POLITECNICO DI TORINO

Department of Management and Production Engineering

Master’s in Engineering and Management

Master’s Degree Thesis

“Applying Artificial Intelligence techniques in Project Management”

Candidate:

KUNNATHUR Mohamad Usama

Student# S249895

Supervisors:

Prof. DE MARCO Alberto (DIGEP)

DEMAGISTRIS Paolo Eugenio (External Teacher)

OTTAVIANI Filippo Maria (External Teacher)

Contents

Table of Figures

Acknowledgments

First of all, I would like to thank Almighty God for granting upon me this opportunity to work on this important topic, allowing me the capability to progress positively and providing me such a friendly supervisor and co-supervisor. This thesis is a result of the guidance and support of many people and I would like to offer my sincere thanks to all of them.

I would like to thank **Prof. Alberto De Marco**, my honored supervisor from DIGEP - Department of Management and Production Engineering, College of Management Engineering, at Politecnico di Torino, for accepting me as a research student and his support.

I want to deeply thank **Prof. Paolo Eugenio Demagistris**, my valued co-supervisor from DIGEP - Department of Management and Production Engineering, College of Management Engineering, at Politecnico di Torino, for trust in me, the insightful discussions, unconditional support, introducing me to new ideas, to the experts and professionals of the work scope.

Last but not the least I would like to than **Filippo Maria OTTAVIANI** my other co-supervisor from DIGEP - Department of Management and Production Engineering, College of Management Engineering, at Politecnico di Torino for friendly coordination throughout the duration of the research, especially your patience, guidance, critical comments and corrections in the thesis work.

I would like to thank my friends and family who have supported me throughout my academic life; without their support, I probably wouldn’t be here. I would like to thank all the other professors who supported me during my studies.

1 Research Purpose

The purpose of this research is to create a model like Earned Value Management (EVM) model in Project Management for better forecasting. Some machine learning concepts will be used to better predict the project duration and minimize the cost. Traditional EVM model formulas to track Projects progress relies on too many assumptions and simplifications. This oversimplification can affect the accuracy of the forecasting which can cause delays and in turn increase project costs. There is also bias related to human perceptions in the EVM model which must be eliminated.

The goal is to develop a new dynamic monitoring and optimization tool to track the progress of the project. Furthermore, to increase forecasting, rescheduling and reallocation process efficiency with new DSS (decision support system). Machine learning an ever-growing field and implementation of it in project management can change how projects are tracked. Machine learning models are trained with data already available so that the computer can make a more informed decision by itself. It should identify all the combinations and provide the best scheduling option which minimizes cost and financial exposures.

Another aspect which will be focused and that is the financial exposure. The aim is to minimize the financial exposure which traditional methods do not account for. This is more practical than theoretical as banks are reluctant to fund the project if the cash flow is not smooth. The financial need should be minimized. There is a trade-off between financial exposure, cost and duration. The goal is to find a balance between these three.

2 Introduction

Projects are subjected to a number of factors that interfere with their efficient completion. Therefore, it is critical to monitor the progress of work in real-time, and systematically analyze any alteration in work schedules and project costs compared to the value previously planned, in order to take preventive measures to reduce the negative impact of perplexing components. The Earned Value Method is one of the most used methods around the world to monitor and control the progress of work. Its primary aim is to measure the progress of the project, predict its cost and duration. Indirectly, EVM may be used to control risk in the context of exceeding project costs and failure to meet the deadline for completion of the investment. The EVM is an analysis based on several measurements using appropriate indicators. Measurements are taken at regular intervals, e.g. at the end of the month, in order to track the trends and variations of these indicators.

Traditional EVM monitoring of project performances is based on the budgeted cost of work performed (BCWP), budgeted cost of work scheduled (BCWS) and the actual cost of work performed (ACWP). The cost variance (CV), schedule variance (SV), or cost performance index (CPI) and schedule performance index (SPI) are then calculated to calculate the project performance. CPI and SPI are usually expressed in a periodic or cumulative way. These terms will be explained in detail later. Although the EVM approach is believed to be the most objective method available for measuring project performance, the method is limited and doesn't account for changes in individual performance values. Unfortunately, the subjective way of assessing work progress is one of the factors that may lead to the intentional falsification of data in order to produce the desired image.

Moreover, the timing of check performance may give false indication due to the time lapse between the work performance check and by the time of corrective actions where some problems may be easily handled by proper direct actions. Changes in baseline due to certain disciplines in the construction must be taken into consideration by updating the schedule. This is the case of the implementation of innovative projects subject to alterations and unforeseen events. The scope of the budget and the schedule are then changed and this, in turn, affects the project completion estimates and leads to difficulty in analyzing EVM indicators.

The goal of this paper consists of introducing Artificial Intelligence methods for constructing better forecasts to predict the final duration of a project. AI methods use historical or simulated data to learn the relation between inputs and one or multiple outputs. In a project control environment, the simulated data contains information with regard to the progress of the project. The proposed methods learn how the performance indicators are related to the project's Real Duration (RD). This knowledge is then applied to new data to come up with an estimate of the project's final duration. Secondly, a generally applicable methodology is put forward starting with the generation of project and progress data. This serves two purposes. The first purpose lies in the nature of the AI methods which learn a relation from existing data. Second, a wide array of data allows us to reach general conclusions. Apart from the data generation, a decision needs to be made on which progress data will be fed to the AI methods. These progress data contain periodic measurements of EVM performance indicators, as well as EVM forecasting estimates. The high volume of data may be correlated, and some data may be irrelevant. The general performance of the AI methods should be assessed and compared against the EVM methods.

3 Concepts

This thesis will use several technical terms and jargons which needs to be explained. The concept of Project Management and Artificial Intelligence will also be discussed as they form the core of the thesis.

3.1 Project Management

An official definition of project management, courtesy of the Project Management Institute, defines the term as: "*the application of knowledge, skills, tools and techniques to project activities to meet project requirements.*"

A more tangible description is that project management is everything you need to make a project happen on time and within budget to deliver the needed scope and quality.

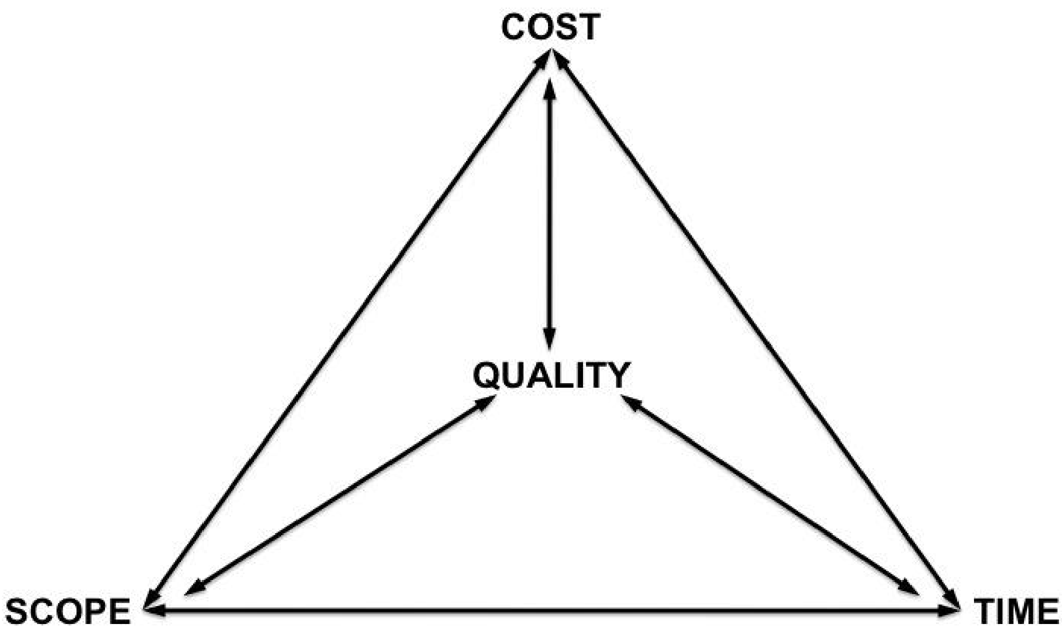


Figure 1: Project Management Triangle of Constraints

A project manager has to make a tradeoff between cost, time and scope while ensuring the required quality. In this thesis, cost & time are focused as well as the other parameter called a financial constraint. Financial constraint is more of a practical term and will be discussed later.

Project management is a set of skills and tools that will help to get the project right in every way. A project manager can employ various concepts such as EVM, ES, CPM, etc. or tools such as MS Project, Excel, etc. This thesis is focused on the EVM model.

3.1.1 EVM model

Earned Value Analysis and Management concept was first developed by the US Department of Defense, in order to assess the programs during the early sixties, from 2005; this concept of project evaluation took place in the general federal project risk management. These days EVA and EVM have become a mandatory need of the US government. OMB (Office of Management and Budget) promotes the utility of EVM to be the preferred mode of management of software projects which are performance-based. It is now a day's used in a variety of industries, consulting and educational establishments including public and private sectors, Including NASA, PMI, Society of Cost Estimating and Analysis (SCEA) and others.

The earned value concept improves upon the standard comparison of budget vs. actual cost which lacks an adequate indicator of progress. Earned value is a value assigned to work which was accomplished during a particular time period. This value can be stated in any appropriate measurable unit such as hours or dollars. Earned value, and Earned Value Analysis (EVA), thus provides progress information that can be compared to the planned budget and actual cost -- to provide additional insight into project status.

There are certain terms associated with the EVM model.

3.1.1.1 Planned Value

Planned Value is a part of the total cost of the project, that should have been spent on the task up to the evaluation time. It is also called Budget Cost of Work Scheduled (BCWS) and denoted as cost units.

3.1.1.2 Actual Value:

Actual Value of Actual cost is the total cost incurred either direct or indirect in the accomplishment of the task, during the evaluation period. It is also called as Actual Cost of Work Performed and expressed in the cost unit.

3.1.1.3 Earned Value:

Earned Value is also termed as Budget Cost of Work Performed (BCWP), which is the approximation of the value of physical work actually performed. It is the relationship between the scheduled cost of the project and the rate with which the resources of the project perform their task, it is mentioned as cost units. In order to calculate the earned value, the total budget is multiplied with the percentage of work performed.

$$BCWP = Percent\ work\ Performed \times Total\ Budget\ Allocated$$

3.1.1.4 Budget at Completion:

Budget at completion (BAC) is the total budget, which is expected to be utilized for project completion, in other words, the expectation of PV to reach its maximum according to the project plan and expressed in cost units.

3.1.1.5 Cost Variance:

Cost variance (CV) is the term used for the difference between the value of the work actually performed for the task or project and actual cost incurred, it the cost performance measurement. Expressed in cost units and can be determined as:

$$CV = BCWP - ACWP$$

3.1.1.6 Schedule Variance:

Schedule Variance (SV), the term is used for the difference between the value of the work actually performed and the value of the work which should have been done on evaluation instance according to the plan. It is the measure of schedule performance and expressed in cost units. It can be determined as below:

$$SV = BCWP - BCWS$$

3.1.1.7 Cost Performance Index:

CPI is a measure of cost performance of task or project; it indicates either task is on budget, overrun or underrun. Measured in cost units and can be determined to divide ACWP by BCWS.

$$CPI = BCWP/ACWP$$

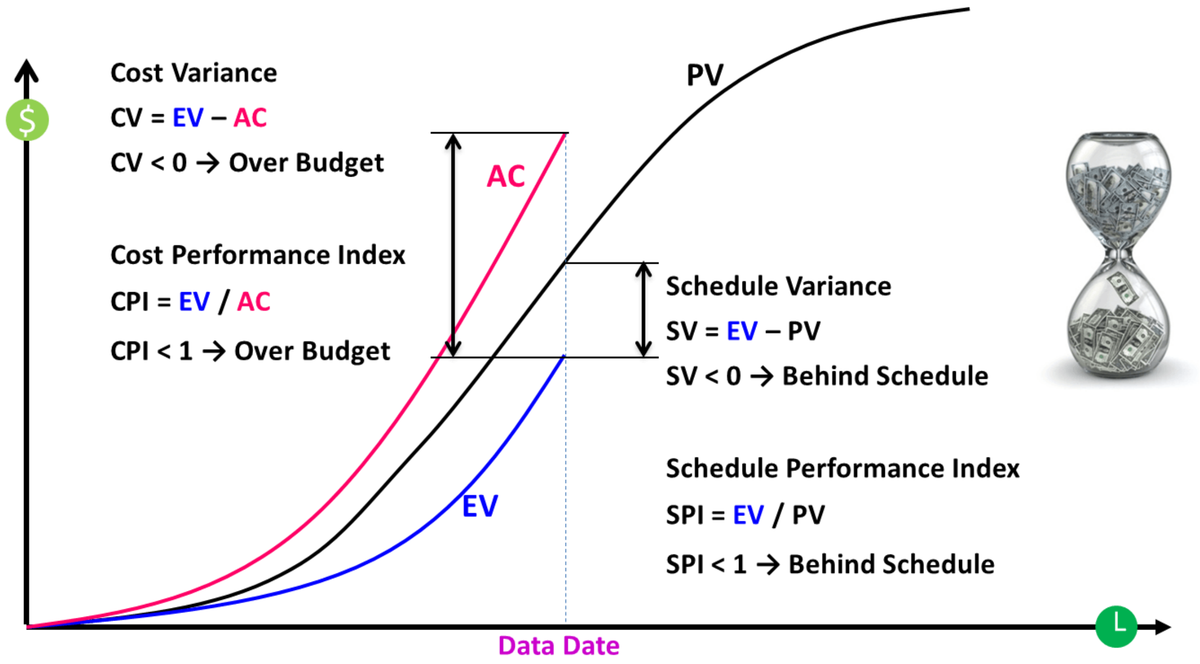


Figure 2: Earned Value Analysis

3.1.1.8 Schedule Performance Index:

Schedule Performance Index determines the schedule performance of a project or task, which is a fraction of the scheduled task which is actually performed despite its scheduled full task.

$$SPI = BCWP/BCWS$$

3.1.1.9 Cost Estimate to Complete:

Cost Estimate to Completion also known as Estimate to Complete is the analysis of the in-progress activity status, according to its cost expenditure, It is the forecast of the remaining expenses of the task or project to complete, ETC indicates further cost expenses on the activity according to its current cost performance. It is denoted as currency units.

3.1.1.10 Cost Estimate at Completion:

CEAC also known as Estimate at completion is the expected value of the task or project on its completion. It is the forecast of activity completion value according to the current performance of the activity.

3.1.1.11 Schedule at Completion:

It is also called budgeted work (BW) is the actually planned duration at completion for activity or project and mentioned in time units.

3.1.1.12 Cost Variance at Completion:

Cost Variance at Completion evaluates the estimation of the cost at the completion of activity or project, either its under-run or over-run, based on its performance.

3.1.1.13 Planned Accomplishment Rate:

Planned Accomplishment Rate is also called the Planned Value rate (PV Rate), which is the average of planned value per specified time period, (usually days, or hours). Denotes planned expenses per time period and expressed in cost units.

3.1.1.14 Actual Time:

Actual Time (AT) is the actual time duration from project initiation to reporting date.

3.1.1.15 Time Variance:

Time variance is the upgraded version of Schedule variance, TV is the measure of the schedule performance evaluation of activity or project in time units. As SV is in cost units so TV has been evolved to overcome the issue and introduce time units for schedule performance. TV may be negative, Positive and Zero indicating project behind schedule, ahead of schedule and on schedule respectively.

3.1.1.16 Earned Schedule:

Earned Schedule (ES) is the time instant on which current earned value was supposed to complete. In time units it is the point on which PV should have been equal to current Earned Value.

3.1.2 Financial Exposure

There is another important aspect that is more important in practice than theoretical and that is financial exposure. The banks want the project cash flow to be smooth and if the financial need is too much then they are reluctant to fund the project. A trade-off should be made on costs and scheduling should be modified not just to minimize the duration and costs but also to give a smooth cash flow. In the below graph there are two cash flows. One is steeper than the other. The steeper one will not be preferred by the bank while the smoother one will be preferred.

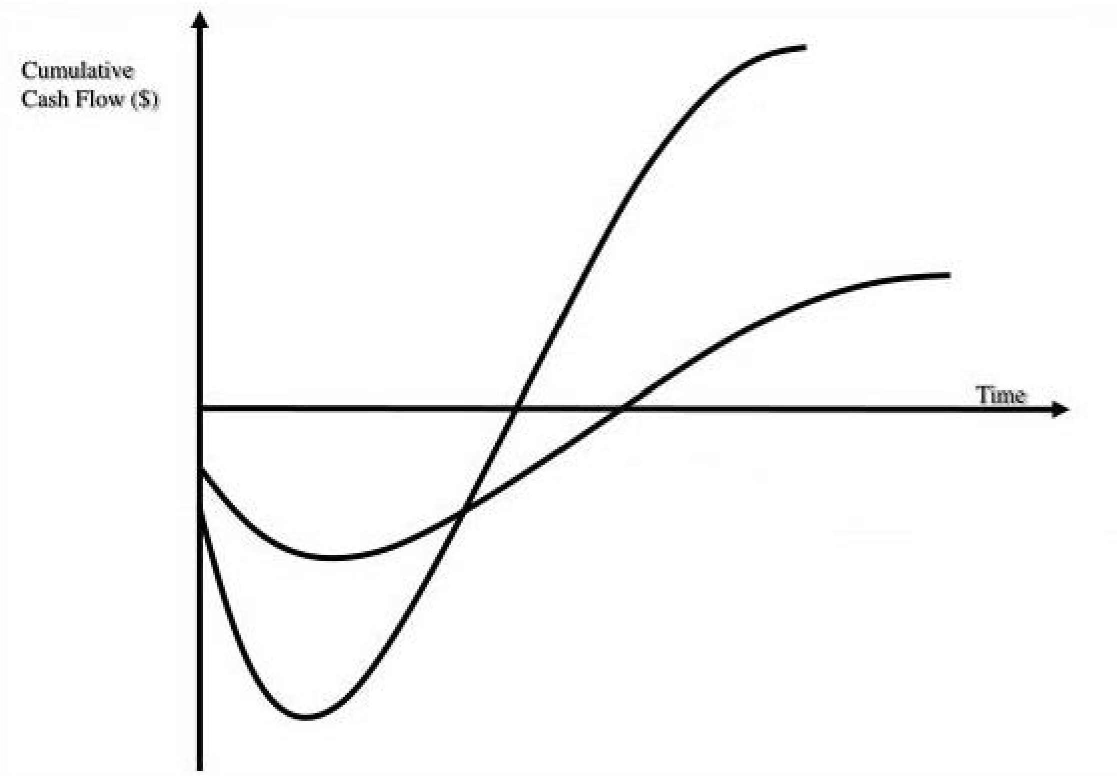


Figure 3: Financial Exposure

3.2 Artificial Intelligence

Artificial intelligence (AI), the ability of a computer or computer-controlled robot to perform tasks commonly associated with intelligent beings. The term is frequently applied to the project of developing systems bestowed with the intellectual processes characteristic of humans, such as the ability to think, discover meaning, generalize, or learn from past experience. Since the development of the computer in the 1940s, it has been demonstrated that computers can be programmed to carry out rather complex tasks—as, for example, discovering proofs for mathematical theorems or playing chess—with great proficiency. Still, despite continuing advances in computer processing speed and memory capacity, there are as yet no programs that can meet human flexibility over wider domains or in tasks requiring much everyday knowledge. On the other hand, some programs have attained the performance levels of human experts and professionals in performing certain specific tasks, so that artificial intelligence in this limited sense is found in applications as diverse as medical diagnosis, computer search engines, and voice or handwriting recognition.

The term "artificial intelligence" has been misused and misunderstood. While for many the term AI means a self-aware robot, who wants to destroy humanity; the experts refer to this specific instance of AI as artificial general intelligence which is a long way to go in the future. However, there are other instances of AI which are being used for good or bad in fields such as medicine, warfare, finances, spying, etc.

3.2.1 The Possible Advantages of AI

Studies show that project managers spend more than half of their time on administrative tasks such as dealing with check-ins and managing updates. Artificial Intelligence systems can handle these relatively easy but time-consuming tasks with ease and can save valuable time. This means project managers can focus more on the complex processes behind their management strategy. They can also spend more time focusing on their employees, which might help them to empower their employees and find further proficiencies. There are not many things that will slow a project down more than a project manager who simply does not have the time to speak to every single team member's need. With time spared using AI-enabled project management systems, the managers can focus more on more important things. Freeing up that time in a day is not only a good way to help projects go forward more effectively, but it can also help make a more comfortable work environment where employees feel supported and know that they have the appropriate resources in place.

3.2.2 Machine Learning

Machine learning is an application of AI that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of programs that can access data and use it to learn for themselves.

The process of learning begins with observations or data, such as examples, direct experience in order to look for patterns in data and make better decisions in the future based on the examples that we provide. The primary aim is to allow the computers to learn automatically without human intervention or help and adjust actions accordingly.

Machine learning is one field of AI that encompasses pretty much all the methods having the biggest impact on the world right now. Simply put, in this method a model is "trained" using huge data and the model tries to find a pattern. (For example, in case of project management if the machine recognizes that if the project is not on track then the critical task is crashed from past data. It can learn this and suggest this in the future.)

Machine learning enables the analysis of massive quantities of data. While it generally delivers faster, more accurate results in order to identify profitable opportunities or dangerous risks, it may also require additional time and resources to train it properly. Combining machine learning with AI and cognitive technologies can make it even more effective in processing large volumes of information.

3.2.2.1 Machine Learning Methods

Machine learning algorithms are often categorized as supervised or unsupervised

3.2.2.1.1 Supervised Machine Learning

Supervised machine learning algorithms can apply what has been learned in the past to new data using labeled examples to predict future events. Starting from the analysis of a known training dataset, the learning algorithm produces an inferred function to make predictions about the output values. The system is able to provide targets for any new input after sufficient training. The learning algorithm can also compare its output with the correct, intended output and find errors in order to modify the model accordingly.

3.2.2.1.2 Unsupervised Machine Learning

Unsupervised machine learning algorithms are used when the information used to train is neither classified nor labeled. Unsupervised learning studies how systems can infer a function to describe a hidden structure from unlabeled data. The system doesn't figure out the right output, but it explores the data and can draw inferences from datasets to describe hidden structures from unlabeled data.

3.2.2.1.3 Reinforcement Machine Learning

Reinforcement machine learning algorithms are a learning method that interacts with its environment by producing actions and discovers errors or rewards. Trial and error search and delayed reward are the most relevant characteristics of reinforcement learning. This method allows machines and software agents to automatically determine the ideal behavior within a specific context in order to maximize its performance. Simple reward feedback is required for the agent to learn which action is best; this is known as the reinforcement signal.

3.2.2.2 Machine Learning Techniques

A machine learning algorithm, also called model, is a mathematical expression that represents data in the context of a ---problem, often a business problem. The aim is to go from data to insight. There are many methods but some of them which will be considered in this thesis are

3.2.2.2.1 Regression

Regression methods fall within the category of supervised ML. They help to predict or explain a particular numerical value based on a set of prior data, for example predicting the price of an automobile based on previous pricing data for similar automobiles. These can be linear regression which is fast but inaccurate and non-linear regression which is slower but more accurate.

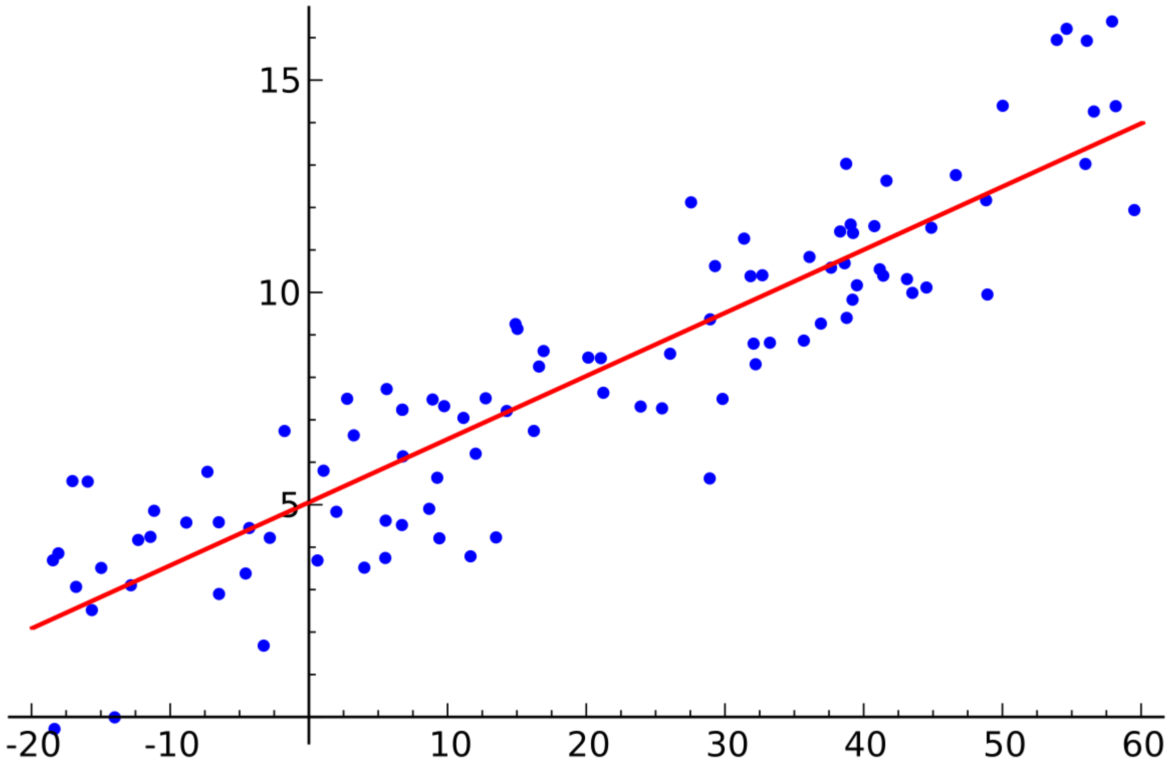


Figure 4: Linear Regression

3.2.2.2.2 Clustering

With clustering methods, we get into the category of unsupervised ML because their goal is to group or cluster observations that have similar characteristics. Clustering methods don't use output information for training, but instead, let the algorithm define the output.

The most popular clustering method is K-Means, where "K" represents the number of clusters that the user chooses to create.

- Randomly chooses K centers within the data.
- Assigns each data point to the closest of the randomly created centers.
- Re-computes the center of each cluster.
- If centers don't change (or change very little), the process is finished. Otherwise, repeat step 2.

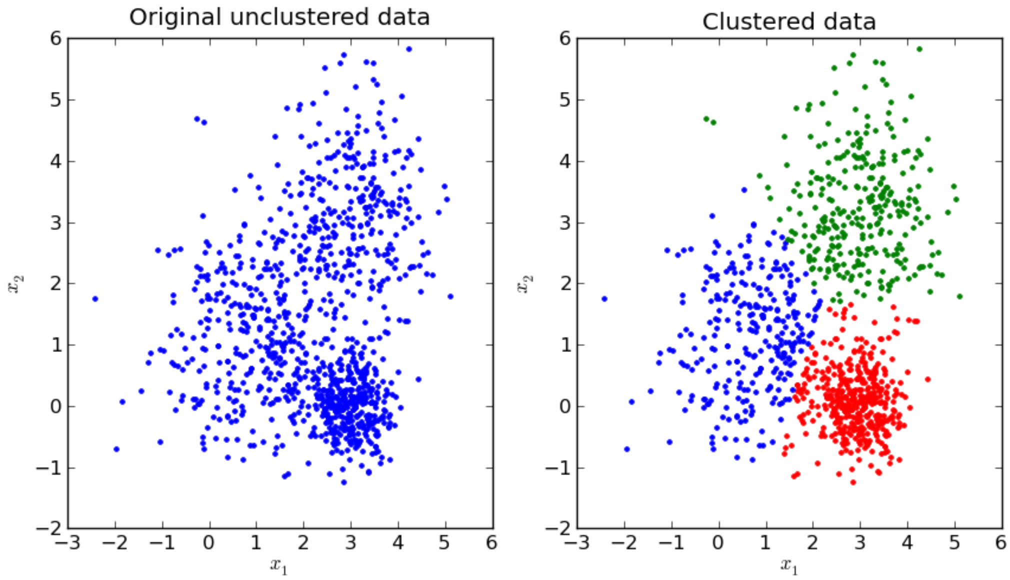


Figure 5: K-Means Clustering

3.2.2.2.3 Dimensionality Reduction

As the name suggests, dimensionality reduction removes the least important information (sometimes redundant columns) from a set of data. In practice, data sets can have hundreds or even thousands of columns (also called features), so reducing the total number is crucial. For instance, images can contain thousands of pixels, not all of which matter to the analysis. Or when testing microchips within the manufacturing process, you might have thousands of measurements and tests applied to every chip, many of which provide redundant information. In these cases, you need dimensionality reduction algorithms to make the data set manageable.

The most popular dimensionality reduction method is Principal Component Analysis (PCA), which reduces the dimension of the feature space by finding new vectors that maximize the linear variation of the data. PCA can reduce the dimension of the data dramatically and without losing too much information when the linear correlations of the data are strong.

3.2.2.2.4 Deep learning

In contrast to linear and logistic regressions which are considered linear models, the objective of neural networks is to capture non-linear patterns in data by adding layers of parameters to the model. In the image below, the simple neural net has three inputs, a single hidden layer with five parameters, and an output layer.

For the best performance, deep learning techniques need a lot of data and a lot of computing power since the method is self-tuning many factors within huge architectures. It quickly becomes clear why deep learning experts need very powerful computers enhanced with GPUs (graphical processing units).

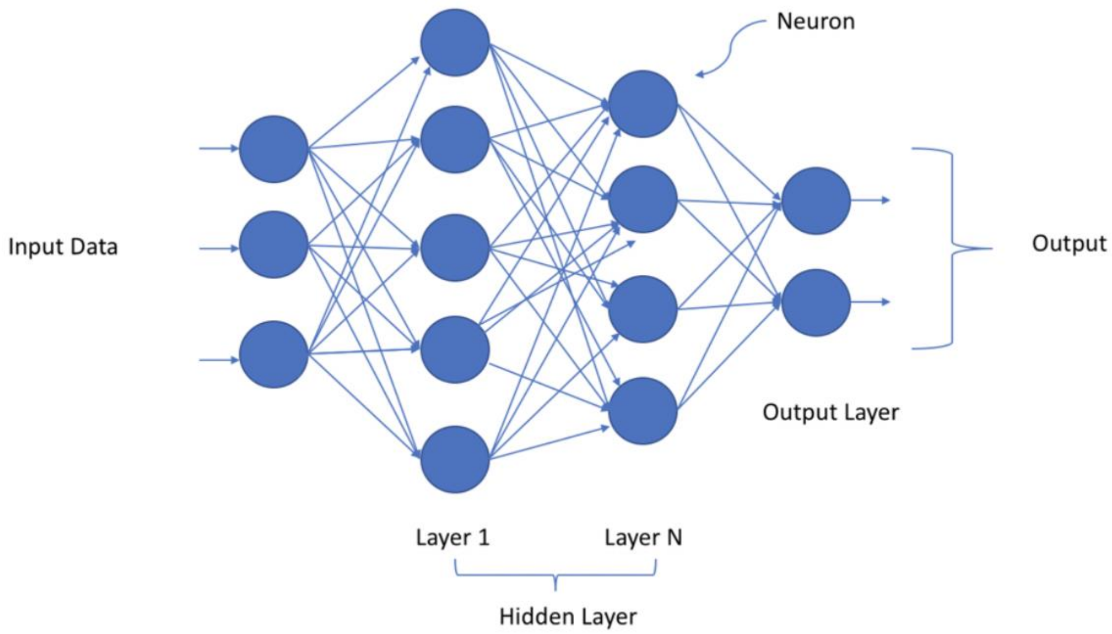


Figure 6: Neural networks

3.2.2.2.5 Ensemble Methods

Ensemble methods use this same idea of combining several predictive models (supervised ML) to get higher quality predictions than each of the models could provide on its own. For example, the Random Forest algorithms is an ensemble method that combines many Decision Trees trained with different samples of the data sets. As a result, the quality of the predictions of a Random Forest is higher than the quality of the predictions estimated with a single Decision Tree.

4 Literature Review

A literature review is a survey of scholarly sources on a specific topic. It provides an overview of current knowledge, allowing you to identify relevant theories, methods, and gaps in the existing research.

Conducting a literature review involves collecting, evaluating and analyzing publications (such as books and journal articles) that relate to your research question. There are five main steps in the process of writing a literature review. A good literature review doesn't just summarize sources it analyzes, synthesizes, and crucially evaluates to give a clear picture of the state of knowledge on the subject.

A literature review is done to show one's familiarity with the topic and scholarly articles. Also, to position oneself relative to other researchers. It can help to show how the research addresses a gap or contributes to a debate.

4.1 Literature Review Process

Though the literature review is different for each case, they follow similar steps. They are:

4.1.1 Search for relevant literature

Using relevant keywords such as **Project Management**, **Earned Value Management**, **Artificial Intelligence**, **Machine Learning**, etc. A list of scholarly articles can be searched. The databases used are:

- Google Scholar
- JSTOR
- EBSCO
- ScienceDirect
- ResearchGate
- University Library

By reading the abstract, one can determine whether the article is relevant to our research. Bibliographies of those articles were also checked. If the same citation appears in many places, then that article is important and was included in our research.

4.1.2 Evaluate Sources

It is not possible to read everything. For this reason, the sources must be evaluated according to the relevance to the thesis. Initially, there were around 50 articles. After careful elimination, the number was brought down to 15. There is a mix of articles that focus on EVM and AI or AI in Project Management or EVM in Project Management. They were selected to grasp the whole argument.

4.1.3 Identify themes and gaps

After reading the literature, there are certain things that were established.

- Trends and patterns (in theory, method or results): do certain approaches become more or less popular over time?
- Themes: what questions or concepts recur across the literature?
- Debates, conflicts and contradictions: where do sources disagree?
- Pivotal publications: are there any influential theories or studies that changed the direction of the field?
- Gaps: What is missing from the literature? Are there weaknesses that need to be addressed?

4.1.4 Outline the structure

There are various approaches to organizing the body of a literature review. These can be:

- Chronological
- Thematic
- Methodological
- Theoretical

For this research, a Methodological approach has been chosen which was appropriate. The results were evaluated on the basis of qualitative versus quantitative research.

4.1.5 Review the articles

The final step is to write a literature review. There were 2 broad topics in which they were classified those were:

- Analysis of EVM variants in the literature
- Analysis of AI for PM in literature

4.2 Analysis of EVM Variants in Literature

(Babar, Thaheem, & Ayub, 2016): Their main argument is EVM is used everywhere while being effective; it only includes schedule and cost. Therefore, not considering the impact of other important aspects of health and safety, stakeholder satisfaction, and quality. Despite its superior formulation, EVM forecasts are still influenced by project risks and uncertainties, leading to inconsistency between EAC results obtained through standard formulae. In order to estimate better EAC, a framework is developed that incorporates various key performance indicators into the risk performance index (RPI). They were able to develop a better model than EVM and thereby exposing the shortcomings of the EVM model.

(Mahdi, Abd-Elrashed, Essawy, & Raed, 2018): This paper discussed the difficulties in implementing EVM in the construction sector in Egypt. Earned value analysis (EVA) is implemented on two different construction projects showing the effect of these difficulties on EVA results. A structured questionnaire and interviews were done for this objective. Their analysis illustrated that the technical problems mainly affect the project's time control, while the project's cost control is affected mainly by financial problems. There are some problems facing the actual cost recording that may impede the implementation of EV as a control technique.

(Ma & Yang, 2012): Due to the long construction cycle, high risk and extreme complexity of construction projects, traditional Earned Value Management appears somewhat limited in application, requiring the introduction of an integrated management control method. Based on the above characteristics of construction programs, this paper improves the traditional Earned Value Method by introducing quality cost. In this way it makes a more complete Earned Value Method assessment criteria system and introduces quality performance index into traditional method performance assessment criteria system to combine schedule management, cost management and quality management. The paper then conducts a case study to compare traditional Earned Value Management with Quality Earned Value Management, which shows Quality Earned Value Management that is a scientific and effective management technique and method.

(Baumann, Dziados, Mariusz, & Kapliński, 2014): The considerable degree of complexity found in construction work, as well as project susceptibility to unpredictable conditions determines the need for ongoing progress monitoring and continuous time-cost analysis during the execution of work. Financial and material analysis, using the Earned Value method applied to the construction of an Underground Gas Storage Facility, including project risks that occurred during the project, helped to identify the advantages and limitations in application of this method of monitoring work progress.

(Kim, Jr, G., & Duffey, 2003): Some problems that may impede the implementation of EVM are High cost, complicated and burdensome paperwork - Poor understanding of EVM - Distrust and conflict between project managers, project consulting and government -Pressures to report only good news.

(Nkiwane, Meyer, & Steyn, 2016): Lack of EVM as a contractual requirement diminished the directive use of EVM, since the parties of a contract cannot have a common understanding of performance in terms of earned value metrics. Discussions in any official correspondence across project role-players/parties cannot be in terms of EV metrics. This confine EVM to one party in a project, and so only that one party's decisions can be based on EV metrics. In this case, the earned value's input into such decisions was implicit, at best. This was also attributable to a limited understanding of EVM, and in particular of the future tense performance metrics such as To-Complete performance index (TCPI).

(J. & Christensen-Day, 2010): Earned value management is ignored and even discouraged for fixed-price contracts since the customer has transferred the risk of cost to the contractor and earned value is understood to be less valuable. Since the customer's cost risk is already lessened by transferring the risk to the contractor while it is required on high-value cost plus projects. On cost-plus contracts, the customer undertakes most of the risk. Earned value, when applied effectively provides early visibility into cost and schedule performance issues and for that reason it can be used as a tool that the customer uses to mitigate their risk.

(Lipke, 2003): He criticized the classic SPI indicators might give false and unreliable time forecasts near the end of the project. He presented the concept of Earned Schedule (ES) that converts the monetary indicators into a time dimension and produces a revised schedule performance indicator [SPI(t)] that can measure the performance along with the whole life of the project.

4.3 Analysis of AI for PM in Literature

(Leitch): This paper discusses the importance of AI in engineering. Though this paper is old (1992), this was selected to highlight the importance of AI which was recognized even more than 2 decades ago. It also discusses the various application of Artificial Intelligence in various branches of engineering.

(Wauters & Vanhoucke, A Nearest Neighbour extension to project duration forecasting with Artificial Intelligence, 2017): In this paper, they provide a Nearest Neighbor based extension for project control forecasting with Earned Value Management. The k-Nearest Neighbor method is employed as a predictor and to reduce the size of a training set containing more similar observations. An Artificial Intelligence (AI) method then makes use of the reduced training set to predict the real duration of a project. Additionally, they report on the forecasting stability of the various AI methods and their hybrid Nearest Neighbor counterparts. A large computer experiment is set up to assess the forecasting accuracy and stability of the existing and newly proposed methods. The experiments indicate that the Nearest Neighbor technique yields the best stability results and is able to improve the AI methods when the training set is similar or not equal to the test set. Sensitivity checks vary the amount of historical data and number of neighbors, leading to the conclusion that having more historical data, from which a relevant subset can be selected by means of the proposed Nearest Neighbor technique, is preferential.

(Wauters & Vanhoucke, A comparative study of Artificial Intelligence methods for project duration forecasting Expert Systems with Applications): This paper presents five Artificial Intelligence (AI) methods to predict the final duration of a project. A methodology that involves Monte Carlo simulation, Principal Component Analysis and cross-validation is proposed and can be applied by academics and practitioners. The performance of the AI methods is assessed by means of a large and topologically diverse dataset and is benchmarked against the best performing Earned Value Management/Earned Schedule (EVM/ES) methods. The results show that the AI methods outperform the EVM/ES methods if the training and test sets are at least similar to one another. Additionally, the AI methods report excellent early and mid-stage forecasting results. A robustness experiment gradually increases the discrepancy between the training and test sets and demonstrates the limitations of the newly proposed AI methods.

(Zhenyou, 2014): This study has successfully produced global optimal reduced models intelligently predicting software cost estimation by employing neural networks with back-propagation learning algorithms combined with genetic algorithms (GA-NN) to determine the most significant explanatory variables among the 16 COCOMO cost drivers. The optimal reduced models and their found significant factors can offer powerful supports for the project managers to make the right decisions in the early stages of the projects. The interesting finding in the reduced models and GA-NN simulation is as following: personnel attributes are the most significant ones, followed by project attributes and product attributes. Computer attributes are the least important in determining the software development effort. That means human capabilities and project management skills are still the most significant factors in the software development process. Flexibility, objectivity, correctness and computational economy are desirable features that make neural networks attractive as a learning-oriented estimator for the software development effort. In the GA-NN experiments, the local optimal subsets of the variables were found with different parameters of various crossover rates and mutation rates, and the nine significant cost drivers among the local optimal subset is then identified as global optimal sets in various stimulations, which can conclude that the reduced models are reliable and the effort estimation is accurate and robust.

(Stylianou & Andreou, 2011): This paper proposes a procedure for software project managers to support their project scheduling and team staffing activities – two areas where human resources directly impact software development projects and management decisions – by adopting a genetic algorithm approach as an optimization technique to help solve software project scheduling and team staffing problems. This paper presented an approach to solving the problem of software project scheduling and team staffing by adopting a genetic algorithm as an optimization technique in order to construct a project's optimal schedule and to assign the most experienced employees to tasks. The genetic algorithm uses corresponding objective functions to handle constraints and the results obtained when using either one of the objective functions show that the genetic algorithm is capable of finding optimal solutions for projects of varying sizes. However, when the objective functions were combined, the genetic algorithm presents difficulties in reaching optimal solutions especially when having a preference to assign the most experienced employees over the project's duration. Through observation of a number of executions, it was noticed that in this case the genetic algorithm couldn't reduce idle "gaps" or was not able to produce a conflict-free schedule. One possible reason for this observation is due to the competitive nature of the objective functions, and a definite improvement to the approach will be to use multi-objective optimization rather than using the aggregation of individual objective functions, which is how the genetic algorithm presently works.

(Soltanveis & Alizadeh, 2016): This paper is written on estimating software development costs, budgets and resources. The error rate, at the estimating costs, has a sizable portion in the success or fail of a project. In general, it is used from similar project histories for project estimation. One of the challenges in this approach is missing values. in this research, first, for handling missing values the K nearest neighbor (KNN) algorithm and Mean Imputation has been used, then for effort prediction, the parametric model-based methods, the nonlinear and polynomial regression(quadratic) is used. The proposed method is performed on the CMI dataset and the results show that the combination of KNN and nonlinear regression (quadratic) has the best response, signifying accuracy improvement and relative error reduction, in comparison with other approaches. According to the results, and by comparing this combined approach with recent researches, it can be stated that the proposed method possesses a higher accuracy on the normalized dataset. In addition, it can complement the current research methods. The proposed method is performed on the CMI dataset and the results show that the combination of K-Nearest Neighbor and nonlinear regression methods has the best response, signifying accuracy improvement and relative error reduction, in comparing with other approaches.

(Semenkina, Popov, & Semenkina): – The paper deals with the scheduling problem relevant in many fields, such as project management, lesson scheduling or production scheduling. We propose to implement a hierarchical problem structure that puts the traveling salesman problem at the top and replaces the nested resource-constrained project scheduling problem with a simulation model. The paper considers using the adaptive parameters control method for Ant Colony Optimization. The performance comparison with such algorithms as LinKernighan heuristic, Genetic Algorithm and Intelligent Water Drops Algorithm is made and competitive results are demonstrated. The adaptive method shows competitive results that are better than the one averaged by settings. This result also shows that it is necessary to move towards some kinds of adaptive methods that allow tuning algorithm parameters during their work for all bionic algorithms. This property is especially important when solving scheduling problems in the context of operational production planning when a solution needs to be found very quickly based on the current state of the manufacturing process.

(Cheng, Tsai, & Sudjono, 2010): This paper mentions that conceptual cost estimates are important to project feasibility studies and affect on the final project success. Such estimates provide substantial information that can be used in project assessments, cost budgeting, engineering designs and cost management. This study proposes an artificial intelligence approach specifically the evolutionary fuzzy hybrid neural network to improve conceptual cost estimate precision. This approach first integrates neural networks and high order neural networks into a hybrid neural network. This paper presents comprehensive descriptions of the proposed Evolutionary Fuzzy Hybrid Neural Network and its application in conceptual cost estimation for construction projects. The EFHNN mechanism integrates HNN, FL, and GA. In the proposed EFHNN, HNN includes both traditional neural (linear) and high order neural networks; FL uses fuzzification and defuzzification layers to sandwich the proposed HNN; and GA optimizes FHNN parameters. The proposed EFHNN is innately different from various GA-

FL-NN approaches, even the previously proposed EFNIM, due to unique HNN layer connection types, modification of FL membership functions, and GA-optimized parameters. Therefore, EFHNN is able to address problems in greater depth with its large number of HNN models, fuzzy concepts and GA optimization. This study proposes two distinctive construction cost estimators. The overall construction cost estimator was created to estimate the total cost in the absence of categorized engineering plans. Category estimators, depend on on additional data inputs, were designed to evaluate engineering costs within categories.

(Feng & Li, 2012): In this paper Genetic algorithm optimizing BP was proposed to aim at handling locality minimum and low convergence speed. The method is based on analyzing the basic fundamental states on how to use a genetic algorithm to improve the ability of BP. After optimizing, the GA-BP model has been built up. The result of GA-BP model can get lower forecast errors and iterations. For these reasons, GA-BP model is appropriate for construction cost estimation. Aiming at the existing deficiency of BP network mode, this paper puts forward the improved network mode (GA - BP). GA - BP is a kind of fast and effective method. it conducts genetic selection for many points in solution space, and searches for optimization point by the improved BP algorithm, which can not only overcome the problems of easily falling into a local minimum point and the poor generalization ability in BP algorithm, but also overcome the long-time training which is caused by searching for the optimal solution in a similar random form. The simulation results show that GA - BP model gets great improvement in generalization ability and has higher reliability. It is more suitable for the practical application of engineering cost estimation than BP model.

4.4 Conclusion of Research Review

After reviewing all the works of literature, it can be concluded that though EVM is a good model objectively for measuring project performances, it does have some limitations like subjectivity and dynamic scheduling which will be discussed in the next section. There are various pieces of literature that use AI or Machine Learning in Project Management effectively. They used various techniques and methodologies and usually they proved to be more accurate than an EVM model. Looking at the literature, our thesis topic has potential as AI can surpass the traditional EVM model.

5 Research Methodology and Analysis

5.1 The CONs of the EVM model

There are some limitations to the EVM model as inferred by the literature review. They are:

- The method is limited and does not directly account for variation in individual performance values.

- The subjective way of reviewing work progress is one of the reasons that may lead to the intentional fabrication of data in order to present the desired picture.
- The Time-lapse between checking and implementation of corrective actions can have some problems.
- It does not account for financial exposure which is a practical parameter as it determines whether a bank will fund a project or not.
- Dynamic scheduling can also have problems because of the aforementioned reasons.

5.2 Objectives

- Represent the project in arrays. Arrays can be easily manipulated in programming languages.
- Transform the arrays to get the necessary variables such as WBS x WBS = Sequencing
- Schedule the project.
- Try various combinations of scheduling to
- Minimize the time and cost under constraints (like capacities) via static and dynamic scheduling.
- Add financial exposure as a variable to minimize
- Find the optimal scheduling.

5.3 Representing project using matrices

A generic project is taken then the project is broken down into the following characteristics:

5.3.1 Work Breakdown Structure (WBS):

The Work Breakdown Structure (WBS) is developed to establish a common understanding of project scope. It is a hierarchical description of the work that must be done to complete the deliverables of a project. Each descending level in the WBS represents an increasingly detailed description of the project deliverables.

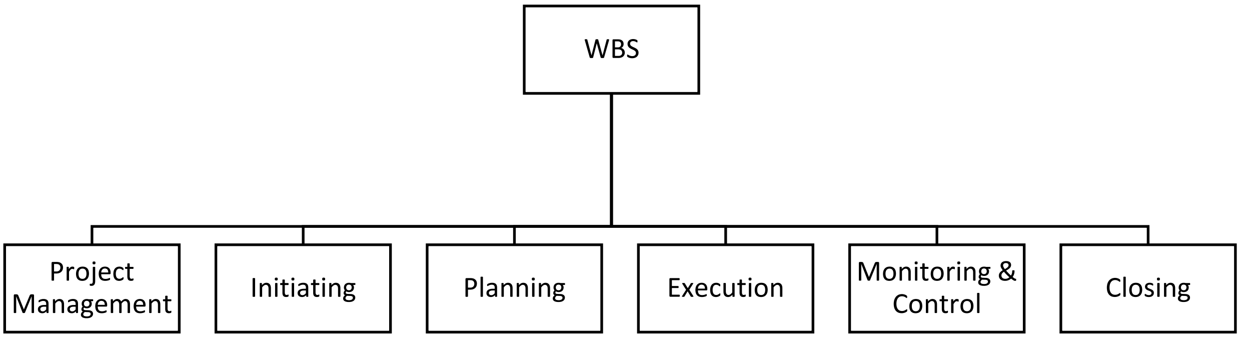


Figure 7: Work Breakdown Structure (WBS)

5.3.2 Organizational Breakdown Structure (WBS):

Organizational Breakdown Structure, or sometimes known as Organization Chart, is a widely used project management tool for representing project organization. It typically begins with the project sponsor, and with all key stakeholders included. In presenting the organization structure, consider the organization or group that is requesting the project and the level of their sponsorship and authority.

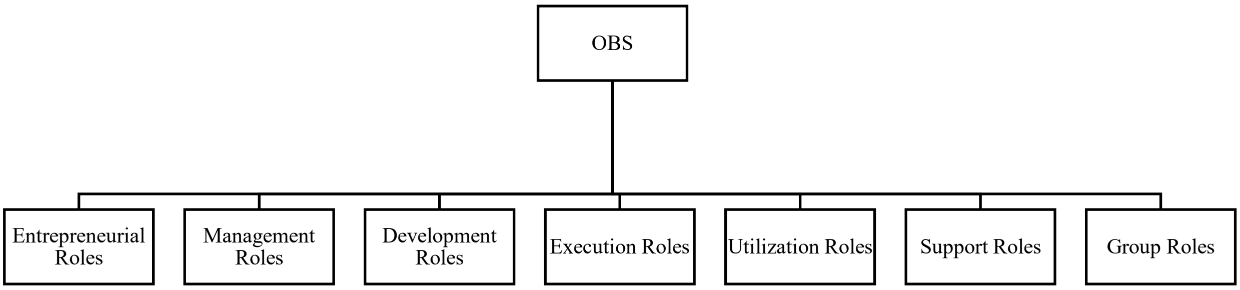
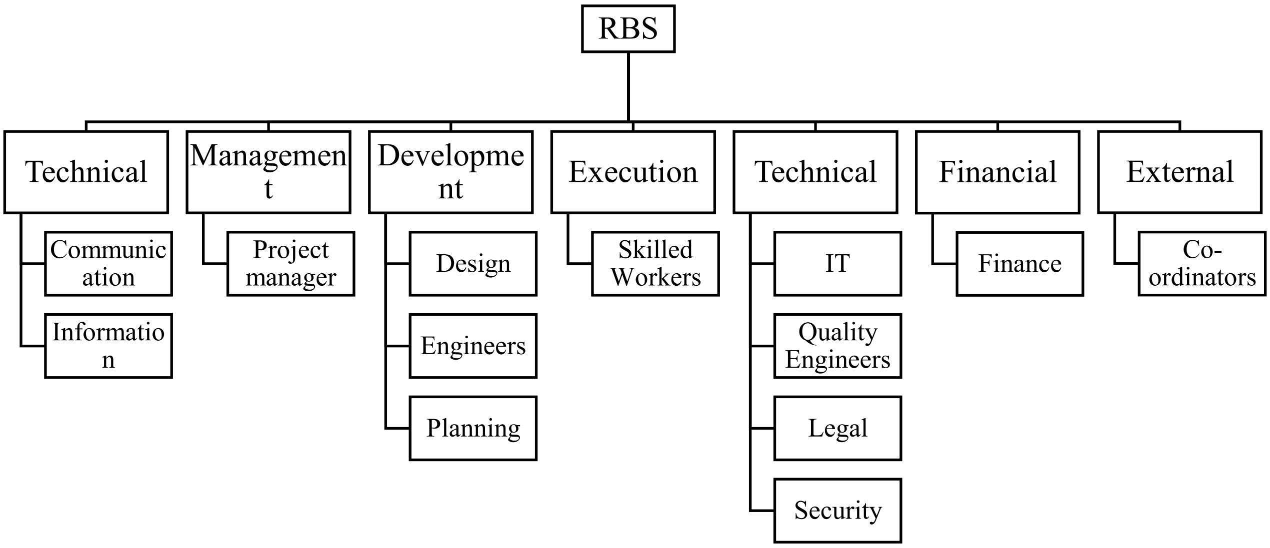


Figure 8: Organizational Breakdown Structure

5.3.3 Resource Breakdown Structure (RBS):

Resource Breakdown Structure (RBS) is a project management tool that provides a hierarchical decomposition of resources, either structured by resource category, types or by IT/business function that has resource needs.



5.3.4 Cost Breakdown Structure (CBS):

Cost Breakdown Structure (CBS) represents a breakdown of the costs of the various components of the Work Breakdown Structure (WBS) including all works or services done by the subcontractors. It is used to continuously compare the actual costs with the budget and integrate it to the cost control system.

Figure 9: Resource Breakdown Structure

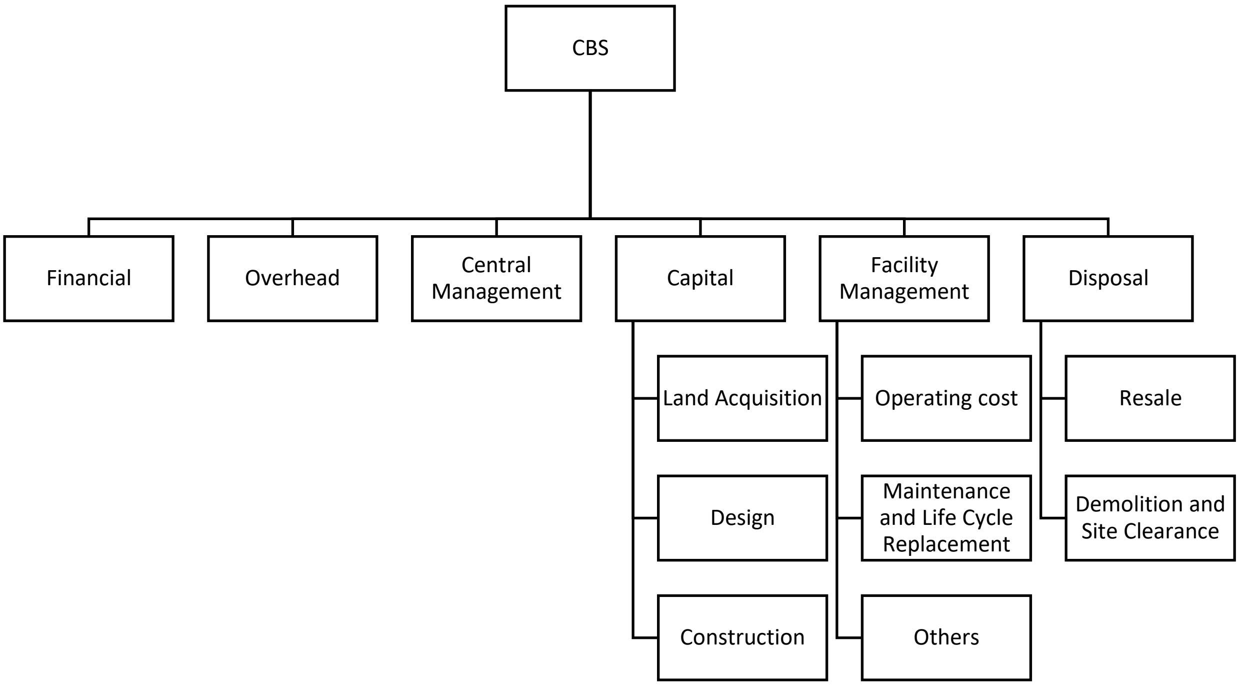
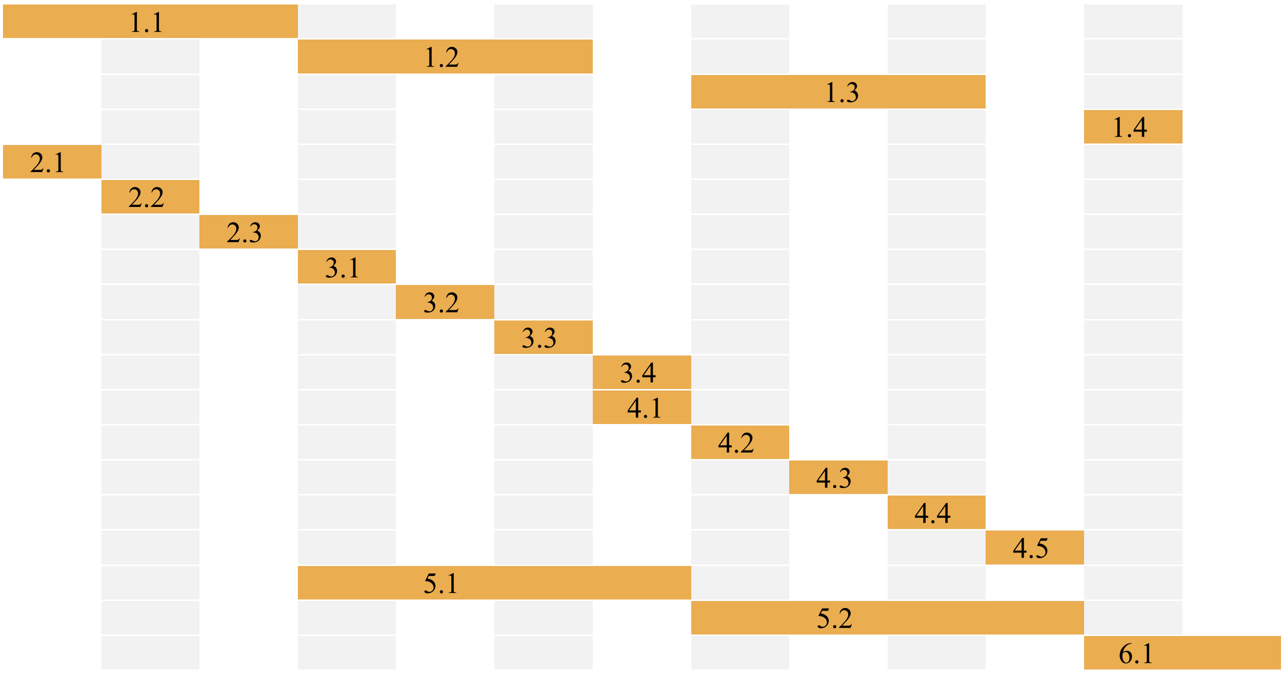


Figure 10: Cost Breakdown Structure

5.3.5 Sequencing

Below is one way the project is sequenced.



To represent it in a matrix form WBS is combined with WBS to give the sequencing.

		1.1	1.2	1.3	1.4	2.1	2.2	2.3	3.1	3.2	3.3	3.4	4.1	4.2	4.3	4.4	4.5	5.1	5.2	6.1	
Project Management	1.1		1																		1
	1.2			1																	1
	1.3				1															1	2
	1.4																				
Initiating	2.1						1														1
	2.2							1													1
	2.3								1									1			2
Planning	3.1									1											1
	3.2										1										1
	3.3											1	1								2
	3.4			1															1		2
Execution	4.1													1							1
	4.2														1						1
	4.3															1					1
	4.4																1				1
	4.5				1															1	2
Monitoring/Control	5.1																				
	5.2																			1	1
Closing	6.1																				
			1	2	2		1	1	1	1	1	1	1	1	1	1	1	1	1	3	21

Figure 11: Sequencing

In the above sequencing, resources are not yet taken into account so final scheduling will be different when resource constraint is also considered.

5.3.6 Resource Requirement Matrix (RRM)

Resource Requirement Matrix is obtained when combining the Organizational Breakdown Structure (OBS) with Work Breakdown Structure (WBS). This gives the resource requirement for each task.

		a.1	a.2	b.1	c.1	c.2	c.3	d.1	e.1	e.2	e.3	e.4	f.1	g.1	
Project Management	1.1	2	1	3		1	1	2	1	1	4	5	3	1	25
	1.2	2	4	1	5		2	1	2	4	2	3		1	27
	1.3			1	5	4	1	3	1	5	3	1	1	3	28
	1.4	1			2	2	3	1	2	2		1	1	3	18
Initiating	2.1	3	3	3	1	1	1	2	2	1		2	1	2	22
	2.2	5	2		2	2	4		4	4	3	2	2	4	34
	2.3	1	3	2	2	1	3	5		4	3	5	3		32
Planning	3.1	3			2	1	1	3	4	2	1	3		1	21
	3.2	2	4				4	3	1	5	1	5	1	3	29
	3.3	4	3	2	2	2		3	3	5		5	3	3	35
	3.4	3	2	1	1	2	4	5	5	2		4	2	1	32
Execution	4.1	2	1	4	4	1	1	5		3	1		1	5	28
	4.2		5	3	2	3	2	1	1	1	2	3	4	2	29
	4.3		1	5		2	2	5	5	3	5	1	2	4	35
	4.4	1	4	5	4	5		4		3	3	2	2	1	34
	4.5			1	3	3	2	1	1	1	4	4		3	23
Monitoring/Control	5.1	2	5		2	2	2	3		1	3	5	1	5	31
	5.2	2	5	5	4	5	1	4	3	3	3		5	5	45
Closing	6.1	3		1	4			3	1	4	2	3	4	1	26
		36	43	37	45	37	34	54	36	54	40	54	36	48	554

Figure 12: Resource Requirement Matrix (RRM)

The above numbers are randomly generated for the sake of simplicity. They don't represent any real-life project. The following table explains the initials pertaining to the resources and also gives their hourly wages. Note that the wages are also generated randomly and are not representative of a real-life scenario.

Table 1: OBS with hourly wage

a.1	**Technical**	**Communication**	**€**	**35**
a.2		Information	€	40
b.1	Management	Project Manager	€	30
c.1	Development	Designer	€	35
c.2		Engineers	€	35
c.3		Planning	€	30
d.1	Execution	Skilled Workers	€	30
e.1	Technical	IT	€	45
e.2		Quality Engineers	€	50
e.3		Legal	€	40
e.4		Security	€	30
f.1	Financial	Finance	€	35
g.1	External	Coordinators	€	20

Combining RRM with the daily wage, Daily Wage Matrix is obtained. It is given as follows.

		a.1	a.2	b.1	c.1	c.2	c.3	d.1	e.1	e.2	e.3	e.4	f.1	g.1	Daily Cos
Project Management	1.1	560	320	720		280	280	480	200	400	800	1200	840	160	6240
	1.2	560	1280	240	1600		560	240	400	1600	400	720		160	7760
	1.3			240	1600	1120	280	720	200	2000	600	240	280	480	7760
	1.4	280			640	560	840	240	400	800		240	280	480	4760
Initiating	2.1	840	960	720	320	280	280	480	400	400		480	280	320	5760
	2.2	1400	640		640	560	1120		800	1600	600	480	560	640	9040
	2.3	280	960	480	640	280	840	1200		1600	600	1200	840		8920
Planning	3.1	840			640	280	280	720	800	800	200	720		160	5440
	3.2	560	1280				1120	720	200	2000	200	1200	280	480	8040
	3.3	1120	960	480	640	560		720	600	2000		1200	840	480	9600
	3.4	840	640	240	320	560	1120	1200	1000	800		960	560	160	8400
Execution	4.1	560	320	960	1280	280	280	1200		1200	200		280	800	7360
	4.2		1600	720	640	840	560	240	200	400	400	720	1120	320	7760
	4.3		320	1200		560	560	1200	1000	1200	1000	240	560	640	8480
	4.4	280	1280	1200	1280	1400		960		1200	600	480	560	160	9400
	4.5			240	960	840	560	240	200	400	800	960		480	5680
Monitoring/Control	5.1	560	1600		640	560	560	720		400	600	1200	280	800	7920
	5.2	560	1600	1200	1280	1400	280	960	600	1200	600		1400	800	11880
Closing	6.1	840		240	1280			720	200	1600	400	720	1120	160	7280
		10080	13760	8880	14400	10360	9520	12960	7200	21600	8000	12960	10080	7680	147480

Figure 13: Daily Wage Matrix

With this matrix one can obtain the total daily cost by resources such as engineers (given in the bottom) or total cost by the tasks (given in the right).

5.3.7 Total Variable Costs

After obtaining various tasks daily costs, their total variable costs can be obtained. Considering a working day has 8 hours and multiplying it with the tasks' duration (also assigned randomly). The total variable cost for each task is calculated

Table 2: Total Variable Costs

Task	**Daily Wage**	**Duration**	**Total Variable Cost (in thousands)**
a.1	6240	20	124.8
a.2	7760	30	232.8
b.1	7760	40	310.4
c.1	4760	100	476
c.2	5760	30	172.8
c.3	9040	20	180.8
d.1	8920	30	267.6
e.1	5440	20	108.8
e.2	8040	30	241.2
e.3	9600	30	288
e.4	8400	30	252
f.1	7360	30	220.8
g.1	7760	20	155.2

5.3.8 Scheduling

Now the project is scheduled. There are some constraints regarding predecessors which are also assigned randomly. Some of the tasks are without constraints so they can be performed in parallel. Resource-based scheduling is performed which means resource constraints are also kept in mind. Maximum 5 resources are available in each category any given day.

Table 3: Project information

Task Name	**Duration (days)**	**Start**	**Finish**	**Cost**	**Constraints**

	New Project	**470**	**01/01/2020**	**19/10/2021**	**€ 5,984,800**	
Project Management		**240**	**01/01/2020**	**01/12/2020**	**€ 1,144,000**	
1.1	PM 1	20	01/01/2020	28/01/2020	€ 124,800	
1.2	PM 2	30	29/01/2020	07/04/2020	€ 232,800	PM 1
1.3	PM 3	40	07/04/2020	02/06/2020	€ 310,400	PM 2
1.4	PM 4	100	02/06/2020	01/12/2020	€ 476,000	PM 3
Initiating		**80**	**01/01/2020**	**21/04/2020**	**€ 621,200**	
2.1	In 1	30	01/01/2020	11/02/2020	€ 172,800	PM 1 SS
2.2	In 2	20	12/02/2020	10/03/2020	€ 180,800	In 1
2.3	In 3	30	11/03/2020	21/04/2020	€ 267,600	In 2
Planning		**110**	**21/04/2020**	**22/09/2020**	**€ 890,000**	
3.1	PL 1	20	21/04/2020	19/05/2020	€ 108,800	In 3
3.2	PL 2	30	20/05/2020	30/06/2020	€ 241,200	PL 1
3.3	PL 3	30	01/07/2020	11/08/2020	€ 288,000	PL 2
3.4	PL 4	30	11/08/2020	22/09/2020	€ 252,000	PL 3
Execution		**300**	**11/08/2020**	**05/10/2021**	**€ 1,534,400**	
4.1	EX 1	30	11/08/2020	01/12/2020	€ 220,800	PL 3
4.2	EX 2	20	02/12/2020	29/12/2020	€ 155,200	EX 1
4.3	EX 3	50	30/12/2020	09/03/2021	€ 424,000	EX 2
4.4	EX 4	60	09/03/2021	25/06/2021	€ 564,000	EX 3
4.5	EX 5	30	25/06/2021	05/10/2021	€ 170,400	EX 4
Monitoring/Control		**370**	**22/04/2020**	**21/09/2021**	**€ 1,649,600**	
5.1	MC 1	150	22/04/2020	17/11/2020	€ 170,400	In 3
5.2	MC 2	220	18/11/2020	21/09/2021	€ 1,188,000	MC 1
Closing		**20**	**22/09/2021**	**19/10/2021**	**€ 145,600**	
6.1	Closing	20	22/09/2021	19/10/2021	€ 145,600	MC 2

With the above information, the matrix of working days is obtained. Holidays are not accounted and only weekends are excluded. This matrix acts as a Gantt chart, a way to visualize the project. The matrix is given below.

		Jan-20	Feb-20	Mar-20	Apr-20	May-20	Jun-20	Jul-20	Aug-20	Sep-20	Oct-20	Nov-20	Dec-20	Jan-21	Feb-21	Mar-21	Apr-21	May-21	Jun-21	Jul-21	Aug-21	Sep-21	Oct-21	
		23	20	23	23	22	23	23	22	23	22	22	24	22	21	24	22	22	23	22	23	23	22	Duration (days)
Project Management	1.1	20																						20
	1.2	3	20	7																				30
	1.3				19	21																		40
	1.4						22	23		11	22	22												100
Initiating	2.1	23	7																					30
	2.2		13	7																				20
	2.3			16	14																			30
Planning	3.1				9	11																		20
	3.2					9	21																	30
	3.3							23	7															30
	3.4								16	14														30
Execution	4.1									16		14												30
	4.2												20											20
	4.3												3	22	21	4								50
	4.4															18	22		20					60
	4.5																		5	15		7	3	30
Monitoring/Control	5.1				8	22	23	23	22	23	22	7												150
	5.2											10	24	22	21	24	22	22	23	22	23	7		220
Closing	6.1																					8	12	20

Figure 14: Matrix of working days

In the above matrix there are two important things to note. Some tasks are split because of resource constraints. The tasks highlighted by orange are in the critical path. That is any delay in those tasks will cause the whole project to get delayed.

5.3.9 Cash flows

There are two important types of costs. Fixed costs and variable costs. Fixed costs don't change along with the duration and variable costs can change.

5.3.9.1 Variable cash flow matrix

After obtaining the above scheduling matrix, a variable cash flow matrix is obtained by combining the total variable cost matrix with the scheduling matrix. The costs are divided proportionately with the number of working days in a month the task is performed. The resulting matrix is as follows.

		Jan-20	Feb-20	Mar-20	Apr-20	May-20	Jun-20	Jul-20	Aug-20	Sep-20	Oct-20	Nov-20	Dec-20	Jan-21	Feb-21	Mar-21	Apr-21	May-21	Jun-21	Jul-21	Aug-21	Sep-21	Oct-21
		23	20	23	23	22	23	23	22	23	22	22	24	22	21	24	22	22	23	22	23	23	22
Project Management	1.1	-124.8																					
	1.2	-23.28	-155.2	-54.32																			
	1.3				-147.44	-162.96																	
	1.4						-104.72	-109.48		-52.36	-104.72	-104.72											
Initiating	2.1	-132.48	-40.32																				
	2.2		-117.52	-63.28																			
	2.3			-142.72	-124.88																		
Planning	3.1				-48.96	-59.84																	
	3.2					-72.36	-168.84																
	3.3							-220.8	-67.2														
	3.4								-134.4	-117.6													
Execution	4.1									-117.76		-103.04											
	4.2												-155.2										
	4.3												-25.44	-186.56	-178.08	-33.92							
	4.4															-169.2	-206.8		-188				
	4.5																		-28.4	-85.2		-39.76	-17.04
Monitoring/Control	5.1				-63.36	-174.24	-182.16	-182.16	-174.24	-182.16	-174.24	-55.44											
	5.2											-118.8	-285.12	-261.36	-249.48	-285.12	-261.36	-261.36	-273.2	-261.36	-273.24	-83.16	
Closing	6.1																					-58.24	-87.36

Figure 15: Variable Cash flow matrix

The above numbers are in thousands of euros. The sign is negative to signify the outflow.

5.3.9.2 Fixed cash flow matrix

Before calculating the fixed cash flow matrix, the overhead cost and contingency budget are calculated. They are as follows.

Total	**-9236.8**
Overhead	-923.68
Contingency	-1385.52
Total	-11546
Revenue	12700.6

The overhead cost is 10 % of the total costs including fixed costs. The contingency cost is 15 % of the total cost. Revenue is 10% of the total costs.

		Jan-20	Feb-20	Mar-20	Apr-20	May-20	Jun-20	Jul-20	Aug-20	Sep-20	Oct-20	Nov-20	Dec-20	Jan-21	Feb-21	Mar-21	Apr-21	May-21	Jun-21	Jul-21	Aug-21	Sep-21	Oct-21
		23	20	23	23	22	23	23	22	23	22	22	24	22	21	24	22	22	23	22	23	23	22
Project Management	1.1																						
	1.2	-300																					
	1.3				-200																		
	1.4						-250																
Initiating	2.1																						
	2.2																						
	2.3																						
Planning	3.1																						
	3.2																						
	3.3																						
	3.4																						
Execution	4.1									-150													
	4.2												-100										
	4.3												-100										
	4.4																						
	4.5																						
Monitoring/Control	5.1																						
	5.2																						
Closing	6.1																						
Overhead		-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99

Figure 16: Fixed cash flow matrix

The overhead costs are divided monthly and will increase if there is any delay.

5.3.9.3 Payment Schemes

There are many payment schemes firms use when paying for a project. In this thesis, 4 common payment schemes are considered and payment scheme matrix is obtained.

They are:

i. Lump-sum at the start
ii. Lump-sum at the end
iii. By progress (percentage)
iv. By scope

The first two are fairly simple. In payment by scope, once a scope is finished a percentage of payment is received. Payment by progress, however, is a bit tricky.

In this payment scheme payments are done at a specific percentage of work completed. In this case every 25%. Keep in mind it is the percentage of work completed, not the percentage of duration. A typical project follows an s-curve model. For this reason, an S-curve is plotted and the progress is estimated from that.

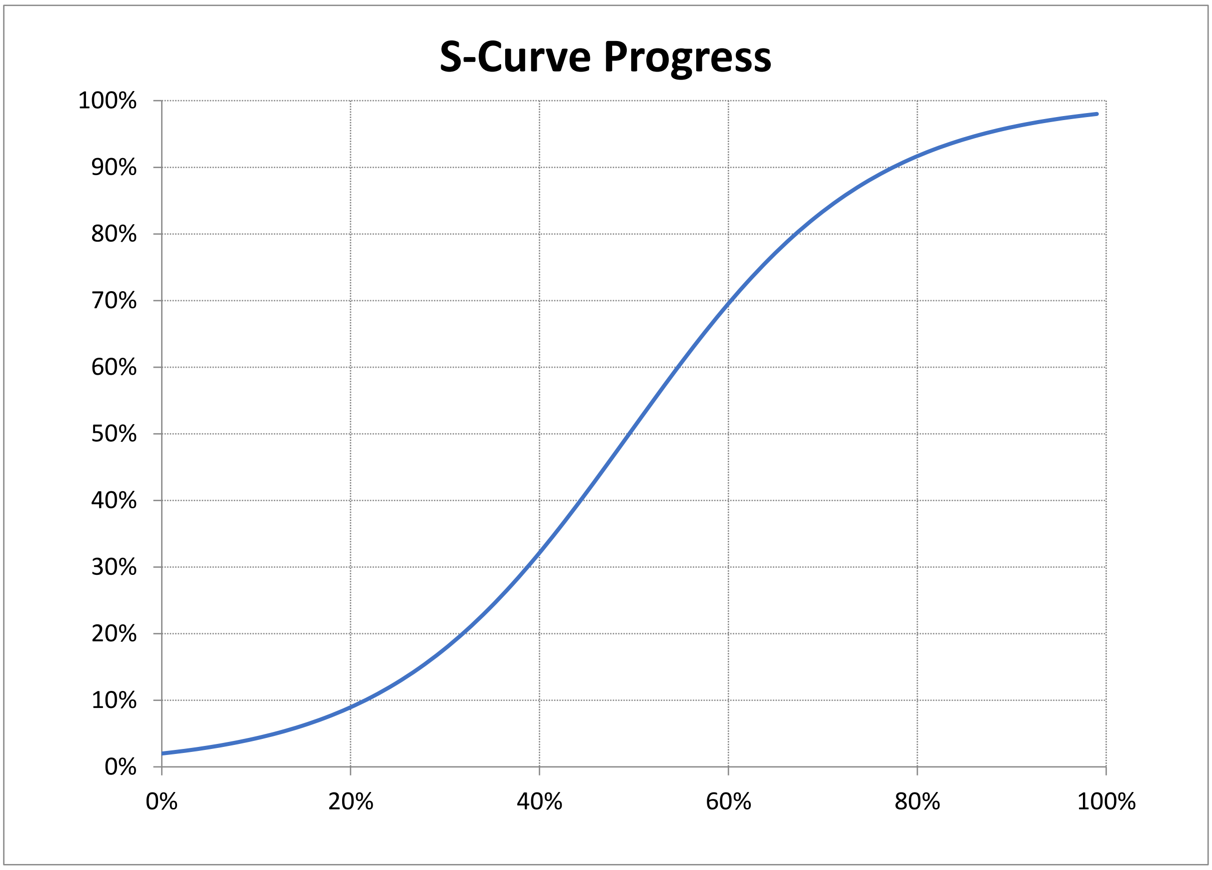


Figure 17: S-Curve

In the above S-curve, it can be seen that when 20% of the time has passed only 10% of the work is done but when 60% of the time is passed, around 70% of work is completed. This follows the real-life trend. This is due to most of the work is performed during the execution phase and because of that the curve rises steeply towards the middle whilst before in the planning stage, not much work is performed.

	Jan-20	Feb-20	Mar-20	Apr-20	May-20	Jun-20	Jul-20	Aug-20	Sep-20	Oct-20	Nov-20	Dec-20	Jan-21	Feb-21	Mar-21	Apr-21	May-21	Jun-21	Jul-21	Aug-21	Sep-21	Oct-21
	23	20	23	23	22	23	23	22	23	22	22	24	22	21	24	22	22	23	22	23	23	22
	5%	9%	13%	18%	22%	27%	32%	36%	41%	45%	50%	55%	59%	63%	68%	73%	77%	82%	86%	91%	96%	100%
Initial	12700.6																					
End																						12700.6
Percentage								3175.15			3175.15			3175.15								3175.15
Scope						4233.53								4233.53								4233.53

Figure 18: Payment Schemes

5.3.9.4 Cumulative Cash flow matrix

Now, all the costs are combined to get the cumulative cash flow matrix. It also includes inflow depending on the payment scheme selected. The below matrix is one example of the matrix with payment by progress chosen.

		Jan-20	Feb-20	Mar-20	Apr-20	May-20	Jun-20	Jul-20	Aug-20	Sep-20	Oct-20	Nov-20	Dec-20	Jan-21	Feb-21	Mar-21	Apr-21	May-21	Jun-21	Jul-21	Aug-21	Sep-21	Oct-21
		23	20	23	23	22	23	23	22	23	22	22	24	22	21	24	22	22	23	22	23	23	22
Project Management	1.1	-124.8																					
	1.2	-323.28	-155.2	-54.32																			
	1.3				-347.44	-162.96																	
	1.4						-354.72	-109.48		-52.36	-104.72	-104.72											
Initiating	2.1	-132.48	-40.32																				
	2.2		-117.52	-63.28																			
	2.3			-142.72	-124.88																		
Planning	3.1				-48.96	-59.84																	
	3.2					-72.36	-168.84																
	3.3							-220.8	-67.2														
	3.4								-134.4	-117.6													
Execution	4.1									-267.76		-103.04											
	4.2												-255.2										
	4.3												-125.44	-186.56	-178.08	-33.92							
	4.4															-169.2	-206.8		-188				
	4.5																		-28.4	-85.2		-39.76	-17.04
Monitoring/Control	5.1				-63.36	-174.24	-182.16	-182.16	-174.24	-182.16	-174.24	-55.44											
	5.2											-118.8	-285.12	-261.36	-249.48	-285.12	-261.36	-261.36	-273.2	-261.36	-273.24	-83.16	
Closing	6.1																					-58.24	-87.36
Overhead		-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99	-41.99
Outflow		-622.55	-355.03	-302.31	-626.63	-511.39	-747.71	-554.43	-417.83	-661.87	-320.95	-423.99	-707.75	-489.91	-469.55	-530.23	-510.15	-303.35	-531.63	-388.55	-315.23	-223.15	-146.39
Inflow									3175.15			3175.15			3175.15								3175.15
Cash flow		-622.55	-355.03	-302.31	-626.63	-511.39	-747.71	-554.43	2757.32	-661.87	-320.95	2751.16	-707.75	-489.91	2705.60	-530.23	-510.15	-303.35	-531.63	-388.55	-315.23	-223.15	3028.76
Cumulative Cash flow		-622.55	-977.57	-1279.88	-1906.50	-2417.89	-3165.59	-3720.02	-962.69	-1624.56	-1945.50	805.66	97.91	-391.99	2313.61	1783.39	1273.24	969.90	438.27	49.73	-265.50	-488.64	2540.12

Figure 19: Cumulative Cash Flow Matrix

With this cumulative cash flow matrix, we obtain the following cash flow.

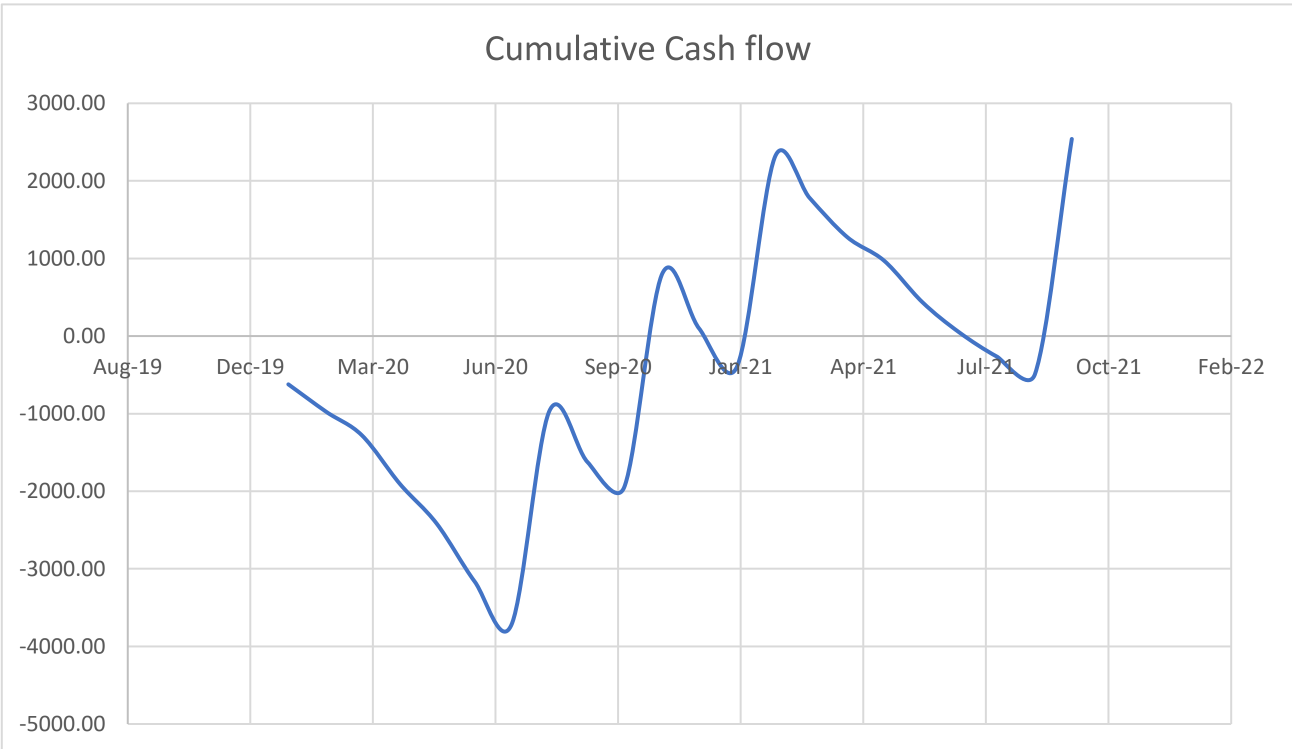
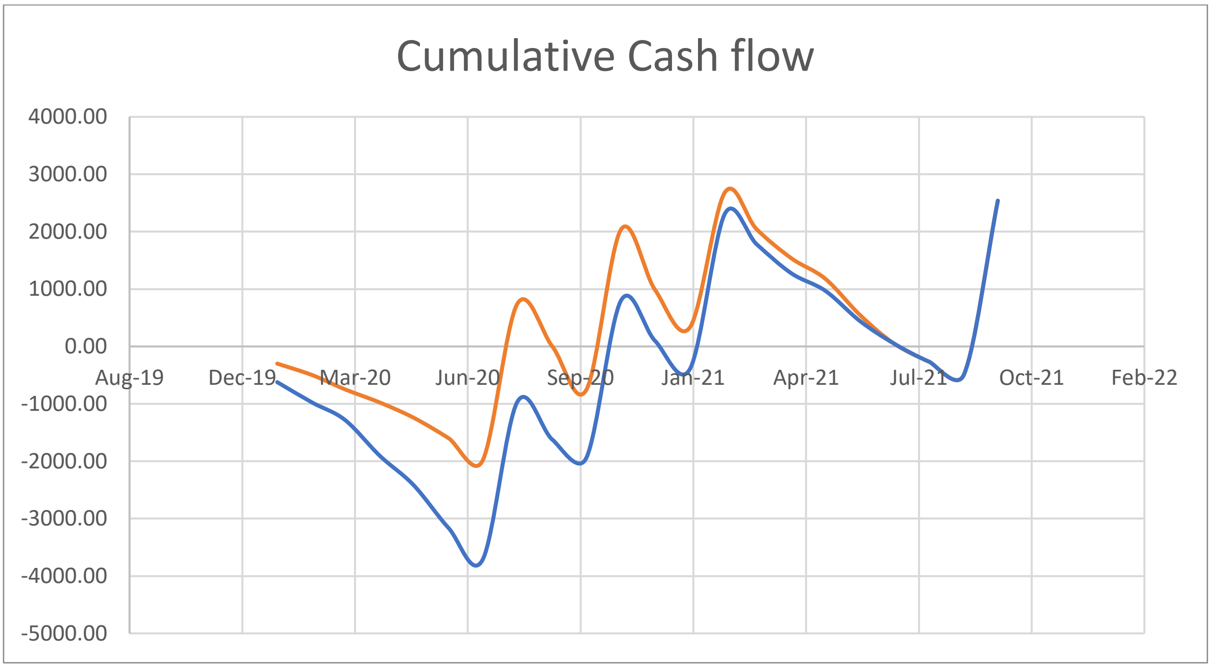


Figure 20: Cumulative cash flow

One of the important things in real life is financial exposure. The financial need which is around 4 million euros in this case is a lot. Traditional project management software does not account for this and as a result it is more difficult to obtain the loan from the bank.

6 Results

In the project, some tasks were rescheduled adhering to the constraints. This new cash flow obtained is as follows. For comparison the initial cash flow is also plotted.



In the above cash flow, it can be seen that the financial need is reduced by almost half to 2 million. The reason is some tasks that were not critical were rescheduled to later dates. This does not affect project duration. The revenue and cost all remain the same. It's just the curve is less steep and as a result the required amount is 2 million which is easier to get than 4 million.

7 Approach

7.1 Research Approach: Qualitative or Quantitative?

Research can be based on the time factor and can be mainly classified into two types; one is qualitative and the other is quantitative. They have their pros and cons. They don't have any apparent superiority over one another and the research approach to be used depends on the type of research to be performed.

Quantitative research is considered to be the best research approach by some because of the appropriate methodology in data acquisition lead to suitable calculated information. The processes which can be used to collect data can be questionnaires, surveys, and coded synthetic monitoring.

While on the other hand qualitative research can be interpreted as the acquisition of knowledge that should be examined, analyzed and completely understood the processes through the medium of a concise point of view. Researchers claimed that qualitative research is mostly based on interviews and interactions which are open-end and on the data of other fields, sometimes can also consist of quantitative data or disruptive statistics.

The broad quantitative approach would be a better choice for the research, but due unavailability of capable resources for the acquisition of real data in order to access real information, a qualitative approach is widely adopted. In this research mostly qualitative and some quantitative approach has been used

7.2 Research Method: Constructive or Empirical?

The research method is a scientific term used for the nature of research conduction, to evaluate how the facts have been derived in the research work, it is a label to research based on the procedure of work done in research.

According to definition constructive research is based on the theories and its evaluation and feasibility or possibility to apply in some technology, it does not need that research must be as per solidity, these type of research may be some case studies, may be based on hypothesis and supposition, while on the other hand Empirical research should be based on reality, must be solid observations, and proofs of theories, in which real data should be utilized to obtain the results.

This research is Empirical as Machine learning needs real-world data.

8 Limitations

- The primary goal was to develop a model to estimate project duration and costs using artificial intelligence. The goal was to accurately forecast duration and costs in order to better track the project.
- In this thesis, Artificial Intelligence concepts were discussed but not applied. The main reason being the lack of real project data. The machine learning methods require a huge amount of data, which is then analyzed and patterns are recognized by the computer.
- Another reason is Artificial Intelligence in Project Management is not in the scope of the master’s thesis. This field is relatively new and not much researches have been done. Whilst doing the thesis, I realized this topic is better suited for the Ph.D. as there are more time and resources.

- The secondary goal of this thesis was to reduce the financial need. In this thesis it was done to some extent by representing the project in matrices but there were some limitations.
- In the financial exposure, only the rescheduling of non-critical tasks was performed. In practice, given the penalty for the delay. A trade-off could be made among cost, duration and financial exposure. For simplicity, in this thesis the financial exposure was optimized without a tradeoff.

9 Conclusion

Even with all the limitations, I learned a lot. The basic methodology (for future) and advantage of AI is:

- Collect real project data
- The data should be organized (structured like a database)
- Apply one of the algorithms mentioned in the thesis.
- Python is generally used but any programming language can be used
- If there are lots of data or more accuracy is needed, then using a cloud service such as Amazon Web Services is better
- The computer analyzes all the data and finds patterns which help it to make a more accurate prediction
- For example, it can predict that projects depending on many external suppliers are 5% delayed or projects done on bureaucratic countries tend to get delayed by 10%. These are some of the things EVM models can't account for.
- In AI the machine basically learns from its mistake so if the prediction is wrong, it will only get more accurate in the future.
- Obviously, it also reduces delays due to human error
- AI in project management does not replace humans. The project managers are still really important. However, it can be a powerful tool for a project manager which can increase productivity.

10 References

Azzeh, M. (2013). Software Cost Estimation Based on Use Case Points for Global Software Development. *International Conference on Computer Science and Information Technology*.

Babar, S., Thaheem, M. J., & Ayub, B. (2016). Estimated Cost at Completion: Integrating Risk into Earned Value Management. *American Society of Civil Engineers*.

Baumann, T., Dziados, A., M. R., & Kapliński, O. (2014). Range of aplication and limitations of the earned value method in construction project estimation.

Benala, T. R., Dehuri, S., Mall, R., & ChinnaBabu, K. (2012). Software effort prediction using unsupervised learning clustering and functional link artificial neural networks.

Cheng, M.-Y., Tsai, H.-C., & Sudjono, E. (2010). Conceptual cost estimates using evolutionary fuzzy hybrid neural network for projects in construction industry. *Expert Systems with Applications*.

Crawford, B., Soto, R., Johnson, F., Misra, S., Paredes, F., & Olguín, E. (2015). Software project scheduling using the HyperCube ant colony optimization algorithm.

Feng, G., & Li, L. (2012). *Application of Genetic Algorithm and Neural Network in Construction Cost.* Zhengzhous, China.

Garcia-Diaz, N., Garcia-Virgen, J., Farias-Mendoza, N., Verduzco-Ramirez, A., Martinez-Bonill, R., E. C.-V., & Macias-Chapula, H. (n.d.). Software development time estimation based on a new Neuro-fuzzy approach.

Ge, Y. (2009). Software Project Rescheduling with Genetic Algorithms. *International Conference on Artificial Intelligence and Computational Intelligence*.

Guo, R., Guo, W., & Wang, W. (2008). Optimization of Multi-project in Limited Resources on the Basis of Improved Genetic Algorithm.

Hamza, D. H., Kamel, D. A., & Shams, K. (2013). Software Effort Estimation using Artificial Neural Networks: A Survey of the Current Practices. *International Conference on Information Technology: New Generations*.

J., C., & Christensen-Day. (2010). Earned Value on Fixed-Price Projects. *AACE international transactions*.

Kim, E., Jr, W., G., W., & Duffey, M. R. (2003). A model for effective implementation of Earned Value Management methodology. *International Journal of Project Management*, 375-382.

Leitch, P. R. (n.d.). Artificial intelligence in engineering. *Control Engineering Journal*.

Lipke, W. H. (2003). Schedule is Different. *The Measurable News*.

Ma, X., & Yang, B. (2012). Optimization study of Earned Value Method in construction project management. *International Conference on Information Management, Innovation Management and Industrial Engineering*.

Ma, Z., & Liu, Z. (2014). BIMbased intelligent acquisition of construction information for cost estimation of building projects. *Procedia Engineering*.

Mahdi, I., Abd-Elrashed, I., Essawy, A. S., & Raed, L. (2018). DIFFICULTIES OF IMPLEMENTING EARNED VALUE MANAGEMENT IN CONSTRUCTION SECTOR IN EGYPT. *International Journal of Engineering Researches and Management Studies*.

Nkiwane, N. H., Meyer, W. G., & Steyn, H. (2016). The use of Earned Value management for initiating directive project control decisions: a case study. *South African Journal of Industrial Engineering*, 192-203.

PMBOK® Guide. (2017). Project Management Institute.

Saleem, M., Hussain, R., Ismail, Y., & Mohsin, S. (2009). Cost Effective Software Engineering using Program Slicing Techniques.

Semenkina, O. E., Popov, E., & Semenkina, O. E. (n.d.). *Adaptive Ant Colony Optimization Algorithm for Hierarchical Scheduling Problem.*

Soltanveis, F., & Alizadeh, S. H. (2016). Using parametric regression and KNN algorithm with missing handling for software effort prediction.

Stylianou, C., & Andreou, A. S. (2011). Intelligent Software Project Scheduling and Team Staffing with Genetic Algorithms.

Vanhoucke, M. (2012). *Project Management with Dynamic Scheduling Baseline Scheduling, Risk Analysis and Project Control.* Berlin: Springer.

Wauters, M., & Vanhoucke, M. (2017). A Nearest Neighbour extension to project duration forecasting with Artificial Intelligence. *European Journal of Operational Research*.

Wauters, M., & Vanhoucke, M. (n.d.). A comparative study of Artificial Intelligence methods for project duration forecasting Expert Systems with Applications. *Expert SystemsWithApplications*.

Zhenyou, L. (2014). Predicting Project Effort Intelligently in Early Stages by Applying Genetic Algorithms with Neural Networks.