Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Techskill : Predicting Student's Technical Skills Using Supervised Machine Learning Algorithms

Ms. Vandana Upadhyay[1] , Dr. Parth Gautam[2]

[1]Research Scholar, Mandsaur University.
[2]Assistant professor Computer science and Applications, Mandsaur University.

The growing demand for Information Technology (IT) professionals across domains such as Data Analytics, Artificial Intelligence, and Programming has highlighted the need for students to acquire specialized technical skills. However, many IT graduates struggle to choose a clear career path due to a lack of guidance and misalignment between academic curricula and industry expectations. This research introduces Techskill, a machine learning-based prediction system designed to guide IT students in selecting suitable technical skills for career development. The model uses key student inputs, including understanding of technical subjects, learning styles, and current competencies, to recommend targeted skills. A survey was conducted among over 500 final-year IT/CS students, collecting data on their interests and technical capabilities. The study also incorporated feedback from HR professionals and a review of existing literature to ensure relevance to industry needs. The Techskill model offers personalized skill recommendations in areas such as Data Analysis, Programming, and Designing (System Analysis and UI). It supports a broad range of users—from students selecting electives to professionals aiming to upskill. The research demonstrates how data-driven methods can enhance career readiness, improve curriculum relevance, and bridge the gap between education and evolving industry demands. All experiments were conducted using Python in a Jupyter Notebook environment, underscoring the practical application of machine learning in educational guidance. Using multiple models, the Random Forest Regressor emerged as the most effective, achieving an R^2 score of 0.9549, which reflects a high prediction accuracy of 95.49%. The model's low Mean Squared Error (8.63) and Mean Absolute Error (1.74) further highlight its strength and reliability.

This research offers meaningful insights for academic institutions and career counsellors, helping them guide students toward career paths aligned with their skills and interests. It also empowers students to make informed decisions about academic projects and specialization areas, enhancing their employability in key IT domains such as Data Analysis, Programming, and System/UI Design.

KEYWORDS: IT technical Skills, Machine Learning, Random Forest Regressor, Learning approaches, Project Experience, Programming Language Expert, Feature Selection.

1. INTRODUCTION

The demand for Information Technology (IT) professionals continues to grow across various domains, including Data and Analytics, Artificial intelligence, cloud computing, cybersecurity, mobile application development, etc [1]. Organizations increasingly depend on highly skilled and specialized IT experts. While this surge in job opportunities is promising for IT graduates, many still face uncertainty when choosing a suitable career path. Therefore, it is essential to develop a system that aids IT students in making informed career decisions based on their technical and individual skills.

In this paper, we introduce Techskill model a prediction system designed to guide IT students in selecting appropriate technical skills for their career path using machine learning techniques. The system tailors its recommendations based on key students inputs, including their level of understanding in different technical subjects, learning approaches and technical competencies (e.g., programming languages, databases,etc). Techskill model is suitable for a range of users: senior students seeking guidance on relevant coursework, job seekers aiming to understand market-demanded skills, and employees looking to benchmark their current skillset against industry standards for career advancement.
To develop techskill model we collected data from more than 500 IT/CS Students, capturing their technical skills levels, interested domain and associated skills. This dataset was employed to train and evaluate supervised machine learning models for predicting the most suitable technical skills — Data Analysis and Visualization, Programming, Designing (System analysis and UI) —for new students. Among the various classifiers tested, the Random forest algorithm score of 74.4 and the lowest error rate, outperforming all other models. All experiments and implementations were conducted using the Python programming language in Jupyter notebook environment.

1.1 Problem Definition
The widening gap between academic curricula and the rapidly evolving demands of the IT industry poses a major challenge for students. Despite completing their academic programs, many students fall short of the technical skills and job readiness expected by employers. Predicting IT technical skills on key student’s inputs, including their level of understanding in different technical subjects, way of learning and technical competencies is an area that remains underexplored, contributing to this disconnect. This research aims to address the issue by developing a predictive model for IT technical skills, thereby improving the alignment between educational outcomes and industry requirements.

1.2 Scope
This research seeks to predict IT technical skills based on key student inputs, such as their understanding of various technical subjects, learning approaches, and technical competencies—an area that remains largely underexplored, contributing to the existing disconnect, the study will develop a predictive model to assist academic advisors and educators in customizing learning pathways. The study aims to:

- Bridge the gap between academic learning and industry expectations by aligning student skillsets with current job market demands.
- Support diverse users, such as final-year students choosing electives, fresh graduates planning job applications, and professionals looking to upskill.

- Build a data-driven framework for personalized skill recommendations by training models on a dataset collected from over 500 IT/CS students.
- Evaluate and compare multiple machine learning models, ultimately identifying the Random Forest classifier as the most accurate and effective for skill prediction.
- Demonstrate practical application of machine learning in educational and career guidance contexts using Python and Jupyter Notebook environments.

The research lays the groundwork for future enhancements in intelligent academic counselling systems, especially those tailored to the evolving landscape of technical education and employment.

2. LITERATURE SURVEY

In recent years, numerous researchers have explored the application of machine learning in higher education to enhance graduate attributes and curricula, thereby supporting employability [8]. To examine how machine learning contributes to continuous quality improvement, we reviewed several prior studies that employed various techniques, including Artificial Neural Networks (ANN), Decision Trees (DT), K-Nearest Neighbors (KNN), Gaussian Naïve Bayes (Gaussian NB), Logistic Regression (LR), Neural Networks (NN), Random Forests (RF), Naïve Bayes (NB), and Support Vector Machines (SVM).

Paper [1] investigates the prediction of optimal career paths for computing students using machine learning, evaluating five algorithms on a dataset from the Saudi IT sector. Among them, XGBoost demonstrated the highest performance. The study recommends future improvements through the use of larger datasets, deep learning techniques, and more precise methods for measuring skills to enhance the accuracy of the training data.

In Paper [2], a predictive model was built using five machine learning algorithms to evaluate the employability of IT students. Among the models, the Decision Tree achieved the highest accuracy, followed by Logistic Regression and Support Vector Machine. The performance of each model was assessed using metrics such as accuracy, precision, recall, and F1 score.

As mentioned in [2], the paper analyzes dropout rates in a first-year Computer Science program over five sessions, finding that Logistic Regression outperformed Decision Trees and SVM in terms of F1 score and recall.

In [3], the paper explores how IoT, AI, and ML enhance student learning and address educational challenges. It presents an ML model for predicting academic success with 97.12% accuracy using XGBoost and highlights the influence of social and demographic factors.

According to [4] , the paper analyses the programming behaviour of 220 master's students to predict academic performance using machine learning models, with SVM and neural networks performing best. The study emphasizes the early identification of struggling students and proposes future improvements for greater accuracy.

As per [5], the paper investigates various algorithms and identifies Logistic Regression as the most effective in predicting final exam outcomes, achieving 68.7% accuracy for students who passed and 88.8% for those who failed at the University of Basra. The research emphasizes the role of machine learning in higher education to improve student success rates and reduce dropouts.

In [6], the author predicted which students are most likely to get work after graduation by using data analytics and machine learning techniques such as SVM, LR, ANN, DT, and discriminant analysis. Also, the features used are hard skills, demographics features, extra/co-curricular activities, and internships the data were obtained from student surveys and institutional databases. The SVM classification algorithm achieved an accuracy of 87.26%.

In [7] the authors sought to determine the key factors influencing graduate employability by employing three classification algorithms: Decision Tree (DT), Artificial Neural Network (ANN), and Support Vector Machine (SVM). The study utilized features such as hard skills, soft skills, demographic attributes, extra/co-curricular activities, university-related factors, and internship experience. Data for the analysis were sourced from institutional databases, with the SVM algorithm achieving an accuracy of 66.096%.

Furthermore, a student employability prediction system was developed by [8] using SVM, DT, RF, KNN, and LR algorithms to predict the students' employability, Institutional databases were obtained, and the hard skills, soft skills, and demographic features were used. The results of this research achieved an accuracy of 91% by the SVM algorithm

In [17] predicted the students' employability based on technical skills the institution databases were collected and the following algorithms were applied SVM, LR, DT, RF, AdaBoost, and NB, the highest accuracy achieved is 70% by the RF algorithm. Finally, the authors in [18] developed a model using various machine learning methods DT, RF, NN, and Gaussian NB to forecast candidate hiring by employing different statistical measures on feature selection such as hard skills, demographic features, and professional experience, the highest accuracy was achieved by Gaussian NB with 99%. Table I depicts and summarizes the relevant studies according to their adopted features, dataset sources, ML models, output features, and accuracy of the best-adapted model to answer RQ1.

In summary, the existing literature highlights various approaches to skill prediction and career guidance using machine learning. However, most studies focus narrowly on academic performance or lack real-world industry input. There remains a gap in integrating student learning styles, technical understanding, and current market demands. This research addresses that gap by combining academic and industry perspectives in skill prediction. The findings from previous studies serve as a foundation for developing a more comprehensive and practical model.

TABLE I. COMPARISON OF RELATED STUDIES

Reference	Year	Adopted Features Categories	Dataset Sources	ML Model	Output Features	Accuracy
Hugo [9]	2018	Hard skills, Demographic features, Extra/co-curricular activities, Internship	Student surveys and Institution databases	SVM, ANN, LR, Discriminant analysis, DT	Employability: {Employed, Not Employed}	SVM 87.26%
Othman et al. [10]	2018	Hard skills, Soft skills, Demographic features, Extra/co-curricular activities, University features, Internship	Institutional databases	DT, ANN, SVM	Employability: {Employed, Not Employed}	SVM 66.0967 %
Alghamlas and Alabduljabbar [11]	2018	Hard skills, Soft skills	Student surveys and Recruiter surveys	DT, NB, NN	Matching to industry-required skills	Naïve Bayes 69%
Dubey and Mani [12]	2019	Hard skill, Demographic features, Extra/co-curricular activities	Student surveys	LR, DT, RF, KNN, SVM	Hiring: {Hired, Not hired}	LR 93%
Kumar and Babu [13]	2019	Hard skills, Soft skills, Demographic features, Extra/co-curricular activities,	Student surveys	DT, Gaussian NB, SVM, KNN	Getting a job: {Yes, No}	DT & SVM 98%

		University features				
Casuat [21]	2020	Hard skills, Soft skills, Demographic features	Institution databases	DT, RF, SVM, KNN, LR	Employability: {Employed, Less Employed}	SVM 91%
Casuat & Festijo [14]	2020	Hard skills, Soft skills, Demographic features	Graduate surveys and Institution databases	SVM, RF, DT	Employability: {Employed, Less Employed}	SVM 91.22%

3. METHODLOGY

This section outlines the methodology adopted in our study, including the machine learning algorithms used and the evaluation metrics applied. Fig 1 illustrates the research process, which consists of the following steps:

1. Data collection & understanding
2. Data preprocessing
3. Splitting the dataset into training and testing sets to train and evaluate the model respectively
4. Building the techskill model using five different ML classification algorithms
5. Model evaluation
6. Generating outcomes from the proposed model to predict technical skill of qualified IT students.

   Each step in the process was carefully designed to ensure the reliability and accuracy of the prediction model. During data preprocessing, missing values were handled, categorical data was encoded, and feature scaling was applied where necessary. The classification algorithms used include Decision Tree, Random Forest, K-Nearest Neighbours, Support Vector Machine, and Logistic Regression. To evaluate the performance of these models, metrics such as Accuracy, Precision, Recall, F1-score, and R^2 Score were calculated. The methodology ensures a structured and data-driven approach to identifying the most suitable technical skills for IT students, aligning both academic inputs and industry relevance.

Process Flow - In this methodology below steps are followed:

Fig 1: The Research Methodology

3.1 Dataset: Collection & Understanding

The dataset used in this research was collected through a structured survey conducted among final-year IT and Computer Science (CS) students. To ensure the relevance and comprehensiveness of the survey, we began by reviewing existing literature to identify the technical skills commonly explored in prior research. Additionally, we consulted HR managers from various multinational companies to gain insights into the most in-demand technical skills in the current job market. Based on these inputs, we designed a well-rounded set of parameters for the survey. The final questionnaire was created using Google Forms and distributed online to a diverse group of IT/CS students to gather their responses effectively. A brief description of each feature selected, and its sample value is described in Table below.

As part of data understanding below points were considered:

- Understand the structure and key characteristics of the dataset.
- Identify important variables and data types.

TABLE II. DESCRIPTION OF DATA FEATURES

Features	Options
Computer Languages and Tools	C++ ,Java, PHP,Python, SQL ,R ,Spark ,Tableau ,Power BI
Technical Subjects	Machine Learning, Data Warehouse, Data Mining, Artificial Intelligence, Linux, Web Technology, Software Engineering, RDBMS ,Big Data ,Cyber Security, Cloud Computing ,Operating System
Learning Approach and Interest/Aspiration	Prefer to learn new technical skills ,hours per week spend on skill enhancement activities, rating proficiency in problem-solving skills, Interested area, Participated in coding competitions or hackathons, technical projects have completed so far, comfortable with debugging and optimizing code ,Technical skills etc

3.2 Data Pre-processing

Data pre-processing is an important phase in the machine learning pipeline that prepares raw data for modelling. Here goal is to make sure the data is clean, consistent, and suitable for use with machine learning algorithms. Key steps in data pre-processing include:

3.2.1 Data Cleaning:

Steps which are used to clean data are below:

- Handling Missing Values: Address missing data by removing incomplete records, imputing values using statistical methods (such as mean, median, or mode), or employing models that can handle missing inputs.
- Noise Removal: Identify and correct or remove noisy data caused by inaccuracies during data entry or collection.
- Outlier Detection**:** Detect and manage outliers.data points that significantly deviate from the norm using statistical analysis or visualization tools.
- Duplicate removal: Checked for and eliminated duplicate records to maintain data integrity.
- Feature consistency check: Verified consistency across related fields to avoid logical errors in the dataset.
- Replace null with correct values, removed personal information from the dataset. Standardization of formats: Ensured consistency in data formats such as dates, text cases, and numerical units.

```
In [16]: data.isnull().sum()

Out[16]: Name                                                                       1
         Gender                                                                     0
          Course                                                                    0
         College                                                                    4
         city                                                                       0
          Certifications                                                           12
         C                                                                          0
         C++                                                                        0
         Java                                                                       0
         PHP                                                                        0
         Python                                                                     0
         SQL                                                                        0
         R                                                                          0
         Spark                                                                      0
         Tableau                                                                    0
         Power BI                                                                   0
         Machine Learning                                                           0
         Data Warehouse                                                             0
         Data Mining                                                                0
         Artificial Intelligence                                                    0
         Linux                                                                      0
         Web Technology                                                             0
         Software Engineering                                                       0
         MERN                                                                       0
         Big Data                                                                   0
         Cyber Security                                                             0
         Cloud Computing                                                            0
         Operating System                                                           0
         How do you prefer to learn new technical skills?                           0
         How many hours per week do you spend on skill enhancement activities?      0
         How do you rate your proficiency in problem-solving skills?                0
         Interested area                                                            0
          Participated in coding competitions or hackathons                         0
         How many technical projects have you completed so far                      0
         How comfortable are you with debugging and optimizing code?                0
         Technical skills                                                           0
         Unnamed: 36                                                              500
         dtype: int64
```

Removed the null values refer the below

output.

```
In [17]: data.duplicated().sum()

Out[17]: 0
```

3.3 Exploratory Data Analysis (EDA)

We used Exploratory Data Analysis (EDA) with different types of charts to understand the data better and see how different features relate to the target variable.

First, we did univariate analysis, which looks at one feature at a time to understand its values and how they are spread out. Then, we did bivariate analysis to explore the relationship between two features, using tools like scatter plots or correlation charts. After that, we performed multivariate analysis to study how several features interact with each other at the same time.

To clearly show these findings, we used visual charts like Vertical & Horizonal bar graphs, , and clustered bar graphs, which make the data easier to understand and interpret.

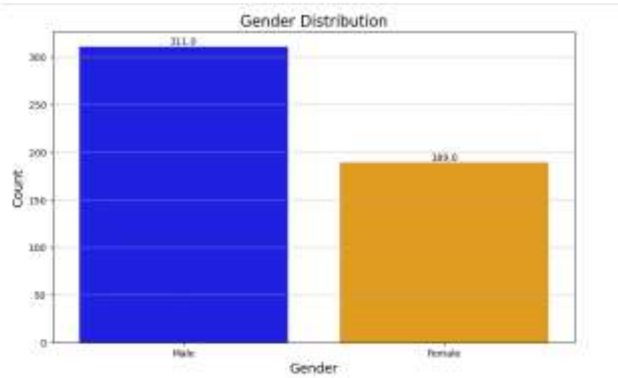

Fig 2: Gender Distribution

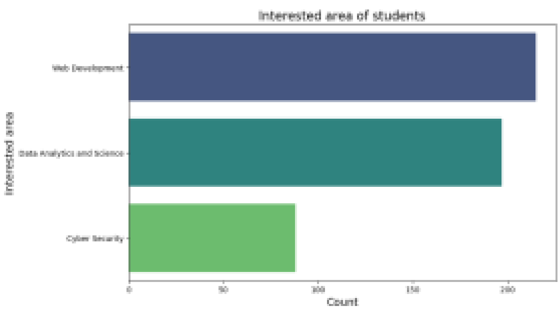

Fig 3: Student's Interested area Distribution

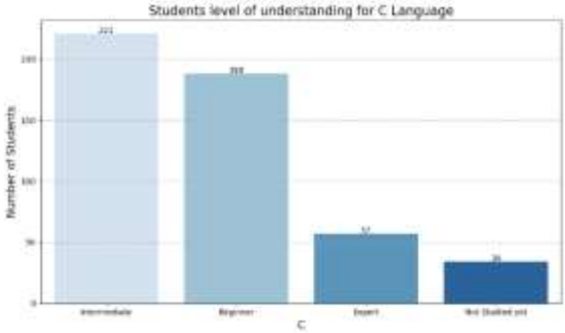

Fig 4: Students level of understanding for C Language Subject

Similar analysis we did for all other subjects like - C++,Java, PHP, Python, SQL, R, Spark, Tableau, Power BI, Machine Learning -

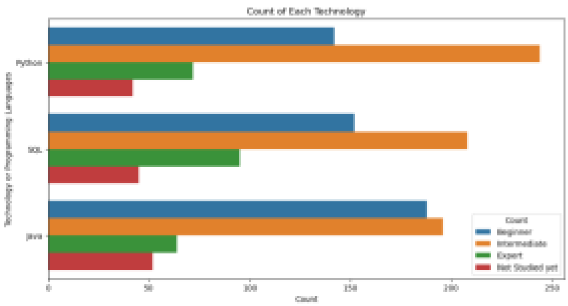

Fig 5: Count of Java , SQL and Python Technology

3.4 Model Selection ,Training and Hyperparameter Tuning

Machine learning model selection involves choosing the most suitable algorithm based on the dataset, problem type (classification or regression), and performance metrics. It helps ensure optimal accuracy, efficiency, and generalization on unseen data.

During model training, a dictionary called param_grids is used to store the hyperparameters for different machine learning models. This dictionary works with grid search methods like GridSearchCV to test various hyperparameter combinations and find the best one for each model. Below is a breakdown of how it is applied during model training:

3.4.1 Linear Regression Model: "Linear Regression is one of the easiest and most widely used methods for predicting values." It works by assuming that there is a straight-line relationship between the input features independent variables and the value we want to predict dependent variable.

No hyperparameters are specified for tuning because linear regression typically doesn't require complex parameter adjustments. It's used as-is for training and evaluation.

3.4.2 Random Forest Model: "Random Forest is a machine learning method that builds many decision trees during training and then takes the average of their results to make predictions.**"** It works well with large datasets and is especially useful when the relationship between the features and the target is complex or not straightforward.

Several hyperparameters are defined for tuning:

- n_estimators: "Number of trees in the forest. Values to try include 50, 100, and 200."
- max_depth: "Maximum depth of each tree. Either unlimited (None) or constrained (example- 10 or 20)".
- min_samples_split: "The minimum number of samples required to split an internal node, tested with values 2, 5, and 10."
- min_samples_leaf: "The minimum number of samples required to be at a leaf node, tested with values 1, 2, and 4. These parameters help control the complexity and performance of the Random Forest model."

3.4.3 Ridge Regression : Ridge regression is a type of linear model that uses something called L2 regularization. This means it adds a penalty to the model if the coefficients (the numbers that multiply each feature) are too large. The penalty is based on the squared values of the coefficients. By doing this, Ridge regression helps keep the model simpler and prevents it from overfitting—especially when some features are highly related to each other (a situation known as multicollinearity).Only the regularization strength alpha is tuned, with values of 0.1, 1.0, 10.0, and 100.0. This parameter controls how much regularization is applied, helping to prevent overfitting by shrinking coefficients.

3.4.4 Lasso Regression: Lasso (Least Absolute Shrinkage and Selection Operator) regression is a linear model that uses L1 regularization, adding the absolute values of the coefficients to the loss function. This helps prevent overfitting and automatically selects important features by reducing some coefficients to zero. As a result, it creates a simpler and more interpretable model. Similar to Ridge, the regularization parameter alpha is tuned over the same set of values. The goal is to find the best level of regularization, which can shrink some coefficients to zero for feature selection.

3.4.5 Histogram-Based Gradient Boosting Regressor: Hist Gradient Boosting Regressor is a gradient boosting method that builds decision trees using histogram-based techniques. It speeds up the process by grouping continuous data into bins instead of handling each value separately. This makes it faster and more memory-efficient, especially for large datasets.

Several hyperparameters are specified for tuning:

- max_iter: The number of boosting iterations,with values 100 and 200 to test how many trees to grow.
- learning_rate: Controls the step size for updating the model at each iteration, with values 0.01, 0.1, and 0.2. Smaller values slow down learning but can improve accuracy.
- max_depth: The maximum depth of each tree, with options to set it to unlimited (None) or restrict it to 10 or 20.

3.5 Model Improvement

Key Functions which are used for model improvement:

GridSearchCV: Searches through the hyper parameter grid and performs cross-validation to find the best combination of parameters.

R^2 score: Measures the goodness-of-fit for regression models, indicating how well the predictions match the actual values.

4. RESULTS:

Below are the machine learning models and their respective scores:

MODEL NAME	SCORE
Linear Regression Model	0.37443547906725
Random Forest Model	0.74449108852306
Ridge	0.29022717286054
Lasso	0.32389578038443
HistGradientBoostingRegressor	0.65292008708802

Below table depicts best model and its accuracy
result score:

Overall Best Model Name	RandomForestRegressor (max_depth=20)
R^2 score	0.9549
MSE	8.63
MAE	1.74

The dataset used in this study comprised 500 samples with 36 features related to students, including variables such as gender, age, subjects level of understanding, learning approaches and technical skills. After evaluating several machine learning models, the Random Forest Regressor emerged as the most effective. It achieved an R^2 score of 0.9549, indicating a prediction accuracy of 95.49%. Moreover, the model's Mean Squared Error (MSE) was 8.63, and the Mean Absolute Error (MAE) was 1.74, further demonstrating its strong predictive performance.

5. CONCLUSION

This research successfully demonstrated the effectiveness of machine learning techniques in predicting IT technical skills based on a diverse set of student-related features such as subjects level of understanding, learning approaches and technical proficiency. By employing various models, the Random Forest Regressor stood out with an R^2 score of of 0.9549, indicating a prediction accuracy of 95.49%. Additionally, the low values for Mean Squared Error (8.63) and Mean Absolute Error (1.74) reaffirm the model's robustness and reliability.

This study provides valuable insights for academic institutions and career counsellors, enabling them to better guide students toward careers that match their skills and interests. It also supports students in making informed choices about academic projects and areas of focus, boosting their employability in specialized IT fields like Data Analysis and Visualization, Programming, and Designing (System Analysis and UI). The high accuracy of the prediction model highlights the effectiveness of data-driven methods in enhancing career preparedness and bridging the gap between education and industry needs.

6. REFERENCES

[1] H. A. Dossari, Z. A. Qahtani, F. A. Nughaymish, M. Alkahlifah and A. Alqahtani, "A Machine Learning Approach to Career Path Choice for Information Technology Graduates," Engineering, Technology & Applied Science Research Vol. 10, No. 6, 2020.

[2] O. E, U. O.M and m. F.I., "STUDENT DROPOUT PREDICTION USING MACHINE LEARNING," FUDMA Journal of Sciences (FJS), December 2023.

[3] O. Ojajuni, "Predicting Student Academic Performance Using Machine Learning," Springer Nature Switzerland AG 2021, 2021.

[4] S. ELIA, "Predicting Student Performance in Programming Courses Using Test Unit Snapshot Data," Stockholm, Sweden, 2024.

[5] A. S. Hashim, "Student Performance Prediction Model based on Supervised Machine Leaning Algorithms," in 2nd International Scientific Conference of Al-Ayen University (ISCAU), 2020.

[6] L. S. Hugo, "Predicting Employment Through Machine Learning.," https://www.naceweb.org/career-development/trends-and-predictions/predicting-employment-through-machine-learning/), accessed Oct. 18, 2022.

[7] S. W. S. I. Y. a. C. P. K. Z. Othman, "Classification Techniques for Predicting Graduate Employability," Int. J. Adv. Sci. Eng. Inf. Technol., vol. 8, no. 4–2, p. 1712, , sep 2018.

[8] Y. H. E. Y. Gehad ElSharkawy, "Employability Prediction of Information Technology Graduates using Machine Learning Algorithms," (IJACSA) International Journal of Advanced Computer Science and Applications, 2022.

[9] P. C. #. E. A. #. U. H. F. #. S. R. #. Khalil Ahammad #1, "A Comparative Study of Different Machine Learning Techniques to Predict the Result of an Individual Student using Previous Performances," International Journal of Computer Science and Information Security (IJCSIS),Vol. 19, No. 1, Jan 2021.

[10] S. L. D. M. M. S. S. Muhammad Imran (*), "Student Academic Performance Prediction using Supervised Learning Techniques," International journal of Emerging Technologies in learning, July 2019.

[11] *. E. W. b. B. B. c. R. M. B. Qiang Hao a, "What are the most important prediction of computer science students' online help-seeking behaviors," Computers in Human Behavior, 2016.

[12] A. C. A. Madhuri T. Sathe, "Comparative Study of Supervised Algorithms for Prediction of Students' Performance," I.J. Modern Education and Computer Science, feb 2021.

[13] H. S. a. S. S. G. b. M. R. F. S. a. Ahmad A. Kardan a, "Prediction of student course selection in online higher education institutes using neural network," Computers & Education ELSEVIER, Jan 2013.

[14] Z. Z. a. C. L. b. J. Z. a. Q. Z. a. Lina Gao a, "Deep cognitive diagnosis model for predicting students' performance," Future Generation Computer Systems ELSEVIER, Aug 2021.

[15] H. A. F. Y. Ahajjam Tarika*, "Artificial Intelligence and Machine Learning to Predict Student Performance during the COVID-19," The 3rd International conference on big data and Business intelligence, Poland, March 2021.

[16] M. T. Gouri Palsokar, "A Study on the Perceived Gap in Between Industry and Academia with Reference to the Curriculum of Post-Graduate Courses In Pharmaceutical Sciences in india," www.researchgate.net/publication, Jan 2018.

[17] G. ElSharkawy, Y. Helmy and E. Yehia, "Employability Prediction of Information Technology Graduates using Machine Learning Algorithms," (IJACSA) International Journal of Advanced Computer Science and Applications, 2022.

[18] K. V. V. Vijayalakshmi, "Comparison of Predicting Student's Performance using Machine Learning Algorithms," International Journal of Intelligent Systems and Applications, 2019.

[19] M. Imran, "Student Academic Performance Prediction using Supervised Learning Techniques," International journal of Emerging Technologies in learning, 2019.

[20] M. T. Sathe, "Comparative Study of Supervised Algorithms for Prediction of Students Performance Applications," I.J. Modern Education and Computer Science, 2021.

[21] M. Berke, "Predicting Academic Achievement with Machine Learning Algorithms," Journal of Educational Technology, 2020.

[22] B. Albreiki, "A Systematic Literature Review of Student' Performance Prediction Using Machine Learning Techniques," Education sciences ,MPDI, , 2020.

[23] M. D, "Performance Analysis of the Impact of Technical Skills on Employability," International journal of Perform ability Engineering, 2021.