Journal of University Teaching & Learning Practice

Volume 18
Issue 8 *Standard Issue 4*

Article 09

2021

Teamwork-performance prediction by using soft skills and technological savvy skills

Hoi Yan Lin
Hong Kong Polytechnic University, Hong Kong, 13901480r@connect.polyu.hk

Jia You
Hong Kong Polytechnic University, Hong Kong, csyjia@comp.polyu.edu.hk

Recommended Citation

Lin, H., & You, J. (2021). Teamwork-performance prediction by using soft skills and technological savvy skills. *Journal of University Teaching & Learning Practice, 18*(8). https://doi.org/10.53761/1.18.8.9

Teamwork-performance prediction by using soft skills and technological savvy skills

Abstract

In today's connected world, forming teams of people to execute projects is seen as a challenge in government agencies and public and private organisations alike. For large enterprises, a small group of thoughtful and committed people performing different roles could essentially change the world. At the same time, however, it is hard to select an effective team with appropriate skills who can work collaboratively. In this project, as a starting point, the study's objectives required formulating skillset models and designing the theoretical framework to investigate project members' capabilities. This study used three undergraduate courses' data as input to find the skill features required in team project assignments. Additionally, possessing soft skills and Technological Savvy skills can help minimize member underperformance. Our case study of Predicting Teamwork Performance (PTPA) system also discovered some attributes that directly influence team projects and evaluating the results pointed out team members' strengths and weaknesses. Therefore, this theoretical framework can help team leaders recognize the skillsets necessary for project management.

Practitioner Notes

1. The technological savvy skill is a specific ability that can be measured and defined. Measurement in our case study showed how well the technical skill variables related to teamwork performance.
2. Teammates with good programming, creativity, communication and logical skills have associated technological savvy skills.
3. The TSS model is a vital asset to consider when developing project management roles. The study also identified soft skills that are adaptive to team processes.
4. Although technological savvy skills are essential in team projects, soft skills help teammates work well no matter how much or little technical knowledge they have.
5. Balancing technological savvy skills and soft skills are primary determinants in how well teams perform.

Keywords

Soft Skills, Technological Savvy Skills, Underperformance, Predicting Teamwork Performance

Introduction

Forman (1994) suggested that individualistic and sociocultural factors can influence cognitive learning and broader social benefit in a community of practice, challenging in higher education. According to Fine and Hallett (2014), they said ongoing social relations could support organisational practices. They examined the role of shared awareness and memory, performance, and differentiation to belong in corporate life. They observed, "group culture could be capable of recognising the tensile strength of organisations in the face of forces of uncertainty that might disrupt allegiance". The argument from Liber (1994) said team-based learning could build-up complexity in the learning experience and strengthen students' skills in the real world. Recent publications from Aranyossy et al., (2018) and Luțaș et al. (2020) indicated that project success needed to consider time, cost, quality, and effectiveness. The teamwork cooperation element is also one of the significant components in education. Team projects are the trend to enhance student learning in practical ways in higher education, and project management issues are typical problems that professional engineers face in the industry. In this study, we observed that team project success factors include knowledgeable people with leadership skills. Technical skills and soft skills are essential issues to support project-based collaborative learning. Students with a good blend of those skills find it easy to reach their learning goals.

Prerna Sindwani (2019) described that engineering graduates in India lacked technical skills, and about 30 percent of engineering applicants failed the aptitude test. As Luțaș et al. (2020) stated, competent project managers should possess education in the project field and soft and technical skills and knowledge to perform specific tasks in the team project, such as competence to perform mechanical, information technology, mathematical, or scientific tasks. Cleland (1994) also claimed that project managers should perform technical tasks of the project and possess project management knowledge. Therefore, required technical skills include knowledge of the technology involved in a successful project, and technical competencies are an essential factor of the organisation's tacit knowledge assets and organisational capability (E1-Sabaa, 2001; Koskinen et al., 2003; Söderlund, 2005). To address this point, the project-based collaborative effort takes place in technical skills by designing the team projects in this study.

Historical research of Callaghan et al. (1994) listed several changes in the teaching approach in the group projects. These included active individual learning, guided personal growth, and group learning to enhance creativity and social skills. Fine (2012) identified that building social structures can shape (a) individuals in social identities, (b) social capital by strengthening group ties based on individuals fulfilling one's goals, and (c) in making small-scale networks. He emphasised that small group interaction can be the "micro foundation of civic society". Those changes are part of skills in project management as well. Bowers et al., (2000) and Cannon et al., (1997) also divided team competencies into three categories — *knowledge, skills, and attitudes*. In terms of project management, Koskinen et al., (2003) and E1-Sabaa (2001) agreed that soft skills are the main components of project management. Soft skills are needed in all jobs in all businesses, particularly in leadership positions. Indeed, soft skills are personal attributes that interact with others in teamwork. These skills make it easier to form relationships with people, create trust, and lead teams. As previous research indicates, Technological Savvy skills are essential in business projects, and they augment soft skills that help minimise underperformance of members in project management teams. Hence, the objective of this project is "Finding the skill sets of Technological Savvy skills and soft skills that minimise underperformance of members in project management teams". Group-based student projects are a preliminary study on teamwork experiences by the evolution of team skills.

To reach the objectives above, a technological savvy skill model (TSS) and a soft skill model (SSM) were developed to deal with this goal by assessing project-based group work. For the soft skill model, initially, attribute evaluation was applied to identify the correlation attributes which are "Good and Very Good Leadership Skills", "Good and Very Good Communication Skills", and "Good and Very Good Logical Skills" since those attributes have a significant impact on the soft skill model. For the TSS model, the attribute evaluation was based on the correlation attributes of "Good and Very Good Programming/Technological Savvy Skills" and "Good and Very Good Logical Skills". The observation was that the conditional probability between two skills affected the project grade, and classification accuracy was applied to find the soft skill attributes that influenced project complexity. The contributions can be summarised as 1) analysing the datasets revealed that savvy technical skills and soft skills are essential skills for a successful project, and 2) a TSS model and SSM model were used for the first time to identify such a phenomenon (which provided the reference models for the different nature of group-based student projects).

Method

In our case study, the model accuracy was experienced when the conditional probability between two node values was P (Leadership skill =0 | Logical skill =0) = 1. The solving method of the posterior probability is P (Logical skill | Leadership skill) = P (Leadership skill | Logical skill) * P (Logical skill)/P (Leadership skill), here P (Leadership skill | Logical skill), P (Leadership skill), P (Logical skill) and P (Logical skill | Leadership skill) are based on Bayes' theorem, with the conditional probabilities illustrated in 2.1.

$$P\ (Logical\ skill \mid Leadership\ skill) = \frac{P(Logical\ skill\ ) * P(Leadership\ skill \mid Logical\ skill\ )}{P(Leadership\ skill)} \tag{2.1}$$

The notation of the classification problems states in 2.2.

$$\mathrm{P}\ (S_k|x) = \frac{P(S_k) * P(x|S_k)}{P(x)} \tag{2.2}$$

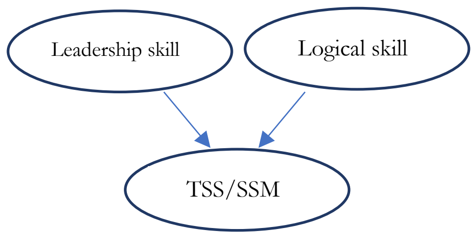
Where,

- $P(S|x)$ is the posterior probability of a soft skill (target variable) that is the given predictor x.
- $P(S)$ is the prior probability of a soft skill.
- $P(x|S_k)$ is the likelihood, which is the probability of the predictor x given a soft skill.
- $P(x)$ is the prior probability of the predictor x.
- k is the notation to distinguish between different soft skills of at least two soft skills in the classification scenario such as good leadership skill / very good leadership skill, good logical skill / very good logical skill).

For example, if a student has excellent leadership skills in a team project, that student can have soft skills. A graphical representation is shown in Figure 1.

Figure 1

The relationship of variables in the conditional probability

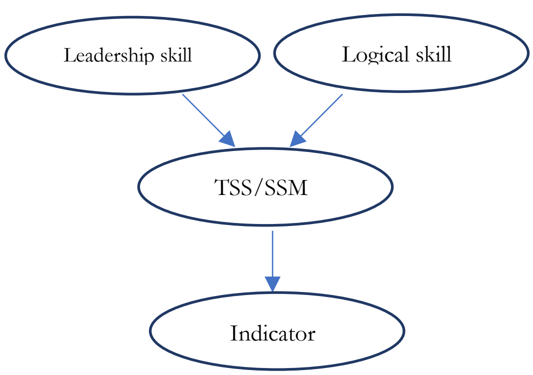


Model specification

The models are an extension of a Bayesian network (1997) defined as an acyclic graph. The consideration of the nodes in the acyclic graph is represented as the random variable $X_{i,\ldots\ldots}X_n$ and for each X_i is a conditional probability distribution. The edges indicate the impact of one node on another. Figure 2 shows the relationship between a prediction indicator and soft skills.

Figure 2

The relationship between a prediction indicator and soft skills



The data are classified in terms of performance as Good, Pass and Marginal for evaluating teamwork performance and finding the proper skills that influence project performance. The next step is to perform attribute evaluation using the attribute evaluator with correlation evaluation to make the features suitable to the model and eliminate noise features. The searching method used is "Ranker" in Weka to identify which attributes have the most significant impacts on team performance.

Hence, the Soft Skill Model (SSM) is specified by the conditional probabilities P (G, P, M), where G is a good class, P is the pass class, and M is the marginal class. Any three events of G, P and M, are independent, then, P (G, P, M) = P(G) P(P) P(M) and the SSM model can be defined as SSM (n, P (G, P, M)). Let G and P be the passing grades, and they are independent, thus,

$$\mathrm{P}(\mathrm{y}|x_{1,\ldots\ldots,}x_n) = \frac{P(x_1|y)P(x_2|y)\ldots\ldots P(x_n|y)P(y)}{P(x_1)P(x_2)\ldots.P(x_n)}$$
$$\mathrm{P}(\mathrm{y}|x_{1,\ldots\ldots,}x_n) = \frac{P(y)\prod_{i=1}^{n} \quad P(x_i|y)}{P(x_1)P(x_2)\ldots.P(x_n)}$$
$$\mathrm{P}(\mathrm{y}|x_{1,\ldots\ldots,}x_n) \propto P(y)\prod_{i=1}^{n} \quad P(x_i|y)$$

So, the maximum probability is $y = max_y\, P(y) \prod_{i=1}^{n} \quad P(x_i|y)$

Feasibility study

A new area of machine learning is growing in status; a machine playing chess is a well-known example. It is practically possible to implement into high-end appliances. In real life, data has thousands of dimensions, and it is not easy to find the right data features for machine learning that can handle the high dimensionality of data.

In this study, the first step was to analyse students' academic records and extract information about their English skills, leadership skills, communication skills, technology savvy skills, logical skills, and hardware skills. Nearly 473 records were collected from students of three project-based subjects (each from an undergraduate program) on their academic performance in the subjects relating to programming/technological study, hardware development, and generic IT study throughout three academic years. Then, these subjects were mapped to the relevant skills. We used those data to train a machine learning model.

The teamwork-performance prediction system was developed to apply the models and display the performance of team members of student project teams. The system is a web application that is hosted by the University's Information Technology Services Centre. Each team member is requested to complete two assessments for their projects and encourage each team member to examine their skills.

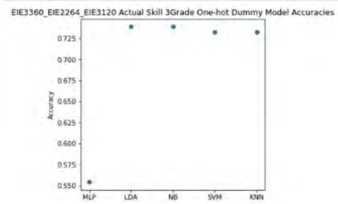
The regression and support vector machine algorithms give a higher accuracy rate than the other machine learning algorithms in the TSS model by inputting the data to the machine learning algorithms. Thus, the regression and support vector machine algorithms were used for data training. Because we were feeding a small dataset, the SSM model used Naïve Bayes (2011) to train the data as Naïve Bayes provided acceptable accuracy scores (see figure 3) compared with the other machine learning algorithms. Also, the Bernoulli Naïve Bayes was selected because our data had multiple features, and each one was assumed to be a binary variable.

Member 1's score is combined with Member 2's input testing data for the TSS model. The code for importing the regression algorithm and support vector machine algorithm comes from the sci-kit learn library. The result is only two and four group underperformances in figure 4.

Figure 3

The accuracy scores of the Naïve Bayes algorithm

```
accuracy_data = {'MLP': accuracy_mlp, 'LDA': accuracy_lda, 'NB': accuracy_nb, 'SVM': accuracy_svm, 'KNN': accuracy_knn}
accuracy_names = list(accuracy_data.keys())
accuracy_values = list(accuracy_data.values())
plt.figure(figsize=(5, 5))
plt.scatter(accuracy_names, accuracy_values)
plt.ylabel('Accuracy')
plt.title('EIE3360_EIE2264_EIE3120 Actual Skill 3Grade One-hot Dummy Model Accuracies')
plt.show()
```




```
accuracy_nb = accuracy_score(y_test, y_pred_nb)
accuracy_nb
```

```
0.7388535031847133
```

Figure 4

The testing data for each task and the codes for regression and support vector machine algorithms

```
test_x
```

Member 1 Project Proposal	Member 1 Project Preliminary Demo	Member 1 Peer Review	Member 1 Final Report and Logbook	Member 1 Presentation	Member 1 Final Project Demo	Member 1 Project Management	Member 2 Project Proposal	Member 2 Project Preliminary Demo	Member 2 Peer Review	Member 2 Final Report and Logbook	Member 2 Presentation	Member 2 Final Project Demo
0.750	0.666667	1.0	0.833333	0.70	0.866667	0.70	0.750	0.666667	1.0	0.833333	0.70	0.866667
1.000	0.533333	1.0	0.783333	0.66	0.800000	0.46	1.000	0.533333	1.0	0.800000	0.66	0.800000
0.875	0.466667	1.0	0.666667	0.64	0.740000	0.44	0.875	0.466667	1.0	0.708333	0.64	0.326667
0.875	0.666667	1.0	0.791667	0.94	0.890000	0.26	0.875	0.666667	1.0	0.775000	0.94	0.890000
0.875	0.666667	1.0	0.841667	0.64	0.833333	0.46	0.875	0.666667	0.8	0.841667	0.64	0.833333
0.750	0.780000	1.0	0.841667	0.96	0.903333	0.50	0.750	0.780000	0.8	0.841667	0.96	0.903333
0.875	0.666667	1.0	0.758333	0.70	0.866667	0.44	0.875	0.666667	0.8	0.775000	0.72	0.866667
0.750	0.666667	1.0	0.833333	0.70	0.866667	0.50	0.750	0.666667	1.0	0.833333	0.70	0.866667
1.000	0.533333	1.0	0.800000	0.66	0.800000	0.56	1.000	0.533333	1.0	0.783333	0.66	0.800000
0.875	0.466667	1.0	0.708333	0.64	0.326667	0.68	0.875	0.466667	1.0	0.666667	0.64	0.740000
0.875	0.666667	1.0	0.775000	0.94	0.890000	0.36	0.875	0.666667	1.0	0.791667	0.94	0.890000
0.875	0.666667	0.8	0.841667	0.64	0.833333	0.36	0.875	0.666667	1.0	0.841667	0.64	0.833333
0.750	0.780000	0.8	0.841667	0.96	0.903333	0.46	0.750	0.780000	1.0	0.841667	0.96	0.903333
0.875	0.666667	0.8	0.775000	0.72	0.866667	0.64	0.875	0.666667	1.0	0.758333	0.70	0.866667

```
from sklearn.linear_model import LogisticRegression
lr = LogisticRegression()
lr.fit(train_x, train_y)
test_y = lr.predict(test_x)
test_y
```

```
array([0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0], dtype=int64)
```

```
from sklearn.svm import SVC
svm = SVC()
svm.fit(train_x, train_y)
test_y = svm.predict(test_x)
test_y
```

```
array([0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 1, 0, 1, 0], dtype=int64)
```

In the case study of technological savvy skill performance prediction, the normalisation of project attributes was calculated for obtaining a consistent input for data training, i.e., Normalised project report (NPR) = actual report score/proportions of the report, Normalised presentation (Npresent) = actual presentation score/proportions of presentation, Normalised project demo (NPD) = exact demo score/proportions of project demo, Normalised peer assessment (NPA) = actual peer assessment score/proportions of peer assessment, Normalised project proposal (NPP) = exact proposal score/proportions of the project proposal, and Normalised attendance (NA) = actual attendance score/proportions of attendance. Thus, $N(X_t) \in$ NPR, $N_{present}$, NPD, NPA, NPP, NA. The output training data determined underperforming members by comparing their actual project scores. The input and output training data for each task, which combined member 1 and member 2 scores, are shown in figure 5.

Sample

The datasets from the subjects to be involved in this project included a small-scale assessment. The results of these assessments provided data for understanding the impact of the SSM model and TSS model. The dataset is from:

Figure 5

An example of training data

train_x

Member 1 Project Proposal	Member 1 Project Preliminary Demo	Member 1 Peer Review	Member 1 Final Report and Logbook	Member 1 Presentation	Member 1 Final Project Demo	Member 1 Project Management	Member 2 Project Proposal	Member 2 Project Preliminary Demo	Member 2 Peer Review	Member 2 Final Report and Logbook	Member 2 Presentation	Member 2 Final Project Demo
0.750	0.700000	0.8	0.750000	1.00	0.680000	0.50000	0.750	0.700000	0.8	0.750000	1.00	0.680000
0.875	0.900000	0.8	0.166667	1.00	0.802400	0.30625	0.875	0.900000	0.8	0.000000	1.00	0.802400
0.875	0.800000	0.8	0.916667	1.00	0.920000	0.64125	0.875	0.800000	0.8	0.916667	1.00	0.920000
0.750	0.726667	1.0	0.783333	0.90	0.890000	0.66000	0.750	0.726667	0.8	0.783333	0.90	0.890000
0.750	0.600000	0.8	0.833333	1.00	0.640000	0.48375	0.750	0.600000	0.8	0.833333	1.00	0.640000
0.750	0.666667	1.0	0.791667	0.70	0.866667	0.52000	0.750	0.666667	1.0	0.791667	0.70	0.866667
0.875	0.677000	0.8	0.750000	1.00	0.875200	0.40625	0.875	0.677000	0.8	0.750000	1.00	0.875200
0.750	0.533333	1.0	0.858333	0.64	0.850000	0.40000	0.750	0.533333	1.0	0.858333	0.64	0.850000
0.875	0.000000	0.8	0.750000	1.00	0.780000	0.39875	0.875	0.600000	0.8	0.750000	1.00	0.680000
0.875	0.553333	1.0	0.750000	0.70	0.776667	0.42000	0.875	0.540000	1.0	0.733333	0.70	0.770000
0.750	0.677000	0.8	0.750000	1.00	0.720000	0.32750	0.750	0.677000	0.8	0.750000	1.00	0.720000

train_y

0	1
1	1
2	1
3	0
4	1
5	0
6	1
7	0
8	1
9	0
10	1

Computer Programming course (EIE2264) is a level-2 subject that gives higher diploma students essential programming skills. This subject provides an environment for students to develop logical skills through a small-scale software development group project and software development project with TWO students as one group. This sample can be relevant to the TSS model.

Network Technologies and Security course (EIE3120) is a level-3 subject that gives the top-up program students a valuable opportunity to exploit their personal development through a case study project with THREE students as one group. This subject provides an environment for students to develop communication skills and an innate desire to learn. This sample can be appropriate to the SSM model.

Integrated Project course (EIE3360) is a level-3 subject which is a software development project lasting for 13 weeks with TWO students as one group. This subject allows students to understand their learning habits and develop their leadership and project management skills/knowledge through an integrated group project. This subject leads to leadership improvement through their peers' assessment in two stages of evaluation. This sample can apply to the TSS and SSM models.

The total number of records was 473. We converted the numerical dataset into the binary-valued qualitative attributes for Bayesian classifiers so that the dataset contained 316 rows and 30 attributes on which to conduct training. The converted dataset for training is shown in figure 6.

Then, the input training data was scaled to a zero mean for better convergence of the training algorithm (see figure 7). The output test data is the 100-times normalised group score rounded to an integer split from the whole data set. It is used for verifying the accuracy of the training algorithm.

Figure 6

The converted dataset with binary-values

```
x_train, x_test, y_train, y_test = train_test_split(
...     x, y, test_size=0.33, random_state=42)
```

```
x_train
```

	English Poor	English Satisfactory	English Good	English Very Good	English Excellent	Creativity Poor	Creativity Satisfactory	Creativity Good	Creativity Very Good	Creativity Excellent	...	Communication Poor	Communication Satisfactory	Commu
424	0	0	0	1	0	0	0	1	0	0	...	0	0	
371	0	0	1	0	0	0	0	0	0	1	...	0	0	
381	0	0	0	1	0	0	1	0	0	0	...	0	1	
422	0	0	0	1	0	0	0	1	0	0	...	0	0	
141	0	1	0	0	0	0	0	1	0	0	...	0	1	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	
106	0	0	1	0	0	0	1	0	0	0	...	0	1	
270	0	0	0	1	0	0	0	1	0	0	...	0	0	
348	0	1	0	0	0	0	0	0	1	0	...	0	0	
435	0	0	0	1	0	0	0	1	0	0	...	0	0	
102	0	1	0	0	0	0	0	1	0	0	...	0	0	

316 rows × 30 columns

Figure 7

The scaled data for training

```
x_train_scaled = preprocessing.scale(x_train)
x_train_scaled
```

```
array([[ 0.39964228,  0.35775601,  0.22298824],
       [-0.62560036, -1.22994949,  0.22298824],
       [ 0.39964228,  0.35775601,  0.22298824],
       [ 0.39964228,  0.35775601,  0.22298824],
       [ 0.39964228,  0.35775601,  0.22298824],
       [ 0.39964228,  0.35775601,  0.22298824],
       [-3.50909526, -4.50459209, -2.60152951],
       [ 0.39964228,  0.35775601,  0.22298824],
       [-0.68967802, -1.22994949,  0.22298824],
       [ 0.39964228,  0.35775601,  0.22298824],
       [ 0.39964228,  0.35775601,  0.22298824],
       [ 0.39964228,  0.35775601,  0.22298824],
       [ 0.39964228,  0.35775601,  0.22298824],
```

Assumptions

Two assumptions were made: the first was the nature of skillsets associated with teamwork performance, while the second was the direct relationship of significant skills (i.e., programming skills, logical skills) between an outcome variable like the indicator "Good" performance, "Pass" performance and "Marginal" performance. Referring to McIntosh's (2012) theory, the assumption is that the potential values of outcome variable Y have an equivalent probability as a function of values of independent variables is shown as equation 4.1,

$$\mathrm{P}(Y_{ith} = 1/\mathrm{X}=\mathrm{x}) = \mathrm{P}\,(Y_{ith} = 1) \qquad (4.1)$$

P (Y_{ith} =1) is the probability of the conditional success for the i^{th} subject given a certain value of X. With relevance to Plane et al., (1997), these probabilities follow a binomial distribution. Based on Hanusheck et al., (1997) and Aldrich et al., (1986), "a logistic distribution is S-shaped and the values ranging from 0 to 1" as W_1X_1 change is shown in equation 4.2.

$$\mathrm{P(Y)} = \frac{\exp\,(W_0+W_1X_1)}{1+\exp(W_0+W_1X_1)} \qquad (4.2)$$

From the nature of the likelihood function and its properties (with reference to the discussion in Christensen, 1990; Collett, 1991; Hanusheck et al., 1997; Fienberg, 1983), the underlying relationship between an independent variable and a dependent variable should follow a probability function that is S-shaped in nature. The maximum likelihood estimated by logistic regression measures the effects of the independent variable upon the probability function. As Hanusheck et al. (1997) stated, the approximated values would likely have observed Y with the highest probabilities. Minimising the sum of the squared errors between the observed and predicted values is required. The python library was used to handle the regression model. The simulation theory was based on St. John's (1990) study, which acquired the effect of academic ability. He provided an excellent illustration of how logistic regression categorical variables apply in categorical dependent variables.

Soft Skill Model

Our small dataset assumes class independence and limited features because of soft skills, including leadership, communication, and logical skills. Thus, no pair of features depends on a student with good leadership skills associated with a good grade. The features are then supposed to be independent. The second is that each feature (soft skill) affixed the same weight, such that good leadership skill and logical skill alone cannot predict the grade accurately.

- C_i is the total number of class labels.
- N is the total number of attributes.
- $\{ T_{1,\ldots\ldots,} T_n \}$ is the total number of values that correspond to the attributes $\{ A_{1,\ldots\ldots,} A_n \}$.
- T_{max} is the maximum of $\{ T_{1,\ldots\ldots,} T_n \}$

T_i is the total number of training instances and counting the number of existences of each class label, and attribute can be processed as E(T_i). The probability of each class label can be found in the training dataset like E (C_i). For the conditional probability of each attribute, the probability of every single attribute is conditioned to find each class label like "Good" teamwork performance, "Marginal" teamwork performance. So,

$$\begin{aligned} \text{SSM} &= \textstyle\sum_{i=1}^{N} C_i(T_i) \\ &= C_i(T_1) + C_i(T_2) + \ldots\ldots\ C_i(T_n) \\ &\leq\ C_i\ (T_{max} +\ T_{max} + \cdots ..\ T_{max}) \\ &\leq \text{N}(C_i)(T_{max}) \end{aligned} \tag{5.1}$$

Then, calculating conditional probability is given by N$(C_i)(T_{max})$. The conditional probabilities and prior probabilities are collected and then retrieved during the classification process. When we classify the test data, the model has to find each instance's possibilities in each class label. Hence, the SSM model is determined by finding the conditional probabilities for a feature value, which means K classes and n variables are K * n. The features presence and independence of each class are as $P(x_{1,} x_2\,|C) = P(x_1|C)\ P(x_2|C)$. It is appropriate to apply this in the binomial distribution. Additionally, Cohen's kappa (1985) theory was used to maximise the overall accuracy as $K_{max} = \frac{P_{max} - P_e}{1 - P_e}$, where P_{max} is the model's maximum overall accuracy.

Technological Savvy Skill Model

The Technological Savvy Skill Model (TSS) for teamwork performance prediction framework is shown as equation 6.1.

$$\text{TSS} = \frac{w_i}{\sum_k w_k} \left(T_{ts} \right) \tag{6.1}$$

Where i = {1…6} is the number of skills sets including English skills, Creativity skills, Programming skills, Leadership skills, Communication skills and Logical skills in a project, a qualification matrix Q (m_x n) expresses the project role i ∈ N ($0 \leq i \leq m$). The weight of each skill is $W_1X_1 + W_2X_2 + W_3X_{3\ldots\ldots}W_6X_6$. Each X ∈ {1….5} of scale, $X \in \mathbb{R}^6$ and $W \in \mathbb{R}^6$. Thus, the savvy technical skill, T_{ts}, represents the exact role of the team member in the stages of a project.

To find the technological savvy skill on student interactions, X_t is a one-hot encoding of students' skill, and maps the categorical value to numerical vectors and $T_{ts} = \{q_{ts,}a_{ts}\}$ that represents the combination of which skills are needed in the project $T_{ts} = \{q_{ts,}\ a_{ts,}\}$ and $X_t \in \{0, 1\}$. For instance, if each attribute has 5 levels (1-5), a one-hot vector with five elements represents it. Level-1 is represented by (1,0,0,0,0), Level-2 by (0,1,0,0,0), and so on. Then, for 6 attributes with 5 levels for each attribute, the input numbers are 6x5=30 inputs. The 6 attributes (30-dimensional binary vectors) are input, and their corresponding target scores are the desired output.
The same principle applies to the case in which each attribute is binary (has/not-has). The output y_{ts} is a vector of length equal to the number of skills, where it is assumed, each student has more than one technical skill, a_{ts+1} that can be found from the entry in y_{ts} corresponding to $q_{ts+1.}$

A comprehensive evaluation of the technological savvy skill prediction framework was used in the students' case. The dataset was from three years of department student cases, including project marks and academic results applied to the TSS model. We achieved a mean correlation between prediction indicator and technological savvy skill of 78.3 percent.

Figure 8

The correlation between prediction score and savvy technological skills

	PredScore	Eng_VG	Eng_Excel	Creat_VG	Creat_Excel	Prog_VG	Prog_Excel	Leader_VG	Leader_Excel	Comm_VG	Comm_Excel	Log_VG	Log_Excel
PredScore	1.000000	0.465343	0.259520	0.477932	0.297675	0.411236	0.322963	0.443230	0.310066	0.404122	0.303563	0.370301	0.292990
Eng_VG	0.465343	1.000000	-0.480148	0.763085	-0.426730	0.715968	-0.404486	0.706799	-0.352429	0.604356	-0.352429	0.658409	-0.465968
Eng_Excel	0.259520	-0.480148	1.000000	-0.592628	0.940897	-0.540495	0.886463	-0.513747	0.886463	-0.487953	0.886463	-0.357188	0.970467
Creat_VG	0.477932	0.763085	-0.592628	1.000000	-0.592628	0.862051	-0.475333	0.908948	-0.526136	0.865581	-0.526136	0.598439	-0.575125
Creat_Excel	0.297675	-0.426730	0.940897	-0.592628	1.000000	-0.540495	0.886463	-0.565962	0.944060	-0.540363	0.944060	-0.412978	0.970467
Prog_VG	0.411236	0.715968	-0.540495	0.862051	-0.540495	1.000000	-0.627743	0.816835	-0.475333	0.819353	-0.526136	0.598439	-0.575125
Prog_Excel	0.322963	-0.404486	0.886463	-0.475333	0.886463	-0.627743	1.000000	-0.446845	0.831613	-0.470234	0.887742	-0.328711	0.916179
Leader_VG	0.443230	0.706799	-0.513747	0.908948	-0.565962	0.816835	-0.446845	1.000000	-0.599497	0.862167	-0.548613	0.680404	-0.549247
Leader_Excel	0.310066	-0.352429	0.886463	-0.526136	0.944060	-0.475333	0.831613	-0.599497	1.000000	-0.521308	0.943871	-0.437449	0.916179
Comm_VG	0.404122	0.604356	-0.487953	0.865581	-0.540363	0.819353	-0.470234	0.862167	-0.521308	1.000000	-0.572382	0.566372	-0.524404
Comm_Excel	0.303563	-0.352429	0.886463	-0.526136	0.944060	-0.526136	0.887742	-0.548613	0.943871	-0.572382	1.000000	-0.383080	0.916179
Log_VG	0.370301	0.658409	-0.357188	0.598439	-0.412978	0.598439	-0.328711	0.680404	-0.437449	0.566372	-0.383080	1.000000	-0.400782
Log_Excel	0.292990	-0.465968	0.970467	-0.575125	0.970467	-0.575125	0.916179	-0.549247	0.916179	-0.524404	0.916179	-0.400782	1.000000

In figures 8 and 9, the Pearson correlation measured how well the skill variables were related to teamwork performance. The implication for those teammates is that good programming, creativity, communication, and logical skills are strongly associated with savvy technological skills. In other words, if the prediction score is more excellent than 50%, teammates with creativity, programming, communication and logical skills are relatively high on savvy technological skills (see figure 10).

Teammates with a higher value of creativity skills indicated high predictability of project scores with a value of 0.5, illustrating that savvy technological skills need very good creative skills. Thus, teammates with good creativity skills are the best predictor of the technological savvy skill model (TSS). However, teammates with excellent English skills imply a negative correlation with the TSS.

Figure 9

The correlation between prediction score and savvy technological skills
(with randomly selected samples)

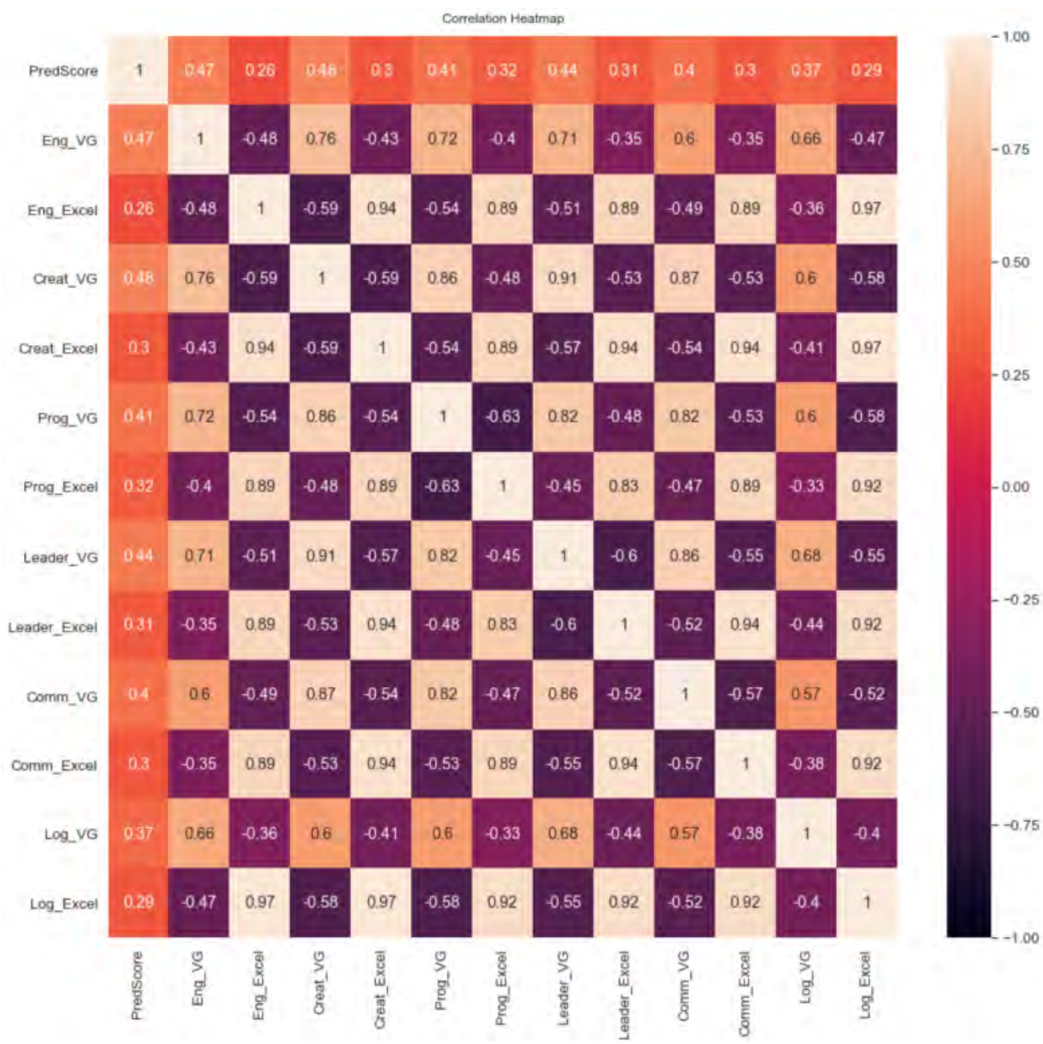
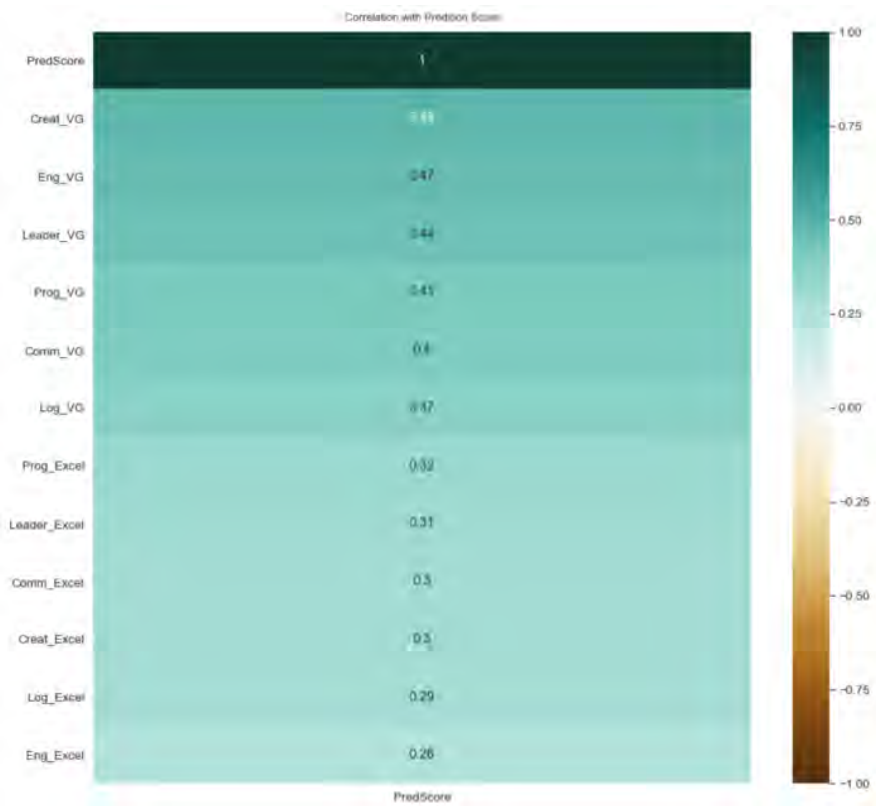


Figure 10

The coefficient of multiple skills with the prediction score



With reference to the inspections of Efron and Gong (1983), bootstrap plots were used to estimate the uncertainty of the dataset of students' marks. The subsamples were repeated, and the range between 0 and 500 was applied to compute the values of the sampling distribution (see figure 11).

In figure 11, uncertainty estimation was discovered in the median from a dataset with 13 elements. A subsample of 13 pieces was generated, and the median was calculated. Fifty data points were considered during each sampling, and the bootstrap procedure was performed 500 times. The corresponding histograms were generated for the mean, median, and mid-range. The mean histogram approximates a typical distribution even though the underlying data distribution is skewed, and the related statistics also have the most insignificant variance.

The data were divided into a training phase and a prediction phase for classification accuracy to find the confidence interval and error about the model's performance with corrected or incorrect predictions. Following the classification accuracy, the interval range can be calculated as in equation 6.1.1.

$$\text{Interval} = \sigma\sqrt{\frac{accuracy(1-accuracy)}{n}} \tag{6.1.1}$$

The interval range is the estimated value of the confidence interval, error n is the sample size, and σ is the standard deviations from the Gaussian distribution. The generally used number of standard deviations and corresponding significance level from the Gaussian distribution is shown in table 1.

Figure 11

The bootstrap plot with 500 uniform random numbers

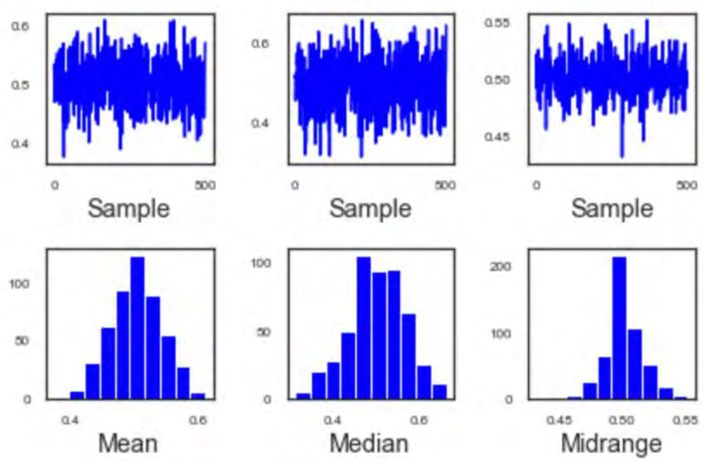


Table 1

The confidence level and standard deviations from mean

Confidence Level	Standard deviation from mean
90%	1.64
95%	1.96
98%	2.33
99%	2.58

Assuming a model with 20% error, (error = 0.2) and a prediction phase data with 88 examples (n = 88). Then, the estimation of 95% confidence interval (σ = 1.96) is used for finding the confidence interval such that,

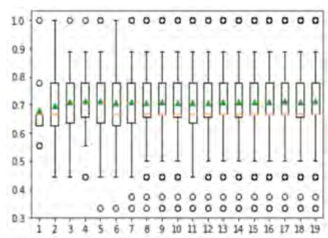
```
interval = 1.96*sqrt((0.2*(1-0.2))/88)
print ('%.3f' % interval)
0.084
```

The calculated value of the confidence interval is 0.084, which shows the classification error of the model is 20% +/- 8.4%. 95% interval is estimated as the confidence interval. The k-fold cross-validation procedures assessed the configuration performance on a dataset as well.

Figure 12

10-fold cross-validation result

```
>1 mean=0.6806 se=0.041
>2 mean=0.6979 se=0.028
>3 mean=0.7116 se=0.021
>4 mean=0.7156 se=0.019
>5 mean=0.7136 se=0.020
>6 mean=0.7083 se=0.017
>7 mean=0.7091 se=0.016
>8 mean=0.7087 se=0.016
>9 mean=0.7099 se=0.015
>10 mean=0.7079 se=0.014
>11 mean=0.7076 se=0.013
>12 mean=0.7075 se=0.012
>13 mean=0.7098 se=0.012
>14 mean=0.7096 se=0.011
>15 mean=0.7100 se=0.011
>16 mean=0.7119 se=0.011
>17 mean=0.7129 se=0.010
>18 mean=0.7119 se=0.010
>19 mean=0.7132 se=0.010
```



In figure 12, 10-fold cross-validation was applied with different repeats to validate the dataset's model performance. Twenty repeated procedures gave a more accurate evaluation of the mean model performance—the logistic regression model from the python library was used as a solver for evaluating a score by cross-validation. Estimating and computing the score were parallelised over the cross-validation splits. A square bar and whisker plot summarise the distribution of the scores for each number of repeats. The line inside and the line on the tail of the square bar indicate the median of the distribution. The triangle inside/outside the square bar represents the arithmetic mean. An asymmetric distribution was obtained, which means the model indicates nearly a central tendency.

Furthermore, a repeating process appears as positive accuracy of about 71% (see the left-hand side output in figure 12). Consequently, it seems the mean merged around a value of about 71%. Stable model performance and the standard error appear to decrease with increasing repeats and stabilise the value of 0.011 in the 14 or 16 repetitions. As a result, the repeated 10-fold cross-validation can diminish the error in the model's performance.

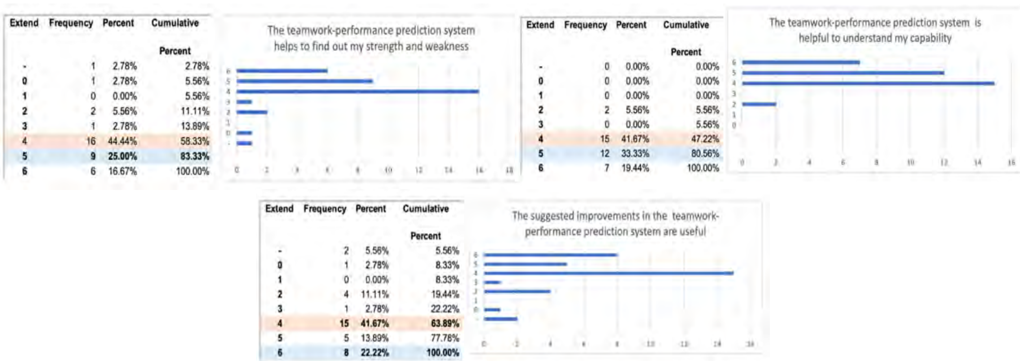
Findings

The TSS model demonstrated exemplary performance in the previous sections. The descriptive statistics are provided in this section. The participating students were requested to complete the online user acceptance survey at the end of the semester. The survey asked students to check the box that best described their observations, the notation being Strongly disagree=1, Disagree=2 (tend to disagree more than not), Disagree=3, Agree=4, Agree=5 (tend to be more agreeable), Strongly agree=6 and Not applicable=0. The questions comprised aspects of recommended skill improvements and establishing self-capability. Over 60% of students agreed that they understood their self-capability, strengths, and weaknesses. Around 50% of students thought the recommendations for skill improvements from the prediction system were practical (see the survey results in figure 13).

The online survey also looked for open-ended questions to indicate students' feedback and the impacts on the working procedures. Table 2 shows the summary of those answers.

Figure 13

The survey results about the self-capability and recommendations of skill improvements



Extend	Frequency	Percent	Cumulative Percent
-	1	2.78%	2.78%
0	1	2.78%	5.56%
1	0	0.00%	5.56%
2	2	5.56%	11.11%
3	1	2.78%	13.89%
4	16	44.44%	58.33%
5	9	25.00%	83.33%
6	6	16.67%	100.00%

Extend	Frequency	Percent	Cumulative Percent
-	0	0.00%	0.00%
0	0	0.00%	0.00%
1	0	0.00%	0.00%
2	2	5.56%	5.56%
3	0	0.00%	5.56%
4	15	41.67%	47.22%
5	12	33.33%	80.56%
6	7	19.44%	100.00%

Extend	Frequency	Percent	Cumulative Percent
-	2	5.56%	5.56%
0	1	2.78%	8.33%
1	0	0.00%	8.33%
2	4	11.11%	19.44%
3	1	2.78%	22.22%
4	15	41.67%	63.89%
5	5	13.89%	77.78%
6	8	22.22%	100.00%

Table 2

Summary of the open-ended questions

How do you feel about discussing their grade with your team members?
1. There is not much argument on the grade since we have admitted what the system told us.
2. Enjoyable. I love the whole process of solving the problem. It fulfils our curiosity.
3. I think it may help us to understand each other for a little bit.
4. It is great. We all know the strengths and weaknesses of each other and support each other.

Overall, the statements from students are quite favourable opinions. They thought the study procedures could be advantageous in understanding team project requirements and recognising their team members' capability. Besides, the statements show the high-level effectiveness of the teamwork-performance prediction system in increasing course performance.

The impacts of SSM model

The calculation of P (G, P, M) for each x_i in X and y_j in y is shown in figure 14. From the testing data, we discovered that the probability of "Pass" had higher accuracy than the probability of "Good" and "Marginal", and the probability was greater than 0.5 (the statistical results are shown in figure 15). Therefore, this can help to classify the accuracy as true positives and true negatives.

As for evaluating the model's reliability, the kappa coefficient was applied to measure how the classifier matched the data ground truth. In the classification problem, the probability of the predictions corresponding with the actual values of class 1 (Pass) and the possibility of the predictions agreeing with the true values of class 2 (Marginal) means the training set's classification accuracy is explicit. The two classifiers are independent and computed by the actual class's proportion and the predicted class's ratio. Let's consider "Marginal" as the positive class, 74% of records (false positives + true positives) to class "Marginal" and 26% of the records (true negatives

+ false negatives) to class "Pass" (see figure 15). Thus, P_e is $P_{e1} + P_{e2} = (P_{e1,target})(P_{e1,predict}) + (P_{e2,target})(P_{e2,predict})$.

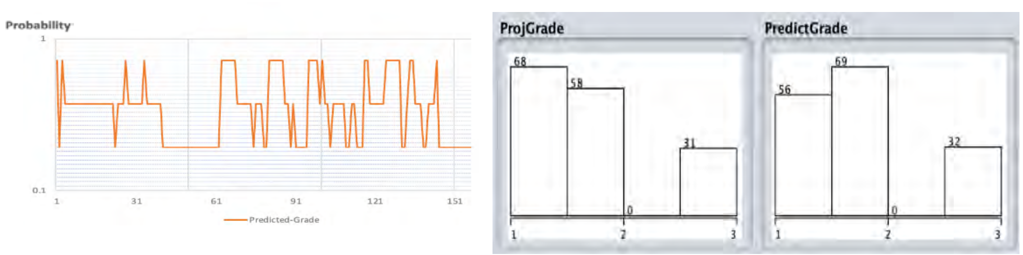
Figure 14

The partial probability results for "G", "P" and "M."

Project-Grade	Predicted-Grade	Marginal Probability	Pass Probability	Good Probability
2	3	0.020179789	0.280543388	0.699276823
1	1	0.935276456	0.059564932	0.005158611
3	3	0.028981988	0.191865645	0.779152367
2	2	0.1871514	0.774989927	0.037858673
2	2	0.019911056	0.920431438	0.059657507
2	2	0.002311207	0.845139974	0.152548819
2	2	0.019911056	0.920431438	0.059657507
2	2	0.002311207	0.845139974	0.152548819
2	2	0.002311207	0.845139974	0.152548819
2	2	0.026169362	0.857212908	0.11661773
2	2	5.49E-05	0.933665356	0.066279749
2	2	0.1871514	0.774989927	0.037858673
1	2	0.107397964	0.780130432	0.112471603
1	2	0.002311207	0.845139974	0.152548819
2	2	0.026169362	0.857212908	0.11661773
2	2	0.010875415	0.544643701	0.444480884
2	2	0.002311207	0.845139974	0.152548819
2	2	0.002311207	0.845139974	0.152548819
2	2	7.22E-05	0.870257727	0.129670065
2	2	0.019911056	0.920431438	0.059657507
2	2	0.019911056	0.920431438	0.059657507
1	2	0.002311207	0.845139974	0.152548819
1	1	0.97341496	0.025042569	0.001542471
2	2	0.1871514	0.774989927	0.037858673
2	2	0.019911056	0.920431438	0.059657507
2	2	6.21E-06	0.834931776	0.165062019
3	3	0.002113739	0.156212046	0.841674215
2	2	0.462668645	0.511926501	0.025404855
2	2	0.045915508	0.941759315	0.012325177

Figure 15

The probability distributions of the predicted grade



After manipulating the confusion matrix, class 2 is 90% true negatives, and only 7.6% is a false negative, which means the difference is significant. The accuracy of all corrected predictions is higher than the misclassification.

To find the significant attributes that can affect the soft skill model, the performance metrics were extended to compare the different characteristics in the dataset. In figure 17, the attributes of "G-Log" and "VG-Log" (Good and Very Good Logical Skill) shows a higher accuracy score of 88%. The precision for class 2 is true positives and suitable for our model as the "recall" is True Positive/True Positive + False Negative. The precision is 80%, and an excellent "F1" score (over 85%) means low false positives and low false negatives. Moreover, the remarkable result of the two classes almost correctly identified the positive cases from all the actual positive cases. The result tells us those team members with good and very logical skills can maximise the team project's soft skills.

Figure 16

The confusion matrix between the predicted grade, actual grade, and class 1 & class 2

Predicted Grade	1	2
Actual Grade		
1	0.742424	0.34
2	0.075758	0.90

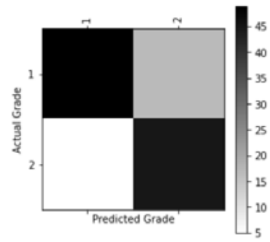


Figure 17

The performance metrics of the "G-Log and VG-Log" attributes

Accuracy Score: 0.875

	precision	recall	f1-score	support
1	0.80	1.00	0.89	16
2	1.00	0.75	0.86	16
accuracy			0.88	32
macro avg	0.90	0.88	0.87	32
weighted avg	0.90	0.88	0.87	32

In figure 18, the attributes of "G-Comm" and "VG-Comm" (Good and Very Good Communication Skill) have a higher accuracy score of 81%, and the precision is greater than 70%. The recall ratio is over 70% positive observations in the actual class. That means fewer false positives and a higher (80%) weighted average of precision and recall. The result tells us that team members with excellent communication skills can maximise the project team's soft skills.

Figure 18

The performance metrics of the "G-Comm and VG-Comm" attributes

Accuracy Score: 0.8125

	precision	recall	f1-score	support
1	0.72	0.93	0.81	14
2	0.93	0.72	0.81	18
accuracy			0.81	32
macro avg	0.83	0.83	0.81	32
weighted avg	0.84	0.81	0.81	32

In Figure 19, the "G-Lead" and "VG-Lead" (Good and Very Good Leadership Skill) attributes have an accuracy score greater than 75%, and the precision is greater than 70%. That means lower false positives, and the precision and recall weighted average is over 70%. The result tells us those team members with good and very leadership skills can maximise the project team's soft skills.

Figure 19

The performance metrics of the "G-Lead and VG-Lead" attributes

```
------------------------------------------------------
Accuracy Score:  0.75
------------------------------------------------------
              precision    recall  f1-score   support

           1       0.67      0.67      0.67        12
           2       0.80      0.80      0.80        20

    accuracy                           0.75        32
   macro avg       0.73      0.73      0.73        32
weighted avg       0.75      0.75      0.75        32
```

According to the above analysis, three observations emerge:

1. Communication skill is one of the soft skills that affect project performance.
2. Leadership skill is the capacity to motivate others, an essential personality trait to influence project performance.
3. Logical skill is the project-related technical skill that is considered crucial among the soft skills.

Discussion

The research results observed that students who used the prediction system in team projects understood the strengths and weaknesses of practical assignments. The study had the opportunity to establish the necessary skills for project management which is also the initial stage of Project-Based Learning (PBL). Students learn by actively engaging in personally meaningful projects such as final year projects. According to the results, the TSS model references teammates with good programming, creativity, communication, and logical skills in the technical nature of team projects. Montfort et al. (2009) mentioned that engineers are required to accomplish critical analysis when dealing with a new problem. As Leppävirta et al. (2011) stated, conceptual knowledge must be established through the iterative process. Thus, reaching a more profound conceptual understanding needs to go through appropriate activities such as logic-based thinking to analyse a situation forthcoming a solution. Therefore, engineering students lacking adequate professional development to train PBL is the main challenge.

The impact of SSM shows that considerable skills like good logical skills can maximise the team project's soft skills. During Project-Based Learning, students can develop communication skills, leadership skills, logical skills, creativity skills, programming skills, synthesised knowledge, and propose a solution to complete the projects. Meanwhile, group-based student projects can reinforce group and individual work skills, including planning and managing time and building communication skills. In terms of the academic field of higher education, Ryan and Ryan (2013) indicated that higher education could provide a rich resource to the systematic development of reflection across whole programs. From the teachers' point of view in higher education, teachers can master the students' knowledge and skills delivered in the courses and allocate the learning resources to the right students. Also, both teachers and students recognise that teamwork is a great benefit to students' future careers.

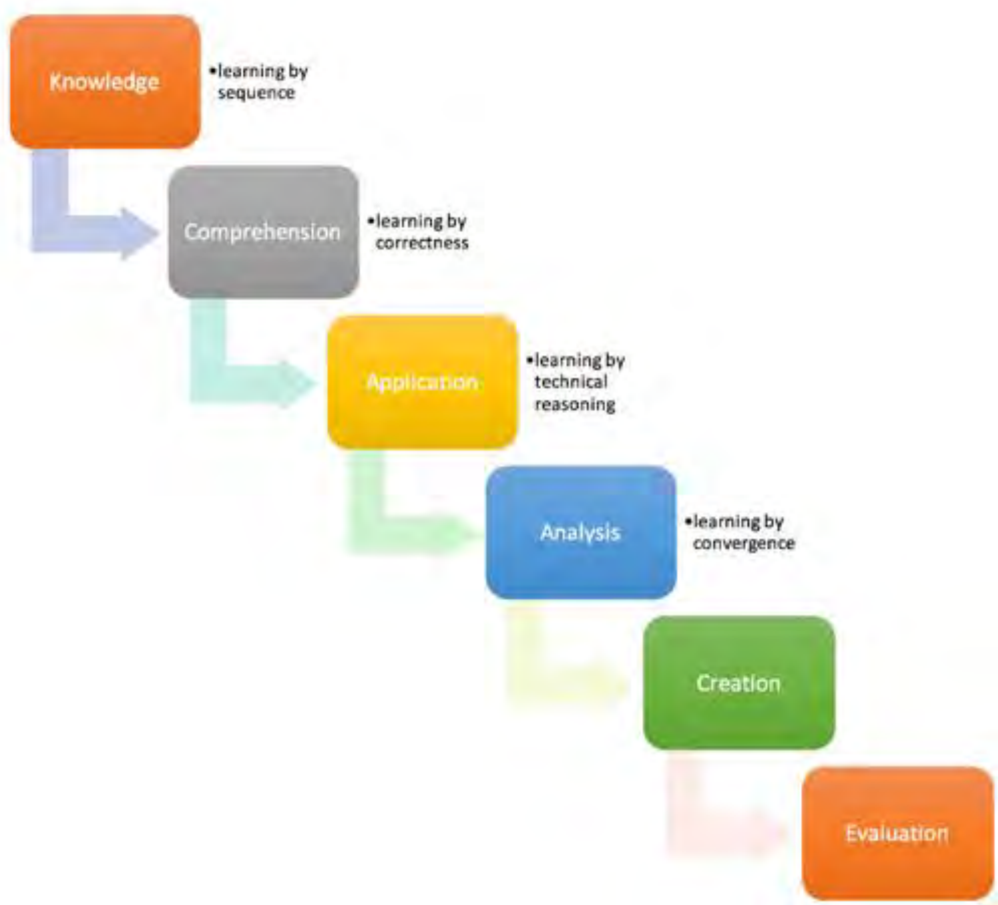
Last, the TSS and SSM models are the bridge in project management to help students develop the growing skills and are significant in finding weaknesses. The fact is that the teamwork performance system affected students' learning outcomes as well.

Conclusion and future research

To conclude, the case study demonstrated that well-rounded Technological Savvy Skills are a critical component in project teamwork. The technological savvy skill is a specific ability that can be measured and defined. Measurement in our case study showed how well the technical skill variables related to teamwork performance. The implication is that teammates with good programming, creativity, communication, and logical skills have associated savvy technological skills. The TSS model highlighted the task allocations and relevant technical skills acquired through formal education or training programs. We believe that the TSS model is a vital asset to consider when developing project management roles through our performance results.

Figure 20

The conceptual learning pattern for the second stage of research



The study also identified soft skills that are adaptive to team processes. Soft skills are more about communication and leadership qualities, such that communication, logical, and leadership skills are soft skills that determine a project's success. Although savvy technological skills are essential in team projects, soft skills help teammates work well no matter how much or little technical knowledge they have. Therefore, balancing savvy technological skills and soft skills are primary determinants in how well teams perform. All the results and impacts are the initial study from the small datasets in the departments. The project will continue to develop as a university-level system.

In the previous sections, the task performance should minimise "underperformance" in the stages project. Nevertheless, we also consider underperformance as a new parameter after the initial stage of performance review. For future investigation, each underperforming member is supposed to conduct the training before joining the next phase of team development. For example, teammates learn information by sequence and apply knowledge by correctness so that the teamwork performance represents the project output.

According to Kepa et al. (2002), training procedures also impact a team's performance. They said every internal process affects team performance, so further research needs to understand the learning patterns better. Figure 20 shows the conceptual framework for the learning pattern model.

The brief learning pattern descriptors appear as:

1. **Classification** - people attend structured learning and learn step-by-step composition and break the tasks into different sub-tasks. They engage in learning activities to include the areas of social skills and academic assignments. Individual's work as instructional methodologies to classify the tasks, distribute tasks, work as a group, work in order, organise tasks in detail and outline the plan.
2. **Precision** – people learn as they correct mistakes and through the practical experience of the open-ended problem. For example, individuals gain knowledge and skills by investigating complex questions or exploring real-world issues and challenges. They enjoy accomplishing the tasks accurately, explaining the facts in detail and measuring the tasks' results precisely.
3. **Work alone** - it means individuals are dynamic learners, and they combine the tasks within their learning contexts, such as taking apart ideas, processes, and things and using the latest tools/technologies to solve problems. These contexts may be referencing a point for ensuring knowledge has been gained. For example, individuals are required to assemble the tasks, build across their learning patterns and demonstrate their experiences in content or expectations. They can figure out the problems and fix the issues by operating concretely and representing graphically.
4. **Confluence -** individuals intended to apply their abilities as performances, knowledge, and skills acquired in learning (Wikipedia, 2021). For instance, individuals transfer information concerning new things from others. They may have acted carefree while brainstorming in group discussions and learned knowledge and skills for new situations (i.e., they have chances to imagine incredible ideas). They always learn to embrace change, and learning is essential to growing as an individual.
5. **A network of cooperation** – individuals learn from peers in larger groups or attempt to learn something together to impact learning positively. For example, an individual can have time when teammates contribute equally and discover the others' strengths in a small group.

Finally, teachers need to create teamwork, and students collaborate with their strengths to team up with the right partners to work together in teams and in team member roles that benefit higher education in evolving the PBL.

References

Aldrich, J.H., and Nelson, F.D. (1986). Linear Probability, Logit and Probit Models (3rd edition). *Beverly Hills, CA.: Sage Publications*.

Aranyossy, M., Blaskovics, B. & Horváth Á. A. (2018). How universal are IT project success and failure factors: Evidence from Hungary. *Information Systems Management*, 35(1), pp.15-28.

Bowers, C. A., Pharmer, J. A., & Salas, E. (2000). When member homogeneity is needed in work teams: A meta-analysis. *Small Group Research*, 31(3), 305–327.

Callaghan, M., Knox, A., Mowatt, I. & Siam, G. (1994). Empirical projects and small group learning, in: H.C. Foote, C.J. Howe, A. Anderson, A.K. Tolmie & D.A. Warden (Eds). *Group and Interactive Learning*, pp.165-170 (Southampton, Computational Mechanics Publications)

Cannon-Bowers, J.A. and Salas, E. (1997). A framework for developing team performance measures in training, in Brannick, M.T., Salas, E. and Prince, C. (Eds), Team Performance Assessment and Measurement – Theory, Methods, and Applications, *Lawrence Erlbaum Associates*, *Hillsdale, NJ*.

Christensen, R. (1990). Log-linear Models. *New York, NY.: Springer-Verlag*.

Cleland, D. I. (1994). Project Management – Strategic Design and Implementation (2nd ed.) *New York: McGraw-Hill*.

Collett, D. (1991). Modelling Binary Data. *England, London: Chapman and Hall*.

Dey, L., Chakraborty, S., Biswas, A., Bose, B., & Tiwari, S. (2016). Sentiment analysis of review datasets using Naïve Bayes' and K-NN classifier, *Int. J. Inf. Eng. Electron. Bus.*

Domingos, P. & Pazzani, M. (1997). On the optimality of the simple Bayesian classifier under zero-one loss. *Machine Learning,* 29 (2/3), 103-130.

El-Sabaa, S. (2001). The skills and career path of an effective project manager. *International Journal of Project Management,* 19(1), pp.1-7.

Efron and Gong. (1983). A Leisurely Look at the Bootstrap, the Jackknife, and Cross Validation, *The American Statistician*.

Fienberg, S. E. (1983). The Analysis of Cross-Classified Categorical Data (rev. ed.), *Cambridge, MA: Massachusetts Institute of Technology.*

Forman, F. (1994). Peer collaboration as situated activity: examples from research on scientific problem solving, in: H.C. Foote, C.J. Howe, A. Anderson, A.K. Tolmie & D.A. Warden (Eds), *Group and Interactive Learning*, pp.3-8 (Southampton, Computational Mechanics Publications).

Fine, G. A., & Hallett, T. (2014). Group cultures and the everyday life of organisations: Interaction orders and meso-analysis. *Organisation Studies*, 35(12), pp.1773-1792.

Fine, G. A. (2012). Group culture and the interaction order: Local sociology on the meso-level. *Annual Review of Sociology*, 38, pp.159-179.

Friedman, N., Geiger, D., & Goldszmidt, M. (1997). Bayesian network classifiers.

Geiger, D., Heckerman, D. (1996). Knowledge representation and inference in similarity networks and Bayesian multi-nets. *Artificial Intelligence 82*, pp.45–74.

Hanusheck, E. K., and Jackson, J. E. (1997). Statistical methods of social scientists. *New York: Academic Press.*

Kepa Mendibil Telleria Derek Little Jill MacBryde. Managing processes through teamwork, *Business Process Management Journal*, Vol. 8, Iss 4, pp338 – 350, 2002

Koskinen, K. U., Philanto, P., Vanharanta, H. (2003). Tacit knowledge acquisition and sharing in a project work context. *The International Journal of Project Management,* 21(4), pp.281-290.

Leppävirta, J., Kettunen, H., & Sihvola, A. (2011). Complex problem exercises in developing engineering stu- dents' conceptual and procedural knowledge of electromagnetics. *IEEE Transactions on Education*, 54(1), pp.63- 66. https://doi.org/10.1109/TE.2010.2043531

Liber, O. (1994). Managing resource based learning, in: H.C. Foote, C.J. Howe, A. Anderson, A.K. Tolmie & D.A. Warden (Eds), *Group and Interactive Learning*, pp.183-186. (Southampton, Computational Mechanics Publications).

Luțaș, M., Nistor, R., Radu, M. & Beleiu, I. (2020). Perceptions Regarding the Profile of an Ideal Project Manager. *Amfiteatru Economic,* 22(54), pp.608-622.

Mchugh, M. L. (2012). Interrater reliability: The kappa statistic importance of measuring interrater reliability theoretical issues in measurement of rater reliability. *Biochem Med (Zagreb)*.

McIntosh, Joshua Grant. (2012). The Impact of Curricular Learning Communities on Furthering the Engagement and Persistence of Academically Underprepared Students at Community Colleges. *Higher Education – Dissertations* 57. https://surface.syr.edu/he_etd/57

Montfort, D., Brown, S., & Pollock, D. (2009). An investigation of students' conceptual understanding in related sophomore to graduate-level engineering and mechanics courses. *Journal of Engineering Education*, *98*(2), pp.111-129. https://doi.org/10.1002/j.2168-9830.2009.tb01011.x

Plane, D. R., and Oppermam. (1997). Statistics for Management Decisions. *Dallas, TX.: Irwin-Dorssey.*

Prerna Sindwani. (2019). More than a third of India's engineers are unemployable, and 20% lack coding skills, *Business Insider, India*. https://www.businessinsider.in/careers/news/over-a-third-of-engineering-graduates-lack-employability-skills/articleshow/72178876.cms

Ryan, M & Ryan, M 2013. Theorising a model for teaching and assessing reflective learning in higher education. *Higher Education Research and Development*, pp.32(2), pp. 244-257

Smeeton, N.C. (1985). Early History of the Kappa Statistic. Biometrics. 41 (3): 795. *JSTOR 2531300.*

Söderlund, J. (2005). Developing project competence: Empirical Regularities in Competitive Project Operations. *International Journal of Innovation Management*, 9(4), pp.451-480.

St. John, E. P. (1990). Price response in enrollment decisions: An analysis of the High School and Beyond Senior Cohort. *Research in Higher Education* 31(2): pp161-176.

Webb G.I. (2011). Naïve Bayes. In: Sammut C., Webb G.I. (eds) Encyclopedia of Machine Learning. Springer, *Boston, MA*.