A rule-based model for software development team composition: Team leader role with personality types and gender classification

Abdul Rehman Gilal a,h, Jafreezal Jaafar a, Mazni Omar c, Shuib Basri a, Ahmad Waqas b

a Department of Computer and Information Sciences, Universiti Teknologi PETRONAS, Malaysia
b Department of Computer Science, Sukkur Institute of Business Administration, Pakistan
c School of Computing, Universiti Utara Malaysia, Malaysia

ARTICLE INFO

Article history:
Received 22 September 2015
Revised 27 January 2016
Accepted 22 February 2016
Available online 27 February 2016

Keywords:
Human aspects
Personality types
Myers–Briggs Type Indicator
Team leader
Gender
Software development
Team composition
Rule-based

ABSTRACT

Context: Recent studies have established the fact that the supply of handy and successful software has decreased to 6%. The past studies have also attributed this supply failure to software development team composition factor. To overcome this problem, it is also suggested in the past studies that the soft skills of team member must be considered along with the hard skills.

Objective: Keeping in view this problem, this study aimed to look for in-depth understanding of team-lead role with personality types of member. This study also included gender to see its diverting impact on personality types and job role, since past studies have also raised many issues pertinent to these two variables.

Method: This study used the experimental data to develop the rule-based model for software development team composition by keeping gender as major effecting variable with personality. There were three independent predictor variables: Team leader role, Personality types, and Gender; and one outcome dependent variable: team performance. Additionally, personality types of team members were measured by using Myers–Briggs Type Indicator(MBTI) instrument. This study divided the experiments into two stages. The first stage was descriptive examination of factual figures of data for model development. Whereas, the second stage was predictive experiments of data for developing the model.

Results: The findings revealed that each gender emerged compatible with different types of personality for the same role. For instance, descriptive analysis part of this research highlighted that feeling(F) personality males were appropriate for team leader role, on another hand, thinking(T) personality females were suitable for the team lead role.

Conclusion: The conclusion can be drawn with the claim that the personality types of software development team roles fluctuate by gender type. Besides, this study revealed and ensured that gender should be kept in the consideration when composing teams based on personality types.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

This study is the continuation of our previous presented work [1] in which the role of programmer was determined with generated rules (if–then) applying MBTI test for making effective software development team composition. Moreover, the work was conducted to address the need of software development which is gaining the exponential fame into the modern world of technology. Therefore, the current study focuses on personality types to produce the rule-based modeling for software team composition by including team leader role of software development applying MBTI personality test.

The demand of software is increasing exponentially in every field. But, the detrimental decrease is observed to successful software development [2]. Many factors are identified that can impact the overall development process. In which, inappropriate team composition is one of the important factors [3–5]. According to Bell [6], team composition is a configuration of members based on their attributes that can strongly influence on the team processes and outcomes. She further maintained that past studies have categorized the team composition research into three dimensions: (1) characteristics of team members (e.g., number of team members, members’ abilities, demographics, and personality traits), (2) measurement of these characteristics, (3) and the analytical
perspective used to approach team composition. Normally, team composition is based on the technical fragment of the work. However, software development integrates the technical part of work with social norms. For example, Capretz and Ahmed [7] asserted that team can be ideal if the hard (technical) skills are combined with soft (non-technical: personality) skills. In the same vein, Dingsøyr and Dybå [8] maintained that isolation of either skills (i.e., technical or social) can be one of the reasons of poor development. Additionally, it is also believed that the consideration of technical skills of developers can be advantageous as long as software developers are evaluated by their personalities, which is a soft skill, to evaluate their temperaments to work under cooperative principles with other team members [9].

Although plethora of research has been carried out in the past to explore the key importance of team composition and personality types in the software engineering, but what personality types are handy and beneficial for an ideal and effective teamwork is still a question for the researchers [10–12]. For instance, an extrovert (E) personality preference is suggested for programmers by Gorla and Lam [13] and, whereas, Capretz and Ahmed [7] suggested an introvert (I) personality trait for programmers. It may be because now a days software development complexity is increasing which demands different personality profiles [14], or different personality profiles are required for male and female developers [15]. Therefore, it is suggested in the past studies that maturity is required in the personality based research in software development.

Consequently, this paper addresses and contributes in the field of software engineering (SE), which offers a rule-based model of effective software team composition by using rough set’s technique. This model composes three predictor variables i.e., team role, personality types and gender to determine team effectiveness. To the best of researchers’ knowledge, the combination of the three variables in determining software team performance by employing rule-based techniques provides a novel approach in SE domain. Therefore, the implications of the paper can be concluded as follows:

i. This research would benefit and enable managers to sort out effective team members equipped with soft skills, which is personality types, in order to compose an effective team for software development.

ii. The findings of this research have produced a rule-based model for the software team composition, which is standard technique where results come in if-then form understandable by human.

Basically, this section of the paper presented the problem, objective, and implication of the study. The following section has highlighted the software engineering literature conducted on the study variables: MBTI, team leader, and gender. In the same vein, the methodology section discusses the experimental part of the study, in which data collection and analysis are highlighted in detail. Moreover, the results and discussion section was divided into two separate sections based on the nature of the analysis: descriptive and Predictive. Once after the results and discussion section, the model development is presented with explanation and proposed algorithm in section number 6. Last of all, the validity threats of the study are discussed in the threats to validity section for explaining the proper use of the model.

2. Related work

One of the considered reasons of all for weak performance and poor results in software project development is the composition of ineffective teams [4,16–19]. A plethora of studies on team composition and personality types in software engineering have been done in the past, but the issue pertinent to a suitable personality composition for effective teamwork is still being questioned [8,20,21]. Additionally, many models of personality composition have been suggested in the past that have failed to win the general consent of researchers. This ambiguity is raised from the different models and theories suggesting different personality types and team composition. For instance, Gorla and Lam [16] suggested an extrovert (E) personality type for programmers while Capretz and Ahmed [7] suggested an introvert (I) personality for programmers. Moreover, the model developed by Capretz and Ahmed [7] was non-empirical, as the model was constructed based on advertisements of jobs. Additionally, the model given by Gorla and Lam [16] was based on empirical data, but authors only focused university students that narrowed the scope of the model. Therefore, the results of that model cannot be generalized to industries. These contradictory suggestions have created many complexities and ambiguities in deciding what personality types are suitable for team organization. Additionally, Cruz and da Silva [22], McLeod and MacDonell [23], and Wiesche and Krcmar [24] asserted that the models suggested in the past for team compositions appeared to be less effective and less lucrative when they were implemented to obtain desired results. According to them, this situation has further intensified the uncertainty among practitioners regarding the effectiveness of these models. In the same vein, Varona et al. [14] also maintained in their review study that increasing complexity in software development requires new personality profiles. They further identified that huge understanding is also required for personality in software engineering.

2.1. MBTI and software development

For the last 50 years, the MBTI has been used as a source for identification of personality preferences and personality type of an individual. This personality type indicator is used for making theories of Jung applicable and useful in everyday life [25]. An individual’s personality type in MBTI is assessed on four dimensions: social interaction (extroversion (E) and introversion (I)), decision making (thinking (T) and feeling (F)), information gathering (sensing (S) and intuition (N)), and dealing with the external world (judging (J) and perceiving (P)) [16]. Both Katharine Cook Briggs and her daughter, Isabel Briggs Myers are considered as pioneer of MBTI who have not only extensively studied the work of Jung but they also explored and inter-related different theories of human behavior i.e., theory of psychological types into practical use. The MBTI test allows individual personality type preferences to be classified according to the 16 types with the results reported as a combination of four-dimensional pairs, which are Introversion (I) and Extroversion (E); Thinking (T) and Feeling (F); Sensing (S) and Intuitive (N); and Judging (J) and Perceiving (P). The 16 possible personality combinations are formed from these four dimensions. For the purpose of this study, MBTI personality type instrument is used to assess personality types amongst software team members. The MBTI instrument is chosen, because it is widely used and accepted amongst researchers in software engineering domains [22,26,27].

2.2. Team leader and software development

Leadership is one of the key elements to successful project completion. Based on the definition from Team Software Process (TSP), a team leader can be defined: a team leader is responsible to bring the management in software development projects and he is responsible for outcomes of the development projects [28]. In the same vein, Ruano-Mayoral et al. [28] further maintained that team leader is also responsible for guiding, motivating team members, handling teams and customers issues, and dealing management. They also mentioned that team leader is also required to follow the deadlines to produce the projects from assigned resources.
Moreover, following requirements are mostly demanded in team leader role within software development teams [28,29]:

- Guidance to developers for successful project development
- Management of tasks between developers for successful project development.
- Provide technical and social support to team members.
- To maintain the communication between team members and phases of software development.
- To report higher authorities at different phases of development.
- Encourage the team members to solve issues.
- Work closely with developers, project managers, and business analyst to synchronize the requirements of the projects (or requirement gathering and sharing).

The requirements can vary from company to company but are common in nature. Moreover, each of the requirements of leader role can be categorized under the umbrella of MBTI. For example, “to maintain the communication with team members” is one of the responsibilities of team leader that highlights the Extrovert (E) trait of personality. On the contrary, Introvert (I) people are less gregarious as compared to E trait. In the same vein, Gorla and Lam [16] conducted a study for finding best personality for software development roles. They confirmed that personality of team leader impacts on the project development results. Their results further showed that INTuitive (N) and Feeling (F) leaders are more productive than Sensing (S) and Thinking (T). However, they did not mention for which gender leader N and F are compatible. In fact, female leaders are more towards F trait than male leaders that hinders meeting deadlines. Therefore, this study further investigates the gender impact over personality of team leader role.

2.3. Gender and software development

In social sciences, many research studies have explored personality and gender either collectively or separately, to address the grave problems of teamwork in organizations and have achieved the acute success as well. However, this problem is still persistent in the field of software development since few researchers have ever tried to test personality and gender collectively to test the suitability of the team handy for software development. In this regard, Richards and Busch [30], Gilat et al. [2], and Rehman et al. [31] also assert that maturity level is yet to find in software development research. In the same vein, Trauth [32] also recommends that the need of improvement is required in the theoretical work on software development.

The study on gender conducted by Gilat et al. [2] comes among the few studies that studied the personality type together with gender. This study investigated the performance variation among software development team members caused by genders’ personality types. For instance, the male-dominated teams create the reasons for the females for being ineffective in teams if the personality type of female is with “E” trait. Furthermore, the study revealed that the female-leader are more convenient with only female or majority-female (i.e., having female in majority) groups. Whereas, male-leaders are acceptable with all kind of team compositions. Critically, this study was just based on tabulated calculation and could not give any statistical or predictive evidences. However, the study also recommended to further research on gender with personality types to reach appropriate conclusions.

Richards and Busch’s [30] study explored the gender and culture parameters to find their effects on the performance in IT workplaces. This study focused the knowing and doing gaps in software development workplaces. The researchers tried to find the effectiveness of diversity on overall performance of the team.

Moreover, authors acknowledged that these results are too weak to generalize that was one of the limitations of the study. But, that gap of the study can be overcome by inclusion of personality in software development. Because, it is also believed that inclusion of personality can help to achieve efficiency, productivity, and quality [14].

3. Methodology

This study used the experimented data [46] to develop the rule-based model for software development team composition by keeping gender as major effecting variable with personality. There were three independent predictor variables: 1. Team leader role, 2. Personality types, and 3. Gender; and one outcome dependent variable: team performance. Additionally, personality types of team members were measured by using MBTI instrument. Moreover, the model was trained and validated on the student population from Universiti Utara Malaysia (UUM). Total 46 teams were composed from software engineering class of final year students where each team had one leader (lead programmer) and four programmers (It should be noted here that though teams were formed with team leader and programmer roles but this study only discusses the team leader role). In this study, students were required to develop a web-based project with certain requirements by following extreme programming (XP) methodology. Moreover, they were asked to submit their projects within 16 weeks. Once after the projects were submitted, an assessment of projects was conducted based on the given requirements by a requirement engineer. Additionally, based on the assessment of requirement engineer, the teams which got 80% or above marks were considered “effective” and, whereas, less than 80% were called “ineffective”. In order to gain the general consent of the developed model, it was further validated from industrial datasets carrying data from three different companies. The dataset from university was divided into two elementary sets of training, and testing. Moreover, training and testing sets were divided with 70% and 30% standard ratios [33–35].

In order to avoid the biasedness, this paper follows the same methods of experiment used in our previous paper on programmer role [1]. Therefore, this study also divided the experiments into two stages. The first stage was descriptive examination of factual figures of data for model development. Whereas, the second stage was predictive experiments of data for developing the model. Moreover, the first stage was set to understand the basic relations and behaviors of the data which is very important for gaining a general consent for the model. In addition to this, descriptive analyses, frequency analyses, graphs, and tables were used to explore and discuss the basics of data by using SPSS (version 20.0) and Microsoft Excel (version 2013). On another hand, decision tree and rough set approaches were used for predictive model development. Decision tree, C4.5 algorithm was used to gain the results. Moreover, Waikato Environment for Knowledge Analysis (Weka 3.6) toolkit was used to experiment decision tree; in which j48 is a java implementation of C4.5. Similarly, for rough set experiments the SAVGeneticReducer and JohnsonReducer were used, as these are the implementation of Johnson and Genetic algorithms in ROSETTA because it is a toolkit for analyzing the data based on rough sets algorithms.

Johnson Algorithm (JA) technique was selected for model development. Genetic Algorithm (GA) was excluded because both JA and GA formed the same results of prediction accuracy, but GA produced more rules than JA. It is, therefore, JA was chosen because it was considered less complex than GA. Furthermore, selection of technique was based on the prediction accuracy obtained from testing of experimented model [36]. The following Table 1 shows the results and accuracy obtained from testing of trained model.
4. Descriptive discussion

The discussion of descriptive analysis was conducted on training set by grouping gender on effective-projects and ineffective-projects’ results. Subsequently, 32 teams were shortlisted for the training set and remaining 14 were kept for testing set. Thus, total 14 males and 18 females were available in the training dataset for team leader role. Moreover, based on the results of requirement engineer, 13 teams could manage to obtain 80% (or above 80%) in training set and 19 teams could not achieve the benchmark. Therefore, in training set, total 13 teams were called effective-teams and 19 were ineffective-teams. Though these results do not show higher differences, but indeed, the effective and ineffective teams’ results will impact on the overall claims. It is also asserted that both effective-projects and ineffective-projects’ results are taken into use for model development. Because, ineffective-projects’ results are stronger than effective-projects’ results. For instance, the female leader in ineffective-projects is 61.11% which is higher than its effective-projects results.

Leader role was further investigated based on overall personality traits appearance in effective-projects and ineffective-projects results. The results were considered effective when the development projects by student teams met the basic requirements and secured benchmark level of project development. Moreover, the leader role was also closely investigated based on MBTI pairs. In effective-projects’ results, while exploring the first pair “IE” of MBTI, it was found that male-leader with an extrovert trait of personality are highly suitable for team leadership. Whereas, introvert trait behavior is not suitable for male-leader role in software development based on this dataset. Moving to the second pair of MBTI, which is “SN”, it can be observed that sensing leaders are suitable for leadership in both genders (i.e., male and female). But, male-leader was not generally found in the effectiveness as compared to female-leaders in intuitive trait. Hence, it could be inferred that sensing personality trait is more confident pair in male-leader and female-leader which is ideal for both traits (i.e., sensing and intuitive). The third (i.e., “TF”) and fourth (i.e., “JP”) pairs were found almost the same in frequency analysis (Fig. 1).

The training set also contained the data with ineffective-projects that means teams which could not achieve the desired results or team could not produce up to the requirements. Fact says that ineffective-projects were about 60% in overall results. Therefore, ineffective-projects’ results were also considered for model development. Considering the first pair “IE” of MBTI in ineffective-projects, it was observed that male and female genders were found similar as in effective-projects (i.e., extrovert male leader are effective and female leader could be effective even if the personality trait is introvert). Therefore, the frequency ratio of IE was measured in both effective and ineffective projects for better understanding. In that frequency analysis of IE, E trait was found favorable for male-leader as it gives the higher rate for effectiveness and less for ineffectiveness in the results. Moreover, female-leader was found fit in both traits as it has very low deviation in effective and ineffective results. The following Fig. 2 gives the details of IE pair of effective and ineffective projects.

In conclusion, this phase of the study has come up with some sort of personality traits from MBTI to shortlist them for the leader role in order to consider or filter for model development. Whereas, “consider” refers to use those traits for making teams and “filter” term refers to avoid those traits for escaping issues in the future. The following Table 2 is showing the statistics of team leader role for considering or eliminating the traits in percentage by classifying with gender.

---

**Table 1**

<table>
<thead>
<tr>
<th>Approach</th>
<th>Technique</th>
<th>Role</th>
<th>Prediction accuracy</th>
<th>Average accuracy of results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision tree</td>
<td>J48</td>
<td>Leader</td>
<td>71.43</td>
<td>64.985</td>
</tr>
<tr>
<td>Rough set</td>
<td>GA</td>
<td>Programmer</td>
<td>58.54</td>
<td></td>
</tr>
<tr>
<td></td>
<td>JA</td>
<td>Leader (17 rules)</td>
<td>71.44</td>
<td>71.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Programmer (22 rules)</td>
<td>70.74</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2**

The statistics for considering or eliminating the traits for Team Leader role.

<table>
<thead>
<tr>
<th>Gender</th>
<th>MBTI traits</th>
<th>Effective percentage</th>
<th>Ineffective percentage</th>
<th>Considered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>I</td>
<td>16.67</td>
<td>37.50</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>83.33</td>
<td>62.50</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>83.33</td>
<td>100.00</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>16.67</td>
<td>0.00</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>33.33</td>
<td>37.50</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>66.67</td>
<td>62.50</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>J</td>
<td>83.33</td>
<td>75.00</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>16.67</td>
<td>25.00</td>
<td>N</td>
</tr>
<tr>
<td>Female</td>
<td>I</td>
<td>57.14</td>
<td>54.55</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>42.86</td>
<td>45.45</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>57.14</td>
<td>100.00</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>42.86</td>
<td>0.00</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>42.86</td>
<td>36.36</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>57.14</td>
<td>63.64</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>J</td>
<td>100.00</td>
<td>81.82</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>0.00</td>
<td>38.18</td>
<td>N</td>
</tr>
</tbody>
</table>
Fig. 2. IE Pair of MBTI for leader Personality in training dataset.

<table>
<thead>
<tr>
<th>Rule no.</th>
<th>Rule</th>
<th>LHS support</th>
<th>RHS support</th>
<th>RHS accuracy</th>
<th>LHS coverage</th>
<th>RHS accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Female AND Introvert AND Judging =&gt; ineffective OR effective</td>
<td>9</td>
<td>5, 4</td>
<td>0.555556, 0.444444</td>
<td>0.28125</td>
<td>0.263158, 0.307692</td>
</tr>
<tr>
<td>2</td>
<td>Extrovert AND Thinking =&gt; ineffective OR effective</td>
<td>8</td>
<td>4, 4</td>
<td>0.5, 0.5</td>
<td>0.25</td>
<td>0.210526, 0.307692</td>
</tr>
<tr>
<td>3</td>
<td>Female AND Thinking =&gt; effective OR ineffective</td>
<td>7</td>
<td>3, 4</td>
<td>0.428571, 0.571429</td>
<td>0.21875</td>
<td>0.230769, 0.210526</td>
</tr>
<tr>
<td>4</td>
<td>Extrovert AND Sensing AND Feeling =&gt; ineffective</td>
<td>6</td>
<td>6, 6</td>
<td>1.0</td>
<td>0.1875</td>
<td>0.315789</td>
</tr>
<tr>
<td>5</td>
<td>Extrovert AND Intuiting AND Feeling =&gt; effective</td>
<td>4</td>
<td>4, 4</td>
<td>1.0</td>
<td>0.125</td>
<td>0.307692</td>
</tr>
<tr>
<td>6</td>
<td>Female AND Introvert AND Intuiting =&gt; ineffective OR effective</td>
<td>3</td>
<td>1, 2</td>
<td>0.333333, 0.666667</td>
<td>0.09375</td>
<td>0.052632, 0.153846</td>
</tr>
<tr>
<td>7</td>
<td>Male AND Sensing AND Feeling =&gt; ineffective</td>
<td>3</td>
<td>3, 3</td>
<td>1.0</td>
<td>0.09375</td>
<td>0.157895</td>
</tr>
<tr>
<td>8</td>
<td>Intuiting AND Perceiving =&gt; ineffective OR effective</td>
<td>2</td>
<td>1, 1</td>
<td>0.5, 0.5</td>
<td>0.0625</td>
<td>0.052632, 0.076923</td>
</tr>
<tr>
<td>9</td>
<td>Male AND Introvert AND Judging =&gt; ineffective</td>
<td>2</td>
<td>2, 2</td>
<td>1.0</td>
<td>0.0625</td>
<td>0.105263</td>
</tr>
<tr>
<td>10</td>
<td>Female AND Perceiving =&gt; ineffective</td>
<td>2</td>
<td>2, 2</td>
<td>1.0</td>
<td>0.0625</td>
<td>0.105263</td>
</tr>
<tr>
<td>11</td>
<td>Extrovert AND Perceiving =&gt; ineffective</td>
<td>2</td>
<td>2, 2</td>
<td>1.0</td>
<td>0.0625</td>
<td>0.105263</td>
</tr>
<tr>
<td>12</td>
<td>Male AND Introvert AND Thinking =&gt; ineffective</td>
<td>1</td>
<td>1, 1</td>
<td>1.0</td>
<td>0.03125</td>
<td>0.052632</td>
</tr>
</tbody>
</table>

Table 2 shows the results based on effective and ineffective results appearance in percentage. Because, it was observed that the gender classification creates the difference in overall appearance of a particular gender. Moreover, a trait is considered (denoted with “Y” in Table 2) if the effective percentage is higher than ineffective percentage (in Excel expression: IF (effective-percentage>ineffective-percentage, "Y", "N"). For example, E trait of male-leader has 83.33 effective-percentage and 62.50 ineffective percentage. That is why, E trait is “Considered” and I trait is “filtered/unconsidered” (denoted with N in Table 2). Therefore, based on the descriptive analysis, E, N, F, and J personality traits were considered effective personality type for male-leader. On the other hand, INTJ personality type has shown the influence in female-leaders. Nevertheless, these were only descriptive results based on the appeared effective and ineffective frequencies. Furthermore, the following section, after descriptive, has comprised of predictive results based on the individual traits with other possible and favorable traits with it.

5. Constructed effective and ineffective rules for team leader role

It is mentioned in the methodology, Section 3, that only 12 rules were obtained from the experiments with JA technique. Among these 12 rules, two types of rules appeared in the results and these were named as Uni-dimension and Bi-dimension. Because, Uni-dimension rules referred to the rules which have one direct decision; either particular personality type is effective or ineffective. For example, if the input is satisfied with the condition of the rule then very straight forward decision will be taken. But, Bi-dimension refers to those decision rules which have shared accuracy in both: effective and ineffective decisions, or rules which have more than one decisions. The following Table 3 shows all rules, Uni-dimension and Bi-dimension, obtained from JA experiments.

Prior to the discussion of rules, it seemed an appropriate to explain the key terms mentioned in the Table 3. Such as, the “rule”, can also be said as redact which is the statement of decision based on “if-then” state driven from dataset. The term “left hand side (LHS) support” is basically referring that how many objects from dataset are matching the if-statement. Whereas, term “right hand side (RHS) support” shows that how many objects are matching “then-statement/then-part” from dataset based on the “if-statement”. Moreover, RHS support will show two numbers if the rule is Bi-dimension, and it must be known that LHS support and RHS support are always equal. For instance, rule no 1 from Table 3 (e.g., Female AND Introvert AND Judging => ineffective OR effective is Bi-dimension rule) has RHS support “5, 4” that means five objects from datasets meet the ineffective then-part and four for effective and the total is equivalent to LHS support. Keeping RHS support in view, the “RHS Accuracy” is obtained from RHS support divided by LHS support; except the then-part possesses two conditions otherwise accuracy is 1. For instance, rule no 1 (i.e., from Table 3) contains two decisions, ineffective or effective, and RHS support shows 5, 4 objects so the accuracy stands as 5/9 = 0.555556 and 4/9 = 0.444444. Lastly, the LHS Coverage refers the overall appearance of “if-part” in dataset by dividing LHS support with total objects of dataset. On another hand, the RHS Coverage term is similar to LHS coverage but it covers the “then-part” of the rule by dividing RHS support with total objects of dataset.
example, in rule no 1 (i.e., from Table T3) the LHS coverage is 0.28125 because the total objects in dataset for the leader role are 32.

Among 12 rules, only six rules could show the effectiveness in the results (i.e., see rule no: 1, 2, 3, 5, 6, 8), whereas eleven (i.e., see rule no: 1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12 from Table 3) out of 12 rules were found ineffective. Moreover, only one rule (see rule number 5) was found Uni-dimension and five (see rule number 1, 2, 3, 6, 8) remaining were Bi-dimension in effective rules. Interest-
ing factor was seemed in these six effective rules that only rule number 5, and 6 were found dominated as effective. On another side, rule number 1, 3, 4, 7, 9, 10, 11, and 12 were dominated by ineffective results. But, rule number 2 and 8 were found undecided results in Bi-dimension rules as they have equal accuracy for effec-
tiveness and ineffectiveness.

It is discussed earlier that effective and ineffective both types of rules were used to develop the model. Thus, effective rules were directly used to compose a team and ineffective rules were used to make a filter to avoid the ineffectiveness in team’s performance. By the way, team-leader composition based on rule number 5 and 6 was direct and efficient. Decision rule 5 was Uni-dimension and it has 0.307692 coverage which is dominated coverage. Hence, one can infer that ENP personality type for team-leader is effective for male and female leader. Moreover, E, N, and F traits were also ver-
ified by the descriptive analysis for male-leader and only N trait was matched in female-leader. So, these predictive results give a hand to weak results obtained from descriptive analysis for female-leader, as these were left for predictive analysis.

Male-leader can only be the extrovert in personality as it is proved by predictive and descriptive results. But, female-leader can perform well if the personality belongs with I trait but the intuitive trait is demanded from the second pair (SN) of MBTI. Rule number 6 (Female AND Introvert AND Intuiting => ineffective OR effective) in Table 3 shows that if female-leader is composed of I and N trait then 67% chances exist for effective performance. In the same vein of effective rules, rule number 2 had shared the equal accuracy with effective and ineffective results but coverage of effectiveness with 0.307692 in effective rules. In which, E trait male or female leader can perform effectively with T personality and that rule is basically from the same class of coverage (i.e., 0.307692). It is more towards effectiveness by gaining higher coverage which is obtained from testing-then-part by dividing total objects into the class listed for then-part of rule. In the end, this rule was selected for using E and T traits together to compose teams for both genders.

In Table 3, rules were sorted descending by the LHS support that means the rule which occurred maximum or higher in dataset will be on top. In which, rule number 1 is arisen the most fre-
cent rule in the dataset by obtaining 9 LHS support. The rule is composed of “Female-leader AND I trait AND J trait” statement by sharing the effectiveness with ineffectiveness. Even, the ineffectiveness is occurred higher with 56% accuracy than effective accuracy with 44%. As effective results sustenance four times in RHS support consequently the coverage remained higher with 0.307692 and 0.263158 coverage was obtained by ineffective results. So based on this accuracy and coverage, one cannot claim the selection of this rule to use or to avoid. But, the presence of I trait in female-leader personality is supported by another effective rule (see rule number 6). Therefore this rule (i.e., rule number 1), which is the composi-
tion of female-leader with introvert trait, is supported by another efficient rule for better results in team performance.

Additionally, rule number 3 from Table 3 just gives the little change in rule number 4, which is 100% of accuracy, with the condition if leader is female then F trait can be adjusted. As the effective accuracy of the rule is distributed with ineffective accuracy with 42% (effective) and 57% (ineffective). Hence, rule number 2 supported that rule with the condition of E trait is in the pair. In

the decision, this rule is considered provided it comes with E trait if the F trait is in the personality of female-leader. Last rule of effec-
tive leaders was rule number 8 which was Bi-dimension with 50–50 accuracy. It was observed that N trait is suitable and sup-
ported by descriptive and other predictive rules. But, P trait has the great ratio in failure/ ineffective teams based on descriptive results. So if no choice than P trait is suitable with N trait in male-leader and female-leader demands I trait if J trait is part of personality.

After effective rules discussion, very straight forward rules, for creating causes for ineffectiveness in team performance, are dis-
cussed. Out of ten ineffective rules (i.e., see rule no: 1, 2, 3, 4, 6, 7, 8, 9, 10, 11, and 12 from Table 3), six rules (4, 7, 9, 10, 11, and 12) were Uni-dimension. It means that these six rules are used to apply for developing the filter list for avoiding inefficiency and boosting the team performance. Based on these rules, the highest ineffective rule is rule number 4 (Extrovert AND Sensing AND Feeling => ineffective) with 0.315789 decision coverage by 6 LHS support. This is critically hazardous personality composition found in the dataset for male and female leaders. Keeping deep discus-
sion on it, this rule (i.e., number 4) was found almost similar with rule number 5 (Extrovert AND Intuiting AND Feeling => effective) which is the most efficient rule with 100% accuracy. But, only sec-
ond pair of MBTI personality (i.e., SN) creates the difference line between them. For instance, if only N trait is changed with S trait then personality type will be 100% unproductive/ lethal and another way around for effectiveness. Moreover, if the leader is male then bounding of S and F traits will be the one of the risky factors as resultant by rule number 7 (Male AND Sensing AND Feeling => ineffective). In the same vein, the rules from 9 to 12 shows the straightforward composition for being ineffective in team. Among them, rule number 10 and 11 gave strong factors for dangerous-
ness of P trait in teams. On another hand, I trait for male-leader showed hopeless for any kind of efficient results in team perfor-
ance. Therefore, these rules were directly applied to filter of the model to compose an effectual team for better performance.

5.1. Validation of rules or prediction accuracy

In order to check the quality of the rules obtained from training experiments, this study used hold-out method to test the predic-
tion accuracy of the rules. For that purpose, testing dataset was used with several classification techniques for understating better results (i.e., Standard Voting, Voting with object tracking, cross-
validation). But, Voting with object tracking technique remained effective than all other techniques. It should be noted that as the JA algorithm results were used to develop the model, therefore, the results of classification technique are based on JA.

The benchmark of the accuracy for this study was set at 70% be-
cause of sample size. It is acceptable benchmark for model develop-
ment by several other researchers. For instance, Bakar [37] states that the model is considered effective if the accuracy reaches up to 70% or above. In the same vein, Hvidsten [38] maintains Bakar’s statements by showing agreement on 70% accuracy acceptance for modeling results. Therefore, this study obtained the results of Vot-
ing with object tracking classifier because it achieved the bench-
mark.

The training set of team-leader role had 32 sample size (i.e., 70%) and its testing set was with 14 size (i.e., 30%). Moreover, the confusion matrix, obtained from testing dataset, is given below in Table 4 that gives the better understanding of the trained model results. Additionally, the results were multiplied with 100 to obtain the results in percentage.

Table 4 explores the results which are confirming 9 ineffective rules those were really ineffective and 1 effective as effective. As mentioned above, the model has used both types of rules, effective and ineffective, for development. Fortunately, the accuracy is also
Table 4
Prediction accuracy table for team leader role.

<table>
<thead>
<tr>
<th>ACTUAL</th>
<th>PREDICTED</th>
<th>0 (ineffective)</th>
<th>1 (effective)</th>
<th>((^\times) 100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (ineffective)</td>
<td>9</td>
<td>1</td>
<td>90%</td>
<td></td>
</tr>
<tr>
<td>1 (effective)</td>
<td>3</td>
<td>1</td>
<td>25%</td>
<td></td>
</tr>
<tr>
<td>((^\times) 100)</td>
<td>75%</td>
<td>50%</td>
<td>71.4286% (Accuracy)</td>
<td></td>
</tr>
</tbody>
</table>

Table 5
Prediction accuracy validation of both roles with industrial data.

<table>
<thead>
<tr>
<th></th>
<th>Industrial accuracy</th>
<th>Academic accuracy</th>
<th>Overall accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>76.92%</td>
<td>71.43%</td>
<td>74.18%</td>
</tr>
</tbody>
</table>

satisfying the development of rules from the training set. So, the obtained 71.43% accuracy is sophisticated, as the acceptance demanded accuracy was 70%. Thus, these rules for team-leader personality can be used to compose a team for better output of team leader role.

5.2. Accuracy validation with industrial data

The validation of predictor rules with industrial data proceeds in the same way as it was conducted under testing set (i.e., 30% from academic data). In which, the “Voting with object tracking” classification technique was used to find the accuracy from industrial dataset. Moreover, the industrial dataset was composed of three different companies with 12 teams. Therefore, sample of 12 members was obtained for team leader role which is nearly equal to the testing set from academic data.

In fact, the results obtained from academic data set (i.e., from training and testing) remained accurate with industrial data. By which, the results can be used to develop the model which can be generalized with academic and industrial software development team composition for leader role only. Table 5 shows the validation accuracy obtained from industrial data sets.

Interestingly, the validation results obtained from industrial data were even better than academic data. As the leader obtained 71.43% from academic testing set but industrial data produced 76.92% accuracy. Even the data sample size was almost similar, therefore, that is considered as acceptable for the deployment in practice or model development. In the end, this study also ensures the claims that personality has the betterment by the graduation of age but it remained same throughout the life [39,40]. Therefore, one can generalized the results of his study with industrial practices if the results are obtained from academic population. Finally, these results boosted the confidence to pave the study to next phase of model development.

5.3. Evaluation of the modeling

It is always required to know the performance of the predictive system as it can produce the different discrimination thresholds. In the same vein, this study used Receiver Operating Characteristics (ROC) curve to evaluate the performance of the system. Moreover, the ROC curve method is widely used in several fields of science. For instance, Park [41] and Kumar [42] maintain that ROC is a common method for evaluating the diagnostic system. They further continue, it is widely used in psychology, engineering, medical, machine learning researches.

This study used the ROSETTA toolkit to obtain the ROC curve value based on the applied classification method. Both rules team leader and programmer (one should refer our previous paper for programmer results and discussing.) produced different results as these were classified separately. Moreover, Fawcett [43] statements were taken into use for selecting the value for the benchmark for the ROC curve. According to him, the developed model is acceptable if it obtains the ROC value equal or greater than 0.5 otherwise unacceptable. He further detailed that the model will be considered perfect if it reaches at 1. Therefore, this study also set 0.5 as the benchmark for the ROC curve. The following Table 6 shows the details of ROC values obtained from experiments.

This study got above average acceptance in evaluation of the model. In which, the leader decision rules in the model are strong enough with 0.81 area under curve. Whereas, programmer also sustained the acceptable curve by 0.62 value. Therefore, overall 0.71 curve area is benchmark of normal and effective development of the model.

6. Proposed model for software development team composition

Based on the results, this study has given a model for software development team composition. In which, six rules (i.e., see rule number 1, 2, 3, 5, 6, and 8 from Table 3) out of twelve were found effective for team leader role. Whereas, remaining rules were used to make the filter for the model. Basically, this filter will participate in two ways: first, it will give confidence to decision makers to make sharp decision, and the second, it will help the model to be more efficient and less complex. Moreover, programmer rules (mentioned in our previous study) were also used in this model. This stands the beauty of rule-based model that whenever alteration is needed just update the rule sets. Therefore, our future research will inject other roles of development (i.e., designer and tester) in this model. The steps to apply the model are given below:

1. Initialize the variables of the model.
2. Transforms the nominal input possibilities to control the overflow of inputs.
3. Input the team information for finding effectiveness or ineffectiveness.
4. Is the inputted team member from ineffective rule list? If yes: jump to the step no 7.
5. Is the inputted team member from the effective list? If No: jump to step no 7.
6. The member is considered effective for team composition, and omit the step 7.
7. The member is considered ineffective for team composition.

The developed model of the study is logically divided into two parts. First part is consisting steps from 1 to 3 for preprocessing of inputs for decision, and 2nd part is built for making decision based on steps 4 to 7. In which the 1st step is supposed to initialize the six (06) variables by any user-defined names. For instance, one for team role, four variables for four pairs of MBTI, and the last for team member gender. In the same vein, the 2nd step of the study is used to give the scope to the inputs of the user by transforming the input possibilities for the particular variable. For example, r\(\text{TeamRole}(\times 1)\) is the transformation of the team member role into the initialized variable \(\times 1\) by giving condition of input 1 as team leader else programmer: r\(\text{TeamRole}(\times 1)=\text{TeamLeader}',1\) or ‘programmer’, else}. Moreover, the 3rd step of the model is for...
receiving the inputs from the user for applying the decision rules to make efficient teams.

After the input of a team is provided to the system, the 4th step will check the inputs of the team one by one (if more than one) from ineffective rules’ list. If the input possesses the personality composition from those ineffective rules list then the input will jump to step 7, which highlights to the user regarding ineffectiveness of the input. Otherwise, step 5 will be applied to see from another list of effective rules to find the composition effectiveness. Fig. 3 is presenting the flow of the model in the chart with possible steps to use.

Likewise ineffective step, the 7 step will be triggered if the input is not matching from the effective rule list. Otherwise, step 6 will be used to show the effectiveness of the input to the user, and the 7 step will then be omitted. In the end, it should be mentioned that step 4th is considered the filter for the study. Because, it reduces the burden of the selection phase (i.e., step 5) by distributing the ineffective rules directly.

7. Threats to validity

In this study, the major concern of validity is the generalization of the model. Because, this study only used the data of Malaysian university and software industry to develop and test the model. It basically leaves the threats for practitioners from other countries who wish to take this model into use. Hence, this model can further be validated from other countries’ data. Moreover, the model was trained and tested based on the small or medium teams which means that the model cannot be generalized with large scale teams. Similarly, the teams were asked to follow the XP methods to develop the projects. Therefore, one concern of validity can exist while practicing the model on another methods: pair programming. Another concern of validity that, yet, MBTI is mostly used instrument in the software engineering domain [22] but, to some extent, it is also criticized on reliability and validity [44]. In the same vein, researchers [45] have also highlighted that personality preferences measures like MBTI are neither valid nor invalid inherently but the results. Hence, the practitioners have to know that undergone reliability and validity changes may cause, to some extent, the elusiveness and variations in future for leaders’ traits. But, this threats to validity can be controlled by adding the new rules in the model. Additionally, this study used MBTI to measure the personality preferences of participants. As a result, this model cannot be used for those who measure the personality preferences from other tests: Five Factor Models (FFM) or Big Five.

8. Conclusion

The conclusion can be drawn with the claim that personality types of software team roles fluctuate by the gender type. Besides, this study revealed and ensured that gender should be kept in the consideration when composing teams based on personality types. The results emanated from descriptive experiments helped to extract the accepting or eliminating traits based on the percentage obtained from frequency analysis. Whereas, the eliminating results were also used in the filter of the model. Moreover, team lead role was prominent with ENFJ personality type for the male leader and female leader was significantly found different with INTJ and ENTJ personality types. Ultimately, the results of this study differs from the study by Gorla and Lam [16] in which N and F traits were mentioned effective for team leader. But, the current study asserts that F trait is suitable for male leader whereas female leader is contrasted with T trait. Therefore, based on descriptive results, the above claim creates the sensitivity in considering male or female leader. Furthermore, the results of predictive experiments also supported to decide the acceptance or elimination of the traits based on rules. It means that predictive results helped to decide the use of traits or avoid of traits in various situations. For future work, this study can be extended by adding other development roles: designer or tester. Because, the results emanated from these studies can either be used for team leader or programmer. Therefore, in order to make a broader model, the extension of rules is required for remaining roles of software development. Moreover, the results of this study should be cross validated with other countries data as the model is trained and tested with Malaysian data only.

References


[27] N.R. Mead, Software engineering education: how far we’ve come and how far we have to go, J. Syst. Softw. 82 (4) (2009) 571–575.


