ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

THE MACHINE LEARNING IN PROJECT MANAGEMENT: A SYSTEMATIC MAPPING STUDY

[1]HAMID CHENNOUK*, [2]BRAHIM EL BHIRI, [3]SAMIR BARA AND [4]EL HOUSSAINE ZIYATI

[1]RITM laboratory, Computer Science and Networks Team ENSEM - ESTC - UH2C Casablanca, Morocco

[2]SMARTILAB laboratory, Moroccan School of Engineering Sciences, EMSI Rabat, Morocco

[3]LRIT, Faculty of Sciences - Mohammed V University, Rabat, Morocco

[4]RITM laboratory, Computer Science and Networks Team ENSEM - ESTC - UH2C Casablanca, Morocco

E-mail: [1]hamid.chennouk@gmail.com, [2]bra.elbhiri@gmail.com, [3]samirbara00@gmail.com, [4]ziyati@gmail.com

ABSTRACT

Machine learning (ML) is a trending research topic that affects a wide variety of disciplines including project management (PM). To provide clear insight of the use of machine learning in project management, identify gaps in current works and help researchers choose best options for their works, we conducted a Systematic Mapping Study review (SMS) following the guidelines presented by Petersen et al. [1]. We investigated studies reported from 2010 to 2019. 137 articles have been analyzed resulting in 33 primary studies. We found that ML is more involved in management of software project, especially, in effort estimation. We also found that a large range of ML techniques have been investigated, and Neural Network is by far the most widely used. We believe the results will be useful to understand the state of art of the use of ML in PM, and shed light on research topics that have not yet been explored.

Keywords: *Machine learning (ML), Project Management (PM), Systematic Mapping Study (SMS)*

1. INTRODUCTION

According to Gartner Analysts, by 2030 , 80 percent of today's Project Management tasks could be eliminated as Artificial Intelligence (AI) takes over [2]. On second thought, it seems logical. In fact, in the era of the fourth industrial revolution, digitalization takes place in most organizations. Consequently, more and more data is generated, especially by projects, which involves the deployment of technologies like machine-learning, which is a branch of AI, that can help with checking for risks, re-baselining plans, evaluating performance, etc. This conclusion is supported by Witten et al., who had previously argued that where there is data there is a possibility of ML application [3].

In this context, the present work aims to research studies that have been carried out about the use of ML in PM, analyzes them, and shed light on challenges that need to be addressed by future works.

As far as we know, there is no similar study in the literature that is interested, without restriction on the project management field and/or type, in the use of ML in PM. Consequently, this work contributes to the literature by drawing up a state of the art regarding subject under stud, shed light to challenges that need to be addressed in future studies, and finally, provide a good reference for the algorithm and technology to choose when implementing an ML solution in PM.

The reminder of this paper is organized as follows. Section 2 describes related work to our proposal. Section 3 breaks down the research methodology adopted to carry out the SMS, presents the results and answers the research questions of the study. Sections 4 covers the limitations of the study. And finally, Section 5 contains a conclusion and presents a set of future work.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2. RELATED WORK

Serval papers were found as a result of our research and were considered relevant to our topic. Below we describe some of them :

- Adriano L.I. Oliveira et al. propose a method for software effort estimating. It apply Genetic Algorithms (GA) to select features and parameters for machine learning techniques used for regression [4].
- In [30] , Anurag Agarwal et al. develop a hybrid method based on GA and neural network (NN) to address the resource constrained project scheduling problem.
- The authors of [12] suggest a method built on Correlation Cascade Neural Network (CCNN) to predict effort from use case diagrams for software project.
- The paper [27] deal with the effort estimation for a software project. Authors deployed Multilayer Perceptron neural network (MLP) to anticipate software effort considering the software size and team productivity.
- Golnoush Abaei and Ali Selamat conduct a survey on software fault detection. the authors analyze the performance of a set of ML techniques. They conclude that Random forest provides the best prediction performance for large data sets and Naïve Bayes for small ones [9].
- Jui-Sheng Chou et al. in [25] use genetic algorithm in several forecasting models to provides an optimization process that predict construction project award prices. The evaluation show that the use of GA and NN deliver better results.
- The works of Ryen W. White and Ahmed Hassan Awadallah focus on the problem of task duration estimation. They train four neural network models namely, Multilayer Perceptiotrainedon (MLP), General regression neural network (GRNN), Radial basis neural network (RBNN) and Cascade correlation neural network (CCNN) to predict how long a task will take to complete. They also show the correlation between that duration and task attributes, context, and history [5]

The main difference between these studies and ours is that, unlike them, we approach project management in its entirety, without any restriction on the project type and without limitation on the project field.

3. RESEARCH STUDY DESIGN

The standard procedures of a systematic mapping study presented in [1] were followed to get a general overview of the of studies published since 2010 on the application of ML techniques in PM. The key point of this kind of review is that it helps, with ease, researchers to summarize publications in any subject of interest, classify them , and provides a visual synthesis of findings. Furthermore, SMS outline gaps in current works and sheds light on future research areas.

As illustrated by Figure 1, the SMS process is accomplished in five stages with specific outcomes:

1) *Definition of Research Questions*. Define the research scope by reformulating the research questions (RQs).
2) *Conducting Search*. Construct search strings from the research questions and target the digital libraries to identify the primary studies.
3) *Screening of Papers*. Select relevant papers to RQs applying inclusion and exclusion criteria.
4) *Keyboarding using Abstracts*. Find most representative keywords of the abstracts to come whit a classification scheme.
5) *Data Extraction and Mapping Process*. Representation and interpretation of the results.

2.1 Research Questions

Research questions are the cornerstone of the SMS. They express the information needs of researchers, generate search strings, allow the definition of inclusion and exclusion criteria, and determine the representative keywords from abstracts. In other words, RQs guide the work by defining the review road-map.

The RQs investigated for this study are synthesized in Table 1.

2.2 Conduct Search

As part of this step, we firstly select digital libraries where the search for papers will be performed. We have chosen ACM, IEEE Xplore and ScienceDirect due to the wide range of studies they are indixing.

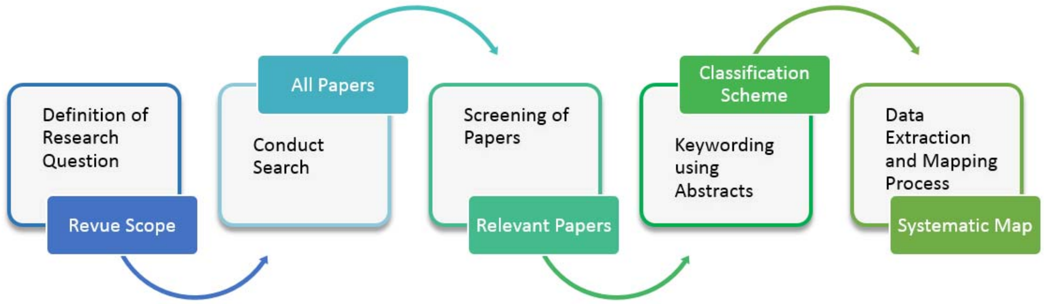

Figure 1. Systematic Mapping Study workflow.

Table 1. Research questions

	Id	Research Questions	Main motivations
Paper perspective	RQ1	What are the bibliographic features of the studies on the use of ML in PM ?	To identify the bibliographic features : Paper Type, Publication Year, Venue and Number of citation
	RQ2	What contribution the studies provide ?	To explore the categories of the works investigated.
Project perspective	RQ3	What are project management fields that benefit from machine learning ?	To shed light to the project management fields that benefit from machine learning.
	RQ4	Which is the type of the projects involved?	To gain knowledge about the type of project that benefit from ML.
ML Techniques perspective	RQ5	Which machine learning techniques have been researched ?	To identify the most commonly used machine learning techniques in project management.
	RQ6	What technologies are used to implement the ML techniques?	To discover the used technology for supporting the proposed solution/model
	RQ7	What are the data sources	Industrial (Real data word) or Academic (simulated data)

Table 2. Digital libraries and associated Queries.

Digital library	Query
ACM	(recordAbstract:("Machine learning" "deep learning" "neural networks") AND recordAbstract:("project management"))
IEEE Xplore	(("Abstract":“Machine learning” OR "Abstract":“deep learning” OR "Abstract":“neuronal networks”) AND "Abstract":"project management")
ScienceDirect	(("Machine learning" OR "deep learning" OR "neural networks") AND "project management")

Table 3. Inclusion and exclusion criteria.

Inclusion
IC1 - Papers published in English
IC2 - Paper published from 2010 to 2019
IC3 - Relevant title (papers deal with the use of ML techniques in PM) **IC4** - Relevant abstract (abstract explicitly mention a ML technique and PM) **IC5** - Journal articles, conference papers, Book chapters Ph.D. thesis
Exclusion
EC1 - Paper not written in English
EC2 - Papers published out of period
EC3 - Irrelevant title **EC4** - Irrelevant abstract **EC5** - Technical reports, Master thesis, and literature reviews

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 4. Primary studies

Reference	Title
[4]	Task Duration Estimation
[5]	Neural Network Models for Software Development Effort Estimation: A Comparative Study
[6]	A Hybrid Approach for Software Cost Estimation Using Polynomial Neural Networks and Intuitionistic Fuzzy Sets
[7]	Integrating Non-parametric Models with Linear Components for Producing Software Cost Estimations
[8]	I-Competere: Using Applied Intelligence in Search of Competency Gaps in Software Project Managers
[9]	A Survey on Software Fault Detection Based on Different Prediction Approaches
[10]	Evaluating Artificial Neural Networks and Traditional Approaches for Risk Analysis in Software Project Management
[11]	Software Effort Estimation Using Machine Learning Techniques
[12]	Software Effort Estimation in the Early Stages of the Software Life Cycle Using a Cascade Correlation Neural Network Model
[13]	Proposing an Enhanced Artificial Neural Network Prediction Model to Improve the Accuracy in Software Effort Estimation
[14]	Solving Problems of Project Management with a Self Enforcing Network (SEN)
[15]	GA-based Method for Feature Selection and Parameters Optimization for Machine Learning Regression Applied to Software Effort Estimation
[16]	Ensemble Missing Data Techniques for Software Effort Prediction
[17]	Evolutionary Fuzzy Hybrid Neural Network for Project Cash Flow Control
[18]	Intelligently Predict Project Effort by Reduced Models Based on Multiple Regressions and Genetic Algorithms with Neural Networks
[19]	Generalized Regression Neural Nets in Estimating the High-Tech Equipment Project Cost
[20]	Risk source-based constructability appraisal using supervised machine learning
[21]	An effective approach for software project effort and duration estimation with machine learning algorithms
[22]	A hybrid model for estimating software project effort from Use Case Points
[23]	Evaluation of construction projects based on the safe work behavior of co-employees through a neural network model
[24]	Work breakdown structure (WBS) development for underground construction
[25]	Optimized artificial intelligence models for predicting project award price
[26]	Integrating uncertainty in software effort estimation using Bootstrap based Neural Networks
[27]	Towards an early software estimation using log-linear regression and a multilayer perceptron model
[28]	Forecasting enterprise resource planning software effort using evolutionary support vector machine inference model
[29]	Competency-based selection and assignment of human resources to construction projects
[30]	A Neurogenetic approach for the resource-constrained project scheduling problem
[31]	Analysis and Evaluation of Project Cost Risk Based on BP Algorithm
[32]	An approach to software development effort estimation using machine learning
[33]	Least Square Support Vector Machine in Analogy-Based software development effort estimation
[34]	SETAP: Software engineering teamwork assessment and prediction using machine learning
[35]	A Neural Network Based Algorithms for Project Duration Prediction
[36]	Software defect prediction using supervised learning algorithm and unsupervised learning algorithm

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Secondly, we reformulate search strings based on keywords identified by applying PICO (Population, Intervention, Comparison and Outcomes) structure suggested in the paper [37] by Kitchenham and Charters :

- Population : In our context refer to project management as application area. Since we are interested in all project management fields we chose "project management" as synonym of this category
- Intervention : Are the machine learning techniques used. We select the following terms : "Machine learning", "deep learning" and "neural networks"
- Comparison : not applicable
- Outcomes : not applicable

The search was performed on the abstract for de ACM and IEEE Xplore library. For ScienceDirect, it was executed on the complete text. The first research launch, an initial set of 137 potential primary studies is obtained.

2.3 Relevant Papaers Selecton

To select the most relevant and appropriate papers to our topic, we have defined a set of inclusion and exclusion criteria summarized in table 3. These criteria was applied, by primary author, to candidate papers abstract in order to decide whether they should be selected or rejected from the study. In case of ambiguity, full text was analyzed and co-authors were involved in the review. Table 4 summarize the primary studies selected.

2.4 Classification Scheme

The classification scheme result is mainly created by keywording of abstracts activity. According to Petersen et al. [1], keywording is a tow-phase procedure that aim to reduce the time needed in constructing the scheme and guarantee the treatment of only relevant papers. During the first phase, abstracts of selected papers were examined to extract their representative concepts and keywords according to the RQs. Then, during the second, full papers text were analyzed to link their contents to the determined concepts.

2.5 Data extraction and Mapping Process

In this section, we answer the research questions for our mapping study based on the extracted information and classification scheme. But first, we summarize the relationship between the studies and the principals concepts of the classification scheme in Table 6.

2.5.1 RQ1 – What are the bibliographic features of the studies on the use of ML in PM

Articles were selected from 3 digital libraries. The Figure 2 shows the relationship between the papers candidate to selection (light turquoise) and those effectively selected (turquoise). 16 papers were taken from ACM, 12 from SienceDirect, and 5 from IEEE. 60.61% of these papers are journal articles and the remaining 39.39% are conference papers (Figure 3).

Regarding the trend of publication in the area being studied, Figure 4 shows the number of annual publications between 2010 and 2019. We observe that that number varies from year to yes, with a peak of publications in 2014, followed by a decrease until 2018 and a trend towards recovery in 2019.

In order to discover the most interesting works from the point of view of their citations, we were interested to the "number of citations". Figure 5 and Figure 6 shows, respectively, total and average annual citation numbers distributed over the years. With regard to total and average citations, top three papers are [15] , [27] and [30], respectively.

A last aspect that has also attracted our attention is the publication sources. All venues published only one paper, except "Automation in Construction" which published 3, [20], [24] and [25]. Considering these findings, we can claim that there is no specialized venue for publishing in the use of ML in PM.

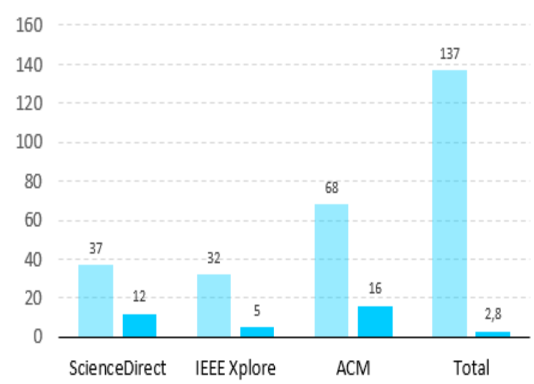

Figure 1. Systematic Mapping Study workflow.

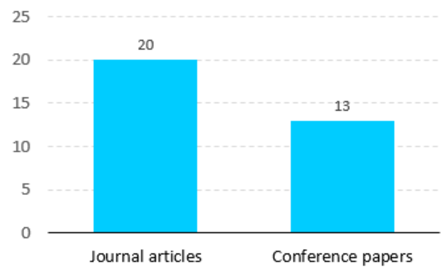

Figure 3: Publications types.

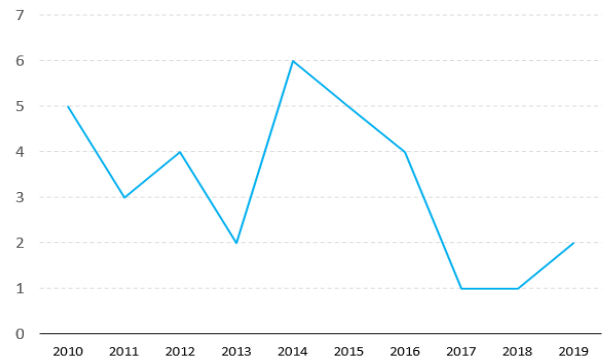

Figure 4: Publications per year.

Table 5. Classification scheme.

Id	Concepts	Description
RQ1	Bibliographic information	Identify the bibliographic features : • Publication Year : Years from 2010 to 2019 • Paper Type : Journal Article or Conference Paper • Source : Venue where the study was published • Number of citation
RQ2	Contribution type	Determine the class of intervention under study, which could be : • Analysis : It is a theoretical study that compare technique, analyze or discusses their uses • Model : The study is based on a set of procedures to obtain outcomes. • Process : The study is based on a series of steps to obtain outcomes. • Method : The study define a particular way or the rules to obtain outcomes. • Framework : The study is based on a framework. • Tool : The study is based on a tool.
RQ3	PM field	Determine the project management field that benefit from machine learning. Possible values are : Cost, Cash flow control, Cost risk, Defect detection, Duration, Effort, Risk, Staffing, scheduling or Work breakdown structure Another way to categorize project management fields is to use knowledge areas defined by Project Management Institute (PMI) in the Project Management Body of Knowledge (PMBOK) [38]. In that case, the potential values will be : Integration Management, Scope Management, Schedule Management, Cost Management, Quality Management, Resource Management, Communications Management, Risk Management, Procurement Management or Stakeholder Management
RQ4	Project type	Identify the type of project. Possible values are : Manufacturing Project, Construction Project, Software Project or Other
RQ5	ML technique	Discover witch ML technique is used by the study. Exhaustive values are illustrated by Figure 12
RQ6	Technology	Designates tools and programming languages used in the study. • Tools : Matlab, Neurosolustions, Minitab, WEEKA, DTREG, IBM SPSS Modeler or Palisade. • Programming languages : Java, R, Python, Perl, Visual Basic.
RQ7	Data sources	Determine the type of data set used to validate the study outcomes. It could be : • Academic : Simulated data set • Industrial : Real data set • Both : Combination of simulated and real data set

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 6. Study map.

	Project type			PM field (PMI)				Contribution type						ML Technique						Technology		Data source		
Ref	Software	Construction	Manufacturing	Cost	Quality	Resource	Risk	Analysis	Model	Process	Method	Framework	Tool	Neural Network	Multilayer Perception	Support Vector Machine	Genetic Algorithm	Decision Tree	Other	Tool	Programming Language	Academic	Industrial	Both
[4]	□	□	□		■			■											■		■		■	
[5]	■				■			■							■				■	□	□		■	
[6]	■			■							■								■	■			■	
[7]	■			■					■					■		■			■	□	□		■	
[8]	■						■						■		■	■			■	□	□			■
[9]	■					■		■						■				■	■	■			■	
[10]	■										■				■	■			■		■		■	
[11]	■				■						■			■					■	□	□	■		
[12]	■				■						■								■	■				■
[13]	■				■				■					■						□	□		■	
[14]	■			□	□	□	□						■	■						□	□		■	
[15]	■				■						■				■		■		■	□	□		■	
[16]	■				■						■							■		■			■	
[17]	■			■									■	■			■			□	□		■	
[18]	■				■						■			■			■		■	□	□		■	
[19]			■	■					■										■	□	□		■	
[20]		■										■				■					■		■	
[21]	■				■				■						■	■			■	■			■	
[22]	■				■				■							■			■	■				■
[23]		■							■					■						■		■		
[24]		■			■						■			■						□	□		■	
[25]		■		■						■				■			■			■			■	
[26]	■				■						■			■						□	□		■	
[27]	■				■				■						■					■		■		
[28]	■				■				■							■			■	■			■	
[29]		■					■		■					■						■		■		
[30]	□	□	□		■						■			■			■		■		■		■	
[31]		■						■						■						□	□	■		
[32]	■				■						■			■					■		■		■	
[33]	■				■						■								■	□	□		■	
[34]	■					■	■					■							■		■	■		
[35]	■				■			■						■	■				■	□	□		■	
[36]	■					■		■										■	■	□	□		■	

■ : presence of the concept in the study, □ : the concept is undefined in the study

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

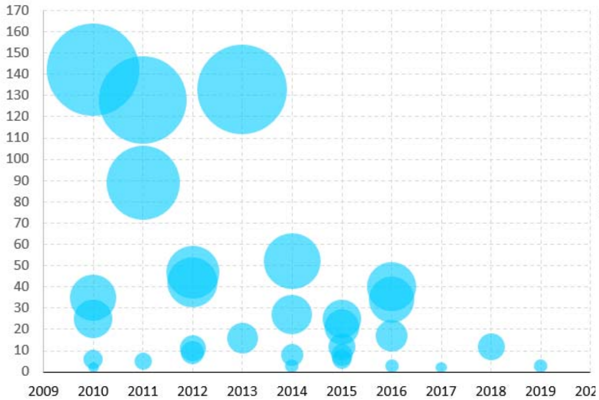

Figure 5: Total citation numbers per year

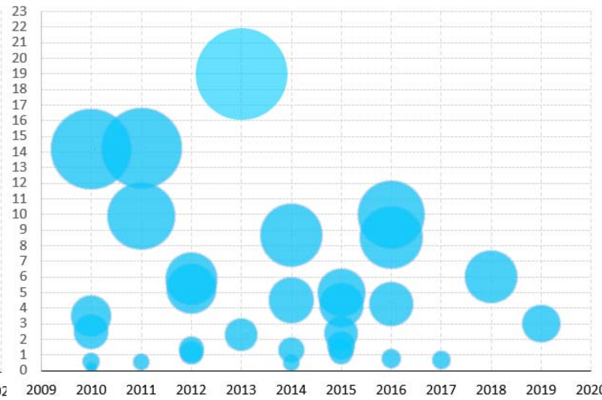

Figure 6: Average citation numbers per year

2.5.2 RQ2 – What contribution the studies provide ?

This question aim to discover the categories of the works investigated. Figure 7 shows that in the first place comes methods, which represent 36.36% of the total of the studies, then comes models with 27.27%, followed by Analysis, tools and Frameworks, with 181.8%, 9.09% and 6.06%, respectively. The last place is taken by processes, which represent 3.03% of the total of the studies.

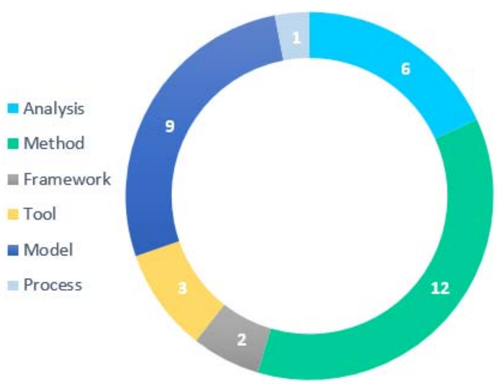

Figure 7: Average citation numbers per year.

2.5.3 RQ3 – What are project management fields that benefit from machine learning ?

We have classified the studies according to the project management field they investigate. The distribution of the fields is illustrated by Figure 8. As clarified by this figure, the predominant field is Effort with a presence of 41.81%, followed by Cost with 11.76%. Risk, Duration and Defect detection arise in the third position whit 8.82%. Then comes Staffing with 2 studies, with represent 5.88%. Finally, Work Breakdown Structure, Scheduling, Cash flow control and Cost risk stand in the last position with a 2.94% of selected studies. It is to be noted that, for a single study [14], the researchers have discussed the use of ML in PM without limiting themselves to a particular field.

Figure 9 shows the fields founded form the perspective of PMI [38]. At the top of the list we find Schedule Management (54.55%), then Cost management (15.15%), then Risk management (12.12%), and then Quality management and Resource management with 9.09% each.

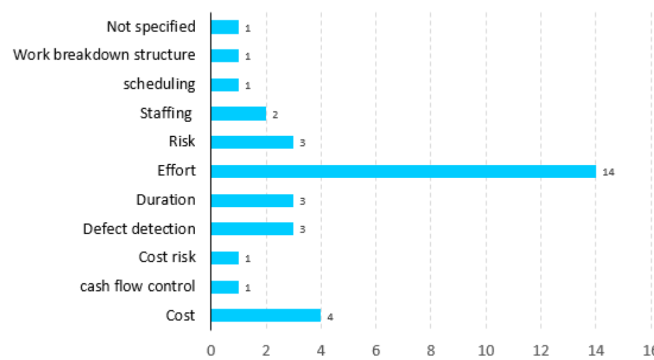

Figure 8: Project management fields.

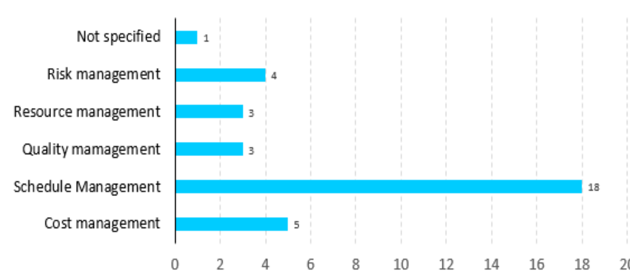

Figure 9: Project management knowledge areas

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2.5.4 RQ4 – Which is the type of the projects involved ?

Figure 10 seek at answering this question. It report our findings on project types that have used ML techniques in their work management. Three types were discovered, Software project represents 72.73% of the selected studies, Construction project represents 18.18% , and Manufacturing project represents 3.03%. For 2 of 33 papers, studies have been carried out independently of the project type.

Regarding the previous question (RQ3), Figure 11 maps project types and their fields . Three remarks are to be made. First, different fields were explored according the project type. Second, from a types perspective, Software project exploits the majority of domains. Third, form a fields perspective, cost is the most involved.

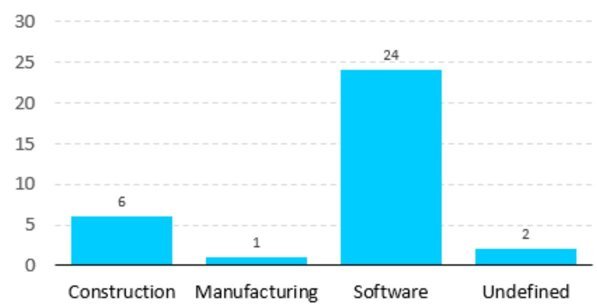

Figure 10: Project types.

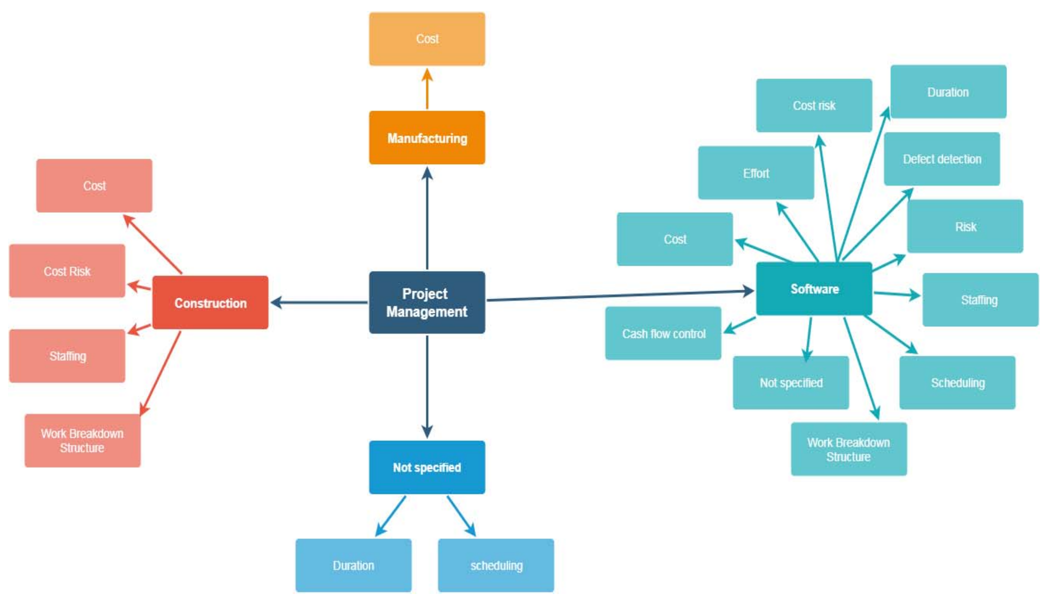

Figure 11: Project types.

2.5.5 RQ5 – Which machine learning techniques have been researched ?

This question deals with the machine learning algorithms researched. It aims at identifying them, their categories, and discovering which among them are most frequently used. As it can be seen from Figure 12, 27 machine learning algorithms were inventoried. One algorithm stands out for its use : Neural Network to classify according to probabilities with a 22,54 % . Others have also to be listed : Multilayer Perception (9.86%), Support Vector Machine (9.86%), Genetic Algorithm (7.04%), Decision Tree (5.63%), Random Forest (4.23%), Naive Bayesian (4.23%). Remaining techniques were sited less than three times.

We have divided machine learning algorithms into categories according to their purpose. Result is shown by Figure 13. It is important to see that the main classes of ML algorithm are presented and that the category the most involved by the studies is Neural and Deep Learning algorithms.

On the other hand, we found it useful to map the technical ML techniques and project fields where they were used. the result is shown in Figure 14.

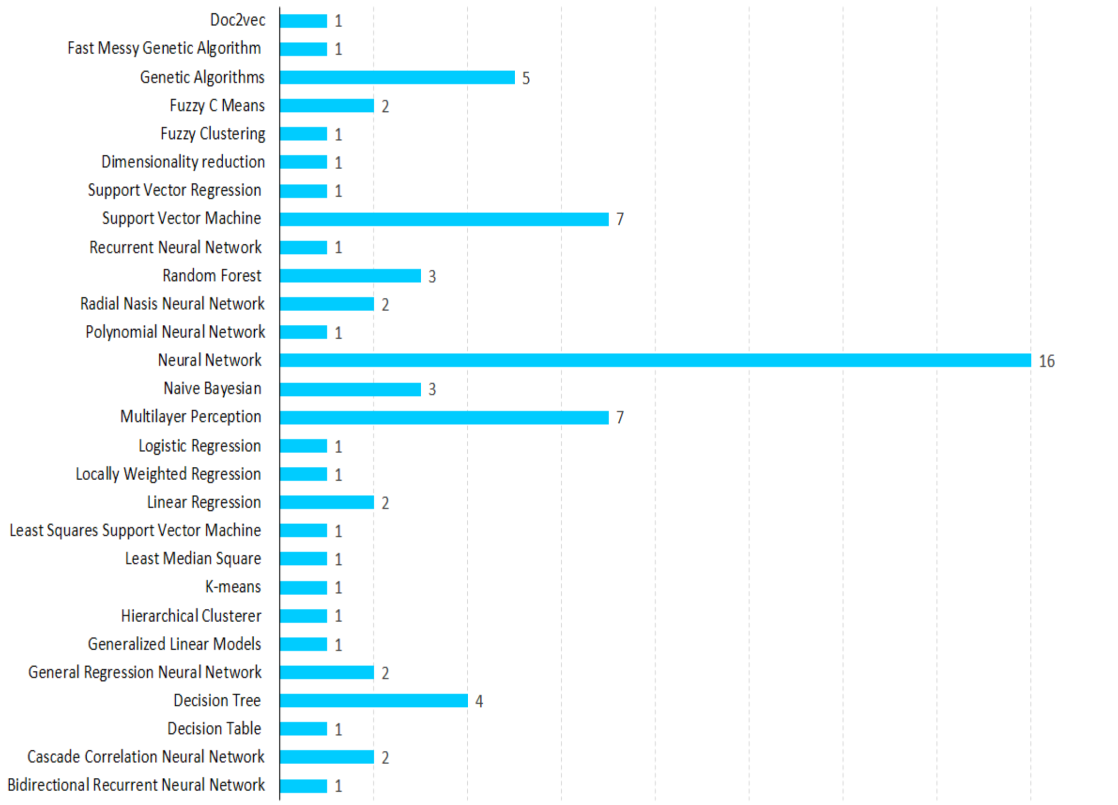

Figure 12: Machine learning techniques.

2.5.6 RQ6 – What technologies are used to implement the ML techniques ?

The number of tools and programming languages involved by the studies are provided in Figure 15. Unfortunately, about half papers (48.48%) didn't report the technologies they used to implement or evaluate their solution.

Figure 16 and Figure 17 shows respectively that, Matlab is the predominant tool and python is the favorite programming languages.

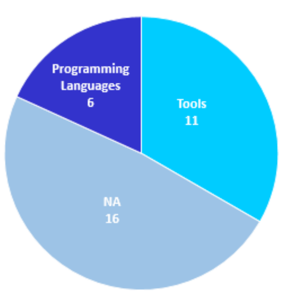

Figure 15: Technologies.

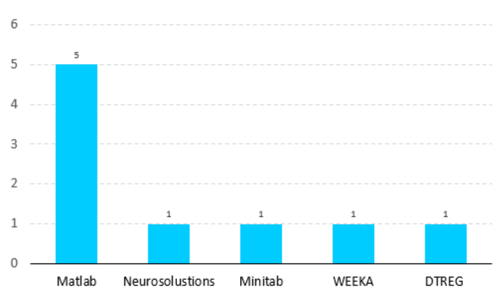

Figure 16: Tools.

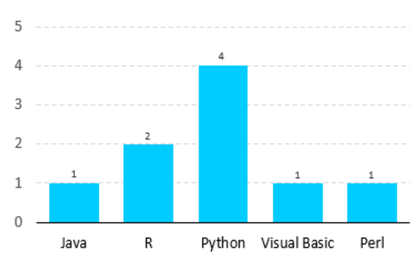

Figure 17: Programming languages.

Journal of Theoretical and Applied Information Technology
31st May 2020. Vol.98. No 10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

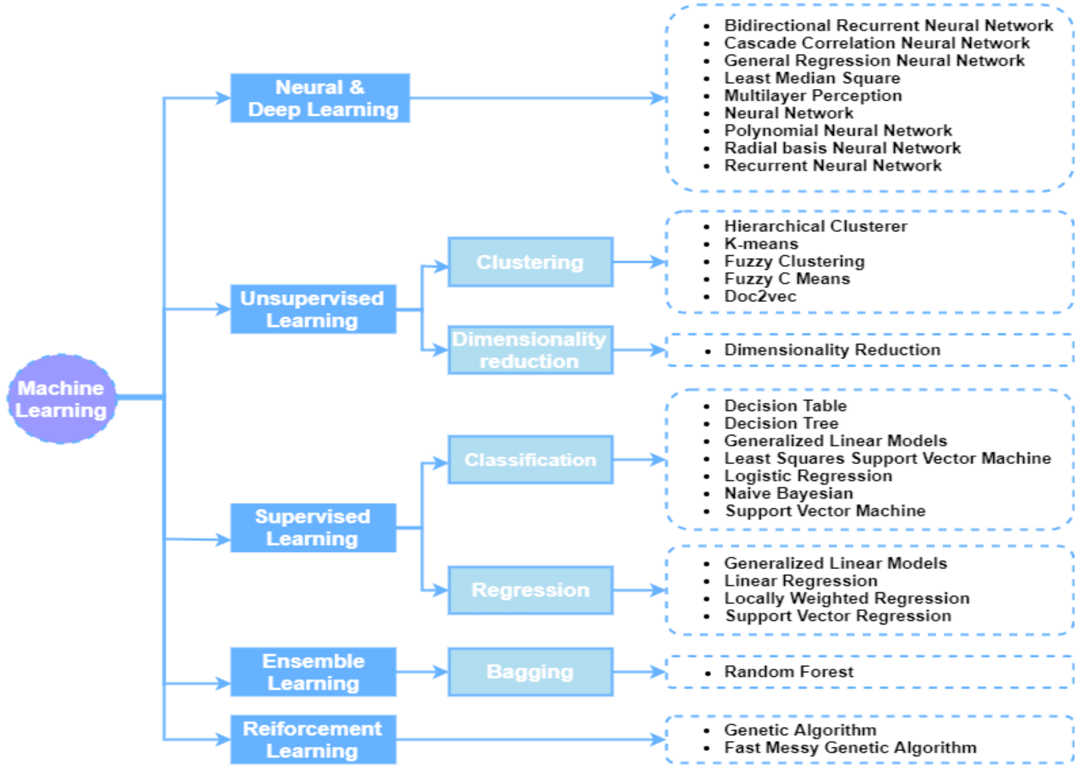

Figure 13: Types of ML algorithms.

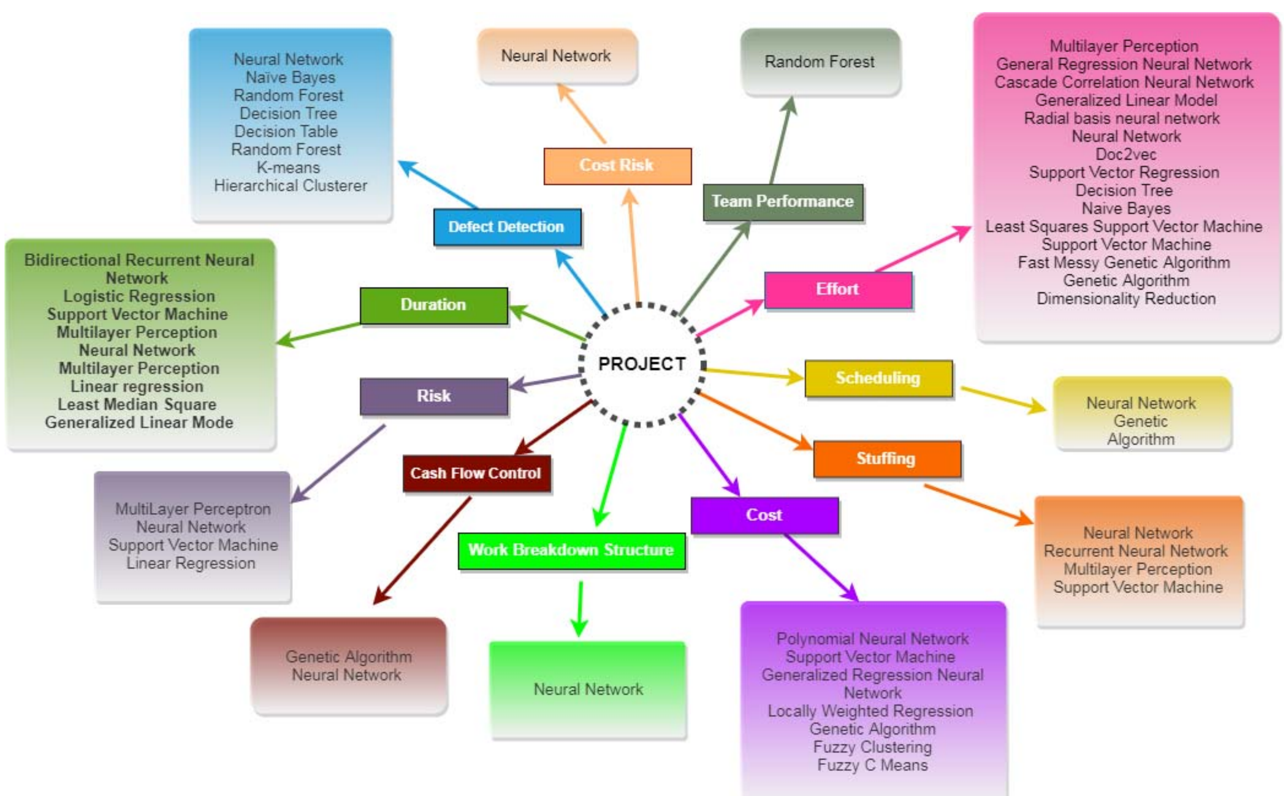

Figure 14: ML algorithms vs Project types.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2.5.7 RQ6 – What are the data sources ?

This question seek to discover the nature of dataset used to validate the studies outcomes. Is it real dataset, simulated dataset or a mixt of both. Figure 18 shows the number of articles found in each category. As we can see, most studies (75,76%) use Industrial dataset in their approach.

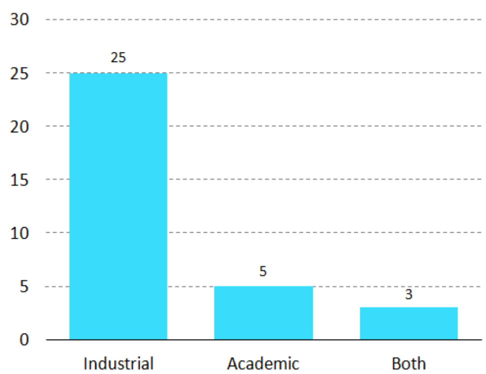

Figure 18: Data sources.

3. THREATS TO VALIDITY

The validity of our study depends on its reliability and the correctness of the outcomes. It can be affected by several threats. Below the said threats are organized in three categories:

The research process :
This group considers the relationship between the research conducted and its results. In our case, we believe that the research process itself is a threat to validity. Indeed, a different choice of digital libraries and/or a change in the selection criteria can probably lead to another initial study and therefore to other results.

The research string :
The research string category's concerns the ability of the papers included in the study to reflect initial expectations of the researchers. These papers are mainly identified by the research string. To reduce the threats related to the selected search string, we applied PICO structure defined by Kitchenham and Charters in [37].

The research executor:
This threat class refers to potential errors of judgment made by the person in charge of analyzing the articles under study. To minimize their impact, the second co-author reviews articles for which there was a mitigated assessment.

4. CONCLUSION

This paper presents the results of a systematic mapping study concerning the use of ML techniques in PM by investigating the scientific publications on this topic between 2010 and 2019. Starting with 137 articles we finish with a primary study of 33 papers. This one, after having been analyzed, allowed us to make the following deductions : (1) Most studies are performed as method; (2) About three-quarters of studies relate to software projects; (3) Effort is the project field the most researched by ML techniques; (4) most of the works uses neural networks; (5) About half papers didn't report the technologies they used to implement or evaluate their solution; (6) most studies uses industrial dataset to train and evaluate their machine learning models.

It is important to point out that, according to our finding (Figure 11 and Figure 15) and from a PMI perspective, there are several project knowledge areas where we haven't yet studied the use ML : stakeholder management, communication management, procurement management and integration management.

Furthermore, projects being temporary endeavors to create a unique product or service whatever their types [38], it will be beneficial for future studies, when addressing this subject , to not be restricted to a particular project type and to address the subject as a whole

REFRENCES:

[1] K. Petersen, R. Feldt, S. Mujtaba, and M. Mattsson, "Systematic mapping studies in software engineering," Proceedings of the12th International Conference on Evaluation and Assessment in Software Engineering, vol. 17, pp. 68-77, 2008.

[2] https://www.gartner.com/en/newsroom/press-releases/2019-03-20-gartner-says-80-percent-of-today-s-project-management, 23/08/2019

[3] I. H. Witten, and E. Frank, Mark A. Hall, and Christopher J. Pal (2016)," Data Mining: Practical Machine Learning Tools and Techniques," Morgan Kaufmann Series in Data Management Systems, 2005.

[4] R. W. White, and A. H. Awadallah, "Task Duration Estimation," WSDM (International Conference on Web Search and Data Mining), pp. 636-644, 2019.

[5] A. B. Nassif, M. Azzeh, L. F. Capretz, and D. Ho,"Neural network models for software

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

development effort estimation: a comparative study," Neural Computing and Applications, vol. 27, pp. 2369-2381, 2016.

[6] A. Kaushik, A. K. Soni, and R. Soni,"A Hybrid Approach for Software Cost Estimation Using Polynomial Neural Networks and Intuitionistic Fuzzy Sets," International Journal of Computer Applications in Technology, vol. 52, pp. 292-304, 2015.

[7] N. Mittas, E. Papatheocharous, L. Angelis, and A. S. Andreou,"Integrating Non-parametric Models with Linear Components for Producing Software Cost Estimations," Journal of Systems and Software, vol. 99, pp. 120-134, 2015.

[8] R. Colomo-Palacios, I. Gonzalez-Carrasco, J. L. López-Cuadrado, A. Trigo, and J. E. Varajao, "I-Competere: Using Applied Intelligence in Search of Competency Gaps in Software Project Managers," Information Systems Frontiers, vol. 16, pp. 607-625, 2014.

[9] G. Abaei, and A. Selamat,"A survey on software fault detection based on different prediction approaches," Vietnam Journal of Computer Science, vol. 1, pp. 79-95, 2014.

[10] C. H. M. S. Timoteo, M. J. S. Valenca, and S. M. M. Fernandes,"Evaluating Artificial Neural Networks and Traditional Approaches for Risk Analysis in Software Project Management," International Conference on Enterprise Information Systems (ICEIS), vol. 3, pp 472-479, 2014.

[11] Monika, and O. P. Sangwan,"Software Effort Estimation Using Machine Learning Techniques," 7th International Conference on Cloud Computing, Data Science & Engineering – Confluence, pp. 92-98, 2017.

[12] A. B. Nassif, L. F. Capretz, and D. Ho, "Software Effort Estimation in the Early Stages of the Software Life Cycle Using a Cascade Correlation Neural Network Model," 13th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, pp. 589-594, 2012.

[13] I. Attarzadeh, A. Mehranzadeh, and A. Barati,"Proposing an Enhanced Artificial Neural Network Prediction Model to Improve the Accuracy in Software Effort Estimation," Fourth International Conference on Computational Intelligence, Communication Systems and Networks, pp. 167-172, 2012.

[14] C. Klüver, "Solving Problems of Project Management with a Self Enforcing Network (SEN),"Computational and Mathematical Organization Theory, vol. 18, pp. 175-192, 2012.

[15] A. L. I. Oliveira, P. L. Braga, R. M.F. Lima, and M. L. Cornélio,"GA-based Method for Feature Selection and Parameters Optimization for Machine Learning Regression Applied to Software Effort Estimation," Information and Software Technology, vol. 52, pp. 1155-1166, 2010.

[16] B. Twala, M. Cartwright,"Ensemble Missing Data Techniques for Software Effort Prediction," Intelligent Data Analysis, vol. 14, pp. 299-331, 2010.

[17] M-. Y. Cheng, H-. C. Tsai, E. Sudjono,"Evolutionary Fuzzy Hybrid Neural Network for Project Cash Flow Control,"Engineering Applications of Artificial Intelligence, vol. 23, pp. 604-613, 2010.

[18] Z. Li,"Intelligently Predict Project Effort by Reduced Models Based on Multiple Regressions and Genetic Algorithms with Neural Networks," International Conference on E-Business and E-Government, pp. 1536-1542, 2010.

[19] J-. S. Chou, and Y. Tai,"Generalized Regression Neural Nets in Estimating the High-Tech Equipment Project Cost," Second International Conference on Computer Engineering and Applications, pp. 281-284, 2010.

[20] D. Kifokeris, and Y. Xenidis,"Risk source-based constructability appraisal using supervised machine learning," Automation in Construction, vol. 104, pp. 341-359, 2019.

[21] P. Pospieszny, B. C-. Chrobot, and A. Kobylinski,"An effective approach for software project effort and duration estimation with machine learning algorithms," Journal of Systems and Software, vol. 137, pp. 184-196, 2018.

[22] M. Azzeh, and A. B. Nassif,"A hybrid model for estimating software project effort from Use Case Points," Applied Soft Computing, vol. 49, pp. 981-989, 2016.

[23] D. A. Patel, and K. N. Jha,"Evaluation of construction projects based on the safe work behavior of co-employees through a neural network model," Safety Science, vol. 89, pp. 240-248, 2016.

[24] E. S-. Irdemoosa, S. R. Dindarloo, and M. Sharifzadeh,"Work breakdown structure (WBS) development for underground construction," Automation in Construction, vol. 58, pp. 85-94, 2015.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[25] J-. S. Chou, C-. W. Lin, A-.D. Pham, and J-. Y. Shao,"Optimized artificial intelligence models for predicting project award price," Automation in Construction, vol. 54, pp. 106-115, 2015.

[26] S. Laqrichi, F. Marmier, D. Gourc, and J. Nevoux,"Integrating uncertainty in software effort estimation using Bootstrap based Neural Networks," IFAC-PapersOnLine, vol. 48, pp. 954-959, 2015.

[27] A. B. Nassif, D. Ho, and L. F. Capretz,"Towards an early software estimation using log-linear regression and a multilayer perceptron model," The Journal of Systems and Software, vol. 86, pp. 144-160, 2013.

[28] J-. S. Chou, M-. Y. Cheng, Y-. W. Wu, and C-. C. Wu," Forecasting enterprise resource planning software effort using evolutionary support vector machine inference model," International Journal of Project Management, vol. 30, pp. 967-977, 2012.

[29] V. Shahhosseini, and M. H. Sebt,"Competency-based selection and assignment of human resources to construction projects,"Scientia Iranica, vol. 18, pp. 163-180, 2011.

[30] A. Agarwal, S. Colak, and S. Erenguc,"A Neurogenetic approach for the resource-constrained project scheduling problem," Computers and Operations Research, vol. 38, pp. 44-50, 2011.

[31] B. Zhu, H. Zhang, and X. Wang," Analysis and Evaluation of Project Cost Risk Based on BP Algorithm," Systems Engineering Procedia, vol. 1, pp. 264-270, 2011.

[32] V-. S. Ionescu," An approach to software development effort estimation using machine learning," Intelligent Computer Communication and Processing (ICCP), pp. 197-203, 2017.

[33] T. R. Benala, and R. Bandarupalli,"Least Square Support Vector Machine in Analogy-Based software development effort estimation," IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE), 2016.

[34] D. Petkovic, M. Sosnick-Pérez, S. Huang, R. Todtenhoefer, K. Okada, S. Arora, R. Sreenivasen, L. Flores, and S. Dubey,"SETAP: Software engineering teamwork assessment and prediction using machine learning," IEEE Frontiers in Education (FIE) Conference, 2014.

[35] W. Han, H. Jiang , X. Zhang, and W. Li," Neural Network Based Algorithms for Project Duration Prediction," 7th International Conference on Control and Automation, pp. 60-63, 2014.

[36] A. Chug, and S. Dhall," Software defect prediction using supervised learning algorithm and unsupervised learning algorithm," The Next Generation Information Technology Summit (4th International Conference), pp. 173-179, 2014.

[37] B. Kitchenham, and S. Charters," Guidelines for Performing Systematic Literature Reviews in Software Engineering," Technical report, EBSE, 2007.

[38] PMI. 2017. A Guide to the project management body of knowledge: (PMBOK® Guide), 6th edition