Implementing Genetic Algorithm-Based Expert Systems to Enhance Academic Career Planning

Abderrahim El Yessefi [1,*], Loubna Cherrat [2], Mohammed Rida Ech-Charrat [3], and Mostafa Ezziyyani [1]

[1]Laboratory of Mathematics and Applications, Faculty of Sciences and Techniques, Abdelmalek Essaadi University, Tangier, Morocco
[2]National School of Commerce and Management, Abdelmalek Essaadi University, Tangier, Morocco
[3]National School of Applied Sciences, Abdelmalek Essaadi University, Tetouan, Morocco
Email: abderrahim.elyessefi@gmail.com (A.E.Y.); cherratloubna2@gmail.com (L.C.); charrat.mohammed@gmail.com (M.R.E.-C.); mezziyyani@uae.ac.ma (M.E.)
*Corresponding author

Manuscript received March 8, 2024; revised June 7, 2024; accepted August 26, 2024; published October 22, 2024

***Abstract*—The integration of technology in education, particularly through the combination of adaptive learning and Artificial Intelligence (AI), has significantly transformed teaching and learning methodologies. This paper addresses the challenges associated with manually planning study paths and modules at universities, including the ordering of relationships between modules across various disciplines. It considers the prerequisites and competencies required for each module and highlights several limitations, such as the extensive time and resources required, subjectivity in decision-making, and the complexities introduced by the LMD system. To address these challenges, a decision support system utilizing genetic algorithms is proposed to increase the efficiency and objectivity of academic career planning. The system analyzes a dataset that includes information on study units, modules, competencies, and learning objectives. Genetic algorithms, which incorporate mutation, crossover, and selection processes, are used to generate coherent and suitable study paths. The system was tested using data from the Faculty of Science and Technology at Tangier, demonstrating its capacity to adapt existing academic pathways and propose enhanced alternatives with improved consistency. The results indicate the potential of the system to enhance the quality of academic career planning within the modular educational framework.**

***Keywords*—artificial intelligence, genetic algorithms, academic career planning, education, decision support systems**

I. Introduction

In recent years, education systems have undergone a profound transformation with the integration of technology, which has fundamentally changed the dynamics of teaching and learning. In particular, the convergence of adaptive learning and Artificial Intelligence (AI) represents a crucial advancement in this transformation. Adaptive learning-which involves using technology to tailor educational content and experiences [1], has emerged as a key driver for enhancing the effectiveness and inclusivity of education on a global scale.

However, at the heart of this technological revolution, the planning of study programs and pathways has become a complex and multifaceted challenge on an international scale. This complexity is heightened by several interconnected factors, necessitating a comprehensive and globally informed approach for educational planning. Notably, the rapidly evolving demands of the global job market require educational planning to address disparities in skill development across regions. Bridging the gap between education and industry needs is crucial for developing a workforce capable of addressing future challenges.

Currently, the planning of academic programs and modules at Moroccan universities is conducted manually by pedagogical authorities, following guidelines set by the ministry and university administration. However, this has several limitations. First, it requires an increased pedagogical workforce and a substantial amount of time to collect and analyze information before decisions can be made. Second, this method lacks objectivity, as pedagogical staff may be influenced by their own experiences. Additionally, the implementation of the Licence, Master, Doctorat (LMD) system in Moroccan universities introduces another challenge. If a student fails one or more modules, they are required to study new modules concurrently with the failed ones in the subsequent semester. Typically, these new modules are constructed based on the content of the previous modules in which the student encountered difficulties. Addressing this complexity is challenging and can potentially lead to delays in students' progression through their studies.

An improved genetic algorithm model can significantly enhance the efficiency of academic scheduling by optimizing course sequences and reducing scheduling conflicts. This technique, which utilizes a variable-length decimal coding scheme and local search operators, addresses the inherent complexities in arranging large-scale academic schedules. By enhancing the genetic operators and refining the fitness functions, this method has been shown to improve average fitness values and reduce convergence times, making it ideal for dynamic and complex academic environments [2].

To overcome these challenges, we have implemented a decision support system, designed to orchestrate academic pathways. This system analyzes a comprehensive dataset that includes information about study units, modules, competencies, and learning objectives. The use of genetic algorithm techniques allows us to construct more coherent and suitable academic plans.

Our system offers several advantages over the manual process. It enhances efficiency by significantly reducing the time and resources required for planning. Additionally, it introduces a higher level of objectivity, achieved through the utilization of genetic algorithms, which are inherently based on objective criteria.

The development of a decision support system for planning academic pathways at Moroccan universities was intended to enhance the overall effectiveness and efficiency of the planning process. This innovation is instrumental in preventing planning gaps and saving considerable time.

This paper makes a significant contribution to the field of educational technology by developing and implementing a genetic algorithm-based system for creating personalized academic pathways.

Traditional learning pathways often follow fixed sequences, failing to accommodate the diverse needs and abilities of individual students. This research addresses the necessity of developing dynamic academic pathways using genetic algorithms, which consider factors such as topic difficulty level, relation degree, and student performance metrics [3]. By leveraging the power of genetic algorithms, our study addresses the complex challenge of course sequencing and curriculum planning in a way that optimizes both educational outcomes and student satisfaction. The proposed system dynamically adapts to individual student preferences and performance data, thus offering a valuable tool for educational institutions seeking to modernize and improve their curriculum delivery methods. This work not only advances the theoretical understanding of adaptive learning systems but also presents a practical solution that can be readily implemented in diverse educational settings.

II. Related Works

The planning and adaptation of university studies have been the focus of various research studies exploring the use of advanced technologies to optimize learning processes. Among these studies, some stand out for their innovative approaches and practical applicability.

Rollande and Grundspenkis [4] developed a graph-based framework for personalized study planning in higher education. They propose a prototype system, the Study Planning System (SPS), which allows learners to design individual study programs. The framework contains four types of graphs that represent different aspects of study programs and courses, facilitating personalized learning paths and study course designs. This approach aims to enhance student engagement and adaptability in learning processes through intelligent tutoring systems and personalized education models.

Sargsyan and Hovakimyan [5] introduced an approach for building adaptive learning tools in pervasive learning systems. They present the Genetic Chooser Algorithm (GCA), which constructs user-adaptable teaching scenarios based on genetic algorithms. These scenarios are developed by considering the quality and quantity of teaching units and the user's knowledge. The paper also explores a course-map building tool within GCA, which helps users track their progress through a course. The focus is on optimizing the learning process and tailoring it to individual learners's needs via genetic algorithms.

Bohadean [6] addresses the optimization of university course scheduling using micro-genetic algorithms. The study focuses on managing both hard and soft constraints in timetable generation, proposing a novel chromosome representation method to minimize infeasible solutions. The research demonstrates the effectiveness of the micro-genetic algorithm in producing feasible and efficient course schedules, with UNITAR International University as the case study. The results show improved performance in term of timetable generation, highlighting the potential of micro-genetic algorithms in solving complex scheduling problems in academic settings.

Perzin [7, 8] focused on optimizing university timetabling using self-adaptive genetic algorithms. Perzin [5] introduces a method for adjusting timetables with minimal changes to accommodate new requirements, featuring a unique encoding for algorithm adaptation and a multicriteria evaluation system. Perzin [8] emphasizes individual student schedule optimization, addressing unique scheduling needs while minimizing clashes. Both papers highlight the efficiency of self-adaptive genetic algorithms in creating effective timetables within complex educational environments.

Hssina and Erritali [9] discusses the development of an adaptive e-learning platform using genetic algorithms. The research focuses on creating individualized learning paths for students in an e-learning environment on the basis of their unique profiles and educational goals. The paper details the approach and architecture of the adaptive system, highlighting how it generates learning paths tailored to student profiles and pedagogical objectives. The study underscores the value of genetic algorithms in automating the customization of educational content, aiming to make learning more adaptive and effective.

Chang and Ke [10] presented a method for creating personalized e-courses using a genetic algorithm. This approach, designed for adaptive learning systems, focuses on efficiently finding appropriate e-learning materials for individual learners. The key feature of this method is the "forcing legality" operation, which not only streamlines the search process but also ensures more accurate results in e-course composition. The paper demonstrates the effectiveness of this approach through experiments, highlighting its potential in adaptive learning environments.

Dhaiouir *et al.* [11] proposed a smart model for classifying and orienting learners in Massive Open Online Courses (MOOCs). This paper addresses the challenge of managing diverse learner profiles in MOOCs by utilizing artificial intelligence and machine learning techniques, particularly neural networks, to categorize learners effectively.

These studies collectively highlight the versatility and effectiveness of genetic algorithms and adaptive systems in personalized education, making significant contributions to the fields of academic career planning and e-learning. However, during our research, we did not find work similar to what we propose in this paper that is, work aimed at replanning university training according to specific criteria. Hence, we hope that this work will add value to the field of adaptation in higher education.

However, during our research, we did not find any studies that specifically address the type of work we propose in this paper, which involves replanning university training according to specific criteria.

III. Conceptual Foundations and Structural Overview

A. Description of the Current State of Studies at the Faculty of Sciences and Technologies of Tangier (FSTT)

Moroccan universities have undergone a significant transformation with the adoption of the Licence, Master, Doctorat (LMD) system, a higher education framework closely aligned with the European model [12]. This strategic shift aims to modernize and standardize higher education by introducing greater flexibility and enhancing the overall quality and competitiveness of academic degrees. The LMD system represents not just a structural change but a comprehensive approach to improving educational provision.

The “Licence” phase, constituting the initial university cycle, spans three years and provides students with foundational training in a specific field. This stage is crucial for preparing graduates to either seamlessly transition into the workforce or pursue further education in the subsequent “Master” cycle. The “Master” level, representing the second cycle, offers advanced and specialized training over two years. It equips graduates to enter the job market as managers or engineers, or to engage in academic research during the following “Doctoral” cycle, thereby fostering a culture of scientific inquiry.

Within the FSTT, the “License” program takes place over three years, with each year organized into two semesters. Each semester challenges students to undertake six modules, ensuring a comprehensive and well-rounded educational experience. Moreover, FSTT offers four distinct career paths during the first two years: MIPC (Mathematical, Informatic, Physics, Chemistry), MIP (Mathematical, Informatic, Physics), GEGM (Electrical Engineering, Mechanical Engineering), and BCG (Biology, Chimestry, Geology). As students progress to the third year, they are presented with a diverse array of over 17 specializations, allowing for a more tailored and personalized academic journey.

This strategic alignment with the LMD system not only demonstrates a commitment to offering a robust and adaptable educational framework but also positions Moroccan universities on the global stage. It fosters international collaboration and ensures that graduates are well-prepared to meet the demands of an evolving global workforce.

B. Disciplines, Modules, and Their Relationships

We focus on the career training options offered at FSTT during the two-year “common core” phase, which enables students to obtain a DEUST (Diploma of University Studies in Science and Technology). The main objective of the common core is to provide students with a solid foundation in sciences, including Mathematics, Physics, Chemistry, and Biology. The DEUST diploma allows students to either continue their studies to obtain a “Licence” diploma in the third year or to apply for admission to engineering programs within the same faculty, as well as to participate in national or international competitions for entry into renowned engineering schools.

Fig. 1 shows the relationships between prerequisites, modules, and competencies. The specific prerequisites and competencies listed in the diagram will vary depending on the module.

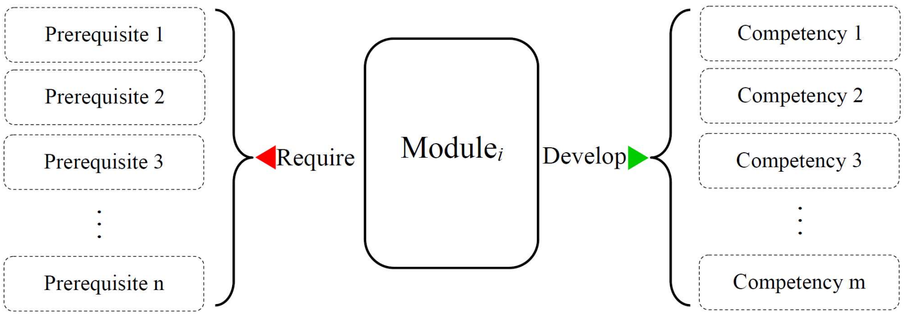


Fig. 1. Illustration of a module outlining its prerequisites and competencies.

C. Methodology

Our approach to adaptive career planning involves methodologies similar to those described by Chang *et al.* [13], who effectively used genetic algorithms to tailor e-course compositions, ensuring a high degree of personalization and efficiency in course delivery. Pallathadka *et al.* [14] confirmed that employing a variety of machine learning algorithms to analyze educational data not only aids in classifying student performance but also significantly reduces failure rates by identifying and addressing learning deficits early.

The goal of this research is to elaborate new, structured academic careers based on a thorough analysis of existing programs within our faculty. The first step involves evaluating these programs for consistency. Each unit or module in the curriculum is defined by a set of prerequisites and specific objectives. Prerequisites are the essential skills needed to start studying a module, while objectives are the competencies that students are expected to acquire by the end of the module.

Throughout each semester, students must possess the required skills to successfully complete the modules and develop the desired competencies. Our approach is to evaluate the curriculum by analyzing the correlation between the skills acquired in one semester and the prerequisites needed for the subsequent semester.

To achieve this, we employ Genetic Algorithms (GAs) as a central part of our methodology. GAs are effective in optimizing complex problems by iteratively selecting, combining, and adjusting potential solutions. Genetic algorithms are particularly suited for handling the multifaceted nature of educational data, allowing for the evolution of highly effective educational strategies through iterative selection, crossover, and mutation of potential solutions. This approach mirrors the dynamic and individualized needs of learners, as demonstrated by Chang *et al.* [13], who successfully used GAs to personalize e-course compositions, thus enhancing learning efficiency and personalization. The flexibility of GAs accommodating various parameters and constraints ensures that the educational models developed are both robust and capable of evolving over time to meet changing educational demands. Therefore, the integration of genetic algorithms into our research methodology supports our objective of developing robust career planning.

IV. IMPLEMENTING THE GENETIC ALGORITHM

A. Genetic Algorithms: Definition

Genetic Algorithms (GAs) are evolutionary algorithms inspired by Darwin's theory of "survival of the fittest" [15]. The key elements of GAs include selection, mutation, and crossover, which are used to generate and evaluate new populations through a fitness function:

- Mutation is a genetic process based on the change (mutation) of one or more genes to produce a fittest individual.
- Crossover aims to create new individuals/sons by combining the characteristics of two parent individuals. This approach is a simulation of biological reproduction where sons inherit the characteristics of both parents.
- Selection in the genetic context is a treatment that aims to select the best individuals from a given population to ensure that the most advantageous characteristics are preserved and transmitted to future generations.

B. Implementation

To implement the genetic algorithm concept, we need to define the necessary elements previously mentioned: Generation, Population, Individual, Chromosome, and Gene.

In our context, these elements are defined as follows:

- **Generation:** In our case, a generation consists of all the academic programs offered within the FSTT (Faculty of Science and Technology).
- **Population:** The initial population is composed of core courses common across all programs. We have identified four distinct pathways (MIPC, MIP, BCG, and GEGM). Each pathway represents a separate individual in the initial population.
- **Individual:** Each pathway (MIPC, MIP, BCG, GEGM) represents an individual within the initial population. Thus, we start with a total of four individuals, each corresponding to one of these pathways.
- **Chromosome:** Each pathway (individual) is divided into four semesters. In this model, each semester is considered a chromosome that characterizes the individual.
- **Gene:** Genes correspond to the specific modules taught within each semester. Therefore, each semester (chromosome) consists of six modules (genes).

Fig. 2 illustrates the proposed model:

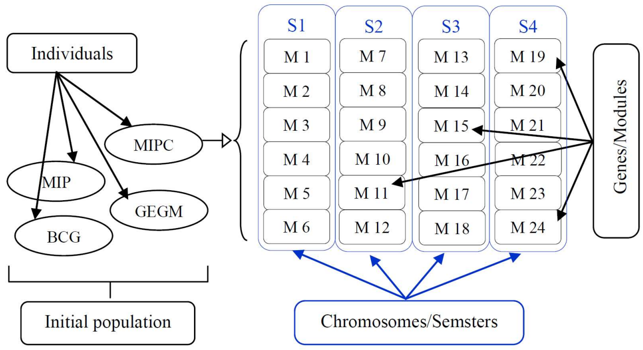


Fig. 2. Model of genetic implementation.

This conceptual frame allows us to effectively apply genetic treatment.

C. Data Encoding

To clarify the process, we start by encoding the studied data to be exploitable by the genetic algorithms.

1) Coding of modules

The initial population consists of four distinct academic careers, each including 24 modules. Although there are several common modules among these careers, our analysis will be applied to the set of modules without considering redundancies. After eliminating duplicate modules, our list comprises a total of 44 distinct modules, numbered M1 to M44. Genetically, these 44 modules serve as the genes used to develop new and more adapted academic careers.

All 44 modules are listed in the Table 1:

Table 1. Encoding genes/modules

Index	Module title	Index	Module title
M1	Electrical and electronic circuits	M23	Organic Chemistry 1
M2	Electricity	M24	Mineral chemistry 1
M3	Analysis 1	M25	Electronic
M4	Algebra 1	M26	Metrology and instrumentation
M5	Algorithmics and Programming 1	M27	Management
M6	Languages and Communication	M28	Electrical engineering
M7	Thermodynamics	M29	Automatic
M8	Point Mechanics and Geometric Optics	M30	Mechanical manufacturing
M9	Analysis 2	M31	Mechanical construction
M10	Algebra 2	M32	Numerical analysis
M11	Structure of matter	M33	cellular biology
M12	Languages and Communication 2	M34	Optics and Radioactivity
M13	Electromagnetism	M35	Cosmology & Internal Geodynamics
M14	Analysis 3	M36	Animal biology
M15	Descriptive statistics/probabilities	M37	External geodynamics
M16	Algorithmics and Programming 2	M38	Plant's biology
M17	Chemical reactivity	M39	Stratigraphy & Paleo-environment
M18	Languages and Communication –3	M40	Structural biochemistry
M19	Solid Mechanics	M41	Microbiology
M20	Quantum Mechanics and Relativity	M42	Organic/mineral chemistry 2
M21	Analysis 4	M43	Metabolic biochemistry
M22	Data structure in C	M44	Database

In reality, at the FSTT, these modules are distributed over the four initial carrers (MIPC, MIP, GEGM, BCG) during the four semesters as Table 2:

In conclusion, the individual will be presented by Fig. 3 as a sequence of modules/genes.

Table 2. Distribution of genes/modules across individuals

Initial Carrers	Semester 1	Semester 2	Semester 3	Semester 4
MIPC	M1	M7	M13	M19
	M2	M8	M14	M20
	M3	M9	M15	M21
	M4	M10	M16	M22
	M5	M11	M17	M23
	M6	M12	M18	M24
MIP	M1	M7	M19	M13
	M2	M8	M14	M20
	M3	M9	M15	M21
	M4	M10	M16	M22
	M5	M11	M17	M23
	M6	M12	M18	M24
GEGM	M1	M7	M19	M28
	M2	M8	M25	M29
	M3	M9	M15	M30
	M4	M10	M16	M31
	M5	M11	M26	M22
	M6	M12	M27	M32
BCG	M33	M36	M38	M40
	M34	M7	M2	M41
	M35	M37	M39	M42
	M11	M17	M23	M43
	M4	M3	M24	M44
	M6	M12	M15	M18

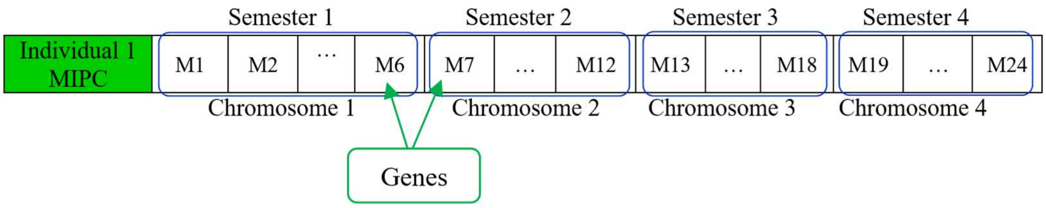


Fig. 3. The encoding of the individual.

2) Coding of skills

To facilitate the analysis and calculation of the resemblance between semesters in terms of skills, we have implemented a skills standardization procedure. This approach requires the consultation of experts in university teaching to validate the process.

Initially, for the 44 modules of the program, each module is characterized on average by 5 preliminary skills and 5 developed skills, amounting to a total of 440 distinct skills. Following a detailed analysis carried out by experts, the prerequisites and skills have been unified into a single list of knowledge, which has been condensed into a set of 138 fundamental skills. Hence, a given skill can be a prerequisite for one module and a skill for another module. The final list contains 138 skills codified in the form "C_i",

where $1 \leq i \leq Nbr_{comp}$

D. Mathematical Modeling

Following the previous coding, we can model our population mathematically so that it becomes easy to process via algorithms:

- **Individul** $\{I_i\}$: each pathway will be represented as P_i where $1 \leq i \leq Nbr_{carr}$ Thus, we have four pathways: $I_1, I_2, \ldots I_{Nrb\_carr}$.
- **Semesters** $\{S_{ij}\}$: Each pathway Pi consists of $Nbr_{semeter}$ semesters, denoted as S_{ij}, where i is the index of the pathway and $1 \leq j \leq Nbr_{semeter}$ is the index of the semester. For example, for Pathway 1, the semesters would be $S_{11}, S_{12}, \ldots, S_{1Nbr\_semester}$.
- **Modules** $\{M_k^{ij}\}$: Each semester S_{ij} contains $Nbr_{modules}$ modules, denoted as M_k^{ij}, where $1 \leq k \leq Nbr_{modules}$. For example, for semester 2 of Pathway 1, the modules would be $M_1^{1,2}$; $M_2^{1,2}$; ... ; $and\ M_{Nbr_{modules}}^{1,2}$.
- **List of Prerequisites** $\{P_{kl}^{ij}\}$: Each module M_k^{ij} has a set of prerequistes, denoted as P_{kl}^{ij}, where $1 \leq l \leq Nbr_{pre_k}$. For example, the prerequisties for module 3 of semester 1 of pathway 2 would be $P_{3,1}^{2,1}$; $P_{3,2}^{2,1}$; ... ; $P_{3,Nbr_{pre_k}}^{2,1}$.
- **List of competencies** $\{C_{km}^{ij}\}$: Similarly, each module M_k^{ij} has a set of competencies, denoted as C_{km}^{ij} where $1 \leq m \leq Nbr_{comp_k}$.... For example, the competencies for module 5 of semester 3 of pathway 4 would be $C_{5,1}^{4,3}$; $C_{5,2}^{4,3}$; $C_{5,3}^{4,3}$; ... $C_{5,\mathrm{NbrComp}_k}^{4,3}$.

Through this mathematical implementation, we can present the individual/path with this expression:

$$I_i = \{ S_{ij} \mid 1 \leq i \leq Nbr_{path} \ and \ \ 1 \leq j \leq Nbr_{semeter}\}$$

with

$$S_{ij} = \{ M_k^{ij} \mid 1 \leq k \leq Nbr_{modules} \}$$

and

$$M_k^{ij} = \{ P_{kl}^{ij} \mid 1 \leq l \leq Nbr_{pre_k} \} \oplus \{C_{km}^{ij} \mid 1 \leq m \leq Nbr_{com\ k} \}$$

The composition operator $\oplus$ means that M_k^{ij} consists of the two sets $\{P_{kl}^{ij}\}$ and $\{C_{km}^{ij}\}$

E. Design of the Fitness Function

To evaluate the fitness of an individual/career, based on the proposed encoding and the methodology we developed above, we consider the fitness value of the individual to be the average of the sums of the similarity rates between the group of skills developed by chromosome $j+1$.Therefore, 3 values of the similarity rate are obtained for pairs of consecutive semesters (M1: S1–S2; M2: S2–S3; and M3: S3–S4).

Pedagogically speaking, a better course is one that has a higher similarity value, which means there is greater coherence between the skills developed during semester j and the prerequisites of semester $j+1$, whereas the student has received all the necessary skills during semester j so that he can study the modules of semester $j+1$.

The fitness function can be presented in the following form:

$$f(Ii) = \sum_{j=1}^{Nbr_{semeter}-1} \frac{\frac{|(\cup_{k=1}^{Nbr_{modules}} C_k^{ij}) \cap (\cup_{k=1}^{Nb\ modules} P_k^{ij+1})|}{|(\cup_{k=1}^{Nbr_{modules}} P_k^{ij+1})|}}{3}$$

To simplify the form of the fitness function, we consider that:

$Gc = (\cup_{k=1}^{Nbr_{modules}} C_k^{ij})$ is the set of skills of all k modules of semester j

$Gp = (\cup_{k=1}^{Nb\ modules} P_k^{ij+1})$ is the set of prerequisites for all k modules of semester $j+1$

$$f(Ii) = \sum_{j=1}^{Nbr_{semeter}-1} \frac{\left(\frac{|Gc(j) \cap Gr(j+1)|}{|Gr(j+1)|}\right)}{3}$$

Fitness Objective: A higher fitness value indicates a better transition and greater consistency between the skills taught and those required in successive semesters, which is desirable in a training path.

In summary, the developed fitness function is used to evaluate the effectiveness of training courses in terms of logical progression and skill building from one semester to the next. It helps identify the most coherent and best structured paths in terms of skills development.

F. Individual Fitness Calculation

On the basis of the collected and normalized data, we applied the fitness function defined above to the four existing individuals/paths, with the 3 values of the similarity rate obtained for pairs of consecutive semesters (M1: S1–S2; M2: S2–S3; and M3: S3–S4).

Thus, the results are as shown in Table 3:

Table 3. Fitness values of the individuals of the first population

Individu	M1	M2	M3	f(I)
I1 (MIPC)	0.37	0.37	0.37	0.37
I2 (MIP)	0.37	0.37	0.30	0.34
I3 (GEGM)	0.37	0,17	0.20	0.24
I4 (BCG)	0.23	0.20	0.13	0.19

Since the initial population is not large enough, the four individuals will all be parents of future generations, with the aim of generating descendants with a higher degree of fitness. To achieve this objective, we implement algorithmic strategies of mutation, crossover and selection.

G. Selection

In the application phase, we consider the fitness threshold to be the average of the fitness values of all the individuals in the population; therefore, we keep only those individuals whose fitness exceeds this threshold. Additionally, we will introduce a second selection criterion: an individual must not have a duplicate gene, meaning that a course cannot contain a duplicate module. The individuals meeting these criteria are considered the fittest and will be used for reproduction (crossover and mutation) to create the next generation.

In our case, to give all four individuals a fair chance, we will start with mutation and crossover for the first time, after which we will apply the selection process to the first child generation. The pseudo code of selection is illustrated by the following Fig. 4.

```
function selection(population, selected,
populationSize, selectedSize):
      totalFitness = 0.0
      for i from 0 to populationSize - 1:
         totalFitness +=
population[i].fitness
      end for
      averageFitness =
totalFitness/populationSize
      selectedSize = 0
      for i from 0 to populationSize - 1:
         if population[i].fitness >
averageFitness and
isUniqueModule(population[i]):
            selected[selectedSize] =
population[i]
            selectedSize++
         end if
      end for
   end function
```

Fig. 4. Pseudocode of the selection function.

H. Mutation

The mutation will be applied to a gene chosen randomly from an individual, which is also selected randomly. However, the genes of the first chromosome are excluded from this process because the fundamental modules of the first semester remain unchanged, while variations are introduced in the following semesters to explore different possible combinations of pathways.

To obtain more possible combinations, we adopt two distinct mutation strategies: intra-individual mutation and inter-individual mutation. In intra-individual mutation, genetic information is randomly modified within the same individual. In inter-individual mutation, genetic changes are carried out between different individuals within the population.

The following figures illustrate the pseudo code for these operations: Fig. 5 represents the mutation process, while Fig. 6 details the crossover mechanism.

```
  procedure mutation(population: array of
Individual, child: array of Individual, N:
integer):
      seed random number generator with
current time
      for i in range(NB_OFFSPRING):
         //Select a random parent index
         parentIndex := random integer in the
range [0, N-1]
   //Copy the selected parent to the child
   child[i] := copy of population[parentIndex]
   //Randomly choose two semesters and one
module in each semester
   semester1 := random integer in the range [0,
NB_SEMESTERS-1]
   semester2 := random integer in the range [0,
NB_SEMESTERS-1]
   module1 := random integer in the range [0,
NB_MODULES-1]
   module2 := random integer in the range [0,
NB_MODULES-1]
   //Swap (permutation) the modules between two
semesters

swap(child[i].semesters[semester1].modules[mod
ule1],

child[i].semesters[semester2].modules[module2]
)
        //Assign a name to the child
         child[i].courseName := "child_" + (i
+ 1)
      end for
  end procedure
```

Fig. 5. Pseudocode of intraindividual mutation.

```
procedure inter_mutation(population: array of
Individual, child: array of Individual, N:
integer)
      initialize random number generator with
current time
      for i from 0 to NB_OFFSPRING - 1 do
   parentIndex := random number between 0 and N -
1 inclusive
          if parentIndex = N then
              parentIndex := N - 1
/Copy the parent into the child
     child[i] := copy of population[parentIndex]
/Choose a random semester and module to mutate
     semester := random number between 0 and
NB_SEMESTERS - 1
     module := random number between 0 and
NB_MODULES - 1
/Choose a random module from the list of modules
newModuleIndex := random number between 0 and 43
     newModule := copy of
moduleList[newModuleIndex]
    //Mutate the selected module
     set
child[i].semesters[semester].modules[module] to
newModule
         deallocate memory used by newModule
    //Name the child
     set child[i].courseName to "child_" + (i +
1)
      end for
   end procedure
```

Fig. 6. Pseudocode of interindividual mutations.

Fig. 7 shows the convergence graph that compares the results of both mutation strategies, intra_individual and inter_individual mutation.

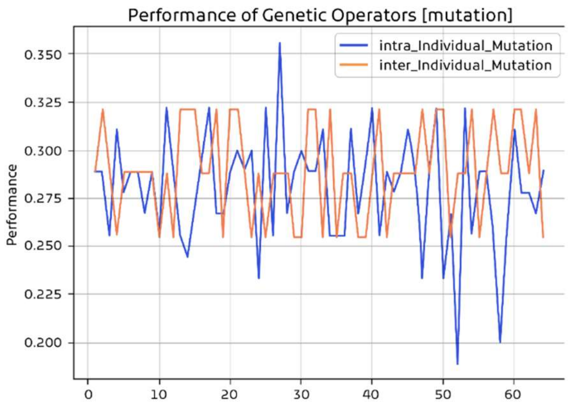


Fig. 7. Comparing performance between intra/inter individual mutation.

I. Crossover

We will initially work the one-point crossover method where a point is randomly chosen. This point determines where the characteristics of the parents will be exchanged to create the children.

For each module in each semester, the function determines whether the module index is before or after the crossover point. If the index is before the crossover point, child1 inherits the corresponding module from Parent 1 and Child 2 inherits it from Parent 2. If the index is after the crossover point, child1 inherits the corresponding module from Parent 2 and Child 2 inherits it from Parent 1. The children's path names are updated to reflect their origin ("Child 1" and "Child 2" in this example).

Thus, two new individuals are created, each possessing a unique combination of modules from their parents. Similarly, we will apply the two-point crossover method to compare the results of each crossover technique. Fig. 8 represent the pseudo code of the crossover function.

```
   function crossover(parent1: Individual,
parent2: Individual, child1: Individual,
child2: Individual):
       pointCrossover := random() %
(NB_SEMESTERS * NB_MODULES) //Randomly choose
the crossover point

       for i from 0 to NB_SEMESTERS - 1:
           for j from 0 to NB_MODULES - 1:
               indexModule := i * NB_MODULES + j
               if indexModule < pointCrossover:

child1.semesters[i].modules[j] :=
strdup(parent1.semesters[i].modules[j])

child2.semesters[i].modules[j] :=
strdup(parent2.semesters[i].modules[j])
               else:

child1.semesters[i].modules[j] :=
strdup(parent2.semesters[i].modules[j])

child2.semesters[i].modules[j] :=
strdup(parent1.semesters[i].modules[j])
               end if
           end for
       end for
   end function
```

Fig. 8. Pseduo code of the crossover function.

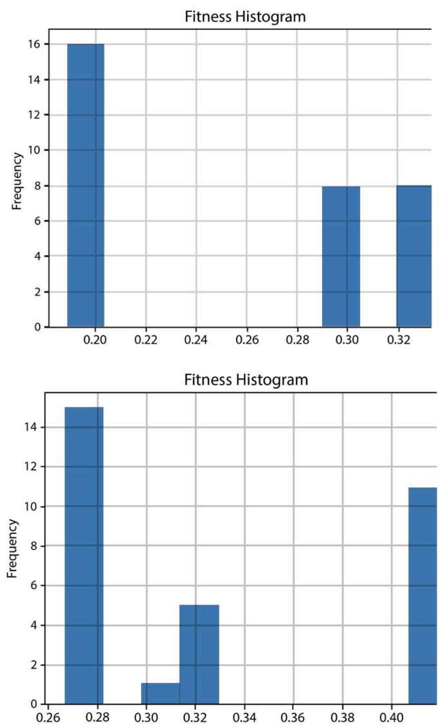
V. Results and Discussions

The implementation of our genetic algorithm for university study planning produced promising results. In this section, we summarize the key findings and discuss their importance.

The selection function effectively identifies individuals with the best fitness namely, those that provide logical continuity between semesters. This ensured the conservation and transmission of the best characteristics of subsequent generations.

The mutation and crossover functions introduce necessary diversity into the population, allowing new combinations of modules to be explored. This increased the chances of discovering more efficient and innovative configurations.

Fig. 9 shows that inter-individual mutation consistently maintains a slightly higher performance level than intra-individual mutation throughout most generations.



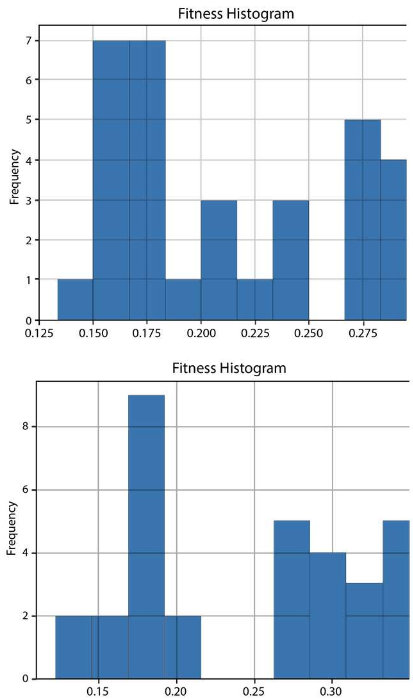


Fig. 9. Histograms fitness of the first four generations.

We combined the 3 functions in the different generations to obtain the best possible results.

Through generations of 32 individuals, we obtain the above fitness histograms:

The algorithm enabled the generation of study plans that were better adapted in terms of coherence between prerequisites and skills. The best results, obtained after approximately 12 generations, had a fitness value of around 0.5, which is quite significant.

The following Table 4 contains the best results. We have also eliminated individuals with any of the three resemblance values that were too small compared to the fitness average.

Table 4. Fitness values of the fittest individuals

Individuals	Generation	M1	M2	M3	f(I)
Child_01	1	0,37	0,4	0,33	0,37
Child_09	1	0,37	0,37	0,3	0,34
Child_10	2	0,43	0,33	0,2	0,32
Child_18	2	0,37	0,37	0,37	0,37
Child_19	2	0,43	0,37	0,30	0,37
Child_07	5	0,43	0,37	0,37	0,35
Child_04	9	0,43	0,3	0,4	0,38
Child_23	10	0,4	0,47	0,33	0,4
Child_04	12	0,37	0,4	0,5	0,42
Child_14	12	0,43	0,37	0,5	0,43
Child_15	12	0,5	0,43	0,57	0,5

The distributions of the modules of some individuals with the best fitness values are as shown in Fig. 10:

```
Modules of the individual child_1:
  Semester 1: M1        M2      M3      M4      M5      M6
  Semester 2: M7        M8      M9      M10     M17     M12
  Semester 3: M19       M14     M15     M16     M11     M18
  Semester 4: M13       M20     M21     M22     M23     M24
```

```
Modules of the individual child_9:
  Semester 1: M1        M2      M3      M4      M5      M6
  Semester 2: M13       M8      M9      M10     M11     M12
  Semester 3: M19       M14     M15     M16     M17     M18
  Semester 4: M7        M20     M21     M22     M23     M24
```

```
Modules of the individual child_4:
  Semester 1: M1        M2      M3      M4      M5      M6
  Semester 2: M7        M8      M9      M10     M11     M12
  Semester 3: M13       M14     M15     M16     M17     M18
  Semester 4: M19       M20     M21     M22     M23     M24
```

```
Modules of the individual child_23:
  Semester 1: M1        M2      M3      M4      M5      M6
  Semester 2: M7        M20     M9      M10     M11     M12
  Semester 3: M19       M14     M15     M16     M17     M18
  Semester 4: M13       M8      M21     M22     M23     M24
```

```
Modules of the individual child_18:
  Semester 1: M1        M2      M3      M4      M5      M6
  Semester 2: M7        M8      M9      M10     M22     M12
  Semester 3: M19       M25     M15     M16     M26     M27
  Semester 4: M28       M29     M30     M31     M11     M32

Modules of the individual child_19:
  Semester 1: M1        M2      M3      M4      M5      M6
  Semester 2: M7        M14     M9      M10     M11     M12
  Semester 3: M19       M8      M15     M16     M17     M18
  Semester 4: M13       M20     M21     M22     M23     M24
```

Fig. 10. Module distribution for the best obtained individuals.

A. Results Validation

To validate the results obtained by our research, two distinct methods were implemented to ensure the robustness and relevance of the educational paths generated by our algorithm. The first method involved creating a specific questionnaire for students to gather their preferences regarding the ideal sequence of certain teaching modules. This survey allowed us to determine the majority preferences for the order of modules. exemplified by the fact that 73% of the participants expressed a desire to study the "Solid Mechanics" (M19) module before "Electromagnetism" (M13), as shown in Fig. 11 below.

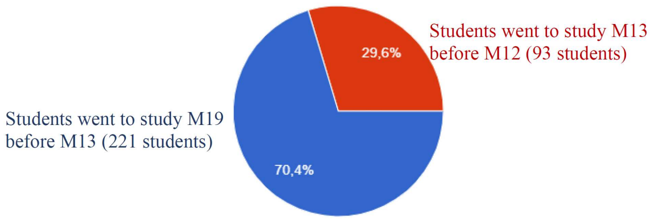


Fig. 11. Percentage of students following their preference regarding the order of the M13 and M19 modules.

These results directly influenced the structure of the educational paths proposed by our algorithm, which placed "Solid Mechanics" (M19) before "Electromagnetism" (M13) (Fig. 9) in accordance with student preferences, by validating the algorithm's ability to align course plans with learner expectations.

The second validation method consisted of analyzing a large academic database provided by the FSTT, comprising over 600,000 student assessment records since 2017. A thorough statistical analysis of these data revealed significant trends in student performance. Notably, students who passed the "Electromagnetism" (M13) module in semester 5 (after initially failing it in semester 3) achieved, on average, better results than those who passed this module in semester 3. Their average scores were also 30% higher than those of other students. This observation supports the effectiveness of our algorithmic model, demonstrating its alignment with actual academic outcomes and reinforcing the importance of a data-driven approach to designing learning paths. These cross-validation methods affirm the validity of our approach and highlight the value of using genetic algorithms for personalized academic planning.

The results clearly show that, compared with traditional methods, the algorithmic approach can significantly improve the planning of studies. The algorithm developed in this paper can also be used to optimize students' individual study plans, potentially reducing the failure rate by avoiding unfavorable academic situations.

Thus, this automation of the course planning process has the potential to significantly reduce the cost of administrative work, allowing academic staff to perfect the results produced by the algorithm and to concentrate on more strategic tasks.

Consistent with findings from Muridan *et al.* [16], our study acknowledges the importance of focused career planning and self-efficacy among students engaged in vocational training, highlighting these critical factors for achieving successful educational and career outcomes.

In conclusion, the results obtained illustrate the effectiveness of the genetic algorithm in optimizing the planning of university studies. After thorough validation, future studies could explore the integration of this system with existing educational platforms for practical and scalable implementation.

VI. Future Work

The current study indicates that using genetic algorithms for university study planning is highly promising. This approach enables the creation of optimal study plans by taking into account academic constraints and available resources.

However, several avenues of development and improvement remain to be explored to take full advantage of this approach.

Incorporation of individual preferences: Incorporating individual student preferences into the algorithm would further personalize study plans, making them not only optimal in terms of academic progression but also aligned with students' interests and career aspirations.

Broadening the input parameters: Broadening the input parameters of the algorithm increases the precision and relevance of the results. This could include data on students' abilities and academic performance, available resources (e.g., number of teachers or classroom availability), and specific study program requirements.

Validation and Testing under Real Conditions: Validating the algorithm in real-world contexts is essential to guarantee its effectiveness and efficiency. This involves pilot testing at universities, collecting feedback, and measuring the impact of the tool on student success and administrative efficiency.

Development of an Intuitive User Interface: Creating an intuitive and user-friendly interface for the system would facilitate its implementation and adoption by administrative staff and students. This interface could also enable direct interaction with the algorithm, providing real-time visualization of study plans and possible modifications.

Integration with university information systems: Integrating the tool into existing university information systems would make its use more fluid and allow automatic synchronization of data, thus facilitating the planning and updating of courses of studies.

Longitudinal studies to measure long-term impact: Conducting longitudinal studies over several academic years would make it possible to measure the long-term impact of the tool on student success and administrative efficiency.

By focusing future efforts on these areas, it is possible to create a revolutionary tool for study planning that is not only effective and efficient but also adaptive and student-centered.

Scalability and Applicability to Large Datasets

Our genetic algorithm (GA) is designed with scalability in mind, making it well suited for application to larger datasets beyond the initial case study presented. The scalability of GAs generally depends on several factors, including the complexity of the fitness function, the size of the population, and the number of generations used for convergence.

Complexity of the Fitness Function: Our GA uses a fitness function that is efficiently computable, even as the size of the data increases. This is crucial because fitness evaluation is the most computationally intensive aspect of GAs. By optimizing the computation of our fitness function, we ensure that our GA can handle larger datasets without a significant increase in computational time.

Population Size and Genetic Operations: The GA is designed to operate with variable population sizes. A larger population can be beneficial for exploring a more diverse set of solutions, which is particularly useful in larger datasets where the solution space is vastly expanded. Furthermore, our genetic operators (crossover and mutation) are tailored to maintain genetic diversity within the population, which is essential for avoiding premature convergence in larger datasets.

Parallel Processing Capabilities: The algorithm is structured to leverage parallel processing. Evaluating the fitness of each individual within the population and applying genetic operators can be efficiently parallelized which significantly reducing the as the dataset size increases.

Adaptation to Varying Data Sizes: Our approach includes adaptive mechanisms that adjust the parameters of the GA (such as mutation rate and crossover probability) on the basis of the characteristics of the dataset. This adaptability enhances its effectiveness in educational settings with varying numbers of students, courses, and scheduling constraints.

Empirical Validation: We plan to validate the scalability of our GA through empirical studies involving datasets of increasing size and complexity. This will not only demonstrate the capability of our GA to scale but also help refine its parameters for optimal performance across different educational contexts.

By addressing these aspects, our GA is poised to be a robust tool for academic planning, capable of handling complex and large-scale educational datasets. This makes it a flexible and powerful solution for educational institutions seeking to implement data-driven decision-making in curriculum planning.

VII. Conclusion

The in-depth exploration of genetic algorithms to optimize study planning at Moroccan universities has paved the way for significant advances in the field of higher education. This study highlights the immense potential of these algorithms to address the complex and multifaceted challenges of academic planning.

The implementation of this technology promises to transform the way educational pathways are designed, moving from a manual, time-consuming, and potentially subjective process to an automated, efficient, and objective criteria-based approach. Key advantages include significant time savings for administrative staff, more precise and personalized student planning, and improved management of available resources.

Furthermore, the integration of this technology into the administrative processes of Moroccan universities represents a significant step forward toward the modernization and optimization of higher education. It offers a viable solution to address complex cases of students with nonlinear academic paths, particularly those who must manage failures in certain modules while progressing in their coursework.

Although the initial results are promising, further research is imperative to refine the algorithm, integrate additional parameters, and validate its effectiveness in real-world conditions. The future of this initiative looks bright, with prospects for continuous improvement and adaptation to the evolving needs of students and institutions.

This research underscores the transformative potential of genetic algorithms in educational planning and management. The successful implementation and outcomes of this study advocate for broader adoption and adaptation of similar methodologies across various educational contexts. Future research could explore the integration of multifaceted educational metrics, such as student well-being and engagement, to further enhance the holistic value of academic programs. Additionally, extending this approach to include real-time data analysis could dynamically refine educational paths in response to ongoing academic performance and changing student needs. By pushing the boundaries of traditional educational frameworks, this work lays the groundwork for a more adaptive, responsive, and student-centered educational system.

In conclusion, this study constitutes a fundamental step toward innovative educational reform, aligning with the needs of the 21st century. This finding demonstrates the potential of advanced technologies to improve not only the quality of education but also to make education systems more responsive, inclusive and adaptable to the digital age.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Abderrahim EL Yessefi is the corresponding author and PhD student who conducted the primary research, implemented the genetic algorithm model, and wrote the initial draft of the manuscript. Dr. Mustafa Ezziyyani supervised the research process, provided critical insights on the methodology, and revised the manuscript for significant intellectual content. Dr. Loubna Charrat served as the co-supervisor, contributing to the conceptual framework and assisting with the refinement of the experimental design. Dr. Mohammed Rida Ech-charrat, as the team leader, facilitated coordination among the research team and provided strategic guidance throughout the project. All authors reviewed and approved the final version of the manuscript. All authors had approved the final version.

References

[1] A. Akavova1, Z. Temirkhanova, and Z. Lorsanova, "Adaptive learning and artificial intelligence in the educational space," *E3S Web of Conferences*, 2023. https://doi.org/10.1051/e3sconf/202345106011

[2] J. Xu, "Improved Genetic algorithm to solve the scheduling problem of college English courses," *Complexity*, 7252719, 2021. doi: 10.1155/2021/7252719

[3] L. Elshania and K. P. Nuçi, "Constructing a personalized learning path using genetic algorithms approach," arXiv preprint, arXiv:2104.11276, 2021.

[4] R. Rollande and J. Grundspenkis, "Graph based framework and its implemented prototype for personalized study planning," in *Proc. 2013 Second International Conference on E-Learning and E-Technologies in Education (ICEEE)*, IEEE, 2013.

[5] S. G. Sargsyan and A. S. Hovakimyan, "Using genetic algorithms in pervasive learning systems," *Open Science Journal of Mathematics and Application*, vol. 2, no. 5, pp. 44–47, 2014.

[6] A. H. Bohadean, "A microgenetic algorithm approach for the soft constraint satisfaction problem in university course scheduling," Master thesis, University UTARA Malaysia, 2013.

[7] R. Perzina, "Solving multicriteria university timetabling problem by a self-adaptive genetic algorithm with minimal perturbation," in *Proc. 2007 IEEE International Conference on Information Reuse and Integration*, 2007.

[8] R. Perzina, "Solving the university timetabling problem with optimized enrollment of students by a self-adaptive genetic algorithm," in *Practice and Theory of Automated Timetabling VI. PATAT 2006*, E. K. Burke and H. Rudova, Eds., Springer, Berlin, Heidelberg, LNCS, 2007, vol 3867. https://doi.org/10.1007/978-3-540-77345-0_16

[9] B. Hssina and M. Erritali, "A personalized pedagogical objectives based on a genetic algorithm in an adaptive learning system procedia computer science," vol. 151, pp. 1152–1157, 2019. doi: 10.1016/j.procs.2019.04.164

[10] T.-Y. Chang and Y.-R. Ke, "A personalized e-course composition based on a genetic algorithm with forcing legality in an adaptive learning system," *Journal of Network and Computer Applications*, vol. 36, issue 1, pp. 533–542, January 2013. http://dx.doi.org/10.1016/j.jnca.2012.04.002

[11] I. Dhaiouir, M. Ezziyyani, and M. Khaldi, "Smart model for the classification and orientation of learners in a MOOC," *International Journal of Emerging Technologies in Learning (iJET)*, vol.17, no. 5, pp. 19–35, 2022. https://doi.org/10.3991/ijet.v17i05.28153

[12] Formation initiale. [Online]. Available: Https://Fstt.Ac.Ma/Portail2023/Formation-Initiale/

[13] C.-C. Chang and G.-C. Lee, "Using genetic algorithms to optimize stopping patterns for passenger rail transportation," *Computers & Industrial Engineering*, vol. 64, no. 1, pp. 412–421, 2013. https://doi.org/10.1016/j.cie.2012.09.011

[14] H. Pallathadka, A. Wenda, E. Ramirez-Asís, M. Asís-López, J. Flores-Albornoz, and K. Phasinam, "Classification and prediction of student performance data via various machine learning algorithms," *Materials Today: Proceedings*, 2021. https://doi.org/10.1016/j.matpr.2021.07.382

[15] S. Khatwani and A. Arya, "A novel framework for envisaging a learner's performance using decision trees and genetic algorithm," in *Proc. International Conference on Computer Communication and Informatics (ICCCI-2013)*, 2013.

[16] N. D. Muridan, M. S. Rasul, R. M. Yasin, A. R. M. Nor, R. A. A. Rauf, and N. A. Jalaludin, "Career planning indicators of successful TVET entrepreneurs," *Sustainability*, vol. 15, no. 6629, 2023. https://doi.org/10.3390/su15086629