Exploring student characteristics as predictors of team effectiveness: A PCA and Machine Learning Approach

Yuxuan Zhang s2729040

University of Twente

Supervisor: Barrios Fleitas, Yeray

Email: y.zhang-22@student.utwente.nl

Abstract

This research focuses on how individual characteristics influence team effectiveness, which are critical to successful educational outcomes at academic institutions. This paper implements PCA technique and develops team predictive models based on student data from University of Twente, providing new insights into the impact of individual characteristics on teams. First, the paper describes data preprocessing methods to address missing data and standardize data formats across years. Second, the paper details the application of decision trees, random forests, and PCA techniques to analyze the effects of different characteristics on team outcomes. Finally, the paper launched a critical analysis of the result. The results reveal that individual previous academic performance has a strong positive correlation with team effectiveness, but Belbin team roles have a smaller impact, likely due to the complex interplay of role within teams. These findings underscore the importance of considering both individual and collective factors in educational team settings.

1. INTRODUCTION

In today's world, teamwork is vital in many areas, from business to education. In educational settings, teamwork helps students reach their academic goals and teaches them how to work with others, solve problems together, and improve their social skills. The success of these teams is important because it prepares students for future jobs where teamwork is valued (Salas, Cooke, & Rosen, 2008), not only improving academic results but also helping students grow in many ways, getting them ready to tackle real-world challenges (Johnson & Johnson, 2009).

However, creating successful teams in schools is often a tough task for teachers. At the start of a course, teachers need to form teams but might not know the best way to do it. There are several ways to make teams: by picking students randomly, letting students choose their teams, or grouping them based on certain qualities (Chapman, Meuter, Toy, & Wright, 2006). Each method has its benefits: random selection is fair and quick, choosing teams can make students feel more connected, and grouping by qualities can bring different skills and strengths together (Bacon, Stewart, & Silver, 1999). But when grouping by qualities, it can be hard for teachers to know which qualities will make a team work well (Oakley, Felder, Brent, & Elhajj, 2004). The big question is: which mix of individual traits leads to the best team performance? Specifically, what should diversity within a team look like to ensure they perform well?

Research concerning team effectiveness and the impact of diversity in educational settings shows divergent results. Some studies suggest that diversity in attributes such as gender, skills, and learning styles can enhance team effectiveness through a broader range of perspectives and problem-solving approaches (Van Knippenberg, De Dreu, & Homan, 2004). Conversely, other studies indicate that excessive diversity may lead to conflicts and complicate team coordination (Bell, Villado, Lukasik, Belau, & Briggs, 2011). This disparity underscores the complexity of the issue and underscores the necessity for further investigation. Additionally, recent studies have begun to explore machine learning as a tool for analyzing these issues in educational contexts, although these methods are still in nascent stages and require additional validation (Olsen & Haslam, 2017).

Considering the aforementioned challenges and gaps in the literature, this study advocates for the application of principal component analysis (PCA) and machine learning techniques to elucidate the dynamics of team effectiveness in educational settings. PCA will be employed to reduce the dimensionality of data, thereby identifying key traits that influence variability in team effectiveness (Abdi & Williams, 2010). Machine learning methods will subsequently model the relationships between these traits and team performance, offering a more comprehensive and robust analysis of crucial predictors (Bishop, 2006). This combination of methods aims to improve on traditional approaches and give teachers useful tools for building successful teams.

The main goal of this study is to find out which student traits are the best predictors of team success in schools. The specific questions we'll look at are:

- What are the key traits that influence how well a team does?
- How well can PCA and machine learning predict team effectiveness?

2. LITERATURE REVIEW

Team effectiveness is a complex construct explored across various disciplines, such as organizational psychology, business management, and education. Despite the lack of a universally accepted definition or measurement approach, several influential theories and models have emerged.

One foundational model was introduced by Hackman in 1987, suggesting that team effectiveness should be assessed through three criteria: the achievement of team goals, the team's capacity for performance improvement over time, and member satisfaction. This framework has been extensively referenced and forms the basis for many subsequent studies.

Salas, Dickinson, Converse, and Tannenbaum in 1992 proposed the "Big Five" framework, which identifies five critical dimensions of team effectiveness: leadership, mission orientation, structural clarity, mutual support, and cohesion. This model emphasizes the significance of internal team dynamics and processes in achieving effectiveness.

Katzenbach and Smith, in their 1993 book "The Wisdom of Teams," argued that team effectiveness should be evaluated based on performance outcomes, learning, and member satisfaction. Their approach underscores the importance of tangible results alongside the developmental and emotional aspects of team dynamics.

The Input-Process-Output (IPO) model, originally proposed by McGrath in 1964, has been a cornerstone for understanding team effectiveness. It posits that effectiveness results from the interplay of team inputs (such as member characteristics and resources), processes (including interactions and behaviors), and outputs (like achieving goals). This model was later expanded into the IMOI (Input-Mediator-Output-Input) model by Ilgen, Hollenbeck, Johnson, and Jundt in 2005, which integrates mediators affecting the process and acknowledges the iterative nature of team effectiveness.

Kozlowski and Ilgen in 2006 highlighted the significance of ongoing team development and adaptation processes, portraying effectiveness as a dynamic, evolving construct influenced by both internal and external elements.

Specific to educational settings, Deeter-Schmelz, Kennedy, and Ramsey in 2002 developed a model focusing on student teams, emphasizing cohesion, leadership, communication, and individual satisfaction as key factors.

Additionally, Tannenbaum, Beard, and Salas in 1992 devised a detailed team evaluation model that considers readiness, environment, processes, and outcomes. This framework has proven particularly valuable for thorough assessments and enhancing team development in specific contexts.

In this study, team effectiveness encompasses three core aspects: motivation, satisfaction, and performance. "Team performance" relates to the team's capability to meet objectives and deliver high-quality outputs, which may include adhering to deadlines, work quality, and achieving set goals (Sundstrom, De Meuse, & Futrell, 1990). "Team member satisfaction" assesses how satisfied individuals feel about their team involvement, focusing on cohesion, effective communication, and mutual support (Cohen & Bailey, 1997). "Motivation" measures the extent to which team members are actively engaged with the team's activities and their collective tasks.2.2 Student characteristics and their impact in team effectiveness

Team effectiveness in educational settings often considers a diverse range of individual student characteristics such as age, gender, nationality, study program or background, Belbin roles, and pre-course qualifications.

The age of students can significantly impact team dynamics and effectiveness. Research indicates that age diversity can bring varied perspectives and experiences to a team, enhancing its problem-solving capacity. However, without careful management, age differences can also lead to conflicts (Zenger & Lawrence, 1989).

Gender diversity within teams is a subject of extensive study with varying outcomes. Some research highlights that gender diversity can enhance creativity and improve decision-making processes (Bear & Woolley, 2011), while other studies suggest it might lead to increased conflicts and reduced team cohesion if not inclusively managed (Horwitz & Horwitz, 2007).

Nationality and cultural diversity can enrich a team by introducing varied viewpoints and problem-solving approaches, enhancing the team's overall capability to tackle complex tasks. Nevertheless, this diversity can also pose communication challenges and affect team unity (Gibson & Vermeulen, 2003).

Differences in study programs or academic backgrounds among students can bring a range of skills and knowledge to a team, proving particularly beneficial in interdisciplinary projects. The diversity in academic expertise encourages a comprehensive approach to problem-solving (Jehn, Northcraft, & Neale, 1999).

Belbin roles, a framework developed by Meredith Belbin, categorizes the typical roles individuals assume in a team, such as coordinators, implementers, and resource evaluators. According to Belbin's theory, a team's effectiveness hinges on a balanced mix of these roles, ensuring that all necessary team functions are adequately covered (Belbin, 2010).

Lastly, previous grades or academic performance can indicate a student's skill and knowledge levels. Teams formed based on academic performance can help ensure an even distribution of skills, although this approach may also lead to perceptions of inequality among team members if not handled carefully (Deeter-Schmelz, Kennedy, & Ramsey, 2002).

Many studies have delved into team formation and effectiveness in educational settings, employing various methods. For instance, Oakley et al. (2004) focused on how selecting team members based on specific characteristics impacts overall team performance. They discovered that teams assembled from a balanced mix of skills and personalities generally experienced enhanced performance and greater member satisfaction.

Chapman et al. (2006) also contributed significantly to this area by examining the effects of different team formation methods—random, self-selected, and based on specific characteristics—on team dynamics and outcomes. Their research indicated that while self-selected teams might exhibit more initial cohesion, those formed from specific characteristics often achieve better long-term performance, thanks to a more effective blend of skills.

Additionally, Olsen and Haslam (2017) applied machine learning techniques to assess team effectiveness in educational environments. Their findings revealed that machine learning could detect complex and non-linear patterns in data, making it a formidable tool for predicting the effectiveness of teams. This innovative approach holds substantial potential for future research and practical implementations in team formation.

By comparing the outcomes of these studies with current research, we can validate and place our findings within a broader context. This comparison not only reinforces the validity of our results but also provides a robust foundation for discussing the implications and suggesting future research directions.

3. METHODOLOGY

The goal of this study is to analyze how individual characteristics of students (such as gender, nationality, age, study program, Belbin roles, previous grades, and grade expectations) influence team effectiveness, understood as the level of motivation, satisfaction and final grade obtained in the team project. For this purpose, a principal component analysis (PCA) and machine learning algorithms (decision trees and random forest) will be carried out. The study uses a cross-sectional design, as data were collected at a single point in time (at the end of the fourth module) for each student cohort over the three years. This design allows for examination of relationships between student characteristics and team effectiveness in a specific context without follow-up over time. Throughout this section the origin and nature of the data analyzed, the methods for preparing and preprocessing the data, and the techniques conducted to analyze the data will be described.

3.1 Data description

The data used in this study were obtained as part of a parallel study conducted by Barrios-Fleitas et al. (2024). The dataset comes from the fourth module, "Data and Information Systems," of the first-year Technical Computer Science (TCS) program at the University of Twente, collected over a period of three years (2021-2023). At this university, credits are organized into learning units called modules, which are groupings of independent courses designed to teach theoretical content while allowing students to apply what they learn through projects. Each module lasts 10 weeks, and while each study program has its own specific modules, there is often sharing of modules between different study programs, promoting multidisciplinary collaboration among students. For instance, modules 2 and 4 of the first year of TCS are shared with students from the Business and Information Technology (BIT) program.

The fourth module, central to this study, focuses on real software development projects with actual clients, creating a simulated real-world environment for team development. It covers key areas such as databases, web programming, and information security. Teams in this module are composed of 5 to 6 members with specific diversity criteria: no more than two members may speak the same language, and ideally, each team includes four Computer Science students and two from other programs, ensuring a varied and interdisciplinary team dynamic.

Building on the dataset described, the original study by Barrios-Fleitas et al. (2024) gathered extensive information about the students, including their personality traits using the 60-item HEXACO scale. However, these personality variables were not included in this study. The decision to exclude these data was due to the complexity they introduced into the analysis. Analyzing the intricate relationships between personality traits and team dynamics, while valuable, would significantly complicate the interpretative process without direct relevance to the primary objectives of assessing educational strategies and team effectiveness within the program. For a detailed overview of the variables that were included in this study, refer to Table 1.

Table 1: List of variables used in this study

	Variable	Description	Possible values
Input variables	Team size	This represents the number of students in a team. Typically, the team size is 6, but this may be affected by students dropping out in the middle of the project.	4, 5, 6, 7
	Gender	This represents the gender of a student. Only male and female were considered in this feature.	Male, Female
	Nationality	This represents the first nationality of a student.	Dutch, German, Bulgarian, Romanian, Chinese, Indian …
	Age	This represents the age of a student.	From 17 to 32
	Study program	This represents the study program of a student. While the sample is composed primarily of TCS and BIT students, students from other programs can also take this module as a minor course.	TCS, BIT, AT-IT-UT, X-CS, X-IBA, X-MST, CREA, M-TM, X-TCS, X-EE, B-AM B-IBA, B-TCS, B-AM B-AT
	Belbin Roles (Role 1, Role 2, Role 3)	Belbin roles are a set of nine team roles that describe typical behaviors that individuals may exhibit when working within a team. The students were given a test to identify their three most dominant roles.	Coordinator, Shaper, Plant, Monitor Evaluator, Resource Investigator, Team Worker, Implementer, Completer Finisher and Specialist
	Previous grades (Mod 1, Mod 2, Mod 3, programing, Projects, Math)	Previous grades are taken from the student's first three modules in the first year; Modules 1, 2, and 3 represent the student's final grades for the first three modules; The variables programming, project and Math represent the average grade obtained by the student prior to the course in specific components of programming, teamwork and mathematics respectively.	On the scale of 1-10.
	Grade expectations (Min, Expected)	Min represents the student’ s minimal acceptable grade, and Expected represents the expected grade of the student.	On the scale of 1-10.
Output variables	Team project grade	The average grade of team project.	On the scale of 1-10.
	Team Motivation	The average of Motivation Level of the students in the team collected by questionnaire in Module 4.	On the scale of 1-5.
	Team satisfaction	The average of Satisfaction Level of the students in the team collected by questionnaire in Module 4.	On the scale of 1-5.

3.2 Data preprocessing

The data used in this study included some missing values, which occurred for several reasons. Some students did not complete the data collection survey or did not finish the module, leaving certain variables (such as the final grade) unknown. Additionally, the university's management system had incomplete records for students who did not pass all components in previous modules, resulting in missing data for the previous grades features. Data collected through questionnaires also had gaps due to students not standardizing their responses, leading to missing data in the Belbin roles and grade expectations features.

Given that our data encompasses first-year students from 2021 to 2023, the extent of missing data varies across cohorts. First-year students in 2021 had up to three opportunities to pass the initial three modules, resulting in fewer missing values for previous grades. Conversely, first-year students in 2023 had only one chance to pass these modules, making it more challenging to obtain complete data, thus increasing the amount of missing data. Except for the variables Gender, Nationality, Age, Study Program, Programming, Project, and Math, which have no missing values, Table 2 presents the percentage of missing values for the remaining study variables in each year.

Table 2: Missing values

	B.Role 1	B.Role 2	B.Role 3	Mod 1	Mod 2	Mod 3	Programmin g	Projects	Math	Min	Expected
2021	0%	0%	0%	9.8%	4.4%	4.4%	4.4%	4.7%	5.1%	9.5%	9.5%
2022	6%	6.3%	24.7%	25.5%	41.2%	41%	22.5%	26.1%	23.1%	5.8%	5.8%
2023	12.5%	12.5%	13.6%	15.1%	7.4%	70.3%	7.1%	9.8%	8.9%	25.2%	25.2%

To address the missing values without significantly reducing the dataset size, we employed the K-Nearest Neighbors (KNN) Imputer. This method imputes missing values by considering the k-nearest neighbors (similar rows) in the dataset, leveraging the inherent structure and relationships within the data to ensure that the imputed values are realistic and consistent with existing patterns (Zhang, 2012).

The KNN Imputer was applied to several numeric features. These features include the final grades for the first three modules (Mod 1, Mod 2, Mod 3), grades for programming exams and projects in Module 2 (Programming, Projects), the average grade for math components across the first three modules (Math), and both the minimum acceptable grade and the expected grade of the student (Min, Expected). These numeric features were suitable for KNN imputation as the algorithm effectively handles numerical data by exploring the variance between data points to identify the nearest neighbors and fill in missing values accordingly.

3.3 Encoding

Since the algorithms used only accept numeric features, it was necessary to encode the categorical data (Role 1, Role 2, Role 3) into numeric data. One Hot Encoder was used to create binary columns for each unique category in the data set, avoiding implicit hierarchy issues that could arise with the use of the Label Encoder (Hastie, Tibshirani, & Friedman, 2009). To keep the dimensionality of the data set manageable, 9 new columns were created for all Belbin roles.

3.4 Principal Component Analysis

Principal component analysis (PCA) is a technique which is normally used for dimensionality reduction. By analyzing the principal components and their corresponding loadings (the correlation between the original features and the principal components), we can identify the features that contribute the most to the principal components. These highly contributing features can be considered the most important or influential features in the dataset (Jolliffe, 2002). The loadings, which represent the contribution of each original feature to the principal components, were calculated to identify most significant features contributing to the variance in the dataset.

3.5 Machine learning algorithms

Since our goal is to analyze the effects of individual factors as well as team factors on individual performance, etc., classification algorithms are more suitable for our study, especially decision tree algorithms. Decision trees split nodes based on feature values that result in the best separation of classes, and the importance of each feature is assessed based on how much it decreases the impurity or increases the information gain at each split (Zvornicanin, 2024). Random forest is the advanced decision tree algorithm. It is composed of multiple decision trees, each of which is built using a random vector sampled independently with the same distribution across all trees in the ensemble (Breiman, 2001). We implemented these algorithms with the scikit-learn library in python which has some built-in functions that can help us analyze feature importance.

3.6 Software used for data analyses

We chose Python as the programming language for our experiments because Python has a bunch of well-researched libraries for machine learning and data processing, for example, scikit-learn. Scikit-learn is a powerful Python library that provides simple and effective tools for data mining and data analysis. In our analysis, we leveraged the scikit-learn library in Python to implement principal component analysis (PCA), decision tree and random forest algorithms. For PCA, we first standardized the dataset using 'StandardScaler' to ensure that all features contributed equally, and then applied PCA to identify the features that contributed the most to the variance of the dataset. We analyzed the loadings to determine the contribution of each feature to the principal components. For our machine learning analysis, we used decision trees and random forest algorithms with scikit-learn to evaluate each feature's importance to the model. We first split the dataset into training and test sets. After training models, we evaluated feature importance using the 'feature_importances' attribute of the tree models, which helped us visualize the contribution of each feature. To further enhance interpretability, we leveraged the SHAP library to create an explainer and generate summary plots to describe the impact of each feature on the model's predictions. SHAP values are a technique for explaining the output of machine learning models. SHAP values quantify the contribution of each feature to the model's predictions to help us better understand which feature has a greater impact on the output (Hamilton, 2024). We also used Numpy, a Python package that provides powerful array processing capabilities to efficiently handle large datasets and complex mathematical operations.

3.7 Ethical considerations

The data for this research comes from the University of Twente's management system, Osiris, as well as data collected from the Academic Skills in Module 4. The data was completely anonymized before we had access to it. We are also committed to destroying the data after the research. We only accessed data that was necessary for our research. Any other data not relevant to our research questions were excluded.

4. RESULTS

In this section, we will analyze which individual factors have a significant impact on team project grade, motivation, and satisfaction by employing decision tree algorithm, random forest algorithm and PCA in python scikit-learn library.

4.1 Results of the PCA analysis

The figure displays the feature loadings for the first two principal components (PC1 and PC2) in a Principal Component Analysis. The loading scores represent how much each original feature contributes to each principal component. These scores help in understanding which features are most influential in the variance explained by the principal components. From it we can see that 'Nationality' has a very high positive loading score, suggesting it is a significant contributor to the first principal component. Several features such as 'Mod 1', 'Mod 2', 'Mod 3', 'Programming', 'Projects', 'Math', 'Min', and 'Expected' have significant contributions to the second principal component. This result shows similarities with the results obtained by our machine learning approach. Both shows that previous grades, especially Module 2, and Grade expectations have a very significant impact on the team.

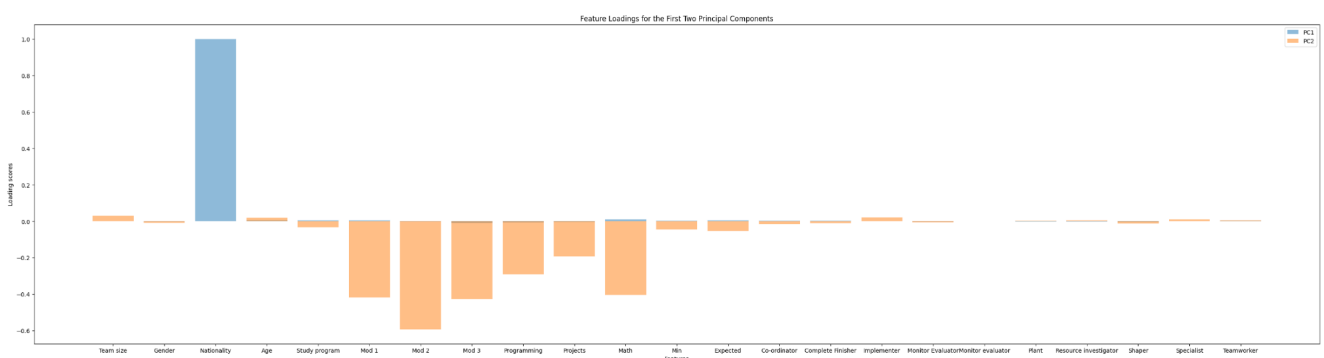

Fig. 1: PCA feature loadings plot

4.2 Results of the decision tree analysis

Table 3 shows results of three decision tree models corresponding to project grade, team motivation and team satisfaction. The project grade and team satisfaction models both show an accuracy of 63%, precision of 64%, and recall of 63%, and team motivation model performs slightly better with an accuracy of 69%, precision of 67%, and recall of 70%. Accuracy indicates the proportion of correct predictions made by the model out of all predictions; precision measures the accuracy of positive predictions and recall measures the proportion of actual positive instances that are correctly identified by the model. Figure 2, 3 and 4 shows all features' SHAP value of each model.

The decision tree model for team project grade shows that features such as 'Mod 2', 'Min', 'Mod 3', 'Expected', 'Mod 1', 'Projects' are the most influential features, reflecting their significant contribution to the model predictions. For team motivation, the most critical factors include 'Projects', 'Mod 2', 'Team size', 'Implementer', and 'Expected', indicating their importance in maintaining team morale and drive. In the case of team satisfaction, the top influencing features are 'Mod 2', 'Expected', 'Mod 1', 'Math', and 'Min', highlighting the key elements that contribute to overall team satisfaction. Across all measures, 'Mod 2', 'Min' and 'Expected' consistently appeared as significant contributors, highlighting their central role in determining various aspects of team effectiveness.

Table 3: Decision tree model results

	Accuracy	Precision	Recall
Project grade model	63%	64%	63%
Team motivation model	69%	67%	70%
Team satisfaction model	63%	64%	63%

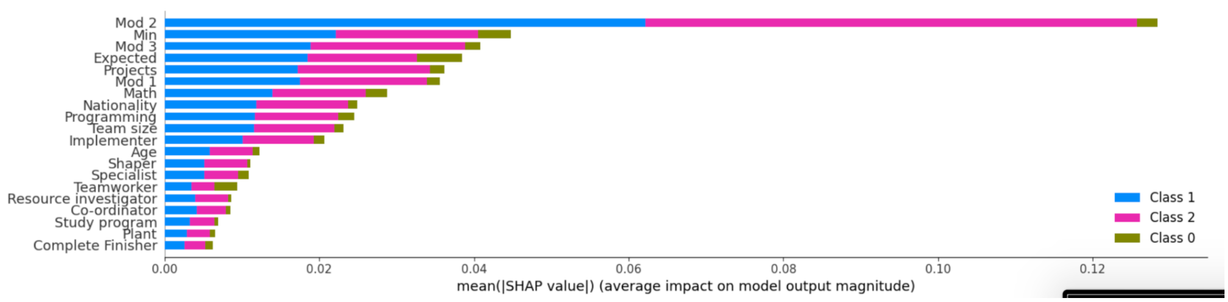

Fig. 1: Decision tree model feature importance for team project grade

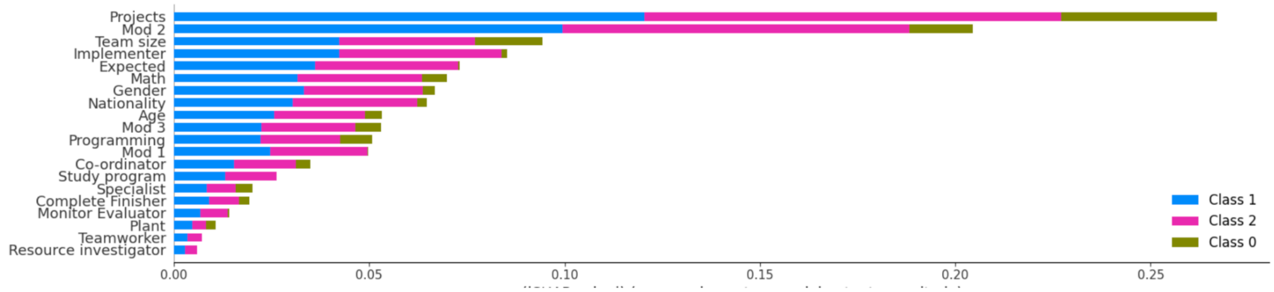

Fig. 2: Decision tree model feature importance for team motivation

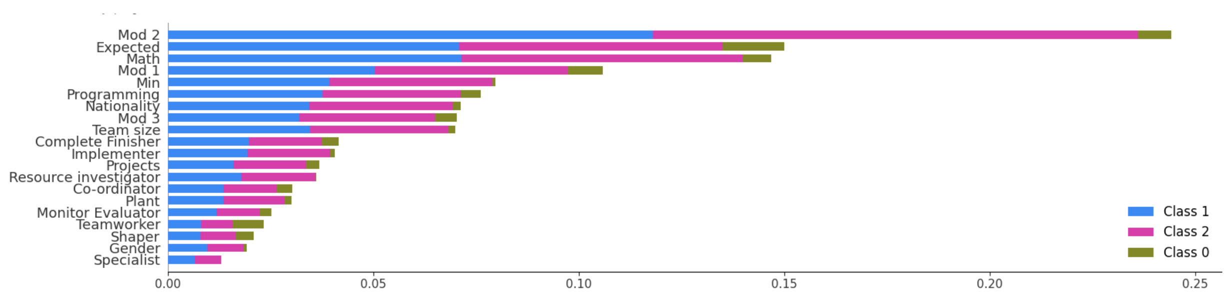

Fig. 3: Decision tree model feature importance for team satisfaction

4.3 Results of the random forest analysis

The random forest models for predicting team effectiveness—specifically project grade, team motivation, and team satisfaction—demonstrate improved performance compared to the decision tree models. The project grade and team satisfaction models both achieve an accuracy of 73%, precision of 71%, and recall of 75%, indicating a high level of reliability in correctly identifying relevant instances. The team motivation model shows an accuracy of 74%, with a precision of 50% and recall of 76%, highlighting its strong ability to identify actual instances of team motivation, though with more false positives. Figures 5, 6 and 7 illustrate the SHAP values for feature importance in each model. For the project grade, key influential features include 'Mod 2', 'Mod 3', 'Expected', 'Projects', 'Min', and 'Programming'. In the team motivation model, the critical factors are 'Mod 2', 'Projects', 'Expected', and 'Math'. For team satisfaction, 'Mod 2', 'Programming', 'Mod 3', 'Projects', and 'Min' stand out as significant contributors. Consistently, 'Mod 2', 'Min' and 'Expected' are pivotal across all models, underscoring their importance in team effectiveness.

Table 3: Random forest model results

	Accuracy	Precision	Recall
Project grade model	73%	71%	75%
Team motivation model	74%	50%	76%
Team satisfaction model	73%	71%	75%

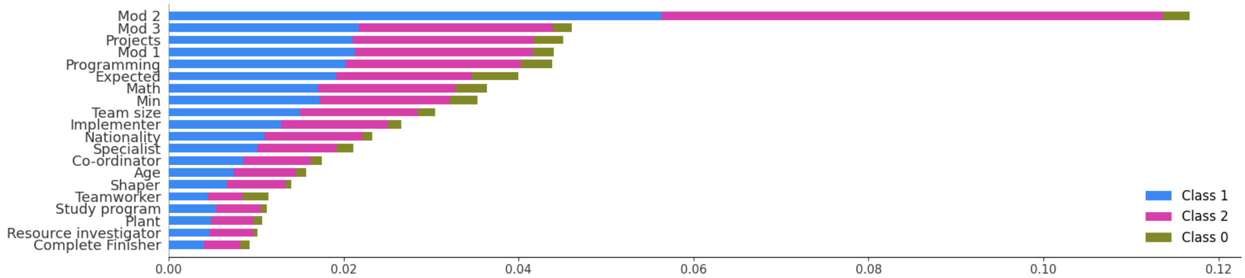

Fig. 5: Random forest model feature importance for team project grade

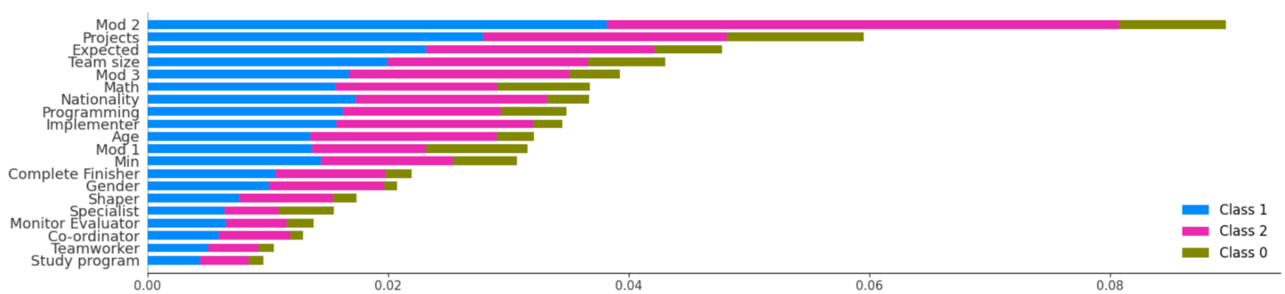

Fig. 6: Random forest model feature importance for team motivation

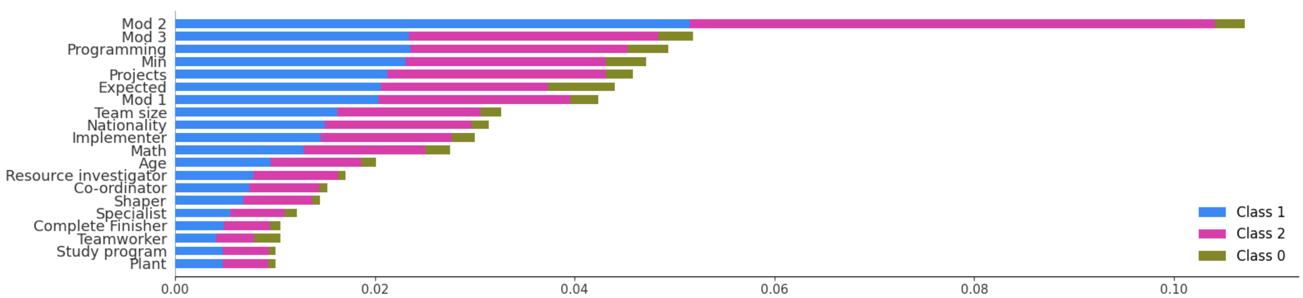

Fig. 7: Random forest model feature importance for team satisfaction

4.4 Comparison of techniques

The results of the three approaches are very similar, with the PCA results emphasizing the contribution of previous grades (Mod 1, Mod 2 and Mod 3). However, the two machine learning approaches particularly highlight the contribution of Mod 2 to team effectiveness and also show that the expected grades (Expected and Min) have a significant impact on models. As machine learning models can handle non-linear relationships and interactions between features, whereas PCA assumes linearity and focuses on variance, potentially overlooking non-linear contributions of feature (Shlens, 2014), plus Random forest models have higher performance. These strengths make random forests a more powerful and reliable tool for analyzing and improving team effectiveness.

5. Discussion

5.1 Results interpretation

The results of all three techniques show that previous grades have a huge impact on team effectiveness. In particular, the Random forest model, the best of the three methods, emphasizes Mod 2 as the most central feature. Figures 8, 9, and 10 below show the correlation between each feature and the three outcomes as well as the p-value, which represents the probability of obtaining an effect at least as extreme as the one observed in your sample data (Di Leo, 2020). We can also see from these matrices that 'Mod 2' has a strong positive correlation with all three outcomes and has a low p-value, suggesting that this is not a coincidence. It makes common sense that Module 2 is about software development, students who achieved good grades on Module 2 are more likely to have solid programming skills. These skills are highly beneficial for succeeding in Module 4, which is centered around web development projects. The high SHAP value of features 'Programming' and 'Projects' can also prove this assumption. Therefore, students who achieve good grades in Module 2 are more likely to also perform well in Module 4 projects. The Random forest model also shows that 'Min' and 'Expected' contribute significantly to team effectiveness. The matrix also shows that these two features have a strong correlation to the three outcomes. This suggests that students with high expected grades may spend more time and effort on the program, resulting in good program grades. However, from the results, Belbin roles do not seem to have much impact on team effectiveness.

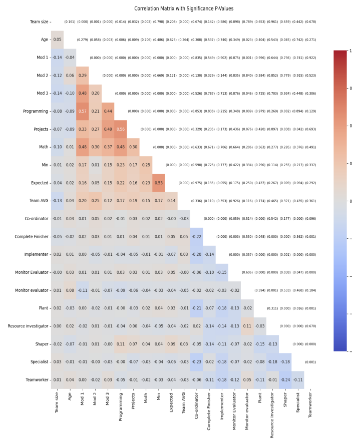

Fig. 8: Correlation matrix for project grade

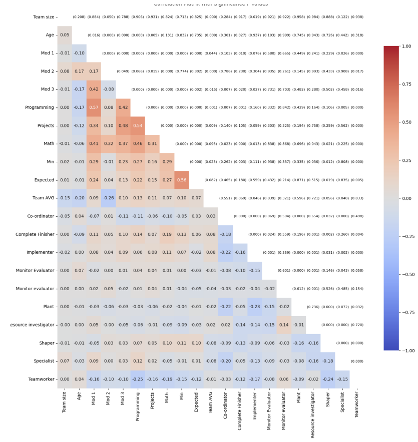

Fig. 9: Correlation matrix for team motivation

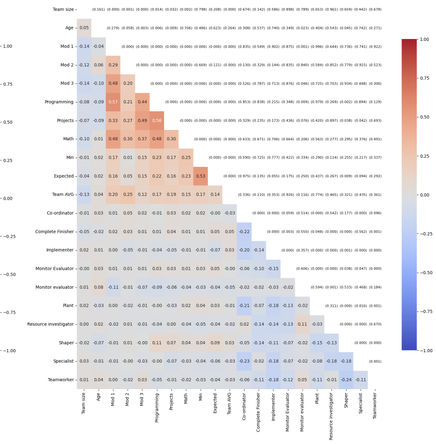

Fig. 10: Correlation matrix for team satisfaction

5.2 Comparison with previous literature reviewed

The existing literature tends to focus more on individual performance rather than team effectiveness. However, our results indicate that students' previous individual grades or academic performance exhibit strong positive correlations with team effectiveness. This is aligned with the existing literature and suggests that high-performing individuals contribute significantly to both their own success and the success of their teams.

However, the effect of individual Belbin roles on team effectiveness seems to be much smaller, which is inconsistent with the existing literature suggesting that Belbin roles are strongly correlated with team effectiveness. This difference may stem from the complex interplay between different roles in a team. In a team context, the impact of an individual's role may be moderated by the roles of other team members, resulting in a more complex dynamic that is not easily captured by models. By focusing on the individual's Belbin role without considering the broader team context, we risk oversimplifying the analysis and thereby missing important interactions that affect team performance. Thus, the impact of Belbin team roles on the model appears to be weaker if not analyzed within the framework of team interactions. Future research should consider these complexities to better understand the role of personality in team effectiveness.

6. Conclusions

6.1 Study limitations

While many insights were gained from this study, there are some significant limitations to this work. The sample size of the dataset may lack the power to generalize the findings to other educational settings or the broader population. The study only involved students from University of Twente, whose experiences are not representative of all student populations. Moreover, some data we used in our analysis was collected from questionnaires, which is not guaranteed to be accurate and may lead to bias and error. Finally, machine learning models oversimplified the complexity of team dynamics, which may ignore the many complex interplays between students. This factor may also have a significant impact on team effectiveness.

6.2 Recommendations and future research

Future research should consider expanding sample size and diversity to increase the generalizability of findings. In addition, the complexity of team dynamics should be incorporated to explore the effects of different combinations of Belbin roles on team effectiveness. When these aspects are explored, we can truly help university administrators develop strategies for grouping students, which may have a real impact on the enhancement of educational practices and team. Thereby testing the results of our research in practice.

6.3 Final thoughts

This study highlights the relationship between team effectiveness and individual characteristics in educational settings. The result shows student's previous grade and grade expectations play significant roles in team effectiveness. While challenges and complexities remain, particularly in Belbin roles don't have as much impact on team outcomes as we expected. This may be due to the fact that we overlooked the interplays of different roles in a team.

By continuing to explore and refine these models and approaches, future research can offer more detailed and actionable insights, ultimately contributing to more effective and satisfying educational experiences for students.

7. References

Salas, E., Cooke, N. J., & Rosen, M. A. (2008). On Teams, Teamwork, and Team Performance: Discoveries and Developments. *Human Factors*, *50*(3), 540-547. https://doi.org/10.1518/001872008X288457

Johnson, D. W., & Johnson, R. T. (2009). An Educational Psychology Success Story: Social Interdependence Theory and Cooperative Learning. Educational Researcher, 38, 365-379. https://doi.org/10.3102/0013189X09339057

Chapman, K., & Meuter, M. (2006). Can't we pick our own groups? The influence of group selection method on group dynamics and outcomes. *Journal of Management Education, 30*(4), 557-569. https://doi.org/10.1177/1052562905284872

Bacon, D. R., Stewart, K. A., & Silver, W. S. (1999). Lessons from the Best and Worst Student Team Experiences: How a Teacher can make the Difference. *Journal of Management Education*, *23*(5), 467-488. https://doi.org/10.1177/105256299902300503

Oakley, B.A., Felder, R.M., Brent, R., & Elhajj, I.H. (2004). Turning Student Groups into Effective Teams.

van Knippenberg, D., De Dreu, C. K. W., & Homan, A. C. (2004). Work Group Diversity and Group Performance: An Integrative Model and Research Agenda. *Journal of Applied Psychology, 89*(6), 1008–1022. https://doi.org/10.1037/0021-9010.89.6.1008

Bell, S. T., Villado, A. J., Lukasik, M. A., Belau, L., & Briggs, A. L. (2011). Getting Specific about Demographic Diversity Variable and Team Performance Relationships: A Meta-Analysis. *Journal of Management*, *37*(3), 709-743. https://doi.org/10.1177/0149206310365001

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics, 2*(4), 433-459. https://doi.org/10.1002/wics.101

Hackman, J. R., Pearce, J. L., & Wolfe, J. C. (1978). Effects of changes in job characteristics on work attitudes and behaviors: A naturally occurring quasi-experiment. *Organizational Behavior and Human Performance, 21*(3), 289-304. https://doi.org/10.1016/0030-5073(78)90055-7

Tannenbaum, S. I., Beard, R. L., & Salas, E. (1992). Team building and its influence on team effectiveness: An examination of conceptual and empirical developments. In K. Kelley (Ed.), *Issues, Theory, and Research in Industrial/Organizational Psychology* (Vol. 7, pp. 117-153). Elsevier Science. https://doi.org/10.1016/S0166-4115(08)62601-1

Katzenbach, J.R. and Smith, D.K. (1993), "The rules for managing cross-functional reengineering teams", Planning Review, Vol. 21 No. 2, pp. 12-13. https://doi.org/10.1108/eb054404

Littlepage, G. E., Schmidt, G. W., Whisler, E. W., & Frost, A. G. (1995). An input-process-output analysis of influence and performance in problem-solving groups. *Journal of Personality and Social Psychology, 69*(5), 877–889. https://doi.org/10.1037/0022-3514.69.5.877

Ilgen, D. R., Hollenbeck, J. R., Johnson, M., & Jundt, D. (2005). Teams in organizations: From input-process-output models to IMOI models. *Annual Review of Psychology, 56*(1), 517-543. https://doi.org/10.1146/annurev.psych.56.091103.070250

Kozlowski, S. W. J., & Ilgen, D. R. (2006). Enhancing the effectiveness of work groups and teams. Psychological Science in the Public Interest, 7(3), 77-124. https://doi.org/10.1111/j.1529-1006.2006.00030.x

Deeter-Schmelz, D. R., Kennedy, K. N., & Ramsey, R. (2002). Enriching our understanding of student team effectiveness. Journal of Marketing Education, 24(2), 114-124. https://doi.org/10.1177/0273475302242006

Tannenbaum, S. I., Beard, R. L., & Salas, E. (1992). Team building and its influence on team effectiveness: An examination of conceptual and empirical developments. In K. Kelley (Ed.), Issues, Theory, and Research in Industrial/Organizational Psychology (Vol. 7, pp. 117-153). Elsevier Science. https://doi.org/10.1016/S0166-4115(08)62601-1

Sundstrom, E., De Meuse, K. P., & Futrell, D. (1990). Work teams: Applications and effectiveness. American Psychologist, 45(2), 120-133. https://doi.org/10.1037/0003-066X.45.2.120

Cohen, S. G., & Bailey, D. E. (1997). What makes teams work: Group effectiveness research from the shop floor to the executive suite. Journal of Management, 23(3), 239-290. https://doi.org/10.1016/S0149-2063(97)90034-9

Zenger, T. R., & Lawrence, B. S. (1989). Organizational demography: The differential effects of age and tenure distributions on technical communication. Academy of Management Journal, 32(2), 353-376. https://doi.org/10.5465/256366

Bear, J. B., & Woolley, A. W. (2011). The role of gender in team collaboration and performance. Interdisciplinary Science Reviews, 36(2), 146-153. https://doi.org/10.1179/030801811X13013181961473

Horwitz, S. K., & Horwitz, I. B. (2007). The effects of team diversity on team outcomes: A meta-analytic review of team demography. Journal of Management, 33(6), 987-1015. https://doi.org/10.1177/0149206307308587

Gibson, C. B., & Vermeulen, F. (2003). A healthy divide: Subgroups as a stimulus for team learning behavior. Administrative Science Quarterly, 48(2), 202-239. https://doi.org/10.2307/3556657

Jehn, K. A., Northcraft, G. B., & Neale, M. A. (1999). Why differences make a difference: A field study of diversity, conflict, and performance in workgroups. Administrative Science Quarterly, 44(4), 741-763. https://doi.org/10.2307/2667054

Belbin, R. M. (2010). Team roles at work (2nd ed.). Routledge. https://doi.org/10.4324/9780080963242

Oakley, B., Felder, R. M., Brent, R., & Elhajj, I. (2004). Turning student groups into effective teams. Journal of Student Centered Learning, 2(1), 9-34.

Chapman, K. J., Meuter, M., Toy, D., & Wright, L. K. (2006). Can't we pick our own groups? The influence of group selection method on group dynamics and outcomes. Journal of Management Education, 30(4), 557-569. https://doi.org/10.1177/1052562905284872

Olsen, J. K., & Haslam, S. A. (2017). The influence of team composition and conflict on team outcomes. Journal of Applied Psychology, 102(3), 450-465. https://doi.org/10.1037/apl0000152

Zhang, S. (2012). Nearest neighbor selection for iteratively kNN imputation. *Journal of Systems and Software, 85*(11), 2541-2552. https://doi.org/10.1016/j.jss.2012.05.073

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Science & Business Media

Jolliffe, I. T. (2002). *Principal Component Analysis*. Springer Series in Statistics (2nd ed.). New York: Springer-Verlag

Zvornicanin, W., & E. (2024). What is feature importance in machine learning? *Baeldung on Computer Science*. https://www.baeldung.com/cs/ml-feature-importance

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5-32. https://doi.org/10.1023/A:1010933404324

Shlens, J. (2014). A tutorial on principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(4), 433-459. https://doi.org/10.1002/wics.101

Di Leo, G., Sardanelli, F. Statistical significance: p value, 0.05 threshold, and applications to radiomics—reasons for a conservative approach. Eur Radiol Exp 4, 18 (2020). https://doi.org/10.1186/s41747-020-0145-

Appendix:

During the preparation of this work the author(s) used ChatGPT in order to generate sample code for data processing. After using this tool, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the work.

During the preparation of this work the author(s) used DeepL for translations. After using this service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the work.

During the preparation of this work the author(s) used Grammarly for grammar and spelling checks. After using this service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the work.

During the preparation of this work the author(s) used txyz.ai to help me review existing literature. After using this service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the work.