ISSN 2521-0122 (Online)
ISSN 2519-7991 (Print)
Vol (8): Issue (4)
October - December 2024

# LGURJCSIT

## Journal of Computer Scienceand Information Technology

Lahore Garrison University Research

**LGURJCSIT ISSN: 2519-7991(Print), 2521-0122 (Online)**

**SCOPE OF THE JOURNAL**

The LGURJCSIT is an innovative forum for researchers, scientists and engineers in all domains of computer science and technology to publish high quality, refereed papers. The journal offers articles, survey and review from experts in the field, enhancing insight and understanding of the current trends and state of the art modern technology.

Coverage of the journal includes algorithm and computational complexity, distributed and grid computing, computer architecture and high performance, data communication and networks, pattern recognition and image processing, artificial intelligence, cloud computing, VHDL along with emerging domains like Quantum Computing, IoT, Data Sciences, Cognitive Sciences, Vehicular Automation. Subjective regime is not limited to aforemationed areas, Journal policy is to welcome emerging research trends in the general domain of computer science and technology.

**SUBMISSION OF ARTICLES**

We invite articles with high quality research for publication in all areas of engineering, science and technology. All the manuscripts submitted for publication are first peer reviewed to make sure they are original, relevant and readable. Manuscripts should be submitted through online journal submission system (https://lgurjcsit.lgu.edu.pk/index-.php/lgurjcsit/user/register).

*To submit manuscripts Online with attach file is strongly encouraged, provided that the text, tables, and figures are included in a single Microsoft word/pdf file. Submission guidelines along with official format is available on the following link:*

*https://lgurjcsit.lgu.edu.pk/index.php/lgurjcsit*

**Contact:** For all inquiries, regarding call for papers, submission of research articles and correspondence, kindly contact at the folloing address:
**LGURJCSIT,** Sector C, DHA Phase-VI, Lahore, Pakistan.
**Phone:** +92-042-37181823
**Email:** lgurjcsit@lgu.edu.pk


***Published by: Faculty of Computer Sciences, Lahore Garrison University, Lahore, Pakistan***

**Volume 8** **Issue 4** **October – December 2024**

**Rana Hammad Hassan, Malik Tahir Hassan**

Digitizing TVET Education through Unified Analysis of Personality and Learning Skills 1- 11

Hassan et al. LGURJCSIT 2024
ISSN: 2521-0122 (Online)
ISSN: 2519-7991 (Print)

LGU Research Journal of Computer Science & IT
doi: 10.54692/lgurjcsit.2024.083613
Vol (8): Issue (4), October – December 2024

# Digitizing TVET Education through Unified Analysis of Personality and Learning Skills

Rana Hammad Hassan*, Malik Tahir Hassan
School of Systems and Technology, University of Management and Technology, Lahore, Pakistan.

Email: rana_hammad@live.com

**ABSTRACT:**

***Personality exerts a pivotal influence on Technical and Vocational Education and Training (TVET) by shaping individuals' learning approaches, social interactions, and practical skill applications. Customizing educational strategies and curriculum designs to suit diverse personality profiles cultivates a more conducive learning environment, enabling each student to optimize their abilities and contribute effectively to their chosen vocational fields. Incorporating personality considerations into TVET programs not only enhances skill acquisition but also fosters personal growth and professional success among learners. This research aims to test our hypothesis that can we predict TVET courses based on personality traits and whether is there any role of demographics (age and gender), and examination performance scores in TVET course prediction. Data for this research were collected from one of the largest TVET training providers and a five-fold cross-validation technique with MDS analysis and Decision Tree methods were used. The result discovered that TVET courses can be predicted based on personality traits and demographics and Examination Scores have a significant role in TVET course prediction. Prediction accuracy of 78% is achieved by SVC, 76% by Naïve Bayes, and 74% by Random Forest classifies.***

**KEYWORDS:** TVET Digitization, Big Five Personality, TVET Systems, ICT Integration in TVET

## 1. INTRODUCTION

TVET plays a crucial role in cultivating a skilled workforce through formal training institutes and industry apprenticeship programs. It serves as a key driver for poverty alleviation by creating diverse pathways for skills development, thereby enhancing career progression opportunities for youth. Our previous exploration [1], [2] revealed that the majority of TVET trainees, accounting for 94%, are young individuals aged between 15 to 25 years, with 86% having completed Matric-level education in Pakistan. This underscores TVET's pivotal role in sustainable development [3], particularly in empowering youth. Many of these trainees are former dropouts living at or below the poverty line, who are acquiring skills to support their families. The United Nations Educational, Scientific and Cultural Organization (UNESCO) Sustainable Development Goals (SDGs) for 2030 [4], including Quality Education (SDG-4), Gender Equality (SDG-5), Decent Work and Economic Growth (SDG-8), and Partnership for the Goals (SDG-17), are highly relevant to TVET. There is a pressing need [5] to introduce digital and modern initiatives to meet the challenges posed by Industry 5.0 in terms of products and services. Personality is undeniably one of the most intriguing facets of human existence, encompassing enduring patterns of thoughts, emotions, and behaviors that distinguish individuals from each other. The integration of personality traits into computing systems and applications marks a transformative era in technology. Over the past decade, the field of personality computing has experienced substantial growth [6], [7]. Specifically, personality computing has found applications in Automatic Personality Recognition, recommendation systems, and human-computer interaction research [8], [9], [10]. By delving into human personality, we can establish connections between personality traits and career paths, skills acquisition, and learning outcomes [11]. This understanding enhances our ability to tailor educational and vocational interventions to individual traits, thereby optimizing personal and professional development. The BFI model and the Myers–Briggs Type Indicator (MBTI) model trace their origins to theories developed in the late 19th and early 20th centuries. Both models aim to categorize individuals based on various personality traits, although they employ different frameworks and approaches. The Big Five-Factor Inventory [12] is a widely used and well-established tool for assessing personality traits, employing a distinct set of questions to measure these dimensions.


***Figure 1: BFI OCEAN Personality Dimensions***

The BFI, also known as the OCEAN model, encompasses five fundamental personality dimensions and is widely recognized in psychological testing [13]. These dimensions—Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism—each capture distinct aspects of personality, as shown in Figure 1.

The rationale behind this research stems from the discrepancy observed despite substantial international funding from organizations such as the United States Agency for International Development (USAID), the Department for International Development (DFID) of the United Kingdom, the British Council, the Japan International Cooperation Agency (JICA), and the World Bank, alongside national funding from both public and private sectors in Pakistan's Technical and Vocational Education and Training (TVET) sector. Despite these investments, the employability rate remains low at 38% [14]. Concurrently, Pakistan faces significant challenges with skill shortages, which contribute to the difficulties encountered in infrastructure development projects [15]. The inadequacy in employability not only impedes economic development [16], [17] but also undermines its equity and sustainability [18]. Previous research conducted by our team assessed the integration of ICT in TVET and emphasized the critical role of digital initiatives in addressing industry challenges.

Data collection for this study was carried out in collaboration with the Punjab Vocational Training Council (PVTC), one of the largest providers of Technical and Vocational Education and Training (TVET) in Punjab, Pakistan [19]. PVTC oversees over three hundred and fifty Vocational Training Institutes (VTIs) across the province, offering training in more than a hundred trades. Enrollment sessions are held biannually, in January and July. Data collection spanned four enrollment sessions from 2019 to 2021. The collected data includes comprehensive information on TVET trades knowledge, VTI details, trainee personal information, and individual BFI personality traits. This data was systematically stored in a database. Throughout all phases, the progress of each trainee's coursework was meticulously monitored, and data was annotated with their actual course progress. This labeled dataset formed the basis for subsequent analysis and predictions in this research. Data collection utilized both an online application and paper-based questionnaire survey forms. A total of 1075 trainee records were cleaned, preprocessed, and labeled with course progress over a period exceeding two years. Approval for data collection and research publication was obtained from PVTC management before conducting the study. To comprehend the challenges contributing to the low employability of TVET trainees, an investigation into the admission and course selection processes of TVET institutes was undertaken. It was discovered that TVET trainees typically gain admission without undergoing academic tests, career counseling, or personality aptitude evaluations. Further exploration involved conducting a comprehensive survey among 901 TVET trainees regarding their course selection process. The findings revealed that 49% of trainees base their course choice on personal knowledge, 25% rely on recommendations from others, and 10% are admitted based on informal referrals. Notably, none of the trainees underwent academic testing, personality assessments, or career counseling during admission. This study aims to analyze the relationship between personality traits and TVET courses. By correlating learning aptitude with personality dimensions, the research seeks to predict dropout rates, student failures, and success rates among TVET trainees more accurately and at an earlier stage. In this research, we utilized several statistical and machine learning techniques to analyze the relationship between BFI personality traits, age, gender, exam scores, and the selection of Technical and Vocational Education and Training (TVET) courses. Five-fold cross-validation [20] was employed to assess the predictive accuracy of our machine-learning models. This method involves dividing the dataset into five equal subsets, where each subset is used as a testing set while the remaining four subsets are used for training. This process is repeated five times, ensuring that all data points are evaluated across different training-testing splits, thereby enhancing the reliability of our predictions. Multidimensional Scaling (MDS) analysis was also utilized to visualize and interpret the high-dimensional relationships within our data, simplifying the understanding of complex patterns. Decision Tree techniques were employed to outline potential outcomes based on our gathered data, providing insights into effective course selection strategies for TVET trainees. Previous studies in TVET [21] have successfully employed k-fold cross-validation to predict student outcomes using various machine learning models such as Decision Trees, Neural Networks, Logistic Regression, and Naïve Bayes. Our specific study found that the Support Vector Classifier (SVC) achieved the highest accuracy of 78%, followed by Naïve Bayes with 76%, and Random Forest with 74% accuracy in predicting TVET trades based on BFI personality traits, age, gender, and exam scores of 75+. These findings underscore the potential of machine learning techniques to optimize course placements and enhance outcomes in TVET programs. The structure of the remaining paper is outlined as follows: Section II offers an overview of the literature review, encompassing relevant theories and studies. Section III delineates the methodology employed in this study, while Section IV presents the data analysis. Section V discusses the hypotheses formulated for this study, and Section VI presents the results. Section VII provides an interpretation of the results, followed by Section VIII which compares the findings with existing literature. Finally, Section IX discusses the conclusions drawn from the study and suggests future research directions.

## 2. LITERATURE REVIEW

Five-fold cross-validation is a robust technique used in data analysis to validate and assess the performance of predictive models by dividing the data into five subsets, with each subset serving as a testing set iteratively while the remaining data act as training sets. This method helps mitigate overfitting and provides a more reliable estimate of model accuracy across different data samples. K-fold cross-validation has been successfully applied in various TVET studies for prediction and analysis. For instance, a case study in Ethiopia [22] involving 7,561 TVET trainees revealed that factors such as Education mode of training experience, Level, Purpose of Assessment, Candidate's category, Age, Sector, Sex, and Employment type significantly influenced students' performance in practical exams. Similarly, factors including Level, Candidate's category, Sex, Age, Employment type, Sector, Purpose of Assessment, and Education mode of training experience were found to be influential for theory exams. In another case study from Benin [23], mathematics and biology marks were used to develop a prediction model for 325 trainees. The study demonstrated that Support Vector Classifier (SVC) and Random Forest algorithms effectively collaborated to provide optimal guidance for learners in the TVET context. Personality traits have increasingly influenced recommendation systems across various domains, with researchers advocating for their integration alongside social characteristics. The Big Five Personality model [24], [25], [26] has proven effective in predicting and enhancing academic performance. In this study, we utilized BFI personality traits to extract knowledge for predicting TVET courses. Despite the absence of a specific BFI dataset for TVET in Pakistan, this research contributes by creating and enhancing such a dataset. Future extensions of this research aim to predict TVET trades for new trainees based on their BFI Personality traits. Personality-driven recommendation systems [27], [28] have demonstrated higher accuracy compared to traditional methods. This study seeks to establish a relationship between TVET trades and BFI personality traits, laying the groundwork for personalized TVET recommendations based on personality traits.

## 3. METHODOLOGY

A research methodology is a set of methods and practices used to collect and analyze data related to a particular research subject. The research methodology for this study is illustrated in Figure 1. Step 1 involves data collection, cleaning, and pre-processing. Step 2 includes initial data analysis and visualizations. In Step 3, hypotheses for the research are formulated, followed by Step 4 where results are evaluated. Step 5 focuses on interpreting the results, and Step 6 concludes with a comparison of findings with existing literature. This section addresses data collection, cleaning, and preprocessing. Section IV covers data analysis, while Section V presents the hypotheses. Section VI discusses the results, and Section VII provides interpretations. Section VIII compares findings with existing literature.

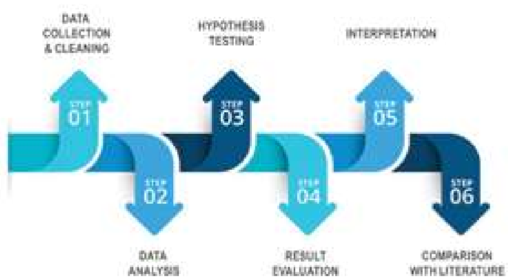


*Figure 2: Research Methodology*

A total of 1075 trainee records from eight VTIs were collected across four admission sessions. The dataset includes trainees' personal information, TVET trade details, and BFI Personality traits. To ensure user security and privacy, each trainee was assigned a unique profile code, and personal details were excluded from the final dataset. Data cleaning involved assessing data quality, removing missing values, eliminating outliers, and resolving unclear entries. After cleaning, 747 trainee records were retained for analysis. These 747 trainees were monitored for course progress and labeled with BFI Personality traits over two years. The study aims to identify personality traits associated with successful trainees in TVET courses. Out of the 747 trainees, 299 were selected based on their achievement of passing the course with a final grade of 75+ marks. Table 1 presents the finalized dataset used for this study.

*Table 1: Dataset D*

| Sr. No. | Trade | |
|---|---|---|
| 1 | Computer Application & Office Professional | 76 |
| 2 | Clinical Assistant | 51 |
| 3 | Auto Mechanic | 23 |
| 4 | Refrigeration & Air Conditioning | 68 |
| 5 | Dress Making | 28 |
| 6 | Computer Hardware & Network Professional | 28 |
| 7 | Computer Application for Business | 28 |
| | Total | 299 |

In a dataset, a data dictionary serves as a structured repository defining metadata and attributes of elements within it. It acts as a reference for comprehending the data's meaning, structure, and relationships. Table 2 presents the data dictionary for Dataset D, where the first column lists dataset attributes or columns, the Description column offers field details, and the Value column displays possible attribute values.

*Table 2: Data Dictionary for Dataset*

| ATTRI-BUTES | DESCR-IPTION | VALUE |
|---|---|---|
| ProfileCode | Profile Code of Trainees | Unique profile code for each Trainee |
| Trade | Trade Name | All Trades Name m-entioned in Table 3 |
| Gender | Gender | Male, Female |
| Age | Age | Age of Trainee bet-ween 14 to 35 years |
| Marks | Marks | Final Exam Marks score between 55 to 99 |
| Grade | Grade | A, A+, B, B+, Fail, Not Eligible for Exam (NE), Not Appear in Exam (NAExm) |
| Score_O | BFI Score_O | Individual Trainee's Openness to Express Score |
| Score_C | BFI Score_C | Individual Trainee's Conscientious Score |
| Score_E | BFI Score_E | Individual Trainee's Extroversion Score |
| Score_A | BFI Score_A | Individual Trainee's Agreeableness Score |
| Score_N | BFI Score_N | Individual Trainee's Neuroticism Score |

Table 1 depicts a dataset where 64% of trainees are male and 36% are female. Figure 3 illustrates the age distribution of these trainees categorized by gender as presented in Table 1.


***Figure 3: Gender Wise Age Distribution***

The age bracket for all trainees ranges from 15 to 35 years, with 94.64% falling between 15 to 24 years of age. Interestingly, male trainees have a maximum age bracket of 19 years, while female trainees reach up to 21 years. This suggests that female trainees typically commence TVET education at an age bracket two years older than their male counterparts. Figure 4 illustrates the distribution of final exam scores from Table 1, dataset. Notably, the highest score achieved by trainees is 78, with fewer students achieving higher scores as the exam score increases. The maximum achieved score by TVET trainees is 98 marks. Figure 5 shows a boxplot of BFI OCEAN score visualization.

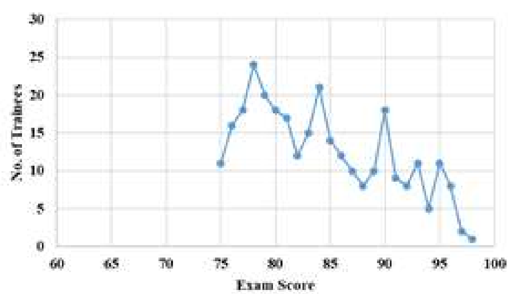


***Figure 4: Exam Score Distribution***

Figure 5 displays the boxplot for BFI OCEAN scores, ranging from 0 to 40 scale. It highlights that trainees scoring 75 or higher on the final exam exhibit distinct scores across all five dimensions of the BFI personality traits. Specifically, the Extroversion scores range from 18 to 26, Agreeableness scores range from 27 to 34, Conscientiousness scores range from 29 to 37, Neuroticism scores range from 15 to 26, and Openness to Experience scores range from 25 to 32. The dataset is divided into 70% training data and 30% test data for analysis, formatted as a Microsoft Excel file.

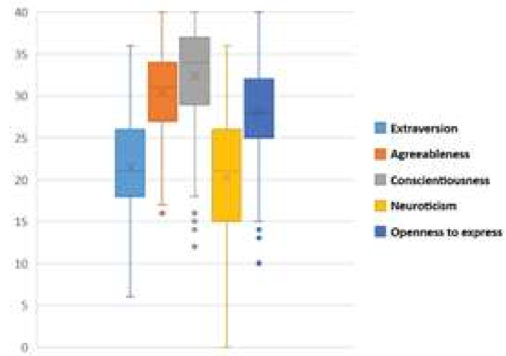


***Figure 5: BFI Score Boxplots***

## 4. DATA ANALYSIS

In machine learning, cross-validation, also referred to as resampling, is a crucial method for evaluating a model's performance on unseen data, thereby assessing its ability to generalize. Among various techniques, 5-fold cross-validation divides the dataset into five equally sized subsets or folds. The model is trained and evaluated five times, using each fold once as the validation set while the remaining four folds serve as the training set. For our study involving 299 TVET trainees (detailed in Table 1), we applied 5-fold cross-validation to predict the Trade Name, utilizing a dataset that includes attributes such as profile code, age, gender, the five BFI OCEAN scores, and examination scores. Figure 6 illustrates the process of 5-fold cross-validation.

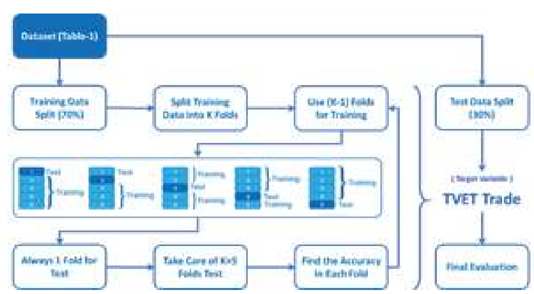


***Figure 6: 5-Fold Cross Validation***

As depicted, Dataset D is divided into training and test sets using a 70:30 ratio. Subsequently, the model is trained on the training set, and its performance is evaluated using a validation set. Performance metrics, including accuracy and standard deviation, are recorded for each iteration. This process is repeated five times, employing each fold once as the validation set. The performance metrics from all five folds are averaged to establish the overall performance of the model. The target variable under investigation is "TVET Trade". For predictions, we employed the following five machine-learning classifiers:

- Logistic Regression
- Naïve Bayes
- Random Forest
- SVC
- Decision Tree

These classifiers were employed to predict the TVET Trade based on dataset attributes, including BFI OCEAN scores, demographic information of trainees, and their examination performance scores. We utilized Multidimensional Scaling (MDS) and the Decision Tree algorithm to determine the optimal features and values from trainees' demographics and examination performance for predicting TVET course outcomes. MDS serves as a valuable tool for visualizing and comprehending intricate data relationships and similarities by projecting them into a reduced-dimensional space while retaining critical structural information. A Decision Tree is a supervised machine-learning algorithm utilized for both classification and regression tasks. It functions by recursively dividing the data into subsets based on input feature values. Each division is guided by decisions rooted in feature values, aiming to create homogeneous subsets relative to the target variable.

## 5. HYPOTHESIS

The fundamental focus of this study is the correlation analysis between skills and personality dimensions. Successful mapping of personality dimensions to skill acquisition would allow for reliable prediction of course success, reduction in dropout rates, and effective influence on career counseling for individuals. This would not only enhance the success rates of TVET programs but also optimize return on investment. Drawing on BFI personality traits, specific age, gender, and examination scores, we have formulated the following hypotheses:

- Null Hypothesis (H0): It is not possible to predict TVET course outcomes based on BFI personality traits, and there is no significant influence of trainees' demographics (age and gender) and examination performance scores on predicting TVET course success.
- Alternative Hypothesis (H1): TVET course outcomes can be successfully predicted based on BFI personality traits, and there is a significant role of trainees' demographics (age and gender) and examination performance scores in predicting TVET course success.

These hypotheses articulate the central aims of exploring the relationship between personality traits, demographics, examination performance, and the success of TVET programs. They provide a framework for testing whether these factors collectively contribute to predicting outcomes in vocational education and training.

## 6. RESULT

In this section, we present the results of the 5-fold cross-validation predictions, as well as the outcomes of the MDS and Decision Tree analyses. Table 3 displays the predictions made by five classifiers on dataset D, quantified by metrics such as Area Under the Curve (AUC), Classification Accuracy (CA), F1 Score, Precision, and Recall [29].

***Table 3: 5-Fold Cross Validation Classifier Results***

| Sr. No. | Models | AUC | CA | F1 | Prec. | Recall |
|---|---|---|---|---|---|---|
| 1 | Logistic Regression | 0.775 | 0.375 | 0.359 | 0.362 | 0.375 |
| 2 | Naïve Bayes | 0.760 | 0.760 | 0.339 | 0.351 | 0.334 |
| 3 | Random Forest | 0.749 | 0.749 | 0.361 | 0.367 | 0.371 |
| 4 | SVC | 0.787 | 0.787 | 0.382 | 0.413 | 0.415 |
| 5 | Decision Tree | 0.609 | 0.609 | 0.306 | 0.307 | 0.308 |

ROC AUC is a widely used performance metric for assessing binary classification models. It measures the model's capability to distinguish between positive and negative classes across various threshold settings. Classification Accuracy (CA) evaluates the overall correctness of predictions. The F1 score provides an alternative method for evaluating machine learning models, focusing on their ability to predict outcomes across different classes rather than overall accuracy, as assessed by the accuracy metric. This score combines two complementary metrics, precision, and recall, into a single value, making it a popular choice in recent studies. It represents the harmonic mean of precision and recall, thereby striking a balance between these two metrics. Precision also referred to as positive predictive value, indicates the ratio of true positive predictions to the total number of instances correctly identified as positive. Recall measures how effectively a machine learning algorithm identifies all positive instances, specifically true positives, among all actual positive examples in the dataset.

The Support Vector Classifier (SVC) attained the highest accuracy at 78%, whereas the Decision Tree achieved the lowest at 68%. Four out of the five classifiers achieved accuracies exceeding 74%. In terms of Classification Accuracy (CA), SVC performed the best with a score of 0.415, while the other four classifiers achieved scores above 0.30. Similarly, all five classifiers obtained F1 scores exceeding 0.30. Regarding Precision, SVC achieved the highest score of 0.41, whereas the remaining classifiers ranged from 0.03 to 0.36. Finally, SVC also achieved the highest Recall score of 0.41, with the other four classifiers maintaining scores around 0.30. In conclusion, SVC showed superior performance with the highest accuracy of 78% and scored highest in CA, Precision, and Recall among the classifiers evaluated, Decision Tree had the lowest accuracy of 68% and Four out of five classifiers achieved accuracies above 74%, demonstrating overall strong performance across metrics.

Figure 7 presents a Radar Graph visualizing the prediction accuracies for each fold in the 5-fold cross-validation performance. The five folds are depicted separately, with each of the five applied classifiers color-coded, showing their respective accuracy values for each fold. Radar graphs are particularly useful for comparing multiple variables across different categories, offering a concise overview of the distribution and consistency of prediction accuracy across folds.

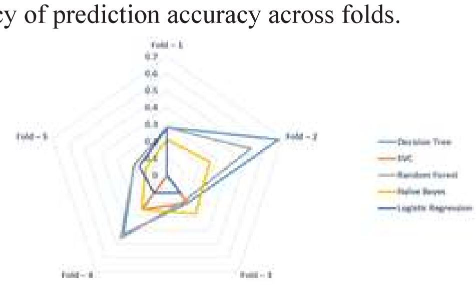


***Figure 7: Fold Wise Accuracy***

To understand the role of TVET trainee's demographic and Examination Score, we have used MDS and Decision Tree. MDS to depict the Trade Wise similarity within the dataset mentioned in Table 3 is illustrated in Figure 8. MDS is a statistical technique used to analyze and visualize the similarity or dissimilarity between objects or cases based on their pairwise distances or dissimilarities. It projects the data points into a lower-dimensional space while preserving the relative distances as much as possible from the original high-dimensional space. In this context, MDS helps to visualize and interpret the underlying structure of trade-wise similarities among the dataset's entries, providing insights into how different trades or categories relate to each other in terms of their attributes or characteristics.

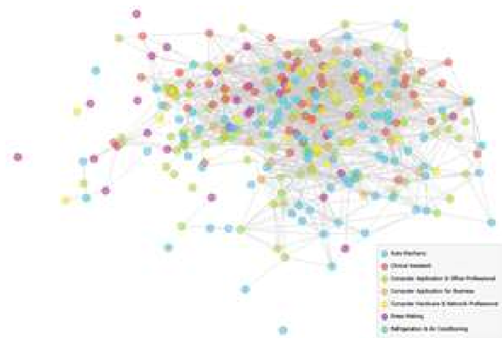

***Figure 8: MDS Analysis***

Figure 9 displays the Decision Tree constructed using the dataset described in Table 3. Decision Trees are hierarchical structures that recursively partition data based on feature values, facilitating clear visualization of decision-making processes within machine learning models. In this instance, the Decision Tree diagram provides insights into how gender influences the classification outcomes of the dataset across multiple folds.

We have visualized the decision tree up to five levels as shown in Figure 9. At the root node of the decision tree, Gender (Male / Female) emerges as the decisive factor. At the second child node, the determining factor shifts to the Result, which represents the performance in examination scores. Moving to the third level of nodes, the BFI personality trait of Agreeableness and the variable Age become decisive factors for both male and female genders. Similarly, at the fourth level of the tree, both genders share the Result of examination scores and Age as common decisive factors. However, at this same level, Female BFI personality trait Extroversion becomes a deciding factor, while Agreeableness remains the decisive trait for males. This distinction indicates gender-specific differences in personality traits. At the leaf nodes of the decision tree for females, BFI traits such as Openness, Agreeableness, and Extroversion are prominent, whereas, for males, traits like

Extroversion, Openness, and Neuroticism play a decisive role. In conclusion, the decision tree analysis reveals that demographic factors such as age and gender, along with examination scores and BFI personality traits, significantly influence the selection of TVET trades. The decision tree analysis highlights the pivotal role of demographic factors (age and ender), examination scores, and BFI personality traits in determining TVET trade outcomes. Gender differences are evident in the selection criteria, particularly in the preference for certain BFI personality traits.

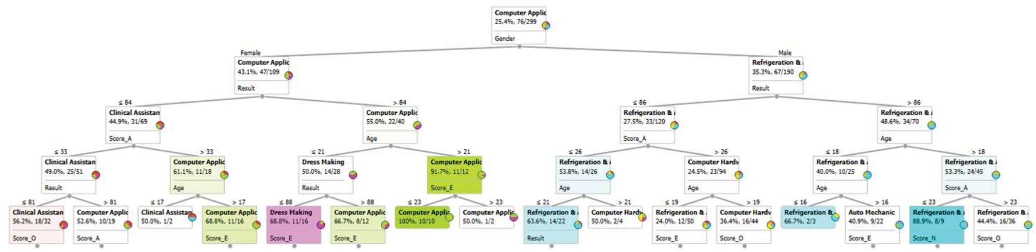

***Figure 9: Decision Tree of Decisive Variables***

Figure 10 displays the standard deviations of five classifiers, reflecting how widely their predictions vary from the dataset's mean. A low standard deviation, exemplified by SVC at 0.96, indicates predictions closely clustered around the mean, suggesting higher consistency. In contrast, Decision Tree and Logistic Regression show slightly higher deviations, signifying more variability in their predictions. Notably, Naïve Bayes exhibits a significantly higher standard deviation, implying greater dispersion of its predictions from the dataset's mean value.

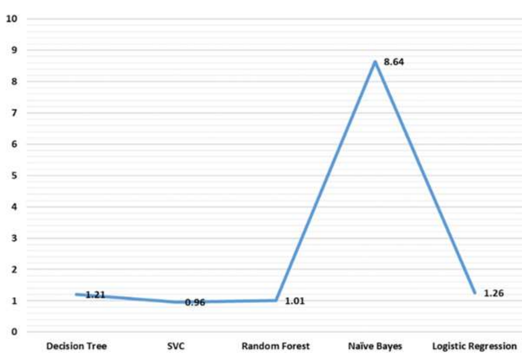


***Figure 10: Standard Deviation***

## 7. INTERPRETATION

In this section, we have examined the relevance of our findings and the implications of the results vis-à-vis our hypothesis. The 5-Fold Classifiers results (Table 3), 5-Fold Accuracy results (Figure 7), MDS Analysis (Figure 8), and Decision Tree (Figure 9) collectively demonstrate a positive correlation between BFI personality traits and TVET trades. Based on these findings, the null hypothesis is rejected, indicating that TVET courses can indeed be accurately predicted using BFI personality traits. Furthermore, Trainees' demographics (age and gender) and their examination performance scores play crucial roles in predicting TVET course outcomes.

Based on our exploration of TVET course prediction using BFI personality traits, Trainee's Demographics, and Examination performance, several conclusions can be drawn from our study and related literature:

- This study represents a pioneering effort in Pakistan and within the TVET sector by successfully mapping learning skills to human personalities.
- Through 5-fold cross-validation, we have effectively predicted TVET courses using BFI personality traits alongside age, gender, and examination scores.
- Our findings are poised to significantly impact various aspects of TVET education, including dropout prediction, career guidance, on-the-job training recommendations, internal assessment enhancements, and tailored course recommendations throughout the TVET training cycle.
- The generic approach we employed for data collection and research can be replicated in other developing countries to gather and analyze TVET-related personality data.
- These results hold practical implications for TVET training providers, policymakers, international funding bodies, researchers, and educators aiming to optimize the performance and return on investment (ROI) of TVET programs.

These conclusions underscore the potential of integrating personality assessments with demographic and performance data to enhance the effectiveness and relevance of TVET education globally. Furthermore, they highlight the transformative impact such predictive models can have on educational outcomes and resource allocation within TVET systems, fostering a more targeted and responsive approach to skills development and career preparation.

## 8. COMPARISON WITH LITERATURE

We conducted Our Model Accuracy (OMA) comparisons against Others Model Accuracy (OTH-MA) using 5-fold cross-validation for a similar TVET dataset, as detailed in Table 4. Notably, we found no comparisons available for Logistic Regression and SVC within similar TVET datasets. However, comparisons for Naïve Bayes, Random Forest, and Decision Tree classifiers were found in existing literature. Our predictions using Naïve Bayes and Decision Tree classifiers outperformed those reported in the literature for similar studies. However, Random Forest yielded better results in other models. In conclusion, SVC achieved a notable accuracy of 78% in our study.

***Table 4: Our Model Accuracy (OMA) comparisons***

| Model Name | OMA | OTH-MA |
|---|---|---|
| Logistic Regression | 0.775 | -- |
| Naïve Bayes | 0.760 | 0.703 [22] |
| Random Forest | 0.749 | 0.950 [23] |
| SVC | 0.787 | -- |
| Decision Tree | 0.609 | 0.194 [21] |

## 9. CONCLUSION AND FUTURE DIRECTIONS

In this study, we evaluated the correlation between TVET learning skills and BFI personality traits and analyzed the influence of TVET trainees' demographics and examination performance scores. We utilized 5-fold cross-validation, MDS analysis, and Decision Tree methods for data analysis. The results revealed a robust relationship between personality traits and learning skills. This study represents an initial exploration into predictive analytics, and its significance lies in its potential dissemination to TVET training providers, policymakers, international funding agencies, researchers, and academia, aiming to enhance the performance and return on investment (ROI) of TVET programs. Future research directions could include:

- Predicting drop-outs in TVET courses.
- Enhancing outcomes of On Job Training (OJT), placements, and career guidance for TVET trainees.
- Improving the quality of internal assessments within TVET courses.
- Developing a framework for a personality-aware TVET course recommendation system.

# REFERENCES


[1] R. H. Hassan and S. M. Awan, "Identification of Trainees Enrollment Behavior and Course Selection Variables in Technical and Vocational Education Training (TVET) Program Using Education Data Mining," *International Journal of Modern Education and Computer Science,* 10(10), pp.14, 2019.

[2] R. H. Hassan and M. T. Hassan, "Identification of Technical and Vocational Education and Training (TVET) trainee's personality attributes which impact skills learning," *in 2024 5th International Conference on Advancements in Computational Sciences (ICACS),* pp. 1–7, Feb. 2024.

[3] United Nations, "US Sustainable Development Goals," United Nations Sustainable Development. Accessed: *[Online],* Available: https://www.un.org/sustainabledevelopment/sustainable-development-goals, Dec. 31, 2018.

[4] R. TVET, "Sustainable Development Goals - Pakistan TVET REFORM Support Programme," Accessed: *[Online],* Available: https://tvetreform.org.pk/sustainable-development-goals/, Jun. 23, 2023.

[5] R. H. Hassan, M. T. Hassan, S. Naseer, Z. Khan, and M. Jeon, "ICT enabled TVET education: a systematic literature review," *IEEE Access*, pp. 1–1, 2021

[6] S. Dhelim et al., "A survey on personality-aware recommendation systems," *Artificial Intelligence Review,* pp. 1–46, 2022.

[7] R. H. Hassan and M. T. Hassan, "A Meta-Analysis of Correlation Between Big Five Personality Traits and TVET Skills Acquisition," *Journal of Xi'an Shiyou University, Natural Science Edition,* 20(07), 2024.

[8] K. K. Jena et al., "E-Learning Course Recommender System Using Collaborative Filtering Models," *Electronics,* 12(1), pp.157, Jan. 2023,

[9] A. J. Fernández-García et al., "Creating a recommender system to support higher education students in the subject enrollment decision," IEEE Access, 8, pp. 189069–189088, 2020.

[10] H. Ning, S. Dhelim and N. Aung, "PersoNet: Friend Recommendation System Based on Big-Five Personality Traits and Hybrid Filtering," *[Online],* Available: https://ieeexplore.ieee.org/document/8675299/, Accessed: Sep. 24, 2023.

[11] P. J. Corr and G. Matthews, "The Cambridge handbook of personality psychology," *Cambridge University Press,* 2020.

[12] D. A. Cobb-Clark and S. Schurer, "The stability of big-five personality traits," *Economics Letters,* 115(1), pp. 11–15, 2012.

[13] S. Algarni and F. Sheldon, "Systematic Review of Recommendation Systems for Course Selection," *Machine Learning and Knowledge Extraction,* 5(2), pp.560-596, Jun. 2023,

[14] 2022 ADB, "Asian Development Bank," Asian Development Bank, [Online]. Available: https://www.adb.org/sites/default/files/linked-documents/52069-001-ssa.pd, Accessed: Oct. 15, 2023.

[15] S. Nisar and R. Asif, "Factors leading to failures of infrastructure development projects in Pakistan: A systematic literature review," *Management Science Letters,* 13(1), pp. 11–22, 2023.

[16] "Labour imbalances," The News. Accessed: Oct. 15, 2023. [Online]. Available: https://www.thenews.com.pk/print/1081706-labour-imbalances

[17] A. Ojha et al., "Barrier Analysis of Effective Implementation of Robotics in the Construction Industry," *in Construction Research Congress 2022,* Arlington, Virginia: American Society of Civil Engineers, pp. 661–669. doi: 10.1061/9780784483961.069, Mar. 2022.

[18] N. Bano, S. Yang, and E. Alam, "Emerging Challenges in Technical Vocational Education and Training of Pakistan in the Context of CPEC," *Economies,* 10(7), pp.153, doi: 10.3390/economies10070153, Jul. 7, 2022,

[19] PVTC, "PVTC Official Website," Punjab Vocational Training Council Official Website. [Online]. Available: http://www.pvtc.gop.pk/

[20] T. T. Wong, "Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation," *Pattern Recognition,* 48(9), pp. 2839–2846, doi: 10.1016/j.patcog.2015.03.009, Sep. 2015.

[21] C. M. Hong et al., "Predicting Students' Inclination to TVET Enrolment Using Various Classifiers," *Pertanika Journal of Science & Technology,* 31(1), 2023, Accessed: Jan. 12, 2024.

[22] A. Messele and M. Addisu, "A model to determine factors affecting students academic performance: The case of Amhara region agency of competency, Ethiopia," *International Research Journal of Science and Technology,* 1(2), pp. 75–87, 2020.

[23] P. HOUNGUE, M. HOUNTONDJI and T. DAGBA, "An Effective Decision-Making Support for Student Academic Path Selection using Machine Learning," *International Journal of Advanced Computer Science and Applications,* 13(11), 2022, Accessed: Jan. 12, 2024.

[24] A. Vinciarelli and G. Mohammadi, "A survey of personality computing," I*EEE Transactions on Affective Computing,* 5(3), pp. 273–291, 2014.

[25] S. Dhelim et al., "A survey on personality-aware recommendation systems," *Artificial Intelligence Review,* pp. 1–46, 2022.

[26] V. Kaushal and M. Patwardhan, "Emerging trends in personality identification using online social networks—a literature survey," *ACM Transactions on Knowledge Discovery from Data (TKDD),* 12(2), pp. 1–30, 2018.

[27] S. Dhelim et al., “Big-Five, MPTI, Eysenck or HEXACO: The ideal personality model for personality-aware recommendation systems,” *arXiv preprint arXiv:2106.03060,* 2021.

[28] R. H. Hassan and M. T. Hassan, “An Ensemble Classifier for TVET Course prediction using Big Five Personality Traits,” *in 2023 25th International Multitopic Conference (INMIC),* pp. 1–6, Nov. 2023,

[29] V. A. Brei, “Machine learning in marketing: Overview, learning strategies, applications, and future developments,” *Foundations and Trends® in Marketing,* 14(3), pp. 173–236, 2020.

## Editorial Policy and Guidelines for Authors

**LGURJCSIT** is an open access, peer reviewed quarterly Journal published by LGU Society of Computer Sciences. The Journal publishes original research articles and high quality review papers covering all aspects of computer science and information technology.

The following note set out some general editorial principles. A more detailed style document can be download at www.research.lgu.edu.pk is available. All queries regarding publications should be addressed to editor at email lgurjcsit@lgu.edu.pk. The document must be in word format, other format like pdf or any other shall not be accepted.

Heading and sub-heading should be differentiated by numbering sequences like, **1. HEADING** (Bold, Capitals) 1.1. ***Subheading*** (Italic, bold) etc. The article must be typed in Times New Roman with 11 font size 1.5 space, and should have margin 1 inches on the left and right.
Length of paper should not be longer than 15 pages, including figures, tables, exhibits and bibliography. Table must have standard caption at the top while figures below with. Figure and table should be in continues numbering. Citation must be in according to the IEEE 2006 style.

# Lahore Garrison University

Lahore Garrison University has been established to achieve the goal of Excellence and quality education in minimum possible time. LGU is situated in the Punjab metropolis city of Lahore is an important milestone in the history of higher education in Pakistan. In order to meet global challenges, it is necessary to touch the highest literacy rate while producing skillful and productive graduates in all fields of knowledge.

## VISION

To be a renowned University in Teaching, Research, Innovation, and Commercialization, providing a conducive environment for the acquisition of latest knowledge so that students may contribute to community support, technical and socioeconomic development.

## MISSION

To play a leading role in technical and socioeconomic development through academic and research excellence while adhering to international quality standards. The University would develop leaders who are multi-disciplinary, value-oriented, creative, and entrepreneurial.

**Contact us:**
**Phone:** +92-042-37181823
**Email:** lgurjcsit@lgu.edu.pk


Published by: Faculty of Computer Sciences, Lahore Garrison University, Lahore, Pakistan.