table 3.2: Course Classification for Artificial Intelligence and Computer Networks

Artificial Intelligence [15 courses]	**Computer Networks [15 courses]**
COS 421: Database Design and Management (57)	COS 421: Database Design and Management (50)
COS 438: Artificial Intelligence II (41)	COS 441: Advanced Computer Networks (36)
COS 461: Organization of Programming Languages (36)	COS 415: Computer Process Control (32)
COS 413: System Modeling and Simulation (31)	COS 242: Data and Computer Communications (31)
COS 341: Computer Architecture (29)	COS 442: Mobile Communication (28)
COS 463: Structured Programming (28)	COS 461: Organization of Programming Languages (28)
COS 105: Introduction to Computer Science for Physical Sciences(27)	COS 341: Computer Architecture (25)
COS 434: Compiler Construction (27)	COS 464: Concurrent Programming (25)

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 04 Apr. 2024, pp: 52-62 www.ijaem.net ISSN: 2395-5252

COS 436: Expert Systems (27)	COS 105: Introduction to Computer Science for Physical Sciences (24)
COS 333: Software Engineering II (25)	COS 201: Computer Programming (23)
COS 415: Computer Process Control (25)	COS 444: Computer Network Security (22)
COS 201: Computer Programming (23)	COS 331: Operating Systems (20)
COS 337: Artificial Intelligence I (23)	COS 104: Introduction to Database System (18)
COS 435: Computer Graphics and Animation (23)	COS 413: System Modeling and Simulation (18)
COS 464: Concurrent Programming (23)	COS 471: Web Application Development (18)

Table 3.3: Course Classification for Theoretical Computer Science

Theoretical Computer Science [14 courses]	
COS 421:	Database Design and Management (48)
COS 415:	Computer Process Control (30)
COS 463:	Structured Programming (26)
COS 341:	Computer Architecture (23)
COS 413:	System Modeling and Simulation (22)
COS 434:	Compiler Construction (22)
COS 333:	Software Engineering II (21)
COS 461:	Organization of Programming Languages (21)
COS 201:	Computer Programming (20)
COS 471:	Web Application Development (20)
COS 105:	Introduction to Computer Science for Physical Sciences (18)
COS 411:	Numerical Methods (18)
COS 419:	Operation Research (18)
COS 444:	Computer Network Security (18)

Though COS 101, COS 102, COS 232, and COS 431can be regarded as thebasics of computer science, they did not meet the cut-off for any of the areas, hence did not appear, except COS 232 in A.I. For more insight, shown in Table 3.4 are all the courses and the area of specialisation where they have the most similarity count i.e

$$C_c = \max_{s=1}^{s=5} C_s$$

Where:
C_c is max similarity count of a course
C_s is similarity count of a course in an area of specialisation

Table 3.4: Courses and Most Similar Area of Specialisation

Course (C_c)	**Area of Specialisation**
COS 421: Database Design and Management (57)	Artificial Intelligence
COS 438: Artificial Intelligence II (41)	Artificial Intelligence
COS 415: Computer Process Control (37)	Systems Engineering
COS 341: Computer Architecture (36)	Systems Engineering
COS 413: System Modeling and Simulation (36)	Software Engineering
COS 435: Computer Graphics and Animation (36)	Software Engineering
COS 441:Advanced Computer Networks (36)	Computer Networks
COS 461: Organization of Programming Languages (36)	Artificial Intelligence

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 04 Apr. 2024, pp: 52-62 www.ijaem.net ISSN: 2395-5252

COS 463:Structured Programming (33)	Software Engineering
COS 242: Data and Computer Communications (31)	Computer Networks
COS 333:Software Engineering II (31)	Software Engineering
COS 331: Operating Systems (29)	Systems Engineering
COS 442: Mobile Communication (28)	Computer Networks
COS 105: Introduction to Computer Science for Physical Sciences (27)	Artificial Intelligence
COS 434: Compiler Construction (27)	Artificial Intelligence
COS 436: Expert Systems (27)	Artificial Intelligence
COS 437: Project Management (27)	Software Engineering
COS 464: Concurrent Programming (26)	Systems Engineering
COS 203: Introduction to Microcomputer Systems (25)	Systems Engineering
COS 201: Computer Programming (23)	Artificial Intelligence/Computer Networks
COS 337: Artificial Intelligence I (23)	Artificial Intelligence
COS 444: Computer Network Security (23)	Software Engineering
COS 202: Software Engineering I (22)	Software Engineering
COS 232: Data Structure (21)	Systems Engineering
COS 102: Computing Practice (20)	Artificial Intelligence
COS 104: Introduction to Database System(20)	Artificial Intelligence
COS 471: Web Application Development (20)	Theoretical Computer Science
COS 231: Assembly Language Programming (18)	Artificial Intelligence
COS 411: Numerical Methods (18)	Theoretical Computer Science
COS 419: Operation Research (18)	Theoretical Computer Science
COS 304: Computer Applications (17)	Software Engineering
COS 335: Automata Theory and Formal Languages (17)	Artificial Intelligence
COS 311: Switching Algebra and Discrete Structures (16)	Systems Engineering
COS 417: Computer System Performance Evaluation (16)	Computer Networks
COS 204: Introduction to Digital System Design (15)	Systems Engineering
COS 101: Introduction to Computer Science for Other Faculties (14)	Systems Engineering
COS 103: Computer Hardware Organization (12)	Systems Engineering
COS 452: Advanced Digital Laboratory (12)	Systems Engineering
COS 351: Laboratory for Digital System Design (11)	Systems Engineering
COS 431: Algorithms (10)	Systems Engineering
COS 439: Games Programming (10)	Artificial Intelligence

The weight of each course does not henceforth affect or influence the resulting computations that follow. Python is also no more used henceforth.

2.3.2 Computation Equation and Procedure

The computation is done in PHP language at the backend of the web application.

It is simple: the result scores of all the courses are summed in the five categories, stored in an array, sorted in descending order, and the three areas of specialization corresponding to the first three scores (highest three) is outputted as the most-suited areas of specialization for the student.

The computations performed are as follows:

a. The average

The arrays contain the scores in those courses.

systems_engineering_array = [cos421, cos415, cos341, cos413, cos331, cos464, cos203, cos242, cos438, cos333, cos105, cos463, cos232].

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 04 Apr. 2024, pp: 52-62 www.ijaem.net ISSN: 2395-5252

software_engineering_array = [cos421, cos413, cos435, cos415, cos463, cos333, cos437, cos461, cos341, cos105, cos444, cos202, cos464, cos201].
artificial_intelligence_array = [cos421, cos438, cos461, cos413, cos341, cos463, cos105, cos434, cos436, cos333, cos415, cos201, cos337, cos435, cos464].
computer_networks_array = [cos421, cos441, cos415, cos242, cos442, cos461, cos341, cos464, cos105, cos201, cos444, cos331, cos104, cos413, cos471].
theoretical_computer_science_array = [cos421, cos415, cos463, cos341, cos413, cos434, cos333, cos461, cos201, cos471, cos105, cos411, cos419, cos444].

The arithmetic average of each array is calculated thus:

systems_engineering_average = $\frac{\sum_{i=1}^{n} systems\_engineering\_array_i}{n}$ ----- (1)

where n is the size of system_engineering_array, in this case 14

software_engineering_average = $\frac{\sum_{i=1}^{n} software\_engineering\_array_i}{n}$ ----- (2)

where n is the size of software_engineering_array, in this case 13

artificial_intelligence_average = $\frac{\sum_{i=1}^{n} artificial\_intelligence\_array_i}{n}$ ----- (3)

where n is the size of artificial_intelligence_array, in this case 15

computer_networks_average = $\frac{\sum_{i=1}^{n} computer\_networks\_array_i}{n}$ ----- (4)

where n is the size of computer_networks_array, in this case 15

theoretical_computer_science_average = $\frac{\sum_{i=1}^{n} theoretical\_computer\_science\_array_i}{n}$ ----- (5)

where n is the size of theoretical_computer_science_array, in this case 14

An associative array is formed thus:
AOS = array("Systems Engineering" => systems_engineering_average, "Software Engineering"
=>software_engineering_average,"Artificial Intelligence" =>
artificial_intelligence_average , "Computer Networks" =>
computer_networks_average, "Theoretical Computer Science" =>
theoretical_computer_science_average);

b. Decision Criterion
The three categories with the highest averages are selected in descending order as the best areas of specialization.

The program implementation is as follows:
All the categories and their corresponding averages are stored in an array, AOS, as shown above. Then, the array is sorted in descending order. Then the array is sliced after the third element and stored in a new array, first_3_areas, as shown below. This then represents the best areas of specialisation.
PHP Program:

```
arsort(AOS); //Descending/reverse sort.

length = 3; start = 0;
first_3 = array_slice (AOS, start, length, true);
first_3_areas = array_keys (first_3);

area_1 = first_3_areas[0];
area_2 = first_3_areas[1];
area_3 = first_3_areas[2];
score_1 = first_3[area_1];
score_2 = first_3[area_2];
score_3 = first_3[area_3];
```

2.3.3 Database Design

The database management system used is MySQL. Only one table is required: the Prediction table. The table contains the student's registration number (primary key); the first, second, and third best areas of specialisation and their corresponding weight scores; and the marks scored in all the individual undergraduate courses.

III. RECOMMENDATIONS

The prediction can be upgraded to be done using a machine learning/artificial intelligence approach instead of the arithmetic computation. The data to be used could be those presented in the excel sheet where a sample data has its label field populatedwith the most similar area of specialisation that the highest course score of each record corresponds to. The code of the normally implemented system is also open-source at github.com/Chinechelum/Final-Year-Project.

Correspondence
All correspondence should be sent to: Stephen Laz-Eke, Graduate, Department of

Computer Science, University of Nigeria, Nsukka, Enugu State, NIGERIA. e-mail: stephen.lazeke.244189@unn.edu.ng, Mobile: +234-816-139-0641.

REFERENCES

[1] Miguéis, V.L.,Freitas, A.,Garcia, P.J.V. and Silva, A. "Early Segmentation of Students According to Their Academic Performance: A Predictive Modelling Approach", Decision Support Systems, Vol. 115, 2018, pp. 36-51, (https://www.sciencedirect.com/science/article/pii/ S0167923618301428), https://doi.org/10.1016/j.dss.2018.09.001.

[2] Techopedia. "Structured Systems Analysis and Design Method (SSADM)" https://www.techopedia.com/definition/3983/ structured-systems-analysis-and-design-method-ssadm, Accessed on 25th Oct., 2021.

[3] Department of Computer Science, University of Nigeria, Nsukka."Postgraduate Study Programme Manual", pp.44–60,http://spgs.unn.edu.ng/wp-content/uploads/sites/35/2016/02/COMPUTER-SCIENCE.pdf, Accessed on 25th Oct., 2021.

[4] Department of Computer Science, University of Nigeria, Nsukka."B.Sc Computer Science Programme", Students' Information Handbook (2019-2024). Mosco Printers.