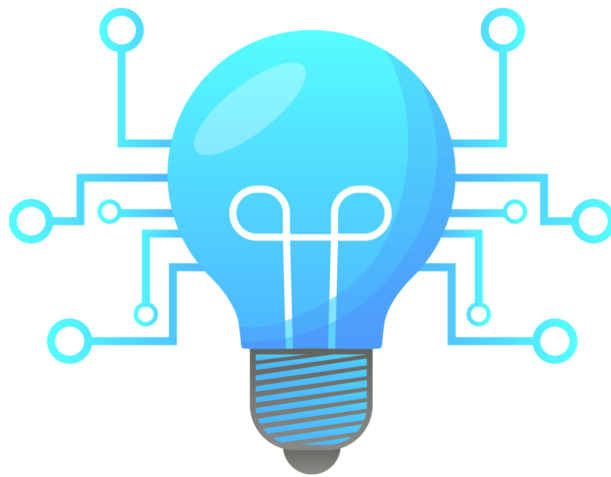


EUROPEAN MASTER IN PROJECT MANAGEMENT

END OF MASTER THESIS

Analysis Of The Use Of Artificial Intelligence In The
Allocation Of Resources In Projects



Student: Figueroa De Jesús, Paola M.

Director: Zuazo Urionabarrenechea, José Ignacio

Co-Director: Toledo Gandarias, Nerea

Course: 2022-2023

Date: Bilbao, October 2023

TRILINGUAL SUMMARY

This research focuses on the relationship between artificial intelligence (AI) and project management, specifically on the task of assigning labor-type resources (people) to projects. The main motivation was to automate the assignment of human resources to different projects with the help of artificial intelligence. The K-Means algorithm (a machine learning algorithm) was selected among different alternatives as the most appropriate to perform the resource allocation. For its validation, different tests were performed by varying the number of employees, projects and competencies in different scenarios. Satisfactory tests were carried out on the allocation of resources, distributing them in a maximum of 10 projects. Additionally, the uncertainty of the results was analyzed with the metrics "Silhouette Index" and "Davis-Bouldin Evaluation Score", finding limitations the higher the number of projects, number of data and their quality. Artificial intelligence is a powerful tool to support a project manager, but in no case to replace him.

Keywords: Artificial Intelligence, Project Management, Resource Allocation, K-Means, Machine Learning

Esta investigación se centra en la relación entre la inteligencia artificial (IA) y la gestión de proyectos, concretamente en la tarea de asignar recursos de tipo trabajo (personas) a los proyectos. La motivación principal fue la de automatizar la asignación de recursos humanos a diferentes proyectos con la ayuda de la inteligencia artificial. Se seleccionó el algoritmo K-Means (un algoritmo de aprendizaje automático) entre distintas alternativas como el más apropiado para realizar la asignación de recursos. Para su validación se realizaron distintas pruebas variando el número de empleados, proyectos y competencias en distintos escenarios. Se realizaron pruebas satisfactorias en la asignación de recursos distribuyéndolos en un máximo de 10 proyectos. Adicionalmente se analizó la incertidumbre de los resultados con las métricas "Índice de Silueta" y "Puntuación de Evaluación de Davis-Bouldin" encontrándose limitaciones cuanto mayor era el número de proyectos, número de datos y su calidad. La inteligencia artificial es una herramienta poderosa como apoyo para un gestor de proyectos, pero en ningún caso para sustituirlo.

Palabras clave: Inteligencia Artificial, Gestión de Proyectos, Asignación de Recursos, K-Means, Aprendizaje Automático

Ikerketa hau adimen artifizialaren (IA) eta proiektuen kudeaketaren arteko harremanean oinarritzen da, zehazki proiektuei lan motako baliabideak (pertsonak) esleitzeko zereginen. Motibazio nagusia adimen artifizialaren laguntzarekin hainbat proiektutarako giza baliabideen esleipena automatizatzea izan zen. K-Means algoritmoa (ikasketa automatikoko algoritmo bat) hautatu zen baliabideen esleipena egiteko aukera egokien artean. Horiek baliozkotzeko, hainbat proba egin ziren, hainbat agertokitako enplegatuen, proiektuen eta gaitasunen kopurua aldatuz. Baliabideak esleitzeko proba egokiak egin ziren, eta gehienez ere 10 proiektutan banatu ziren. Horrez gain, emaitzen ziurgabetasuna aztertu zen "Silueta

Indizea" eta "Davis-Bouldinen Ebaluazio Puntuazioa" metriken bidez, eta zenbat eta handiagoa izan projektuen kopurua, datu-kopurua eta horien kalitatea, orduan eta muga handiagoak aurkitu ziren. Adimen artifiziala tresna indartsua da projektuen kudeatzaile bati laguntzeko, baina inola ere ez hura ordezkatzeko.

Gako-hitzak: adimen artifiziala, projektuen kudeaketa, baliabideen esleipena, K-Means, ikaskuntza automatikoa

TABLE OF CONTENTS

1. INTRODUCTION	1
2. CONTEXT	2
2.1 RESOURCE ALLOCATION IN PROJECT MANAGEMENT	2
2.2 IMPORTANCE OF RESOURCE ALLOCATION	2
2.3 RISKS OF INEFFECTIVE RESOURCE ALLOCATION ON PROJECTS	4
2.4 WHY USE ARTIFICIAL INTELLIGENCE TO ALLOCATE RESOURCES?	5
3. OBJECTIVES AND SCOPE OF THE RESEARCH WORK	9
3.1 OBJECTIVES	9
3.2 SCOPE: WHAT ARE THE CONSTRAINTS OF THIS RESEARCH?	9
3.3 POTENTIAL BENEFITS BY IMPLEMENTING AN AI ALGORITHM FOR RESOURCE ALLOCATION	9
3.4 SUSTAINABILITY OBJECTIVES	11
4. ANALYSIS OF THE STATE OF THE ART	12
5. VALUATION OF ALTERNATIVES FOR RESOURCE ALLOCATION	23
6. WORK METHODOLOGY	26
7. TESTING PHASE – RESULTS	31
7.1 SCENARIO 1	31
7.2 SCENARIO 2	32
7.3 SCENARIO 3	33
7.4 SCENARIO 4	35
8. MODEL EVALUATION	37
8.1 SILHOUETTE INDEX	37
8.2 DAVIES-BOULDIN INDEX	38
9. TESTING DISCUSSION	40
10. PROGRAM LIMITATIONS AND UNCERTAINTY	41
11. CONCLUSIONS	43
12. IMPROVEMENTS AND FUTURES TESTS	45
13. BIBLIOGRAPHY	46
ANNEX 1 : INPUT DATA	49
ANNEX 2: TESTING RESULTS – K MEANS ALGORITHM	50
ANNEX 3: K-MEANS ALGORITHM – CODE	51

TABLE INDEX

TABLE 1: BENEFITS AND RISKS OF AI APPLICATION	19
TABLE 2: ALGORITHMS AND CHARACTERISTICS	24
TABLE 3: EVALUATION OF ALGORITHMS	24
TABLE 4: SCENARIOS	29
TABLE 5: RESULTS SCENARIO 1	31
TABLE 6: RESULTS SCENARIO 2	32
TABLE 7: RESULTS SCENARIO 3	33
TABLE 8: RESULTS SCENARIO 4	35
TABLE 9: SILHOUETTE EVALUATION SCORE	38
TABLE 10: DAVIS-BOULDIN EVALUATION SCORE	38

DIAGRAM INDEX

DIAGRAM 1: K-MEANS ALGORITHM	27
DIAGRAM 2: USE/APPLICATION OF K-MEANS ALGORITHM FOR RESOURCE ALLOCATION	28

1. INTRODUCTION

In project management, task automation using artificial intelligence (AI) is a promising possibility to improve the efficiency of project management in various aspects. The research project focuses on how AI could play a key role in project management, especially in the task of resource allocation.

First, it is contextualized, what is resource allocation in project management? It also makes visible the importance of this task in all projects and how it can affect the course of the project. Then the question arises, why use artificial intelligence for this task? This question is answered throughout this work. A state of the art analysis of the relationship between artificial intelligence and project management has been carried out. Several algorithms are evaluated in order to choose one to carry out an experiment based on several hypothetical scenarios.

As part of the methodology of the work, a series of tests of an artificial intelligence algorithm is performed. Specifically of machine learning, called K-Means. The main objective is the allocation of resources (people, employees) to different projects of a company, which is based on the assessment of the performance of employees in previous projects. It is tested with 4 different scenarios, varying the number of employees, projects and competencies in each scenario. It is concluded that the algorithm is capable of assigning people to projects based on the characteristics provided. However, it is an algorithm that can be used for smaller teams, as more powerful algorithms capable of handling large amounts of data could be used for other scenarios.

At the end, conclusions are drawn based on the results of running the experiment, which makes visible the effectiveness and limitations of implementing AI in project management. It is interpreted that AI is able to transform processes in project management, but also the importance of adaptability and the correct interpretation of these results. The project contributes to the visibility of the emerging role of AI in project management and contributes to future research topics for this evolving field.

2.CONTEXT

2.1 RESOURCE ALLOCATION IN PROJECT MANAGEMENT

Projects are a temporary activity that aims to create something unique, such as a product, service or outcome. To carry out a project, it is necessary to have resources and adequate planning. Resource allocation is essential for the success of the project and requires careful and efficient management. In this section, it will be exposed the importance of resources in projects, the need for proper resource allocation, the associated risks and how artificial intelligence can be used as a support tool for project managers within this task. But first, what are resources?

In the context of project development, resources can include machinery, capital, physical facilities, and many other elements. One may argue that, the most valuable resource of all is the human resources. It is people who will carry out all the necessary tasks to achieve the project goals and deliver the final product that meets the established expectations.

The process of allocating resources involves distributing them in a manner that maximizes their effectiveness and efficiency. This is especially crucial in project management, as the success and progress of a project relies heavily on how resources are allocated in terms of cost and planning. The project manager needs to have the ability to envision their team, assess their capacity, identify instances of overallocation of resources and recognize other factors that could potentially impact the project's progress in relation to personnel.

2.2 IMPORTANCE OF RESOURCE ALLOCATION

As previously stated, resource allocation is a critical responsibility for a project manager. In the words of Cristina Selaru in her work on Resource Allocation in Project Management, "Resources are scarce; therefore, we need to carefully allocate them in order to obtain the desired results." The statement is accurate, as one of the constraints of resources is their availability. Also, they're not unlimited. It will depend on the project planning and how it will be carried on.

For example; in software development projects, there are several positions to be filled, such as programmers, software engineers, computer engineers, data analysts, team leaders, equipment installers, database and networks specialists, among others. Additionally, there are multiple programming languages and project types within the field, thousands of variables to consider, and other determining factors that dictate the type of resource and the skills required per resource to successfully execute the project.

Software development projects are very relevant nowadays, since we are in a digital era where technology is influencing society in important ways. It is crucial that the project manager has knowledge about the necessary methodology to lead this type of project. This is a key factor that can influence the development of the project itself. A basic methodology that can be used with software development projects is:

1. Initiation
2. Analysis
3. Design
4. Implementation
5. Integration
6. Test
7. Termination

These are some generic phases that are outlined in the article [Resource Allocation in Project Management](#).

In software development projects, it is crucial to have well-defined milestones to track deliverables throughout the project. For instance, a deliverable could be a milestone in the planning or could be equivalent to an important phase of the project. [1]

According to the article, methodologies can be a specific way to approach a project, including templates, forms and other resources that can be used. It is important to note that many project managers have come to the conclusion that following a methodology strictly is not always the most convenient, but adapting it to the needs of each particular project is. [1] Methodologies can be very helpful guides to steer the course of projects.

As mentioned, resource allocation is a vital part of project management, but why it is so important? By properly distributing the resources available to you as a project manager, there is a high chance that you will be able to deliver according to the estimated timeline. Furthermore, by properly allocating resources, you will be able to effectively utilize the capabilities of the team members you have selected to the tasks assigned to them. This goes hand in hand with maintaining good communication with the team and clearly outlining the true goals regarding their work. By this, it is possible to keep the project budget on track and also stay up to date with the planning.

Being able to manage and allocate resources involves having the necessary resources available when needed. In an organization, a human resource with the appropriate attitude and behavior can be considered a valuable asset. [2] Additionally, the role of the Human

Resources department in recruiting personnel can also influence the outcome of projects, as good candidate selection can make a difference.

The combination of these factors (and many more) can positively impact the project outcome, enabling the project manager to achieve the expected results. While the importance of resource allocation is highlighted, what are the potential risks of not doing it effectively?

2.3 RISKS OF INEFFECTIVE RESOURCE ALLOCATION ON PROJECTS

There is no “correct” or 100% ideal way to carry out tasks. It depends on many factors within the project itself. Knowing their team and managing effectively is essential for PM’s to achieve the project’s objectives. If resources, specifically people, are assigned ineffectively, there may be negative consequences both during the project and in the final outcome of the product or service. The following are some of the risks that have been identified in connection with this issue.

In first place, we have the issue of costs overruns. This occurs when a project uses more money than it was budgeted. If a project exceeds the budget, it can represent a significant risk, as forecasts were made to try to avoid this type of situations. Deviations may be considered normal depending on the analysis prior to the start of the project, but going well beyond the budget is not positive. The lack of adequate personnel to complete tasks, resulting in delays to start them, the need of more materials than anticipated, hiring subcontracted personnel at last minute to meet deadlines and to address mistakes made by your original team members (meaning by, paying tons of money, more than expected), may be related to costs overruns in a project, which can affect the budget.

Also, the failure to deliver deliverables in a timely manner is a risk associated with inefficient resource allocation in project management. They may not be submitted on time if team members lack the required knowledge or tools to accomplish their assigned tasks. This can have an adverse effect on the project’s progress, directly impacting timeline, schedule delivery dates and planned budget.

Another significant risk related to resources and the previously mentioned hazards is the failure to comply with the plan. Failure to adhere to the plan due to reasons such as late deliveries (due to not effectively allocating resources), delayed project start, or external factors beyond the project environment could have an adverse impact on the scheduled project completion date. However, it cannot always be ensured that everything will go as expected.

As a result, the project outcome may fall short of expectations and lead to produce work that is below the required standard by the team. This can be attributed to a lack of

expertise or necessary tools to meet the standards for the product or service. In turn, this can result in a low-quality end product or service. Failing to deliver as expected can also lead to team discouragement and low morale. Additionally, not meeting these expectations can damage the organization's reputation, which can have a negative impact on future project outcomes.

To effectively build a project team and allocate resources, project managers can assess their team members based on previous experiences, educational qualifications, social skills, and other competences that could be relevant to the project. While these variables and more can be helpful completing this task, in today's world surrounded by technology, project managers can use technological tools to assist them in this process and also in other aspects of managing projects such as planning and communication between stakeholders. It is known that artificial intelligence is a hot topic in the market and it starting to get relevant between project management. By this, how AI is being used within projects? How can artificial intelligence benefit the process of resource allocation in project management?

2.4 WHY USE ARTIFICIAL INTELLIGENCE TO ALLOCATE RESOURCES?

Artificial Intelligence seeks to reproduce human-like ways of perceiving, reasoning, learning and solving problems [3]. Contrary to popular belief, AI does not necessary involve robots; it can refer to computer programs that intend to imitate human behavior, specifically to make decisions based on past experiences. In terms of AI algorithms, this could mean "learning" from data. AI has the potential to extract information and support decision-making research [3]. Artificial intelligence has been in existence for years, but its only recently that its potential to contribute to society has been discovered. Now, it has been given the visibility it deserves. It's a powerful tool that can benefit society in such projects related to the medical field, education, transportation and the environment. Through the thesis, it will be provided a deeper exploration of this field. For this section, several examples will be presented related to the uses of AI in project management and how can it be used for resource allocation tasks.

USE OF AI IN HR PROCESSES

In organizations, AI models are used in terms of HR recruitment processes by becoming a part of digitization¹ of the field. Recruitment, performance evaluation, best practices analysis and many more related to the field are being integrated within HR Information Systems [4] Recommendation engines integrated in HR Systems play a big part, since they can recommend candidates based on data recollected by them. [4] This can contribute to recruiting candidates for the staff who will be valuable for the development of new projects in the organization.

¹ The process of converting information from analog to digital, usually confused with digitalization

USE OF AI IN CONSTRUCTION PROJECTS – ML AND DEEP LEARNING

Artificial Intelligence is being used in various stages of Building Construction Industry (BCI) projects. For example, the modeling of the economic reuse and estimation of residual value of Building Construction and Demolition Waste (BCDW) is being carried out using Artificial Neural Network (ANN) [5]

It has also been used to predict the technical and economic viability of materials and potentially recyclable materials within construction projects. [6] These uses of artificial intelligence algorithms focus on Machine Learning models. Machine learning can be considered a type of artificial intelligence in which the algorithm learns from experience and it's capable of making predictions based on patterns found in data sets. ML models are distinguished by their straightforward ability to automatically retrieve representations of data. [5]

Deep Learning is another type of AI algorithm (or architecture) that is also used in the BCI. It is a computational technique that uses a series of hidden processes with the purpose of exploring data representations and correlating them at diverse levels of abstraction. [7]. Among the application of these algorithms are Convolutional Neuronal Networks² (CNN). Since these algorithms are mainly used for image classification, they have been useful in sorting materials for recycling and double-checking the possibilities of reusing some of the materials in Building Construction and Demolition Waste (BCDW) on BCI projects. [6] Also as a part of Deep Learning, the model Deep Neural Network³ (DNN) has been used specifically in the composition of BCDW and to predict the demolition waste in a Circular Economy with a relative accuracy of 97% [7]. This implies that the model has been highly accurate on the outcome during its usage.

As can be evidenced throughout the article [6] the main purpose of using artificial intelligence in BCI projects it's to encourage the adoption of Circular Economy practices.

USE OF AI IN SOFTWARE DEVELOPMENT PROJECTS

Nowadays, companies are increasingly embracing artificial intelligence. There is a growing demand for developers who specialize in or possess knowledge of this emerging field. It has become a highly discussed topic in the job market. AI based tools are proving to be a life saver for programmers, as they streamline the programming process in diverse areas such as

² Deep learning architecture designed for image processing [8]

³ Feed-forward neural network with many layers of transformations and nonlinearity. The output of each layer feeds the next layer. [9]

debugging, designing strategies, features, documenting code and even with virtual assistants. With a wide variety of tools available today, companies are investing in them to enhance and augment the capabilities of their programmers, ultimately optimizing project development within the organization. Subsequently, some AI tools that can be applied within software development projects will be discussed.

For the main activity on software development, which is programming, numerous tools exist to streamline this process and support programmers in their tasks. GitHub Copilot is one of the most prominent tools nowadays, since its ability to transform natural language into code line suggestions. Powered by Open AI's Codex, this tool joins the ranks of resources aiding programmers in this domain. Also, Microsoft offers the Visual Studio Intellisense platform, which assists programmers during code composition.

The debugging and maintenance aspect is also vital, as it involves identifying errors in the code and ensuring that the code operates at its optimal performance. As tools for this task, here is debugcode.ai, a web platform where you can provide your code and it will verify potential errors and suggest solutions. In this context, there's also Deepcode, which highlights critical bugs in your code. Both of them are useful tools for these processes.

In the field of data mining, there are RapidMiner and Microsoft Azure Machine Learning, which enable programmers to analyze big data and create artificial intelligence models for any organization, responsibly.

The documentation aspect is also one of the most important, as it describes how the code will work, any special considerations that need to be known, and helps maintain code readability and maintainability over time. For assistance in documentation, the most well-known tools in the market today are GPT-3 and Mintlify.

Finally, as part of the research and resource acquisition for programming, there are several libraries and resource platforms that assist programmers in this task. **TensorFlow**, IBM Watson and MS Azure AI Platform are some of them. Their objective is to offer organizations the possibility to implement artificial intelligence in their businesses. They provide various tools and documentation to make the process achievable for the companies.

After discussing various applications of artificial intelligence in different types of projects, how can artificial intelligence be applied to resource allocation in project management? Is it possible to use it for this purpose? Certainly, recommendation systems are part of machine learning. In general, they involve recommending personalized "items" to a user based on data collected about the user. This is a broader topic that will be exposed in other sections of the research. As a recommendation system, a machine learning model could be designed to

recommend which candidates within the company would be most suitable to perform a specific task. While it will never replace the role of a project manager, it can be a valuable supportive tool for resource management and allocation within an organization.

3.OBJECTIVES AND SCOPE OF THE RESEARCH WORK

3.1 OBJECTIVES

The main objective of the research is driven by the following interrogative: *How can we utilize AI for resource allocation in project management?* In order to achieve this, the intention is to explore the utilization of AI in projects in general and subsequently, propose several models or algorithms (based on a preliminary analysis) that can be applied for this task.

It is crucial to evaluate all the gathered options through various comparative analyses to choose the best one for resource allocation. Furthermore, the necessary tests and experiments will be conducted accordingly. The experiment itself will be performed based on various hypothetical scenarios. Eventually, conclusions can be drawn from the upcoming experiment.

3.2 SCOPE: WHAT ARE THE CONSTRAINTS OF THIS RESEARCH?

The scope of this research is to explore AI utilization in projects in general and its application to resource allocation. With this in mind, various hypothetical scenarios will be employed to experiment with a AI algorithm to accomplish this task.

Along the work, a state-of-the-art analysis is performed by reviewing the most recent research based on AI and project management. There's also a valuation of models, in which its analyzed different models that can assess the principal problem: resource allocation in projects. Based on this analysis, a model is selected in order to carry on the experiment, in which 4 different scenarios with different employees, projects and competencies are tested. The research will explore how can AI be used in projects, particularly for resource allocation. Some hypothetical scenarios will be used to conduct the experiment and evaluate different AI algorithms. The data analysis will identify patterns and trends. The results will be evaluated, leading to relevant conclusions. Additionally, limitations and potential areas for improvement in the implementation of the solution will be identified.

3.3 POTENTIAL BENEFITS BY IMPLEMENTING AN AI ALGORITHM FOR RESOURCE ALLOCATION

Resource allocation can become a tedious and time-consuming task due to the consideration of various technical and non-technical factors. These factors include the experience and technical skills of the resource, time management, availability, workload, and other relevant aspects.

Following a thorough brainstorming session, several key benefits of implementing an AI algorithm for resource allocation have been identified.

1. Simplified and standardized process

The resource allocation process will be simpler, as the project manager will not have to manually evaluate potential candidates for their project. This standardized process ensures consistency across the entire company, making it easier for everyone involved.

2. Enhanced decision making

Using the algorithm to make decisions will result beneficial, as these data-driven decisions will be supported by information gathered and analyzed by the algorithm. By relying on real data, these decisions have the potential to generate positive outcomes for the project in progress.

3. Time management

The process will be efficient, resulting in saving time for resource allocation. The project manager will no longer need to manually allocate resources, eliminating the time-consuming nature of this task.

4. Cost savings

Costs will be reduced by selecting an optimal resource. There will be no waste of money (or little) in reassigning the same tasks for another resource, nor will there be time lost. Time wasted results in financial losses as well.

5. Efficiency in work teams

Work teams can become more efficient when tasks are assigned to them based on their skills and knowledge.

6. Potential for achieving optimal results in the project and meeting project objectives

By achieving efficient teamwork, significant cost and time savings can be realized. Moreover, there is a high possibility of attaining outstanding project outcomes while successfully meeting its objectives.

7. Meeting the organizations expectations and strategic objectives

By successfully delivering the anticipated project outcomes and ensuring customer satisfaction, this can effectively support the organization's strategic objectives.

To sum up, the implementation of an AI algorithm for the task of resource allocation in projects holds several potential benefits for the project manager and the organization itself. It is a simple and standardized process, data driven decisions that can result positively for everyone involved. By leveraging AI technology, project managers can streamline resource allocation, make informed decisions and subsequently, drive successful project outcomes.

3.4 SUSTAINABILITY OBJECTIVES

This section discusses how the research work aligns with the Sustainable Development Goals. According to the research, the two goals that aligned more with it are: Goal 8 (Decent Work and Economic Growth) and Goal 9 (Industry, Innovation and Infrastructure).

DECENT WORK AND ECONOMIC GROWTH

Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all.



By implementing AI algorithms for resource allocation, in this case, employees specifically in projects, it can be contributed to Goal 8 in terms of optimizing workforce utilization and enhancing overall project efficiency. It is explored how AI algorithms can be used for allocating employees efficiently into different projects, considering factors such as competencies and skills. This can lead to better workforce management, reduced skill mismatches and improved project outcomes.

INDUSTRY, INNOVATION AND INFRASTRUCTURE

Building resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation.



With this research work, it can be contributed to enhancing the efficiency and effectiveness of infrastructure development and resource allocation in projects. It is expected that the work will follow a line of research in the field of artificial intelligence and project management in general.

4. ANALYSIS OF THE STATE OF THE ART

INTRODUCTION

Nowadays, artificial intelligence it's a hot-topic as it is. Through this analysis, the intention is to explore the state of the art in the application of AI in project management. The purpose of this analysis is to first define the concept of artificial intelligence, how it has been integrated project management, explore some trends and future lines of research in the field of AI in projects.

First, what's the meaning behind state of the art? The state of the art can be considered as the highest or current level of innovation in a particular field. In this case, AI and project management. As part of the research project to analyze and apply techniques or algorithms for resource allocation in projects, it seemed important to first analyze various techniques and learn the basics of AI. Thus, it is able to see how up to date the field is with respect to project management today. The analysis is vital, as it is able to provide a broader perspective of the field. At the same time, it is useful for the organization of ideas that ultimately contribute to the development of the project.

To develop the analysis, the predominant method being used is the search and interpretation of academic articles in academic databases, as well as the definition of terms by dictionaries and relevant sources. References based on electronic journals have also been used.

THEORETICAL FRAMEWORK

¿WHAT IS ARTIFICIAL INTELLIGENCE?

The term Artificial Intelligence is a complex term to define specifically. It has evolved over the years. The term was first introduced by John McCarthy in 1955 [8]. He defined it as "the science and engineering of making intelligent machines" [9]. Then, in 1956, it was at the Dartmouth Summer Research Project on Artificial Intelligence that AI was established as a discipline [9]. Today, AI can refer to the simulation of human intelligence processes by machines. As the author of the [9] article states: AI stands for the imitation by computers of the intelligence inherent in humans.

APPROACHES AND FUNCTIONS OF ARTIFICIAL INTELLIGENCE

In the present day, artificial intelligence encompasses a wide range of applications and capabilities. In this section, various functions of AI will be explored, including expert systems, machine learning, natural language processing, computer vision (also known as machine vision), neuronal networks and the utilization of AI in robotics.

1. MACHINE LEARNING

Machine learning is a very famous approach of AI nowadays and is in rise. As the authors from article [11] stated: “Because of the abundance of data, machine learning has the potential to contribute to solving complex problems in academia and industry”.

As Arthur Samuel defined; Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed [11]. Machine learning approaches can be categorized generally as: supervised learning, unsupervised learning and reinforcement learning [12]. The purpose of ML is to learn from the data that is given and relies on different algorithms to solve problems. It is important to note that, there is no unique algorithm to solving a problem. The choice of algorithm relies on the specific problems aimed to address, variables involved, type of model that would be most suitable and other factors.

Some of the most common algorithms used in Machine Learning approaches are the following:

LINEAL REGRESSION

The lineal regression is a model built on the assumption that the response variable (output) is a linear of function pf explanatory variables (input variables) [11]. This model can be used to assign a class to each input data point, to predict continuous variables [11]. The goal of linear regression is to be capable of establishing a relationship between 2 variables. There are several applications of this model in real life cases, like for example: predicting the future stock prices based on historical data, analyzing the climate patterns and predicting student performance based on variables such as: attendance to class, dedicated study time, grades, etc.

DECISION TREES

Decision trees are algorithms related to classification, which is mainly used for grouping purposes. It is a model where; the nodes and branches are interconnected to form the tree. Each node represents a feature in a category to be classified and each subset defines a value that can be taken by the node. [13] Due to its simplicity and ability to handle multiple data forms, decision trees have been implemented in various fields for precise analysis. For example, they can be used to improve email marketing campaigns for e-commerce, since after applying the algorithm, it can predict customer loyalty & improve response rate of predict email campaigns [14]. It can also be a tool to predict diseases based on symptoms, predict student academic performance [15] and to detect fraudulent activity by analyzing patterns in data.

SUPPORT VECTOR MACHINES (SVM)

Support Vector Machines serve as supervised classifiers which has been proved highly effective in solving a wide range of pattern recognition and computer vision problems [16]. The primary objective of Support Vector Machines is to construct a decision boundary that effectively separates two classes, enabling the prediction of labels based on one or more feature vectors [17]. The authors from the article Applications of Support Vector Machine (SVM) Learning in Cancer Genomics expose that; “This decision boundary, known as the hyperplane, is orientated in such a way that it is as far as possible from the closest data points from each of the classes. The closest points are called support vectors”. SVM has been used in text categorization, computational biology, banking, medical imaging and genomics [17]. In genomics, the algorithms have been used in cancer classification, since it was first used to classify 2 different types of leukemia using gene expression microarray data [17].

2. ARTIFICIAL NEURONAL NETWORKS (ANN)

Artificial Neuronal Networks is and AI approach inspired by the functions of the biological nervous systems, particularly the brain’s neuronal networks. ANN can be comparable machine produced to function the same way the human brain performs a given task of interest [18]. They are highly beneficial models used for clustering, pattern recognition and prediction and find general applications in tasks such as image recognition, natural language processing, and excel in self-learning, adaptivity fault tolerance, and other advanced capabilities [18]. ANN can also be applied in different disciplines such as engineering, medicine, agriculture, mining, technology, business, art, etc [18].

Some of the most common algorithms used in Artificial Neuronal Network approaches are the following:

GENERATIVE ADVERSARIAL NETWORKS (GANs)

Generative Adversarial Networks are generative models which were introduced first by Goodfellow in 2014 [19]. GANs work like this: it comprises a generator and a discriminator. As the authors of the article [19] state: the generator tries to capture the potential distribution of real samples and generates new data samples; and the discriminator is often a binary classifier, discriminating real samples from the generated samples as accurately as possible. Some of the applications of GANs is to generate

photorealistic object images (animals, people faces), indoor and outdoor scenes, translate images from a source domain to the target domain, generate high-definition images, and others [19]. Basically, the objective of GANs is to generate realistic data that emulates the characteristics of the original data distribution.

MULTI-LAYER PERCEPTRON (MLP)

Multi-layer Perceptron is an algorithm that generates a nonlinear function that enables the prediction of output data from given input data [20]. It applies a supervised training procedure using examples of data with known outputs [20]. Is composed of interconnected nodes, also known as neurons. It's an algorithm commonly applied in classification and regression tasks, with various cases such as image recognition, natural language processing and financial forecasting.

CONVOLUTIONAL NEURAL NETWORKS (CNNs)

Convolutional Neural Networks are able to extract features from data with convolution⁴ structures. CNNs draws inspiration from the process of visual perception. CNNs kernels represent receptors that can respond to various features; activation functions simulate the function that only neural electric signals exceeding a certain threshold can be transmitted to the next neuron [21]. Several applications of CNNs include predicting time series data such as electrocardiogram time series (ECG), forecasting weather patterns and predicting traffic flow[21]. Additionally, CNNs are valuable for recognizing images and identifying objects.

3. NATURAL LANGUAGE PROCESSING (NLP)

Natural Language Processing is an artificial intelligence approach in which the focus is to learn, understand, recognize and produce human language content [22]. Initially, the applications were focused on machine translation, speech recognition and speech synthesis [23]. Nowadays, developers have crafted tools that are capable of practical application, including the construction of spoken dialogue systems and speech-to-speech translation. Furthermore, these tools can extract valuable insights from social media and discerning sentiments and emotions related to products and services [23]. Some of the most common Natural Language Processing algorithms are Tokenization, which breaks down text into smaller units. Also, there's Name Entity Recognition (NER), which is able to identify text such as places, people, organizations, etc.

⁴ Mathematical operation that combines two functions to create a third function

4. COMPUTER VISION

Computer Vision refers to an approach that encompasses concepts aimed at mimicking the capabilities of the human visual system [24]. Is able to perform tasks such as object and image recognition and generation, classification, vehicle navigation, face recognition, and many others [24]. In practical scenarios, Computer vision has been applied in the fields of: agriculture, in order to develop smart farming approaches for automation of processes [25]. Also, it has been applied in geology. A dataset named Geo-Fossils I has been developed to provide accessibility to the geology community for the classification of fossils and related materials [26] Another application is in the medical field, specifically in the laparoscopic surgery procedure [27]. The application of CV in the medical field represents a very important milestone, as its contributing with the development of support tools for making real-time automated decisions and surgeon training systems [27].

5. AUTOMATED PLANNING AND SCHEDULING

Automated Planning and Scheduling is an artificial intelligence subfield in which computers automatically plan and schedule actions and events [28]. This encompasses the ability to plan and schedule tasks, allocate resources and manage events. Some fields in which this AI approach is being used are: manufacturing, robotics, IT and Project Management. Some of the benefits of using this approach is the optimization of resources and time-saving [28]. This approach employs a range of AI algorithms to achieve its objectives and currently stands as one of the most actual research domains.

6. CASE-BASED REASONING & EXPERT SYSTEMS

Case-Based Reasoning is an AI approach that draws inspiration from the role of memory and reminders in human cognition. As exposed in the article [29] Case-Based Reasoning (CBR) addresses new problems by remembering and adapting solutions previously used to solve similar problems. Miquel Sánchez-Marré, the author of Principles of Case-Based Reasoning, references a Case-Based Reasoning Cycle, originally introduced by A. Aamodt and E. Plaza, which consists of 4 main steps: *Revise*, “the most similar cases/problems to the one being solved”. In second place, *Reuse*, “the information in that case/problem in order to solve the case/problem”. In third place, *Revise*, “the proposed solution”. Finally, *Retain*, “the parts of this experience to be reused for future problem solving”.

On the other hand, Expert Systems are computer-based systems that simulate human decision making [30]. Essentially, these systems were developed to mimic cognitive processes of an expert in various domains of knowledge. This approach is based on two key elements: knowledge base, which is all the information need that an expert is going to use to decide [30]. Is the information that the system is going to recollect and analyze in order to make and informed decision. The second element is, the inference engine, which simulates human decision-making based on knowledge base and a rule base [30]. This approach is being used in a lot of different fields such as medicine, geology, engineering, etc.

So, how do these two techniques relate to each other? Case-Based Reasoning serves as a problem-solving approach that seamlessly integrates with Expert Systems. The integration aims to improve decision-making and replicate cognitive processes of humans in specific cases.

APPLICATION OF ARTIFICIAL INTELLIGENCE IN PROJECT MANAGEMENT

Thorough the exploratory analysis of advancements, a variety of ways in which artificial intelligence is applied across various domains of project management have been identified. This section will provide an overview of these uses and applications.

As it is known, project management can be divided into some main concepts such as: resource allocation or management, scheduling, decision making, communication, team management, risk management and monitoring and control of projects. These aspects are very important in project management, and managing each of them in an effective way could be beneficial for the project results.

Before delving into the use of AI in project management, it is crucial to highlight some key terms about the implementation of these tools in organizations in order to achieve a successful incorporation of them. According to the article [31], the authors expose the following specific points to consider.

The article [31] discusses how we can identify uses of artificial intelligence in companies. It discusses several points that should be considered before adopting these new technologies. First, it states that organizations do not understand what artificial intelligence is, which leads companies to have exaggerated expectations of what this technological tool is. To mitigate this problem, it is important to educate about what artificial intelligence is, before identifying the possible uses and implementation of AI in the company. The authors emphasize [31] that data is an indispensable element. That is why the criteria for AI application, the

available data and the enterprise must be on the same page when using or implementing AI. It is concluded that, if these three elements mentioned above are not met, the potential of AI is at risk, which means that a wrong implementation may be made and the full potential of this tool may not be exploited.

According to the article Developing purposeful AI use cases – A structured method and its application in PM, the authors expose the use of linear regression to predict future project effort based on cost drivers and software size variables. This means that, with a linear regression approach, one is able to predict the effort needed for the project, based on information from past projects already executed. Also, using the NLP (Natural Language Processing), it is possible to predict the cost of projects, which could be a support system for project decision making [31]. This can improve aspects in terms of estimations, project scope and task change assignments, this specifically in software development projects according to the article [31]. What is common in several articles related to the application of artificial intelligence in project management is that, instead of strictly deciding or acting as a PM or human being, artificial intelligence serves as a decision-making tool, which triggers the task to be performed by the PM.

What has been found in terms of AI in project management has been the implementation of AI-driven bots in various project management software. Support bots and chatbots have been identified that simplify and assist the project manager in fulfilling his duties. Some examples of these applications are:

Bot Stratejos – This bot supports project team members in terms of reminding them of daily tasks, deadlines creating new tasks or reminding them for the completion of certain tasks. It is a bot that contributes to the areas of scheduling and decision making [31].

Redbooth Bot – This is a bot that also shows daily tasks to employees and risks and issues that they might face or that they are currently working on [31]. Which contributes to the scheduling area as well and the risk management part.

PMOtto – This tool possesses the capability to perform tasks typically handled by project managers, offering alerts related to task deadlines, costs and resource allocation, all driven by machine learning algorithms [31]. Functioning as a personalized assistant for the project manager, it provides support within project management software such as Microsoft Project [32], thereby offering potential assistance withing scheduling.

Bot Extensions – There are also other extensions which can be combined with various project management software. In this case, one of the best known is the use of bot extensions for Atlassian tools, such as Jira and Confluence. These bots are dedicated to support the project team in data entry, risk analysis and project monitoring [32] .

Other technological tools for project management based on artificial intelligence include TARA. TARA is a platform that focuses on product development lifecycle and project planning and monitoring. It uses machine learning to automate the definition of project approach, tasks and time planning [32]. It is also capable of creating work teams and monitoring and forecasting projects [32].

Cloverleaf is another AI-based platform that supports gaining insights into how project team member's work. It is a very useful tool for the project manager to see how the team works. Here it is possible to identify strengths and weaknesses of employees in terms of skills [31], which could also allow to put together a good project team. This tool is an excellent support for team management, human resources in projects.

BENEFITS AND RISKS OF AI APPLICATION

After reviewing multiple articles on the implementation of artificial intelligence, it has been possible to recognize a series of advantages and possible risks in relation to artificial intelligence and project management. In this section, benefits and risks will be presented, accompanied by a brief contextualization.

Table 1: Benefits and Risks of AI Application

Benefits	Risks
Automation of repetitive tasks	Employment
Effective problem solving	Data quality/availability
Decision making based on data	Human interpretation
Project scheduling optimization	Resistance to change
Liabile predictions of costs, time, dates, etc.	Uncertainty

One of the benefits is the automation of repetitive tasks. With the help of AI, it becomes possible to automate repetitive tasks that do not require human intervention. Thus, project managers can focus on other, more vital tasks that do require their intervention. Furthermore, there exists the capability of effective problem solving, since artificial intelligence can assist in resource distribution, while also relying on predictive insights in terms of costs and delivery timelines. Another very important benefit is the decision making based on data. The project manager will be able to make decisions based on historical data with the help of AI implementation, resulting in well-informed decisions that can be beneficial for the outcome of projects. Optimizing project scheduling is also achievable, with the help of the AI tools mentioned earlier. To conclude with the benefits, implementing reliable algorithms will result in reliable predictions in terms of costs, time estimates, delivery dates, among others. It is

always important to note that, these and other benefits could be identified by implementing AI for project management, as long as it is applied in a responsible and consistent manner.

On the other hand, although artificial intelligence is seen as ideal, there are also risks in its implementation. In the article Optimizing Project Management using Artificial Intelligence the author exposes some risks in the use of AI and its impact on project management. One of them is employment. When using AI, it is likely that for simpler and repetitive tasks it is not necessary to have the intervention of an employee, but at the same time, as an impact on project management, the algorithm may not correctly assign a task to an employee [33]. This also depends on how the algorithm is applied and the data available for its implementation. This is another risk that can be identified. It is possible that the algorithm may generate inconclusive results for the project due to the lack of data or the quality of the data [33]. Human interpretation can also be a risk in this aspect. A person without knowledge or context of the problem or how the algorithm works could interpret the results incorrectly, which could affect the decision making and consequently the results in the project. Resistance to change is also a very important issue, as the people involved in the organization or project may not agree to use AI as a support tool. This could have a negative impact on the course of the project. Finally, one of the most important risks is uncertainty. Uncertainty could be reflected in several ways, because the result of an AI algorithm is not necessarily correct by itself, as it depends on several factors, such as the available data, the quality of this data, also which model will be applied for the specific case or task to be solved with the implementation of artificial intelligence. It is important to note that decisions should be made not only based on data, but also on experience and intuition, which are not necessarily skills that a machine can mimic.

HIGHLIGHTED STUDIES

Project management is a currently undergoing evolution and the adoption of no novel technologies to align with the advancements of today's world. In this section some studies will be shown that have been done regarding the implementation of artificial intelligence with project management and how they have been able to contribute positively to it.

In the article [34] some challenges that companies face when meeting deadlines are studied, specifically in software development projects. The study focuses on interviewing team members and project managers to identify challenges they face in meeting deadlines. An AI framework is proposed to mitigate these challenges. The framework was developed based on the issues identified and expressed by the company's employees; it was evaluated and 72% of the study participants rated it as a useful tool and were willing to integrate AI in the work area to facilitate the task process and meet deadlines. The results of the study conclude that AI could be used to improve project management and that companies should consider

implementing this tool to meet deadlines. Deadlines are an important part of projects, and if they are not met, the project could be at risk and thus be a failure.

The article [6] delves into the role of AI in enhancing circularity within construction projects. It highlights 13 specific domains within the Business Construction Industry (BCI) where AI has the potential for application. For example, for the selection of circular materials, ANN (Artificial Neural Networks) algorithms could be used, which would be able to decide between the various options of circular materials depending on the scenario and circumstance needed. For the prediction of hazardous materials, a machine learning algorithm is used which is able to categorize the hazardous materials based on the building register [6]. The model has had an accuracy of between 74% and 83% [6]. To estimate the Building Construction & Demolition Waste Generation, machine learning techniques such as linear regression, ANN, and Grey Model have been adopted. The Grey Model and Artificial Neural Network achieved the highest prediction accuracy [6].

It is important to recognize that project success can hinge on both the project manager and the effectiveness of project planning. In the article [35], the authors embark on research to apply AI in software engineering projects. They comment that AI could be used to reduce errors in project planning and thus avoid project failure. In the study, the authors have analyzed several articles related to the implementation of AI in software engineering. One of the papers that have been analyzed by the authors is Towards Effective AI-Powered Agile Project Management. Here, an AI framework is proposed to support agile PM with the aspects of automating repetitive tasks to meet project analytics for estimation and risk prediction [35]. Another article mentioned is: Artificial Intelligence Based Risk Management Framework for Distributed Agile Software Development, where an artificial intelligence framework for risk mitigation is proposed. The authors state that "manual risk management is inefficient, since it depends exclusively on the experience and analysis of the person who carries it out". Therefore, it should be automated. Finally, in the article Analysis of Software Engineering for Agile Machine Learning, the challenges of overseeing machine learning data within the Scrum framework are evaluated. Concerns regarding collected data are analyzed, and recommendations for improving execution of Agile machine learning projects are proposed.

In the pursuit of effective project management and meeting deadlines, the integration of AI emerges as a transformative solution. In one study [34], challenges to meeting deadlines in software development projects were addressed through an AI framework, yielding positive results. This underscores the potential for AI to enhance project management and underscore its importance in avoiding project failure. Similarly, another study [6] showcased AI's application in improving circularity within construction projects, demonstrating its versatility across diverse industries. As AI's influence grows in project management, its capacity to reduce errors in planning and elevate project outcomes is becoming evident. With AI frameworks

bridging gaps and improving decision-making, it's evident that AI is steering projects towards greater success, mitigating challenges and ensuring project deadlines are met efficiently.

TRENDS AND FUTURE RESEARCH

After the research, it is evident that one of the trends identified is bots based on AI (machine learning mainly) and virtual assistants. Bots working as stand-alone software or as extensions to other project management software. In the future, the authors of [35] mention that artificial intelligence will evolve to be able to assist project managers with much more complex tasks than they have been able to automate today. Future areas of implementation include planning, reporting and monitoring [35].

In the area of construction projects, as part of the Optimization of Circular Product Management Infrastructure and Material Flow (as mentioned in the article [6]), the authors expose the potential of AI to optimize circularity management infrastructures, facilitating the efficient recovery of valuable materials, is a promising area for future research exploration. Another area the authors mention is the Optimisation of Circular Business Model. They mention that this is one of the main areas where AI is applied. They mention that; "future research is needed to explore how ai could bring together the key aspects of a successful CBM, such as predictive maintenance, pricing of circular materials, development of trading platforms for secondary materials, demand predictions for circular materials, and intelligent inventory management in BCI".

An interesting area for future research is advanced risk management. Artificial intelligence could be able to identify patterns in similar completed projects and identify potential risks with the current project. Also, it could predict the delivery dates of projects based on other deliveries, which could also be applied to deliveries by team members or external companies that are not in constant contact or control.

According to the papers analyzed, the advances in the implementation of artificial intelligence in project management are repetitive tasks or tasks where human intervention is minimal. Most of the tools are complements to more advanced software and serve as extensions and support to the main tools. It will be interesting to develop complete software that have integrated AI in several areas such as planning, communication, resource allocation, delivery, project closure, lessons learned sessions, among others.

LIMITATIONS

The analysis of the state of the art has been limited to data collection on the basics of artificial intelligence, most common AI models, which AI models have been used in various areas of project management and how this has been done. Some current studies have also been analyzed in relation to the topic.

5. VALUATION OF ALTERNATIVES FOR RESOURCE ALLOCATION

Before choosing a model or algorithm to proceed with the experiment, some factors were evaluated to ensure its effective implementation with different scenarios. This section includes the essential characteristics that were considered, highlighting the chosen algorithm to work with.

The first thing is to analyze, in what circumstances will the model be applied? How specifically is artificial intelligence going to be used for? It is important to do this exercise, since there are an infinite number of algorithms and ways to automate certain tasks, so it is a vital step to analyze well and understand the particular case with which it will be worked with.

Once the case has been analyzed, it is important to consider the amount of data. There are algorithms that work or give better results depending on the amount of data available. This also puts into perspective that depending on the complexity of the case to be solved, an algorithm with more or less capacity should be chosen

Regarding the characteristics considered for evaluating and selecting a model for the experiment, its included:

- 1. Model Implementation Complexity**

As the experiment to be performed is a fairly simple one, the aim is to obtain an algorithm that is easy to implement, but at the same time that meets the results expectations of the experiment.

- 2. Model Execution Time**

Look for an algorithm that is effective in achieving the results after execution. This will also depend on the amount of data available and other factors.

- 3. Interpretable Model Results**

It is important to emphasize that the results of the model can be easily interpreted.

- 4. Model Adaptability**

A model will be considered which can be adapted as the needs to apply it grow.

- 5. Performance of Various Scenarios**

A model that is consistent and performs well in various scenarios depending on the case in which it is to be run.

After analyzing the characteristics to consider, the possible models or algorithms to use are:

Table 2: Algorithms and Characteristics

Algorithm	Approach	Definition	Highlights
Decision Trees	Machine Learning	decision making based on the evaluation of characteristics. the model evaluates conditions until it reaches a final node, which contains the result of the operation.	Easy interpretation of results, adaptability, sequential decision making
K- Means	Clustering/Machine Learning	is capable of dividing a data set into groups, the objective of which is to group all similar points.	Easy implementation, easy interpretation of results, adaptability for different scenarios
Collaborative Filtering	Recommendation Systems	is used to recommend products to users based on their interactions with other products.	Based on user and item filtering, based on user interactions, similar users or elements
K-Nearest Neighbor	Classification/Machine Learning	finds data points closest to a new data point that is not recognized. This information is used to make a final classification decision.	Simple implementation, great for small datasets, adaptability

The decision trees can become very dependent on the quality of the data and it is also a model sensitive to data variations, which is sought to make the model as stable as possible. Collaborative filtering could be an ideal model, since it is a recommendation system. The negative side is that it is a model that is based on user and element interactions, which does not fit with the case to be tested. It is also a model that requires a significant amount of data. Trying to fit it could take a lot of time, which could be saved by choosing some other approach that is more suitable. K-Nearest Neighbors could be a possible solution, only, it does not show the groups of people or users classified, but classifies a new user to a group, which for the particular case to be experimented is not adequate.

Table 3: Evaluation of Algorithms

Algorithm	Model Implementation Complexity	Model Execution Time	Interpretable Model Results	Model Adaptability	Performance of Various Scenarios	Total
Decision Trees	4	4	3	2	2	15
K-Means	5	5	4	4	3	21
Collaborative Filtering	3	4	5	2	2	16
K-Nearest Neighbor	5	4	4	3	3	19

The proposed algorithms evaluated have been selected depending on the overall approach and after reading about possible algorithms to be used. An evaluation was done based on the 5 most important characteristics to be considered and each one of them was rated by algorithms, on a scale from 1 to 5. After grading, the algorithm with the highest score was K-Means.

K-Means, a straightforward algorithm to implement that is a suitable one for the case to be experimented. The main objective is to cluster company employees into different projects based on competency evaluations specific to each project, derived from the employee's performance. Due to the available data and evaluation results, this simple

algorithm is well-suited and can achieve interpretable results. The following sections will discuss what is the K-Means algorithm and its implementation across the different scenarios intended for experimentation.

6. WORK METHODOLOGY

This section will provide an overview of the K-Means algorithm, including its definition, usage and other relevant information. Also, the work methodology in which the experiment was carried on. Various scenarios will be presented to test the algorithm's performance with different datasets. These tests aim to compare the algorithm's effectiveness, enabling to draw conclusions about its application.

WHAT IS THE K-MEANS ALGORITHM?

K-Means clustering is a machine learning algorithm which represents a straight forward unsupervised algorithm⁵ employed for addressing diverse grouping problems [1] It groups N data samples into k categories. [2]

The fundamental concept behind this algorithm involves starting with an initial clustering that may not be optimal, and then relocating each point to its nearest center. The clustering centers are updated by computing the mean of the points assigned to them, and this relocating-and-updating process is repeated until convergence criteria⁶ are met. [3]

MAIN USES OF THE K-MEANS ALGORITHM

Some of the most common uses of this algorithm are: data clustering, image segmentation, personalization of recommendations, market analysis and data understanding. There are various applications where classification plays a crucial role. For instance, it can be used to categorize plants or animals into species. Additionally, it can also be utilized for optimizing stock market strategies [4].

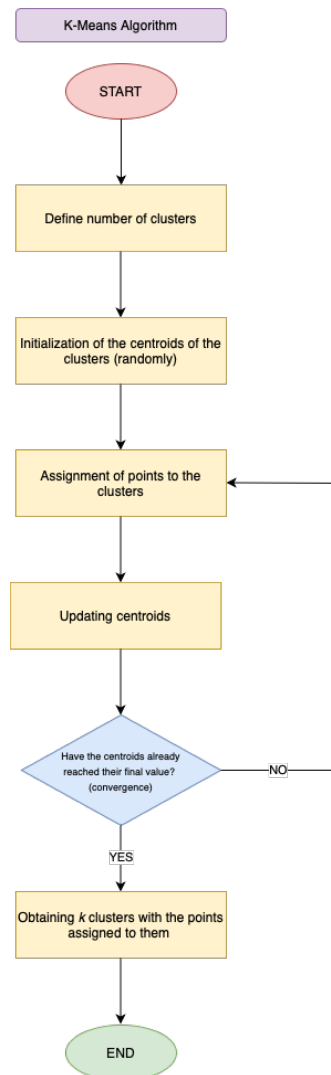
HOW DOES THE K-MEANS ALGORITHM WORK?

- a. You define the k number of groups into which you want to divide the data.
- b. Cluster centroids are randomly initialized.
- c. Cluster points are assigned to the cluster- each data point is assigned to the cluster whose centroid is closest.
- d. The centroids are updated- once all the points are assigned, the centroids are updated/recalculated by taking the measure of all the points belonging to that cluster. These iterations move the centroids of the points to each cluster.
- e. Steps 3 and 4 are iterated until the centroids find a definitive value (a convergence criterion is met).
- f. At the end, the k clusters are obtained with the points assigned to each of them. The points refer to the objects to be analyzed in the algorithm. Each cluster represents groups of data similar to each other.

⁵ uses machine learning algorithms to analyze and cluster unlabeled datasets. These algorithms discover hidden patterns or data groupings without the need for human intervention. [4]

⁶ determines when the algorithm is considered to be finished and has reached a stable solution. Determines when iteration ceases [6]

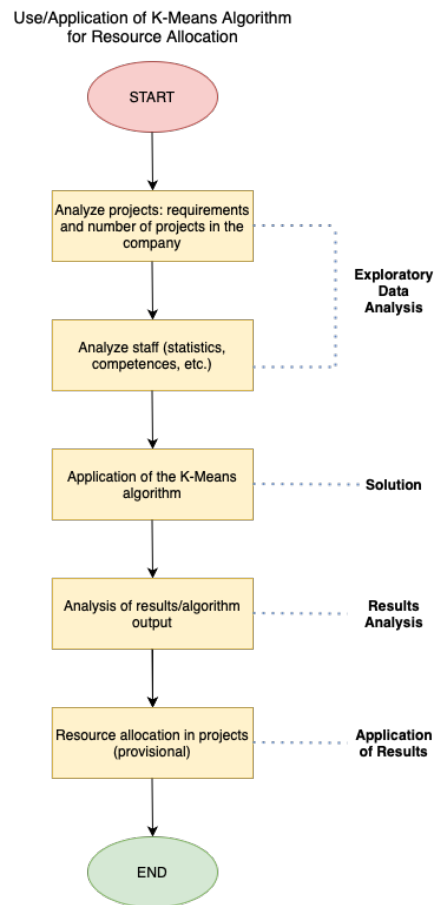
Diagram 1: K-Means Algorithm



ALGORITHM/PROGRAM APPLICATION OBJECTIVES

The purpose of the code or program to be presented is to group all the employees of a company in relation to the available projects. What is sought is to divide all the company's personnel among the projects that exist, based on the employee's qualifications in certain competencies. These competencies are also listed as requirements for each of the company's projects. Some competencies may overlap between projects, while others may not.

Diagram 2: Use/Application of K-Means Algorithm for Resource Allocation



It should be clarified that the allocation of resources according to the algorithm is not definitive, since there are other external factors that may contribute to or affect this allocation, such as availability, workload, etc. These factors are not considered in the research, only competencies that can be assessed.

The application or use of the algorithm is one of sequential steps. In addition, its objective is to serve as a tool to support the project manager to complete the use of duty in relation to the allocation of tasks and resources in the various projects in which is involved. These results are not meant to be definitive, if not, orientated.

TESTING METHODOLOGY

This section will discuss in general the scenarios in which the k-means algorithm was tested. The data used will be contextualized in terms of employees, projects and competencies that employees have been evaluated on.

The scenarios to be evaluated are as follows:

Table 4: Scenarios

	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Employees	65	35	150	300
Projects	6	4	8	10
Competencies	100	67	75	50

To apply the k-means algorithm, various hypothetical scenarios have been developed.

Among other important tasks, a project manager is responsible for assigning tasks/work to their team (resources). It must be visualized per project, which employees are the best fit for the company's projects. This could be based on the employee's performance in various projects or the competencies they possess in comparison to the technical requirements of each project (the algorithm is based on the last concept mentioned)

By implementing this approach, the employee selection process can be enhanced for the project, thereby maximizing performance and achieving optimal results. Ultimately, this will enable the project manager to meet and exceed project expectations.

For this first scenario, there are 65 employees in the company and 6 different projects. In this case, the projects are related to software development and IT. Each project has different requirements to be analyzed. The objective is to distribute the 65 employees among the 6 projects. After this distribution, the task of assigning people to tasks in the different projects will be more manageable for the project manager.

The available data on the company's employees and projects for this case are as follows:

- There is a **projects table** which includes: a unique identifier for each project, project name, project duration, the budget (in euros), the complexity level (1-5), the team size (number of people needed for the project) and the company's project priority (1-3).
- There's also the competencies or requirements necessary to fulfill the company's projects. Every project has a specific set of requirements.
- There is an **employees** table with the following information: employee's unique identifier, employee's name, gender, years of experience and education (1-3)

- There is a **ratings** table that includes all employees with their respective ratings in all competencies, which belong to each of the projects. Each employee must be evaluated in all competencies, which belong to each of the projects. Each employee must be evaluated in a total of 100 competencies, which in this case, are divided across 6 projects.

When the K-Means algorithm is applied, it evaluates the competencies that each employee possesses. It does not include the average performance score of the employee in a specific project, it is only based on technical competencies, which serves as requirements in the different projects of the company.

For the second scenario, there are 35 employees in the company and 4 different projects. In this case, the projects are also related to software development and IT. Each project has different requirements to be analyzed. The objective is to distribute the 35 employees among the 4 projects. After this distribution, the task of assigning people to tasks in the different projects will be more manageable for the project manager. The available data on the company's employees and projects for this case is the same as for the Scenario 1, since the same database is going to be used.

For the third scenario, there are 150 employees in the company and 8 different projects. In this case, the projects are just identified with a unique ID from 1-8. Each project has different requirements to be analyzed. The total of competencies in which all the employees were scored is 75 competencies, which are distributed through all the 8 projects. In this case, the competencies have a unique ID and are also unique to every project. The objective is to distribute the 150 employees among the 8 projects. After this distribution, the task of assigning people to tasks in the different projects will be more manageable for the project manager. The available data on the company's employees and projects for this case was randomized.

In the final scenario, the company has 300 employees and 10 distinct projects. Each project is defined by a unique ID ranging from 1 to 10, and each project has specific requirements for analysis. There are a total of 50 competencies that have been assess for all the employees, and these competencies are distributed across the 10 projects. Each competency is assigned to a unique ID. The goal is to allocate the 300 employees to the 10 projects. Once the allocation is done, the project manager will find it easier to assign individuals to tasks within the different projects, making the process more manageable. The available data on the company's employees and projects for this case was randomized.

7. TESTING PHASE - RESULTS

This section discusses the results of the implementation of the algorithm in the scenarios discussed in the previous chapter. An analysis of the results obtained will be carried out.

7.1 SCENARIO 1:

After applying the K-Means algorithm for the Scenario 1; the obtained results are the following:

Table 5: Results Scenario 1

Clusters	# Employees	Cluster Project ID	Project Name	Avg. Score of Cluster
Cluster 0	6	13	Videogames	2.28
Cluster 1	8	14	Data Analysis	2.28
Cluster 2	14	13	Videogames	2.05
Cluster 3	11	13	Videogames	2.10
Cluster 4	14	12	Mobile Apps	2.2
Cluster 5	11	16	Networks and Systems	2.16

After obtaining and analyzing the results of the Scenario 1, the following conclusions can be reached:

- Firstly, project number 11 (Web Design) does not have an assigned cluster. This indicates that the employees within the company either lack the required skills to handle the tasks associated with this project, or their scores in relation to the project's required skills are comparatively lower than those of other projects.
- In clusters 0, 2 & 3, the Videogames project remains dominant. This implies a significant similarity among employees within each cluster, as their competency scores align closely with the requirements of the Videogames project.
- The cluster with the highest score for the Videogames project is Cluster 0. It means that it may contain the most optimal employees to work on this particular project.
- In Cluster 2, there is no major difference between projects 13 (Videogames) and 15 (Cybersecurity) in terms of score. Employees in this cluster could easily perform tasks in the Cybersecurity project.
- Within Cluster 3, the Videogames project once again prevails. However, another project closely ranked in terms of scores is project 16 (Networks and Systems). This indicates that employees within the cluster could comfortably perform tasks in both projects.
- Cluster 5 is assigned to project 16, and interestingly, there isn't a significant difference in scores compared to Cluster 3 employees (2.10). This suggests that the employees in Cluster 3 could also effectively contribute to project 16 (Networks and Systems).

- The clusters with the most employees are Cluster 2 and Cluster 4, with 14 employees respectively.
- The clusters with the fewest employees are Cluster 0, which has 6 employees, and Cluster 1, with 8 employees.
- The highest averages are: Cluster 0 and Cluster 1, both with an average score of 2.28.
- The lowest averages are: Cluster 2, with an average score of 2.05 in Videogames and 2.04 in Cybersecurity.
- The conclusion drawn is that a majority of the employees possess high scores to undertake the projects: Videogames and Networks and Systems.
- The analyses lead to the conclusion that the employees lack the competencies and requirements needed to carry out tasks in project 11, which is the Web Design. Consequently, it becomes the responsibility of the project manager to explore alternative approaches to fulfill this specific project for the company.

7.2 SCENARIO 2:

After applying the K-Means algorithm for the Scenario 2; the obtained results are the following:

Table 6: Results Scenario 2

Clusters	# Employees	Cluster Project ID	Project Name	Avg. Score of Cluster
Cluster 0	5	11	Web Design	2.22
Cluster 1	10	16	Networks and Systems	2.06
Cluster 2	8	14	Data Analysis	2.16
Cluster 3	12	16	Networks and Systems	2.16

After obtaining and analyzing the results of the Scenario 2, the following conclusions can be reached:

- In first place, project number 15 (Cybersecurity) does not have an assigned cluster. This suggests that the employees in the company either lack the necessary skills to handle the tasks associated with this project, or their scores in relation to the required skills for the project are relatively lower compared to other projects.
- In clusters 1 (average of 2.06) & 3 (average of 2.16), the Networks and Systems project remains dominant. This implies a significant similarity among employees within each cluster, as their competency scores align closely with the requirements of the Networks and Systems project.

- The cluster with the highest score for the Networks and Systems project is Cluster 3. It means that it may contain the most optimal employees to work on this particular project.
- The cluster with the most employees is Cluster 3, with 12 employees
- The cluster with the fewest employees is the Cluster 0, with only 5 employees.
- The highest average is: Cluster 0, with an average score of 2.22.
- The lowest average is: Cluster 1, with an average score of 2.06.
- The conclusion drawn is that a majority of the employees possess high scores to undertake the project of Web Design. Still, there's no major difference between the average score within all the 4 projects evaluated.
- The analysis reveals that the employees do not possess the necessary competencies and requirements for effectively perform in project 15, which focuses in Cybersecurity. As a result, is the responsibility for the project manager to explore alternative approaches to fulfill the specific requirements of this project for the company. One approach could involve evaluating the average score of each cluster in relation to project 15. Based on that evaluation, the project manager can select the appropriate team members from the identified clusters to work on this particular project.

7.3 SCENARIO 3

After applying the K-Means algorithm for the Scenario 3; the obtained results are the following:

Table 7: Results Scenario 3

Clusters	# Employees	Cluster Project ID	Avg. Score of Cluster
Cluster 0	9	5	2.74
Cluster 1	22	6	2.05
Cluster 2	23	8	2.26
Cluster 3	18	5	2.43
Cluster 4	19	4	2.07
Cluster 5	23	2	2.29
Cluster 6	17	5	2.31
Cluster 7	18	8	2.29

Based on the results and analysis of Scenario 3, the following conclusions can be drawn:

- Projects 3, 7 and 1 have no assigned clusters. This may be due to several reasons. Perhaps because there is an uneven distribution of competencies, when calculating the mean, the results are not as competitive with the mean scores of the other projects. That is, the result may depend on the number of competencies or requirements for

each project. The results may vary if each project has more or less requirements. The size of the dataset also influences these results.

- Clusters 0, 3 and 6 correspond to project number 5, indicating a significant similarity among the scores of employees for this particular project. The repetition of the project across multiple clusters does not necessarily indicate an issue. Instead, it suggests that a substantial number of employees share similar scores relevant to the requirements of project 5.
- Clusters 2 and 7 are both associated with project 8, indicating that the employees in these clusters have comparable scores on the project's requirements. The similarity in scores suggests that these employees possess similar scores on the competencies, resulting in their assignment to the same project.
- The clusters with the most employees are: cluster 5 and cluster 2, with 23 employees each.
- The cluster with the fewest employees is cluster 0, with a total of 10 employees.
- The cluster with the highest average is cluster 0, with 2.74. This may be due to the fact that it is a cluster with few employees and this may influence the average score.
- The cluster with the lowest average is cluster 1, with 2.05. It has a total of 22 employees.

After analyzing the test results with scenario 3, it is concluded that projects continue to be repeated in the clusters. This is due to the fact that there is a similarity between the scores of the employees in these projects. Also, the amount of data for each project may have an influence. Each project has a number of competencies that belong to it, and this may affect the results in one way or another, but it is still a normal behavior.

For repeated projects, the project manager could evaluate the second highest score for each cluster, and from there assign employees to other projects.

For projects that have not been assigned to any cluster, the scores on the requirements of each project may not be as relevant compared to the other scores of the other projects. In this case, the project manager could evaluate each cluster and identify which cluster has the highest average score in the project that has not been assigned, and then choose the personnel to complete the resource allocation.

It is noticeable that the amount of data influences the results (compared to scenarios 1 and 2). Further analysis by the project manager would be necessary to identify which employees could perform tasks in the projects that have not been assigned to clusters. Always keep in mind that this is an indicative tool and not definitive.

7.4 SCENARIO 4

After applying the K-Means algorithm for the Scenario 4; the obtained results are the following:

Table 8: Results Scenario 4

Clusters	# Employees	Cluster Project ID	Avg. Score of Cluster
Cluster 0	17	5	2.33
Cluster 1	36	7	2.35
Cluster 2	30	4	2.31
Cluster 3	28	8	2.24
Cluster 4	24	1	2.38
Cluster 5	45	6	2.42
Cluster 6	33	2	2.44
Cluster 7	32	10	2.30
Cluster 8	28	7	2.37
Cluster 9	27	9	2.51

Based on the results and analysis of Scenario 4, the following conclusions can be drawn:

- Among all the projects, only project 3 doesn't have an assigned cluster. This can indicate that employees may have lower scores in meeting its requirements. Scores greatly influence clustering results. Not assigning a cluster allows for evaluating project requirements and assessing employee strengths and weaknesses to improve clustering and project outcomes.
- Clusters 1 and 8 have been assigned the same project, which is project 7. This assignation is based on the requirements of each project. This can be due to that the employees in the cluster are very similar in terms of que scores of que requirements of project 7. This can mean that, employees in both clusters can perform tasks in project 7 and be successful within it.
- Cluster 5 stands out as the largest cluster among the rest, with a total of 45 employees out of the overall workforce of 300.
- Cluster 0 has the fewest employees, with a total of 17 individuals.
- Cluster 9 results as the cluster with the highest average, with a score of 2.51. In contrast, cluster 3 has the lowest average, with a score of 2.24.

In Scenario 4, the number of employees has increased to about 300 employees. Many more employees compared to the other scenarios. The number of projects has remained relatively low and compares favorably with the others, with 10 projects. What has changed is the number of competencies or requirements per project. In this case each project has 5 specific requirements, adding up to a total of 50 requirements overall. Significantly lower compared to

the other scenarios. It implies that every employee needs to have scores on the 50 requirements/competencies.

This data variance could affect the outcome of the clustering. However, it does not mean that there is a clustering that is correct or more correct than... This depends on the quality and quantity of the data. Also, the context in which the solution is being applied.

8. MODEL EVALUATION

In this section, the evaluation of the K-Means model will be conducted to assess its performance compared to the four scenarios previously tested. Evaluating a model serves the purpose of analyzing its performance and accuracy, specifically how well it can execute the designated task it was designed for. This evaluation process allows for identification of strengths, weaknesses and areas for improvement, ensuring optimal results within the specific application context.

It's important to note that the model evaluations are indicative and subjective, since the results will always depend on the context, the type of data and the quality of the data available at the time of applying the artificial intelligence model.

There's an infinite of approaches when evaluating an artificial intelligence model. In the context of my research, two commonly employed metrics for evaluating K-Means models will be utilized. These metrics are the Silhouette Index (or Silhouette Score) and the Davies-Bouldin Index (DBI).

8.1 SILHOUETTE INDEX

The Silhouette Index is an evaluation metric used for evaluating clustering models and is often used as a clustering quality measure.

As the authors from the article [Silhouette Index as Clustering Evaluation Tool](#) said; the Silhouette Index *"defines for each object in the dataset, the measure of how this object is similar to other objects from the same cluster, in comparison with objects of other clusters."*

The results of the evaluation ranges from -1 to 1, where:

- Scores nearing 1 indicate strong clustering, indicating that the samples are well grouped together. It means that the samples are far away from the neighboring clusters. [5]
- A score close to 0 can mean overlapping or ambiguous clusters. It means that the sample is on or very close to the decision boundary between two neighboring clusters. [5]
- And negative values imply that samples have been incorrectly assigned to clusters. [5]

After evaluating the model with the Silhouette Index for each of the scenarios tested, the results are the following:

Table 9: Silhouette Evaluation Score

Silhouette Evaluation Score	
Scenario	Score
1	0.16
2	0.69
3	0.44
4	0.30

Based on the results of the evaluation of the Silhouette Index on the model, it can be concluded that Scenario 2 has the best score of 0.69, indicating a good separation between the clusters. The Scenario 3 also has a high score of 0.44, while Scenarios 1 and 4 have lower scores of 0.16 and 0.3 respectively.

This indicates that the scenario with the most optimal results according to this evaluation is Scenario 2. However, further analysis can be required in order to determine whether the clusters obtained are actually meaningful or not. This results all depend on the context in which the algorithm was applied. The results are orientated, and can be used as reference for identifying strengths and further improvements to it.

8.2 DAVIES-BOULDIN INDEX

The Davies-Bouldin Index (DBI, or DBI Score) is another metric used for cluster validation. The score is defined as the average similarity measure of each cluster with its most similar cluster, where similarity is the ratio of within-cluster distance to between-cluster distance. Clusters which are farther apart and less dispersed will result in a better score. [6]

The minimum score of the DBI is zero (0), with lower values indicating better clustering.

After evaluating the model with the Davies-Bouldin Index for each of the scenarios previously tested, the index score results are the following:

Table 10: Davis-Bouldin Evaluation Score

Davies-Bouldin Evaluation Score	
Scenario	Score
1	1.47
2	0.43
3	0.9
4	1.29

Among the scenarios, Scenario 2 demonstrates the best performance according to DBI, with a score of 0.43. This means that the clusters in here are well separated and compact. Scenario 3 follows as the second lowest with 0.9. It indicates a reasonably good clustering. However, in Scenario 4, there's a high score of 1.29, which can suggest potential overlap within clusters compared to the other scenarios. At last, the highest score is 1.47, corresponding to Scenario 1. This means poor separation between the clusters.

Based on the analysis, Scenario 2 seems to have the best clustering performance, followed by Scenario 3. Scenario 4 has a weaker performance and lastly, Scenario 1 appears to have the highest score, meaning the overlapping of clusters. These results as the other evaluations are orientated, and it can also require further analysis to see if the clusters meet the requirements according to each scenarios needs. It all depends on the context and the data possessed to conduct the application of the algorithm.

9. TESTING DISCUSSION

In the following section, the testing results obtained will be discussed from applying the K-Means model in various scenarios.

First of all, the testing has been performed in 4 different scenarios. The point of this application is to distribute the company's employee workforce by projects, obtaining the best employees to perform tasks in the specific projects; this based on the scores that each one obtained in various competencies that also serve as requirements for each project.

The number of projects, employees and competencies or requirements was variable, as this made it possible to obtain different results in order to be able to perform a comparative analysis of the results of each scenario.

One thing that stood out in the testing phase was that in the scenarios it was common for projects to be repeated in clusters. This means that there is a high similarity between employees, in terms of scores on the requirements of each project. Another reason on why projects may be repeated is the scale on which users are rated. In all the scenarios it has been from 1 to 3, but if it varies between 1 to 5 or 1 to 10, it is possible that there is more variability in the results.

There are several options to address the cases in which there are projects that remain unassigned clusters. One of them is to evaluate the scores of the employees of each cluster with respect to that unassigned project, and from there, to be able to choose which employees are the best fit for those particular projects. Also, several iterations of the model (different initializations) can be performed, which can change the results. This does not mean that the quality will change, as it uses the same data for comparison.

Finally, the model has been evaluated with two different metrics for the K-Means algorithm: Silhouette Index and Davies-Bouldin Index. These evaluations position Scenario 2 as the scenario with the most optimal results. It means that, the clustering in this scenario has been the most correct compared to the others. However, this does not mean that the clustering in the other scenarios is not correct, as the model evaluations are subjective, and each depends on the context in which the algorithm is applied. It also depends on the quality and quantity of data is worked with.

Upon analyzing and interpreting the algorithm results, it is deemed normal to observe multiple projects within each cluster. This variation is contingent upon the available data for comparison. It is important to note that the results and clusters have the potential to change when additional data is introduced into the database. For instance, incorporating more projects, employees or requirements to meet the company's criteria may impact the clustering outcomes.

10. PROGRAM LIMITATIONS AND UNCERTAINTY

Determining the accuracy of something is often challenging, and this applies to the K-Means algorithm. Its performance and outcomes are influenced by several factors, including the characteristics of the data, the quantity and quality of the data, the specific problem being addressed, the chosen approach, and various others. It is crucial to carefully consider these factors and assess the suitability of the algorithm in the given context.

After applying the K-Means algorithm, some limitations have been identified regarding the application of this algorithm for the assignment of employees in various projects of a company.

CLUSTER NUMBERS

The K-Means algorithm is an algorithm that performs best when having to cluster a small number of clusters. In this case, the maximum number of clusters that has been done is 10, but we do not know for sure how the algorithm would behave if there were 20, 30 or 50 projects. For more complex cases, it is very likely that much more complex algorithms such as artificial neural networks would be needed, which could be able to perform these clusters or the problem could be refined from another point of view.

QUANTITY AND QUALITY OF DATA

Data quantity and quality is a crucial factor when using AI algorithms. If there is insufficient data, the algorithm may not provide the most suitable assignment, resulting in analyzing alternative problem-solving approaches. On the other hand, if there are excessive amounts of data, which is challenging to handle, the results may not be optimal too, requiring the consideration of more advanced algorithms.

MISSING DATA

The K-Means algorithm is sensitive to missing data. Failing to address this issue prevents successful clustering. Moreover, the method chosen to handle missing data significantly influences the algorithms results. It is crucial to select an appropriate approach based on the specific context of the application.

CLUSTER INTERPRETATION

While code may run smoothly, without errors, the interpretation of the algorithm results is crucial for making informed decisions and allocating resources effectively. It should emphasize the importance of prior knowledge in their interpretation. It is crucial to provide the user with insights of the data to be analyzed and the problem intended to be solved. By this, users can better understand and extract meaningful insights from the algorithm's outputs.

FACTORS CONSIDERED FOR CLUSTERING

In this case, only employee scores in competencies related to the projects were considered. However, future improvements could involve taking into account additional factors such as performance scores assigned by project managers, employee availability, workload, and other relevant considerations that could impact the clustering outcomes. This might require the use of more advanced algorithms or to change the approach to solve the problem.

UNCERTAINTY

Another factor to consider for this is uncertainty. Artificial intelligence is not something that is 100% certain. It's important to note that the results obtained from this application are based on the information provided to the algorithm. Machines can never fully replace human work. These tools, including AI, are developed to serve as supportive tools for users. However, there is inherent uncertainty when applying algorithms because how can we validate that results are 100% accurate?

To address this uncertainty, it's essential to conduct deeper data analysis on which the algorithm is based. By ensuring data quality, we can have more reliable results for resource allocation in projects. Results can be optimized based on evaluations made with the model, and this results always depends on the information used as input for the algorithm. Users can trust such tools as they eliminate many manual processes and potential flaws, but they should remain aware of both the advantages and limitations applying tools like this.

11. CONCLUSIONS

Project management is an area that today is evolving along with the most current technological tools. As part of this research work, the task of analyzing artificial intelligence as a tool to implement and improve aspects of project management was given. One of the objectives was to evaluate the risks and advantages of allocating resources in the best way.

As part of the analysis of the state of the art, the first step was to contextualize and understand the basics of artificial intelligence. This encompassed the analysis of various AI approaches and algorithms. In addition, the applications of AI in project management were researched and some outstanding studies on this subject were discussed. Future research lines and contributions to the field were also presented.

The main question driving the research was: can the task of resource allocation be performed using artificial intelligence? For this purpose, several AI models were evaluated, from which one was chosen to perform an experiment based on a particular case developed. The K-Means algorithm was the chosen one, which was run in several iterations of scenarios using a baseline case as the starting point. The number of employees, projects and competencies were variable between all the 4 scenarios. The motivation of the baseline case was to group employees of a company in the ongoing projects.

In Scenario 1, one project was repeated in 3 clusters, which means a high similarity between employees qualified to perform tasks on that project. In this case, it would be the task of the project manager to look for alternative ways to find employees who can perform tasks on other projects and choose the best employees to work on that repeated project. Also, for projects that do not have a cluster assigned to them. In Scenario 2, one project did not have a cluster assigned to it. Here, most of the employees had the skills to work on a particular project. However, the difference in average scores between clusters was not significant. In Scenario 3, 3 projects had no clusters assigned, which could mean several things: uneven distribution of competencies, when calculating the mean, the results are not as competitive with the mean scores of the other projects. The amount of data for each project may have an influence. Each project has a number of competencies that belong to it, and this may affect the results in one way or another, but it is still a normal behavior. Finally, in Scenario 4, only one project had no cluster assigned; this may indicate that employees may have lower scores in meeting its requirements. Scores greatly influence clustering results.

After testing the algorithm in the four scenarios, a model evaluation was performed using two metrics: Silhouette Index and Davis Bouldin Score. The results of these evaluations showed that, the K-Means algorithm performed better in Scenario 2, respectively using the two metrics. Scenario 2 consisted of, 35 employees, 4 projects and 67 competencies. The results mean that, the clustering in this scenario has been the most correct compared to the others. However, this does not mean that the grouping in the other scenarios is not correct, as the evaluations of the models are subjective, and each depends on the context in which the algorithm is applied. It also depends on the quality and quantity of data being worked with.

It was noticeable that when performing the clustering, projects were repeated by clusters. It is concluded that, there is a high similarity between employees, in terms of scores on the requirements of each project. Another reason why projects may be repeated is the user rating scale.

After applying the K-Means algorithm for the purpose of allocating resources, the following limitations were identified: The first one being that, the K-Means algorithm is an algorithm that works best when a small number of clusters need to be grouped. In second place, the quantity and quality of data is a crucial factor when using AI algorithms. If the data is insufficient, the algorithm may not provide the most appropriate assignment, so alternative problem-solving approaches will have to be analyzed. Third, it's important to note that the K-Means algorithm is sensitive to insufficient data. If this problem is not taken into account, successful clustering cannot be performed. In fourth place, the interpretation of clusters is another one. The importance of prior knowledge in their interpretation should be emphasized. It is crucial to provide the user with knowledge about the data to be analyzed and the problem to be solved. Also, take into account other factors to compare them. Finally; uncertainty. Artificial intelligence is not a 100% sure thing. It is important to keep in mind that the results obtained from this application are based on the information provided to the algorithm. Machines will never be able to fully replace human work.

After discussing the results and finishing the experimentation process, it is concluded that, artificial intelligence is a powerful tool and capable of doing whatever the programmer decides. It is versatile and adaptable. Clarity about the intended context for implementing this technology is crucial, as a lack of clarity could lead to misguided implementation. Both the quantity and quality of data should be taken into account. Choosing an algorithm or approach can rely on these factors, as well as the specific scenario. It's important to understand that while artificial intelligence is a useful tool, it can also generate variable results and introduce uncertainty. Emphasizing that outcomes will depend on the data available and require accurate interpretation is crucial.

While doing this research work, I realized how erroneously or utopian the term AI has become known, like robots or computers acting as human beings, as people interpret it nowadays. It is simply mathematical functions, which programmers are able to adapt depending on the context in which they are going to be used. It is a tool that should be approached with prudence, as there always be factors that will influence the result. I believe that AI can replace human work in several tasks that are automatic and simple, but in project management, today it is nothing more than a support tool for the project manager. Each project is a different world, which will depend on the context and will have many fluctuating factors which cannot be decided only by a machine. Yes, some tasks can be performed by AI, but for now, we will still need human intervention in some critical tasks that require not only technical and mathematical skills, but also human experience and point of view.

12.IMPROVEMENTS AND FUTURE TESTS

In any research, there is always room for improvement, and this study is no different. After the algorithm completes its function, and the results were analyzed, areas for future enhancement were identified. Additionally, considering and analyzing other factors could make this tool even more comprehensive and effective.

For example, it could be analyzed on other external factors that could be considered for clustering, such as performance scores, employee workload and availability. Is it possible to solve the problem using other algorithms or approaches? Without a doubt, experimenting with different algorithms that could handle this assignment is worth considering. This is one of the questions that could be taken into account for future improvements of the tool.

Another future improvement would be to make this tool a more comprehensive software. This could involve using various AI algorithms to address different questions and potential scenarios that a project manager faces when allocating resources. Additionally, interactive dashboards could be included to display the data and results, making the tool more user-friendly. The goal is to make resource allocation tasks more manageable and less manual for project managers.

Finally, this section acknowledges the potential for improvement and future enhancements. By analyzing the algorithms results, areas for further development have been identified. Experimenting with additional algorithms and approaches to solve the problem is a promising path for future research. Transforming the tool into a more comprehensive and complete software by implementing more algorithms and visual data representation would greatly benefit project managers in this task. These considerations have the potential to pave the way for continued advancements in this field.

13. BIBLIOGRAPHY

- [1] C. Selaru, "Resource Allocation in Project Management," vol. 2, no. 4, 2012.
- [2] G. Huzooree and V. Devi Ramdoo, "Review of Effective Human Resource Management Techniques in Agile Software Project Management," International Journal of Computer Applications, vol. 114, no. 5, pp. 10-15, 2015.
- [3] A. Martinez-Millana, A. Saez-Saez, R. Tornero-Costa, N. Azzopardi-Muscat, V. Traver and D. Novillo-Ortiz, "Artificial intelligence and its impact on the domains of universal health coverage, health emergencies and health promotion: An overview of systematic reviews," International Journal of Medical Informatics, vol. 166, 2022.
- [4] A. M. Votto, R. Valecha and R. H. Raghav, "Artificial Intelligence in Tactical Human Resource Management: A Systematic Literature Review," International Journal of Information Management Data Insights, vol. 1, no. 2, 2021.
- [5] C. Debrah, A. P. Chan and A. Darko, "Artificial intelligence in green building," Automation in Construction, vol. 137, p. 104192, 2022.
- [6] B. I. Oluleye, D. W. Chan and P. Antwi-Afari, "Adopting Artificial Intelligence for enhancing the implementation of systemic circularity in the construction industry: A critical review," Sustainable Production and Consumption, vol. 35, pp. 509-524, 2023.
- [7] L. A. Ankabi, A. O. Oyedele, L. O. Oyedele and R. O. Salami, "Deep learning model for Demolition Waste Prediction in a circular economy," Journal of Cleaner Production, vol. 274, p. 122843, 2020.
- [8] C. Manning, "Artificial Intelligence Definitions," Stanford University Human-Centered Artificial Intelligence, Stanford, 2020.
- [9] C. Collins, D. Dennehy, K. Conboy and P. Mikalef, "Artificial intelligence in information systems research: A systematic literature review and research agenda," International Journal of Information Management, vol. 60, p. 102383, 2021.
- [10] E. Schrijvers, C. Prins and H. Sheikh, "Artificial Intelligence: Definition and Background," in Mission AI. Research for Policy, Springer, Cham, 2023, pp. 15-41.
- [11] F. Maleki, K. Ovens, K. Najafian, B. Forghani, C. Reinhold and R. Forghani, "Overview of Machine Learning Part 1," Neuroimaging Clinics of North America, vol. 30, no. 4, pp. e17-e32, 2020.
- [12] B. Mahesh, "Machine Learning Algorithms - A Review," International Journal of Science and Research (IJSR), vol. 9, no. 1, 2018.
- [13] B. Charbuty and A. Abdulazeez, "Classification Based on Decision Tree Algorithm for Machine Learning," Journal of Applied Science and Technology Trends, vol. 2, no. 1, pp. 20-28, 2021.
- [14] G. Sammour, K. Vanhoof and H. Qabbaah, "Decision Tree Analysis to Improve E-Mail Marketing Campaigns," International Journal "Information Theories and Applications", vol. 26, no. 1, 2019.

- [15] R. Hasan, S. Palaniappan, A. R. A. Raziff, S. Mahmood and U. Sarker, "Student Academic Performance Prediction by using Decision Tree Algorithm," in *International Conference on Computer and Information Sciences (ICCOINS)*, Kuala Lumpur, Malaysia, 2018.
- [16] J. Nalepa and M. Kawulok, "Selecting training sets for support vector machines: a review," *Artificial Intelligence Review*, vol. 52, no. 2, pp. 857-900, 2019.
- [17] S. HUANG, N. CAI, P. PENZUTI PACHECO, S. NARRANDES, Y. WANG and W. XU, "Applications of Support Vector Machine (SVM) Learning in Cancer Genomics," *Cancer Genomics & Proteomics*, vol. 15, no. 1, pp. 41-51, 2018.
- [18] O. Isaac Abiodun, A. Jantan, A. Esther Omolara, K. Victoria Dada, N. AbdElatif Mohamed and H. Arshad, "State-of-the-art in artificial neural network applications: A survey," *Heliyon*, vol. 4, 2018.
- [19] W. Kunfeng, G. Chao, D. Yanjie, L. Yilun, Z. Xinhua and W. Fei-Yue, "Generative adversarial networks: introduction and outlook," *IEEE/CAA Journal of Automatica Sinica*, vol. 4, no. 4, pp. 588-598, 2017.
- [20] H. Taud and J. Mas, "Chapter 27 Multilayer Perceptron (MLP)," in *Geomatic Approaches for Modeling Land Change Scenarios*, 2018, pp. 451-455.
- [21] Z. Li, W. Yang, S. Peng and F. Liu, "A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 12, pp. 6999-7019, 2022.
- [22] I. Zeroual and A. Lakhouaja, "Data science in light of natural language processing: An overview," *Procedia Computer Science*, vol. 127, pp. 82-91, 2018.
- [23] J. Hirschberg and C. D. Manning, "Advances in natural language processing," *SCIENCE*, vol. 349, no. 6245, pp. 261-266, 2015.
- [24] V. Kakani, B. P. Nguyen, H. Kim and V. R. Pasupuleti, "A critical review on computer vision and artificial intelligence in food industry," *Journal of Agriculture and Food Research*, vol. 2, p. 100033, 202.
- [25] V. Dhanya, A. Subeesh, N. Kushwaha, D. K. Vishwakarma, T. Nagesh Kumar, G. Ritika and A. Singh, "Deep learning based computer vision approaches for smart agricultural applications," *Artificial Intelligence in Agriculture - KeAi*, vol. 6, pp. 211-229, 2022.
- [26] N. Athanasios, "Geo Fossils-I: A synthetic dataset of 2D fossil images for computer vision applications on geology," *Data in Brief*, vol. 48, 2023.
- [27] K. Guo, H. Tao, Y. Zhu, B. Li, C. Fang, Y. Qian and J. Yang, "Current applications of artificial intelligence-based computer vision in laparoscopic surgery," *Laparoscopic, Endoscopic and Robotic Surgery - KeAi*, 2023.
- [28] A.I for Anyone, "What is automated planning and scheduling? : AI terms explained- AI for anyone," [Online]. Available: <https://www.aiforanyone.org/glossary/automated-planning-and-scheduling>. [Accessed 2023].
- [29] A. K. Goel and B. Diaz-Agudo, "What's Hot in Case-Based Reasoning," in *AAAI Conference on Artificial Intelligence*, 2017.
- [30] A. Asemi, A. Ko and M. Nowkarizi, "Intelligent libraries: a review on expert systems, artificial intelligence, and robot," *Library Hi Tech*, vol. 39, no. 2, pp. 412-434, 2021.

- [31] P. Hofmann, J. Jöhnk, N. Urbach and D. Protschky, "Developing Purposeful AI Use Cases- A Structured Method and Its Application in Project Management," in *15th International Conference on Wirtschaftsinformatik (WI)*, Potsdam, 2020.
- [32] G. Auth, O. Jokisch and C. Dürk, "Revisiting automated project management in the digital age – a survey of AI approaches," *Online Journal of Applied Knowledge Management*, vol. 7, no. 1, pp. 27-39, 2019.
- [33] V. Prifti, "Optimizing Project Management using Artificial Intelligence," *European Journal of Formal Sciences and Engineering*, vol. 5, no. 1, pp. 29-37, 2022.
- [34] Y. Sheoraj and R. K. Sungkur, "Using AI to develop a framework to prevent employees from missing project deadlines in software projects- case study of a global human capital management (HCM) software company," *Advances in Engineering Software*, vol. 170, 2022.
- [35] T. Crawford, S. Duong, R. Fueston, A. Lawani, S. Owoade, A. Uzoka, R. M. Parizi and A. Yazdinejad, "AI in Software Engineering: A Survey on Project Management Applications," *arxiv - Cornell University*, 2023.
- [36] K. S. Pranata, A. A. S. Gunawan and F. L. Gaol, "Development clustering system IDX company with k-means algorithm and DBSCAN based on fundamental indicator and ESG," *Procedia Computer Science*, vol. 216, pp. 319-327, 2023.
- [37] B. Yang, X. Fu, M. Hong and N. Sidiropoulos, "Towards K-means-friendly Spaces: Simultaneous Deep Learning and Clustering," 2016.
- [38] X. Jin and J. Han, "K-Means Clustering," *Encyclopedia of Machine Learning and Data Mining*, pp. 695-697, 2017.
- [39] M. M. Navarro, M. N. Young, Y. T. Prasetyo and J. V. Taylor, "Stock market optimization amidst the COVID-19 pandemic: Technical analysis, K-means algorithm, and mean-variance model (TAKMV) approach," *Heliyon*, p. e17577, 2023.
- [40] Sci-Kit Learn, "Selecting the number of clusters with silhouette analysis on KMeans clustering¶," [Online]. Available: https://scikit-learn.org/stable/auto_examples/cluster/plot_kmeans_silhouette_analysis.html. [Accessed 2023].
- [41] Sci-Kit Learn, "sklearn.metrics.davies_bouldin_score," [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.davies_bouldin_score.html.
- [42] Y. LeCun, Y. Bengio and G. Hinton, "Deep Learning," *Nature*, vol. 521, pp. 436-444, 2015.
- [43] G. Montavon, W. Samek and K.-R. Müller, "Methods for interpreting and understanding deep neural networks," *Digital Signal Processing*, vol. 73, pp. 1-15, 2018.
- [44] M. Sanchez-Marre, "Principles of Case-Based Reasoning," Secció d'Intel·ligència Artificial Dept. de Llenguatges i Sistemes Informàtics Universitat Politècnica de Catalunya, Barcelona, 2001.

ANNEX 1: INPUT DATA

This annex includes the input data used for the implementation of the K-Means algorithm for the 4 scenarios tested.

For each scenario, there is a csv format file called **comp_ratings.csv**. This is the dataset that includes the initial scores of the employees, the input data for the algorithm. With these scores, the K-Means algorithm is run and the clustering is performed. The columns of this dataset are as follows:

comp_name - is the name of the competency to be evaluated. There are technical competencies such as programming languages, or personal competencies such as creativity or adaptability.

employee_id - is the unique identifier of the employee, it represents the employee who has been rated on X competency.

project_id - is the ID of the project to which X competency to be rated belongs.

comp_rating - is the score given to an employee in X competency. It ranges from 1-3 (1 low, 2 medium, and 3 high) and in the case of this research work, it has been randomly assigned.

The file used in each scenario is located in a separate folder.



Scenarios 1 & 2, Scenario 3, Scenario 4

ANNEX 2: TESTING RESULTS – K MEANS ALGORITHM

This annex includes the complete results of the implementation of the K-Means algorithm for the 4 scenarios tested.



The Testing Results file includes the output of the algorithm implementation. The columns in the file are as follows:

Clusters - is the cluster or group number. The number of clusters per scenario is the number of projects. clusters = projects.

Employee ID - represents the unique employee identifier.

Avg. Score of Emp. - is the average score of that employee. The average comes from the score of the competencies corresponding to that project, per employee.

Cluster Project ID - is the unique project identifier.

Project Name - project name, only applies to Scenarios 1 and 2.

Avg. Score of Cluster - average score of the employees belonging to the cluster.

ANNEX 3: K-MEANS ALGORITHM - CODE

This annex includes the K-Means code used for the 4 scenarios.

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.cluster import KMeans
from mpl_toolkits.mplot3d import Axes3D
from tabulate import tabulate
from sklearn.impute import SimpleImputer
from IPython.display import display

pd.set_option('display.max_columns', None)
pd.set_option('display.max_rows', None)

#Tables for K-Means algorithm
ratings = pd.read_csv('#data', delimiter=';')
employees = pd.read_csv('#data', delimiter=';')

# Merge data from ratings & employees
merged_data = pd.merge(ratings, employees, on='employee_id', how='inner')

employee_ids = employees['employee_id'].unique()[:150]
merged_data = merged_data[merged_data['employee_id'].isin(employee_ids)]

filtered_data = merged_data[merged_data['project_id'].isin([1, 2, 3, 4, 5, 6, 7, 8])]

employee_features = filtered_data.pivot_table(index='employee_id',
columns='project_id', values='comp_rating')

employee_features_filled = employee_features.fillna(employee_features.mean())

kmeans = KMeans(n_clusters=8, random_state=42)
kmeans.fit(employee_features_filled.values)
cluster_labels = kmeans.labels_

# ID for projects of interest
project_ids = [1, 2, 3, 4, 5, 6, 7, 8]

# Filter projects data
filtered_data = merged_data[merged_data['project_id'].isin(project_ids)]

employee_features_filled['cluster'] = cluster_labels
employee_features_filled = employee_features_filled.round(decimals=2)
```



```
# Print employees in each cluster
print("Employees in each cluster:")
for cluster in employee_features_filled['cluster'].unique():
    employees_in_cluster =
employee_features_filled[employee_features_filled['cluster'] == cluster].index
print(f"Cluster {cluster}: {employees_in_cluster}")
```