Agile Project Status Prediction Using Interpretable Machine Learning

Ali Akbar ForouzeshNejad
School of Computing, University of Portsmouth
Portsmouth, United Kingdom
ali.forouzesh@port.ac.uk

Farzad Arabikhan
School of Computing, University of Portsmouth
Portsmouth, United Kingdom
Farzad.Arabikhan@port.ac.uk

Nigel Williams
Faculty of Business and Law, University of Portsmouth
Portsmouth, United Kingdom
nigel.williams@port.ac.uk

Alexander Gegov
School of Computing, University of Portsmouth
Portsmouth, United Kingdom
alexander.gegov@port.ac.uk

Omer Sari
School of Computing, University of Portsmouth
Portsmouth, United Kingdom
Omer.Sari@port.ac.uk

Mohamed Bader
School of Computing, University of Portsmouth
Portsmouth, United Kingdom
mohamed.bader@port.ac.uk

***Abstract*- Monitoring and forecasting the progress of information technology projects stands as a significant challenge in project management. Over the past two decades, agile project management has become a crucial factor influencing project success. Despite this, existing research has not presented a comprehensive model capable of predicting project outcomes based on agile features. In light of this, this study aims to develop a predictive model for information technology project outcomes using agility metrics. The results indicate that metrics related to teamwork and the team's capabilities, along with their collective experience, have the most significant impact on project success. The study employs the Decision Tree as an interpretable model to establish rules and predict project success. The accuracy of the model designed in this study is an impressive 97%, surpassing the accuracy of SVM at 71% and KNN at 82%.**

Keywords- Interpretable Machine Learning; Decision Tree; Agile Project Management; Project Status Prediction

I. Introduction

Agile project management emphasizes flexibility and changeability, facilitating cooperation and communication, providing timely and reliable delivery, and valuing continuous customer interaction. An agile project management approach can also help improve project oversight. This approach allows project managers to continuously review and evaluate the status and progress of the project [1].The agile project approach reduces planning time by dividing projects into short periods called sprints, allowing for performance checks and rapid feedback. This helps identify issues early and predict project success. However, project managers' intuitive monitoring and evaluation can lead to errors[2] .

Further, it is difficult to monitor a large number of projects and human resources at the same time. One of the basic challenges for managers during implementation is to predict the status of projects at the end, through which they can stop failed projects as soon as possible and redeploy the organisation's resources to focus on projects with a higher probability of success.

In this regard, this article formulates a data-driven model based on machine learning algorithms to predict the status of projects at the time of completion. Research has suggested that machine learning algorithms can be used to monitor the status of projects and predict their future status based on agility features and defined principles [3]. Machine learning algorithms can be employed to forecast project timelines and potential delays. These methods analyse available data, uncovering patterns and relationships within various project characteristics and factors. Subsequently, they leverage these insights to predict project timelines and delays.[3] . The initial step in predicting project timelines involves gathering data on past projects, encompassing their specific attributes and completion statuses. Machine learning algorithms are then utilised to construct a model that can predict project completion status.

Critical agile features are first identified in predicting the status of projects in the telecom industry. Then, a model is designed to predict the status of the projects according to the collected data. Overall, compared to other past studies, the main innovations and advantages of this work are that the first interpretable machine learning model is presented to predict the status of projects according to the agility criteria in the telecom industry. In the continuation of this paper, in Section 2, the literature review I discussed, and then the research gap is analyzed in detail. In Section 3, the solution method is described, and in Section 4, the findings are reported. In Section 5, the conclusions are explained.

II. Literature Review

In this section, a literature review is reported, focusing on the sections of Agile Project Status, which is the main concept of this project, as well as Interpretable Machine Learning in Project Management.

A. Agile Project Status

In the field of evaluating and predicting the success of projects, studies have been done with different approaches and different industries. For example, Zhang et al.[4] conducted a study to predict construction project success using a hybrid evolutionary support vector machine model. Their study proposes using an evolutionary support vector machine (ESIM) inference model to predict project success dynamically.

Sharma & Chanda [5] presented a Bayesian network model to predict the probability of success of research and development projects. In their study, it is stated that the success of projects is based on risk management abilities. In this regard, in their study, a quantitative model based on the Bayesian network has been designed to estimate and predict the probability of project success by considering the desired risks.

Lue et al.[6] have investigated the relationship between project complexity and the success of construction projects in a study. In this study, it is assumed that project complexity has a negative relationship with project success. Second, 245 survey questionnaires on project complexity and project outcomes in China were collected based on the literature review and expert interviews. The project complexity was measured using correlation and factor analysis as information, task, technology, organisational, environmental, and goal complexities. Finally, the structural equation modelling technique was used to test the hypothesis and to investigate the effect of different complexities on the project's success. The findings of this study support the hypothesized negative relationship between complexity and success of complex construction projects.

In their study, Yu et al.[7] employed a deep learning technique to forecast the outcome of crowdfunding projects. To create the model they aimed for, they retrospectively acquired a dataset from Kaggle, encompassing past records of Kickstarter campaigns. The proposed MLP model can provide significant results when applied to various crowdfunding platforms that have not been addressed before. The components considered for building the model include the project goal, the number of supporters, the duration of the project, the classification, and the project's location. The findings show that their article's MLP model provided the best result with the highest degree of confidence compared to other models in this field.Also, Xia et al. [8] presented a hybrid model based on a Bayesian network to predict safety performance in construction projects. Predictable risks are defined in five levels based on skill errors, decision-making, violations, unfavorable physical environment, etc.

Yeh & Chen [9] have presented a machine-learning model to predict the success of crowdfunding projects. Their study's primary concern and challenge is finding a model to predict crowdfunding projects before the project starts. In this regard, various features that affect the success of projects are identified, and the success rate of projects is predicted using the artificial neural network (ANN) algorithm. The findings show that the developed neural network model of the research had the best performance compared to logistic regression and linear regression.

Dumitrașcu-Băldău et al. [10] have also presented a model for predicting factors affecting the success of international projects. The main concern raised for their study is the difficulty of implementing international projects. In this regard, the factors affecting the international projects have been identified, and using the neural network algorithm and IBM SPSS Modeler software, the importance of the factors affecting the success of the international projects has been identified, and the probability of success has been estimated.

Hnatchuk et al.[11] Provided a prediction method to evaluate the success of project implementation. A neural network algorithm has been used to design the prediction model. To predict the project, components based on project requirements such as human resources, capital, project location, etc., have been used, and based on the designed algorithm, the success of the project has been estimated in terms of cost, time, quality, and complexity of the project. The findings show that the designed model has a better execution performance due to estimating project complexity and other success components. Princes & Said[12] investigated the impact of project complexity along with the components of leadership integrity, performance readiness, and management sustainability and financial success. The findings show that all components impact financial stability except performance readiness, which does not positively affect financial stability.

B. Whitebox Machine Learning

Moving towards knowledge driven and interpretable modelling and white box approach is now valued significantly more by the field experts. This trend leads to a better understanding of the decision-making process. Project management, by nature, is involved with decision making and an attempt for transparency in decision process could improve the quality of decisions made. Verenich et al.[13] evaluated the performance of white-box machine learning in the prediction process. They implemented the WBML approach on data from several real projects compared to black box approaches. Their results show that the WBML approach can accurately predict performance features comparable to Black Box approaches. It also provides more detailed and precise insights into the factors affecting performance. Velez et al. [14] conducted a study aimed at analyzing Interpretable Machine Learning. Their study discussed challenges related to Interpretable modeling, stating that one of the main challenges is that the performance of configurable systems depends on various system settings, which can complicate the modeling of these systems' performance. Then, Interpretable approaches were evaluated, and the results showed that this approach can accurately model the performance of configurable systems.

Considering the explanations and literature review of the Interpretable ML field, it is observed that the use of these approaches in the development of machine learning models provides better insights to audiences and decision-makers and also gives developers the ability to increase the accuracy of models.

Subramanian et al.[15] used the Interpretable ML approach for identifying processes and governing equations of the dynamics in manufacturing systems. Data from specific projects were used to evaluate both models. Their findings indicate that the Interpretable approach offers greater transparency and interpretability, and the model's accuracy can be improved by better parameter tuning. Also, the validation of Interpretable approaches is more robust and their study shows the high capability of WBMLs in identifying Research Gap and Contribution

The previous sections briefly presented the studies related to this research topic. First, the articles will be briefly described in a table format to examine the research gap. Then, based on them, innovations and research gaps will be discussed. For examining the papers, four project success features below were selected.

Efficiency, effectiveness, and endorsement play crucial roles in project success. Efficiency refers to accomplishing tasks with minimal resource consumption, while effectiveness is achieving desired outcomes or objectives and Endorsement involves the support and approval of stakeholders, which can lead to a project's acceptance and ultimate success by realization of planned and emergent benefits. In addition, project complexity, defined as the presence of numerous diverse, interconnected elements, can be measured by considering differentiation and interdependency, and this definition can be applied to various aspects of project management[16]. Table 1 shows a summary of the literature review.

As can be seen in Table 1, studies related to predicting the probability of success of projects are often in the construction industry, so a study that is presented with a case study of the telecom industry focusing on the development of an agile approach is rarely found. Furthermore, the algorithms used are not white-box; therefore, the structure and rules of checking the status of projects in their studies are unclear. According to the explanations given, the contributions of the current research are as follows:

- This study is the first research that deals with the design of a project success prediction model based on agility features.
- This study is the first that provides a complete structure based on the components of Endurance, Effectiveness, and Efficiency in predicting the success status of projects.
- This study is the first work that considers the components of complexity affecting the success of projects.
- This study is the first model that predicts the success of projects in the telecom industry.
- This study uses white-box algorithms, which clearly express the ability to analyse prediction rules.

III. Method

The method used in this study is the decision tree algorithm. A decision tree algorithm, as an interpretable model, is a hierarchical decision-making method that provides decisions for prediction or classification by using tree structure and branching of data features.

The decision tree algorithm is known for its interpretability in machine learning because its rules and decisions are understandable and justifiable[17] . It constructs the model step-by-step, transparently selecting the best feature to split at each step, and continues this process until a stopping condition is met. The resulting decision tree is a hierarchy of nodes and branches, with each node representing a decision based on an attribute. The algorithm directs data through these nodes to reach the final result. The decision tree's advantages include high interpretability, flexibility, and the ability to work with various types of data [18] .

IV. Result

This section defines the features used to evaluate the success rate of projects based on agile project management criteria. It then explains the relationships and importance of various variables in predicting success. Finally, a prediction model is created using the decision tree algorithm.

A. Features for predicting the success status of projects.

The features for evaluating and predicting the status of the success of the projects, which were identified based on the principles of agile project management, are categorized into four categories: Efficiency (Table 2), Effectiveness (Table 3), and Endurance (Table 4), and also Complexity (Table 5).

B. Analysis of features and their importance

Data related to 360 projects in the telecom industry have been used to build the project status prediction model. The data could be accessed from the Organization Database. These projects have different values of each indicator and are in different stages of project progress, based on which a comprehensive model can be designed.

An examination of the correlations among features has been conducted, with the outcomes displayed in Fig.1., the stronger the relationship between the two features. This Correlation helps to recognize the relationship between. For example, it can be seen from Fig.1 that team size has stronger relationships with story point completed, cycle time with BAC, and lead time with length of sprint. features and shows which features must be increased to enhance project success.

One of the other important points is to identify the impact of features on the goal of the model, which is the level of project success. In this regard, the following results, Fig. 2, have been obtained by checking the importance of features in the designed algorithm.

TABLE: SUMMARY OF THE LITERATURE REVIEW

Case study	Method	Features				Aim	Authors
		Complexity	*Efficiency*	*Effectiveness*	*Endurance*		
Commercial construction projects	The combined method of support vector machine (SVM) and principal component analysis (PCA)		*			Project cost performance forecasting	Son et al.,[19]
Construction projects	Support vector machine and genetic algorithm		*			Predicting project success	Cheng et al., [20]
Financing projects	Regression algorithms		*			Predicting project success	Li et al., [21]
Research and development projects	Bayesian networks	*	*			Predicting the probability of project success	Sharma & Chanda, [22]
Construction projects	Structural equations and questionnaires	*				Examining the relationship between project complexity and Project Success	Luo et al., [23]
Construction projects	Human Factors Engineering and Biz Network			*		Project safety performance prediction	Xia et al., [24]
Crowdfunding projects	Neural network and decision tree and random forest				*	Predicting project success	Yu et al., [25]
Crowdfunding projects	Artificial neural network			*		Predicting the success of a crowdfunding project	Yeh & Chen, [26]
International IT projects	neural network		*			Estimating the importance of the factors affecting the success of international projects	Dumitrașcu-Băldău et al, [10]
IT projects	neural network		*	*		Predicting the characteristics and evaluating the success of project implementation	Hnatchuk et al., [26]
Indonesian Project Management Institute	Structural equations and questionnaires	*				The effect of project complexity, leadership integrity, performance readiness, and management stability on financial stability	Princes & Said, [12]
Telecom industry	Interpretable machine learning (Decision Tree)	*	*	*	*	Predicting the success status of projects considering complexity and agility criteria	This study

The determination of a feature's significance in Decision Tree models hinges on the degree to which each feature diminishes impurity within the decision trees. This measure is derived from Equation 2, where Importance(Xi) denotes the significance of feature X_i within tree T.

$$Feature\ Imortance\ (Xi) = \frac{1}{T}\sum_{1}^{T} Imrotnace\ \ (Xi) \qquad (1)$$

It can be seen that Sprint Goals Achieve is the most important indicator of the success of projects in the telecom industry. After that, the Total Sprint Goals index has a high impact on the success of projects. The noteworthy point is that the better the defined Sprint Goals are realized in a project, the more likely the project will be successful in the specified time. Another noteworthy point is that the quality of the work and operational team is also very important. Team Expertise Score and Team Moral features are also very important features. In other words, the team's experience has a greater impact on the components of cost and time in the project's success with an agile approach. Another very important indicator in determining the success level of projects is Velocity, which shows the capacity and capability of the team.

C. *Project status prediction model with DT*

By analyzing the features and collecting data, the forecasting model of the project status should be designed. The target column in this model is the status of the project in three levels: "failed," "delayed," and "success." Projects that are completely out of order and are no longer followed up in the organization are labeled "failed," projects that have always been running but are delayed are labeled "Delayed," and projects that are completed on time are labeled "Success."

According to these components, a decision tree model was designed according to the 360 data related to the projects with labels. 70% of the data is considered for building the training model, and 30% is included for testing the model. The built model was designed in different nodes and depths, and the best result was the tree with a depth of 8 levels. Fig.3 shows the decision tree.

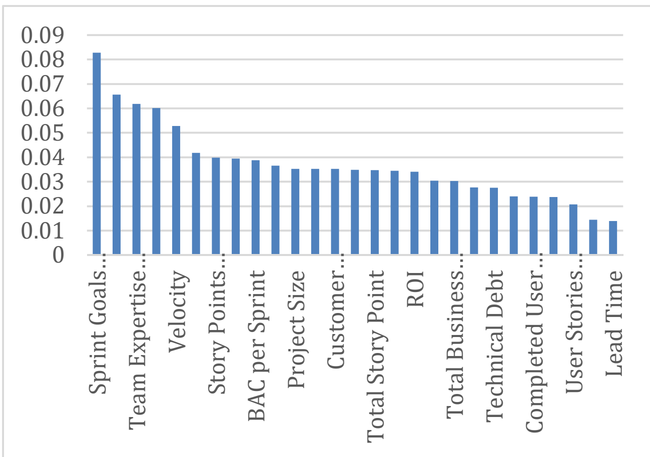

Figure.2. Importance of features in determining the status of the project

TABLE 2. EFFICIENCY CRITERIA

Criteria	**Description**
Team Moral	The development team's collective spirit, motivation, and job satisfaction level during a sprint.
Customer Satisfaction Score (CSAT)	Agile measures the degree of customer satisfaction with the deliverables produced in a sprint
Team Size	the total number of individuals actively involved in working on tasks during a particular sprint
Team Expertise Score (TES)	Team Expertise Score (TES) in Agile quantifies the development team's collective technical skill and experience level in a sprint.
Technical Debt	the implied cost of additional rework caused by choosing quick but suboptimal solutions during a sprint
Time to Market (TTM) day	is the period from the project's inception to its delivery to the end-users.

TABLE 3. EFFECTIVENESS CRITERIA

Criteria	**Description**
Lead Time	Lead Time is the duration from a task's introduction to completion within the Agile workflow.
Cycle Time	Cycle Time measures the period a task takes from start to completion
Length of Sprint (day)	the fixed duration, typically 1-4 weeks, during which specific work has to be completed and made ready for review
Sprint Number	the sequential order of a sprint within the project timeline, indicating its position in the workflow
Number of User Story	the count of individual functionalities or features to be developed within a sprint
Completed User stories	the number of user stories fully developed, tested and accepted within a sprint
BAC per Sprint	the total budget allocated for the completion of a specific sprint's tasks
Cost performance index (CPI)	a measure that calculates the efficiency of the budget utilization during a sprint
Schedule Performance Index (SPI)	the efficiency of time utilization relative to the planned schedule in a sprint
Total Story Point	the sum of the estimated effort required to complete all user stories within a sprint
Story Points Completed	the total estimated effort for all user stories successfully finished in a sprint.
Number of bugs reported	the count of defects or issues identified during a sprint's development and testing phase

TABLE 4. ENDORESMENT CRITERIA

Criteria	**Description**
Total Sprint Goals	The comprehensive list of objectives that a team aims to achieve during a specific sprint
Sprint Goals Archive	The team successfully met the number of set objectives during a particular sprint.
Total Business Value	the overall value that all completed user stories bring to the Business during a sprint
Business Value Delivered	the value the completed user stories contribute to the Business in a sprint.
Velocity	the total amount of work a team can complete in a single sprint.
Backlog Management Index (BMI)	a metric that assesses the efficiency of backlog handling and issue resolution in a sprint
ROI	a measure of the financial value gained from the development effort expended on a sprint.

TABLE 5. COMPLIXITY CRITERIA

Criteria	**Description**
Technical Debt	the implied cost of additional rework caused by choosing quick but suboptimal solutions during a sprint
User Stories	natural language descriptions of features or functionalities from the perspective of end-users
Dependencies	tasks or user stories that rely on each other's completion or progression within a sprint.
Time to Market (TTM) day	is the period from the project's inception to its delivery to the end-users.

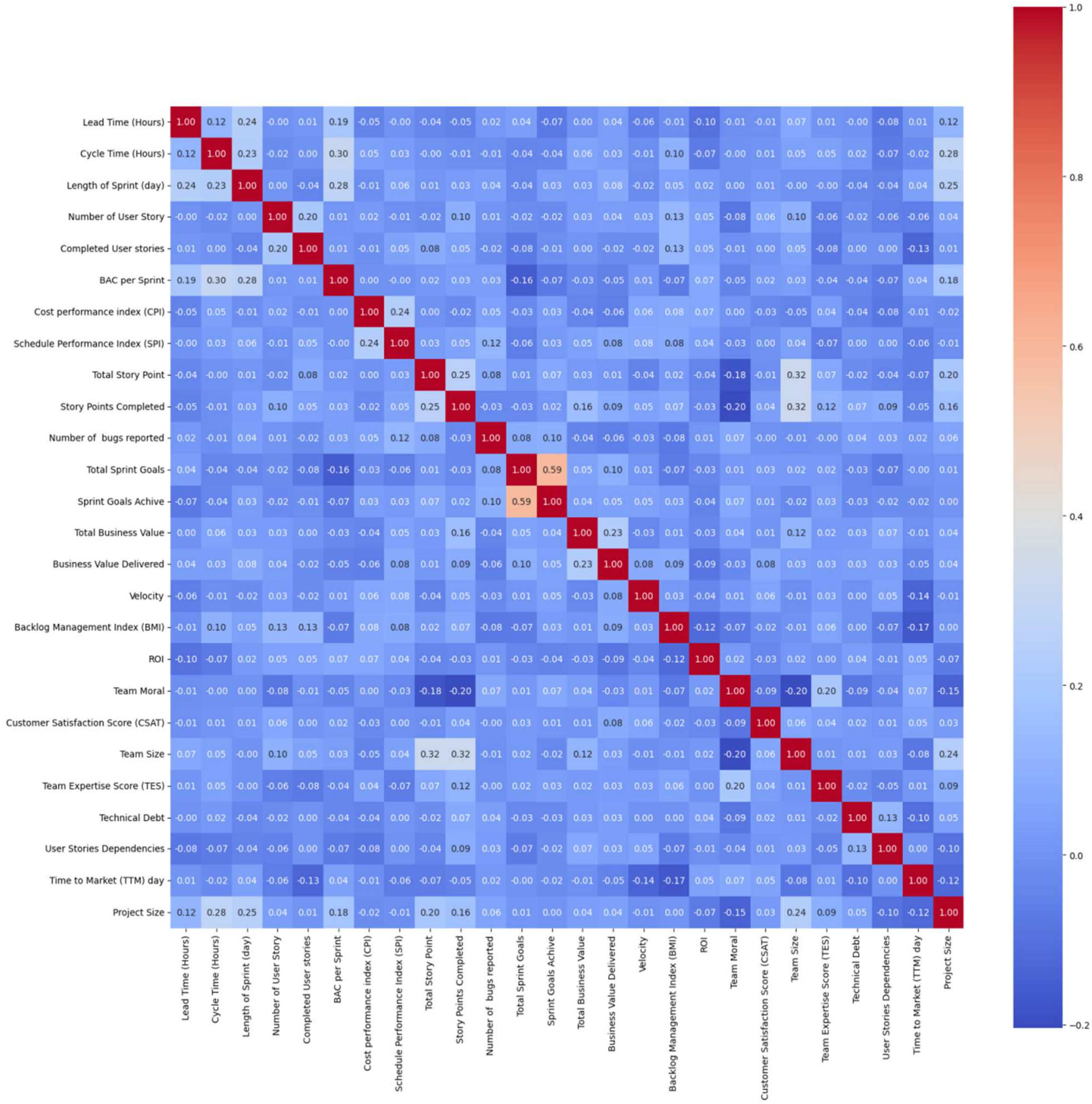

Figure 1. Correlation of criteria

The three different statuses of the project include Class 0 or "Delayed," Class 1 or "Failed," and Class 2 or "Success." As it is clear from Fig.3, there are 26 rules to predict the project's status. For example, some of these rules are:

If Sprint Goals Achieve <= 0.464 and Team Moral > 0.236, then "Project Success."

If Sprint Goals Achieve > 0.464 and Sprint Goals Achieve <= 0.739, and Number of Bugs reported<= 1.5 and Team Experience Score <= 6.35, then "Project Success."

If Sprint Goals Achieve > 0.464 and Sprint Goals Achieve <= 0.739 and Number of Bugs reported<= 1.5 and Team Experience Score > 6.35 and CSAT <= 6.0, then "Project Delayed"

Now that the model is designed, the accuracy of the model should be measured. Precision, recall, and f1-score indexes check the model's accuracy.

$$Precision = \frac{\sum_{i=1}^{l} TP_i}{\sum_{i=1}^{l} TP_i + FN_i} \tag{2}$$

$$Recall = \frac{\sum_{i=1}^{l} TP_i}{\sum_{i=1}^{l} TP_i + FP_i} \tag{3}$$

$$F1^{Score} = \frac{2 * (\text{precision} * \text{recall})}{(\text{precision} + \text{recall})} \tag{4}$$

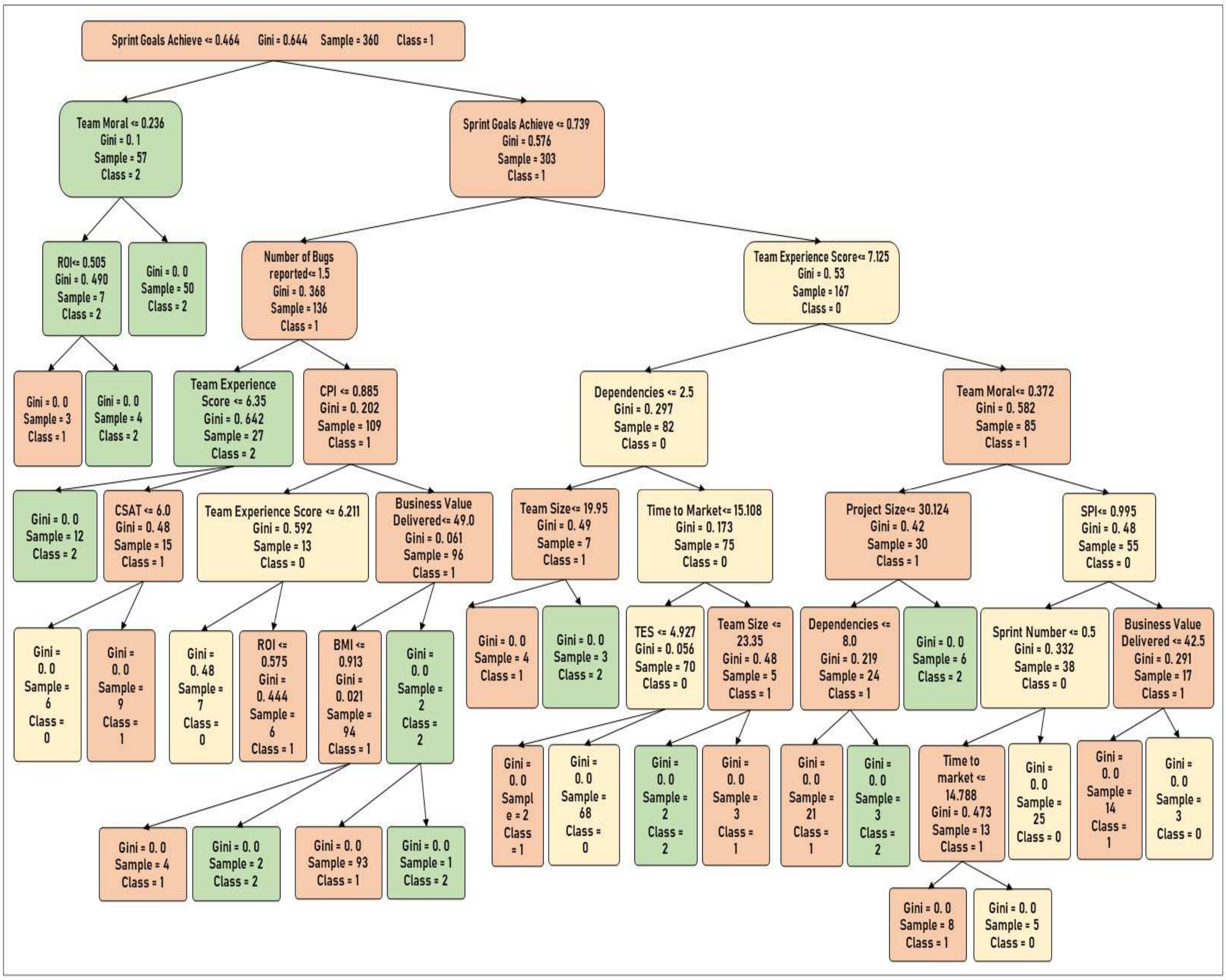

Figure 3. Project status prediction tree structure

In these formulas, TPi is the data labeled Pi, and the predicted value is the same; FPi if the data is labeled Pi but the predicted value is different. Table 6 shows the indices' value compared to the K-Nearest Neighbors (KNN) and Support Vector Machine (SVM) algorithms, which have been used as benchmark methods to compare the validity and accuracy of the developed model.

TABLE 6. THE PROPOSED MODEL PERFORMANCE MEASURES

Model	Precision	Recall	F1 Score
Decision Tree	70.9	60.9	70.9
SVM	0.72	0.70	0.71
KNN	0.82	0.81	0.82

It can be seen that the DT algorithm is by far the best-implemented algorithm, according to the data of this article. Additionally, the DT algorithm offers greater transparency and interpretability compared to other classification algorithms.In order to further validate the effectiveness of the decision tree (DT) algorithm and ensure robust results, we have conducted cross-validation in addition to the simple training/testing split.

V. CONCLUSION

This work presents a model to predict the project's status in the telecom industry. The project status has been divided into three classes, indicating that the project is failing, delayed, or completed on time and succeeding. Following the agile project management criteria, the evaluation and forecasting features of the project status were identified in the categories of Endurance, Effectiveness, and Efficiency. Then, following these features, some variables were identified as complexity. In total, 28 features have been defined, and data related to them were collected for 360 projects. In the review and analysis of the features, it can be seen that the existence of an experienced and strong team to advance the project is the most important indicator. After examining the features using the Decision Tree algorithm, which is a white-box algorithm, the rules related to predicting the status of the projects were identified, and it was observed that the accuracy of this model was 97%, in contrast to algorithms such as SVM with 71% accuracy and KNN with 82% accuracy. In the continuation of this study, it is suggested that this project status prediction model is implemented according to the agile project management structure in different industries, and the model's effectiveness in those industries should also be checked. It is also possible to use hybrid methods such as FIS and DEA to identify the project status from past data and predict the project's status in the future according to those labels, combined with machine learning algorithms.

REFERENCES

[1] J. C. Alvarenga, R. R. Branco, A. L. A. Guedes, C. A. P. Soares, and W. da S. Silva, "The project manager core competencies to project success," International Journal of Managing Projects in Business, vol. 13, no. 2, pp. 277–292, Mar. 2020, doi: 10.1108/IJMPB-12-2018-0274/FULL/PDF.

[2] B. Taha Jijo and A. Mohsin Abdulazeez, "Classification Based on Decision Tree Algorithm for Machine Learning," Journal of Applied Science and Technology Trends, vol. 2, no. 01, pp. 20–28, Mar. 2021, doi: 10.38094/jastt20165.

[3] N. Basurto, A. Jiménez, S. Bayraktar, and Á. Herrero, "Improving the Prediction of Project Success in the Telecom Sector by Means of Advanced Data Balancing," Cybern Syst, vol. 54, no. 3, pp. 306–320, 2023, doi: 10.1080/01969722.2022.2030007.

[4] X. Zhang, S. Hu, D. Chen, and X. Li, "Fast covariance matching with fuzzy genetic algorithm," IEEE Trans Industr Inform, vol. 8, no. 1, pp. 148–157, Feb. 2012, doi: 10.1109/TII.2011.2172453.

[5] S. K. Sharma and U. Chanda, "Developing a Bayesian belief network model for prediction of R&D project success," Journal of Management Analytics, vol. 4, no. 3, pp. 321–344, Jul. 2017, doi: 10.1080/23270012.2017.1304291.

[6] L. Luo, Q. He, J. Xie, D. Yang, and G. Wu, "Investigating the Relationship between Project Complexity and Success in Complex Construction Projects," Journal of Management in Engineering, vol. 33, no. 2, p. 04016036, Mar. 2017, doi: 10.1061/(ASCE)ME.1943-5479.0000471/ASSET/25786573-0514-4C82-AB8A-FD030DE84F48/ASSETS/IMAGES/LARGE/FIGURE7.JPG.

[7] P. F. Yu, F. M. Huang, C. Yang, Y. H. Liu, Z. Y. Li, and C. H. Tsai, "Prediction of Crowdfunding Project Success with Deep Learning," Proceedings - 2018 IEEE 15th International Conference on e-Business Engineering, ICEBE 2018, pp. 1–8, Dec. 2018, doi: 10.1109/ICEBE.2018.00012.

[8] N. Xia, P. X. W. Zou, X. Liu, X. Wang, and R. Zhu, "A hybrid BN-HFACS model for predicting safety performance in construction projects," Saf Sci, vol. 101, pp. 332–343, Jan. 2018, doi: 10.1016/J.SSCI.2017.09.025.

[9] J. Y. Yeh and C. H. Chen, "A machine learning approach to predict the success of crowdfunding fintech project," Journal of Enterprise Information Management, vol. 35, no. 6, pp. 1678–1696, Nov. 2022, doi: 10.1108/JEIM-01-2019-0017/FULL/XML.

[10] I. Dumitrascu-Băldău, D. D. Dumitrascu, and G. Dobrota, "Predictive Model for the Factors Influencing International Project Success: A Data Mining Approach," Sustainability 2021, Vol. 13, Page 3819, vol. 13, no. 7, p. 3819, Mar. 2021, doi: 10.3390/SU13073819.

[11] Y. Hnatchuk, O. Pavlova, and K. Havrylyuk, "Method of Forecasting the Characteristics and Evaluating the Implementation Success of IT Projects Based on Requirements Analysis.," in IntelITSIS, 2021, pp. 248–258.

[12] E. Princes and A. Said, "The impacts of project complexity, trust in leader, performance readiness and situational leadership on financial sustainability," International Journal of Managing Projects in Business, vol. 15, no. 4, pp. 619–644, May 2022, doi: 10.1108/IJMPB-03-2021-0082/FULL/XML.

[13] I. Verenich, M. Dumas, M. La Rosa, and H. Nguyen, "Predicting process performance: A white-box approach based on process models," Journal of Software: Evolution and Process, vol. 31, no. 6, p. e2170, Jun. 2019, doi: 10.1002/SMR.2170.

[14] M. Velez, P. Jamshidi, N. Siegmund, S. Apel, and C. Kastner, "White-box analysis over machine learning: Modeling performance of configurable systems," Proceedings - International Conference on Software Engineering, pp. 1072–1084, May 2021, doi: 10.1109/ICSE43902.2021.00100.

[15] R. Subramanian, R. R. Moar, and S. Singh, "White-box Machine learning approaches to identify governing equations for overall dynamics of manufacturing systems: A case study on distillation column," Machine Learning with Applications, vol. 3, p. 100014, Mar. 2021, doi: 10.1016/J.MLWA.2020.100014.

[16] D. Baccarini, "The concept of project complexity—a review," International Journal of Project Management, vol. 14, no. 4, pp. 201–204, Aug. 1996, doi: 10.1016/0263-7863(95)00093-3.

[17] A. A. ForouzeshNejad, "A hybrid data-driven model for project portfolio selection problem based on sustainability and strategic dimensions: a case study of the telecommunication industry," Soft comput, vol. 28, no. 3, pp. 2409–2429, Feb. 2024, doi: 10.1007/S00500-023-08445-W/TABLES/18.

[18] M. Y. Cheng, Y. W. Wu, and C. F. Wu, "Project success prediction using an evolutionary support vector machine inference model," Autom Constr, vol. 19, no. 3, pp. 302–307, May 2010, doi: 10.1016/J.AUTCON.2009.12.003.

[19] H. Son, C. Kim, and C. Kim, "Hybrid principal component analysis and support vector machine model for predicting the cost performance of commercial building projects using pre-project planning variables," Autom Constr, vol. 27, pp. 60–66, Nov. 2012, doi: 10.1016/J.AUTCON.2012.05.013.

[20] M. Y. Cheng, C. C. Huang, and A. F. Van Roy, "Predicting project success in construction using an evolutionary gaussian process inference model," Journal of Civil Engineering and Management, vol. 19, no. SUPPL.1, Dec. 2013, doi: 10.3846/13923730.2013.801919.

[21] Y. Li, V. Rakesh, and C. K. Reddy, "Project success prediction in crowdfunding environments," WSDM 2016 - Proceedings of the 9th ACM International Conference on Web Search and Data Mining, pp. 247–256, Feb. 2016, doi: 10.1145/2835776.2835791.

[22] S. K. Sharma and U. Chanda, "Developing a Bayesian belief network model for prediction of R&D project success," Journal of Management Analytics, vol. 4, no. 3, pp. 321–344, Jul. 2017, doi: 10.1080/23270012.2017.1304291.

[23] L. Luo, Q. He, J. Xie, D. Yang, and G. Wu, "Investigating the Relationship between Project Complexity and Success in Complex Construction Projects," Journal of Management in Engineering, vol. 33, no. 2, p. 04016036, Mar. 2017, doi: 10.1061/(ASCE)ME.1943-5479.0000471/ASSET/25786573-0514-4C82-AB8A-FD030DE84F48/ASSETS/IMAGES/LARGE/FIGURE7.JPG.

[24] N. Xia, P. X. W. Zou, X. Liu, X. Wang, and R. Zhu, "A hybrid BN-HFACS model for predicting safety performance in construction projects," Saf Sci, vol. 101, pp. 332–343, Jan. 2018, doi: 10.1016/J.SSCI.2017.09.025.

[25] P. F. Yu, F. M. Huang, C. Yang, Y. H. Liu, Z. Y. Li, and C. H. Tsai, "Prediction of Crowdfunding Project Success with Deep Learning," Proceedings - 2018 IEEE 15th International Conference on e-Business Engineering, ICEBE 2018, pp. 1–8, Dec. 2018, doi: 10.1109/ICEBE.2018.00012.

[26] Y. Hnatchuk, O. Pavlova, and K. Havrylyuk, "Method of Forecasting the Characteristics and Evaluating the Implementation Success of IT Projects Based on Requirements Analysis.," in IntelITSIS, 2021, pp. 248–258.