COGNITIVE SCIENCES AND HUMAN DEVELOPMENT

Predicting Post-Internship Employability Using Ensemble Machine Learning Approach

Aiza Azlin binti Kahlik & Abdulrazak Yahya Saleh*
Faculty of Cognitive Science and Human Development, Universiti Malaysia Sarawak, Sarawak, Malaysia.

ABSTRACT

Graduate employability is crucial for both students and higher education institutions. While academic performance has traditionally been a key predictor of employability, its predictive power is limited, necessitating the exploration of additional factors influencing post-internship job placement. This study investigates the impact of internship-related variables on graduate employability, such as duration, training performance, and prior work experience. Employing a machine learning approach on a dataset comprising student records from Universiti Malaysia Sarawak spanning from 2019 to 2021, we compared the performance of various algorithms, including ensemble methods. Feature selection and repeated K-fold cross-validation optimised model performance. Results indicate that stacking outperforms traditional models, achieving an accuracy of 91%. Particularly, internship duration and training performance emerged as significant predictors of employability. These findings underscore the importance of robust internship programs in enhancing graduate outcomes. Future research could explore the competencies developed during internships and their correlation with job success.

Keywords: graduate employability, machine learning, internship, career readiness, employability prediction, ensemble methods

ARTICLE INFO
Email address: ysahabdulrazak@unimas.my (Abdulrazak Yahya Saleh)
*Corresponding author
https://doi.org/10.33736/jcshd.7518.2024
e-ISSN: 2550-1623

Manuscript received: 1 August 2024; Accepted: 5 September 2024; Date of publication: 30 September 2024

1 INTRODUCTION

The transition from academia to the professional world is a complex journey fraught with challenges, particularly in securing employment (Tamrat, 2023; Webb et al., 2022). Despite increasing higher education attainment rates, youth unemployment remains a persistent global issue (Nisha & Rajasekaran, 2018). This phenomenon underutilises human capital and erodes economic growth and societal stability (Herbert et al., 2020). As a result, higher education institutions (HEIs) have increasingly emphasised employability as a core outcome, seeking to equip graduates with the necessary skills and competencies to thrive in the job market (Hassock & Hill, 2022). Given the perceived importance of internships, it is no surprise that the number of bachelor's degree holders who have undergone an internship during their studies has been increasing in proportion and absolute numbers, as shown in Figure 1. In 2010, 51,293 bachelors, or 69.0%, had gone for an internship. By 2019, that number had doubled to 106,502 graduates, making up a proportion of 88.4% of all bachelors who have undergone an internship.

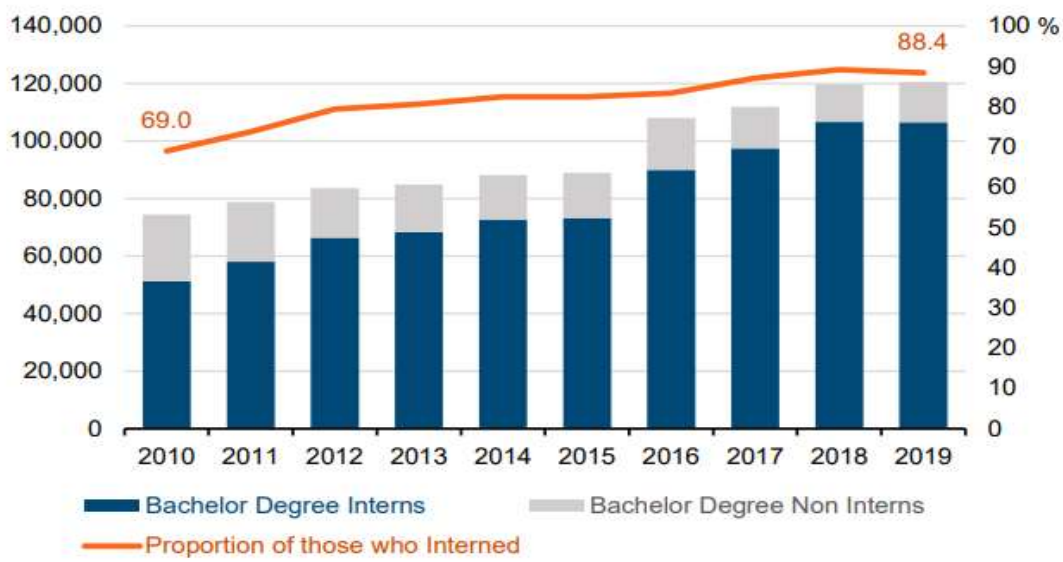

Figure 1. Number and proportion of bachelors by internship status, 2010-2019 (Ministry of Higher Education, Malaysia (2021)).

Internships have emerged as a pivotal strategy to bridge the gap between theoretical knowledge and practical experience (Baker & Fitzpatrick, 2022; Rogers et al., 2021). Beyond theoretical knowledge acquisition, internships provide valuable exposure to real-world work environments (Kim et al., 2022; Oberman et al., 2021). Students gain first-hand experience navigating workplace dynamics, professional etiquette, and the application of academic knowledge in practical settings. Additionally, internships can enhance students' self-confidence and serve as a stepping stone for entering the job market upon graduation (Perusso & Baaken, 2020). This study investigates the impact of internships on student employability by analysing student performance data and internship characteristics. By providing opportunities to apply classroom learning in real-world settings, internships can enhance graduates' employability (Margaryan et al., 2022). However, the effectiveness of internships in predicting post-graduation employment outcomes remains a subject of ongoing research and debate. While studies have shown that internships can positively impact career development (Grillo, 2023; Del Rio Rajanti, 2024), a comprehensive understanding of how specific internship factors influence job placement still needs to be improved.

Previous research has explored the relationship between internships and employability, often employing statistical or machine-learning techniques (e.g., Haque et al., 2024; Vo et al., 2023). While these studies have provided valuable insights, they have limitations. For instance, some studies have relied on limited datasets (Saidani et al., 2022), while others have focused on specific populations (Casuat et al., 2020). Additionally, the predictive power of these models needs to be more consistent (ElSharkawy et al., 2022). To address these gaps, this study uses a comprehensive machine-learning approach to investigate the impact of internship-related variables on post-internship job placement. By employing a diverse set of algorithms and a larger dataset, we aim to enhance the prediction of graduate employability.

Specifically, this research will investigate how to incorporate specific attributes and employee parameters to enhance the predictive model's accuracy and relevance for assessing post-internship students' employability. Additionally, we aim to determine how machine learning techniques can be effectively utilised to develop a robust predictive model for assessing students' employability through analysing and interpreting their internship performance. Finally, this study will evaluate the effectiveness and accuracy of the newly developed employability prediction model in predicting post-internship employment outcomes.

The findings of this study have the potential to inform the design and implementation of more effective internship programs, enabling HEIs to better prepare students for the workforce. Additionally, the developed predictive model can provide valuable insights for students to make informed decisions about their internship experiences and career paths.

2 METHODS

This section provides a systematic overview of employability prediction model architecture, structured around three key dimensions: machine learning models, data sources, and performance evaluation metrics. Traditionally, statistical sampling and surveys have been used to predict employability rates. However, the surge in machine learning (ML) has led to a significant shift in research methodologies. Most recent studies (approximately 95%) have adopted supervised ML algorithms due to their superior predictive performance, with limited exploration of unsupervised methods.

Most studies have used institutional data, primarily from registration units or relevant departments. To evaluate model performance, various metrics have been employed, including precision, recall, F1-score, and accuracy. However, a notable proportion of studies have omitted performance evaluation altogether. Building upon this foundation, the current study employs a comprehensive approach, as depicted in Figure 2, by applying a diverse range of ML algorithms and ensemble methods to predict student employability. Leveraging institutional datasets, we explore the predictive power of KNN, SVM, XGBoost, AdaBoost, CatBoost, LightGBM, Neural Network, and ensemble methods. By identifying critical predictive features, we aim to enhance the accuracy and reliability of employability predictions.

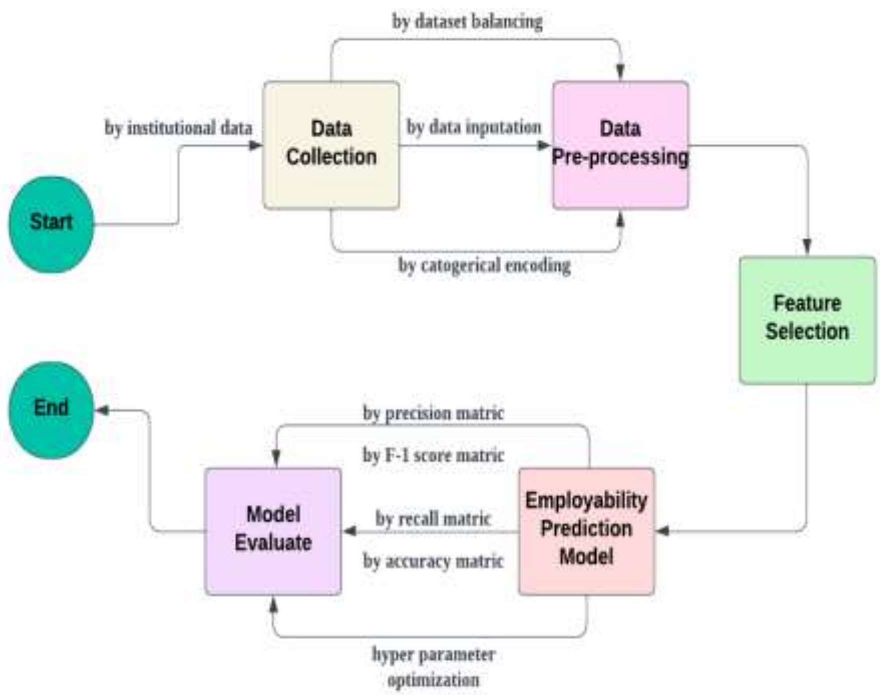

Figure 2. Employability prediction model architecture.

2.1 Data Collection

This study employed a dataset comprising student records from Universiti Malaysia Sarawak (UNIMAS) spanning from 2019 to 2021. The focus was on students who had completed work experience and were in their later stages of study or early careers. The data was collected from academic records, internship databases, and the UNIMAS institutional repository, encompassing information on 1066 students. The dataset encompassed a variety of attributes relevant to employability prediction.

Each student record included a unique identifier, gender (male or female), educational status (undergraduate or postgraduate), field of study, and specialisation. Academic performance was captured by student performance (grades or GPA). Professional experience, internship duration, and training performance were also recorded. Additionally, the dataset included assessments of appearance, public speaking skills, physical condition, alertness, confidence, presentation abilities, and communication skills. The target variable was included in model development, indicating whether a student received a job offer post-internship.

2.2 Data Pre-processing

The dataset comprised 17 complete and distinct attributes. Numerical variables included Work Experience (Work_Exp), Internship Length (Intern_Len), Training Score (Score_Training), Appearance, Speaking Ability (Speaking), Physical Condition (Physc_Cond), Alertness, Confidence, Presentation Ideas (Present_Ideas), Communication Skills (Comm), and Student Performance (Stud_Perf), scaled from 1 to 5. Categorical variables encompassed Educational Status (Edu_Status), Gender, Field of Study (Field), Department (Dept), and Specialisation (Specialised). To facilitate machine learning analysis, categorical variables were transformed using the one-hot encoding (Edu_Status, Gender) and ordinal encoding (Field, Dept, Specialised). This pre-processing ensured data compatibility and optimised model performance.

Table 1. Features for predictive modelling.

Type	Features
Numerical Features	Work_Exp, Intern_Len, Score_Training, Appearance, Speaking, Physc_Cond, Alertness, Confidence, Present_Ideas, Comm, Stud_Perf
Categorical Features	Edu_Status, Gender, Field, Dept, Specialized

2.3 Feature Selection

Feature selection is a critical step in machine learning model development. It aims to identify the most relevant features that contribute to predicting the target variable. Reducing dimensionality and eliminating irrelevant information enhances model performance, interpretability, and computational efficiency.

In this study, three complementary feature selection methods were employed. First, Univariate Selection assessed the individual correlation of each feature with the target variable, 'Offer_Recv'. Features with the strongest statistical relationships were selected, including 'Intern_Len', 'Score_Training', 'Work_Exp', 'Stud_Perf', and 'Appearance'. Second, Recursive Feature Elimination (RFE) iteratively removed features with the least contribution to model performance based on a specified evaluation metric, identifying 'Intern_Len', 'Score_Training', 'Appearance', 'Physc_Cond', and 'Stud_Perf' as crucial predictors. Finally, Tree-Based Feature Importance evaluated features based on their contribution to decision splits within tree-based models, highlighting 'Score_Training', 'Appearance', 'Alertness', 'Comm', and 'Stud_Perf' as highly important in predicting 'Offer_Recv'.

The results from these methods were combined to determine the final set of features, prioritising features that consistently appeared across multiple techniques. This approach ensured that the subsequent modelling process included only the most relevant and informative attributes. This method simplifies the selection process and ensures that only the most relevant features, in terms of their direct impact on predicting job offer outcomes ('Offer_Recv'), are retained. By focusing on straightforward relationships between individual features and the target variable, Univariate Selection enhances interpretability while maintaining computational efficiency, aligning well with the study's objectives and dataset characteristics.

2.4 Machine Learning Training Models

To enhance graduate employability prediction, this study investigates the impact of internship-related variables on post-internship job placement. Employing a machine learning approach on a UNIMAS dataset (2019-2021), we compared the ensemble methods with traditional algorithms (K-Nearest Neighbors (KNN), Support Vector Machines (SVM) with linear, RBF, and polynomial kernels, LightGBM, CatBoost, AdaBoost, XGBoost, Neural Networks). Feature selection optimised model performance. The primary objective was to accurately predict employment status (Emp_Status), a multi-class classification problem. Ensemble methods were employed to improve predictive accuracy by combining multiple models. Bagging reduces variance by creating multiple models through bootstrap sampling, leading to more stable and reliable predictions. Boosting, such as XGBoost, LightGBM, and CatBoost, sequentially builds models, focusing on correcting errors from previous iterations and enhancing the model's ability to capture complex patterns in the data. XGBoost excels in scalability and optimisation, while LightGBM prioritises speed and efficiency. CatBoost specialises in handling categorical features, making it suitable for datasets with mixed data types. Boosting iteratively trains base models, assigning higher weights to misclassified instances. This study's AdaBoost.M1 algorithm focuses subsequent models on difficult-to-predict cases, improving overall accuracy. Random forests construct multiple decision trees by randomly selecting features at each split. This approach reduces overfitting and enhances generalisation. The ensemble's prediction is based on the majority vote of individual trees. Stacking combines predictions from multiple base models as input to a meta-model, improving overall predictive performance by leveraging the strengths of different algorithms. These ensemble methods were selected for their ability to improve predictive accuracy and handle complex classification tasks, making them suitable for studying student employability prediction. Traditional algorithms served as a baseline for comparison. K-Nearest Neighbors classifies data points based on similarity to their nearest neighbours, which can effectively capture local patterns in the data. Support Vector Machines find optimal hyperplanes to separate data into classes, providing a clear decision boundary. XGBoost is a gradient-boosting algorithm known for its efficiency and performance and can handle complex datasets. Feature selection was employed to identify the most predictive attributes, enhancing model performance and reducing dimensionality.

2.5 Model Evaluation

To assess the performance of each boosting model, Repeated K-fold cross-validation was employed. In this method, the dataset is divided into $k = 10$ folds, and the process is repeated multiple times to ensure robust evaluation.

Accuracy, precision, specificity, and F1-score are fundamental metrics employed to evaluate the performance of classification models, including those predicting employability. These metrics provide a quantitative assessment of the model's ability to classify instances into their respective categories correctly. As shown in Figure 3, specific terminology is used to categorise predictions to evaluate the accuracy of a classification model. A true positive (TP) occurs when the model correctly predicts a positive outcome. Conversely, a true negative (TN) signifies a correct prediction of a negative outcome.

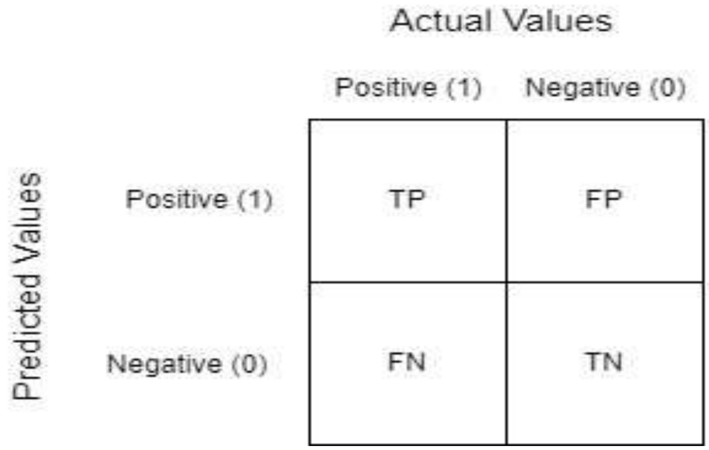

Figure 3. The confusion matrix.

In contrast, false positives (FP) represent instances where the model incorrectly predicts a positive outcome when the actual outcome is negative. On the other hand, false negatives (FN) occur when the model incorrectly predicts a negative outcome when the true outcome is positive. These terms are crucial for calculating performance metrics and understanding the model's strengths and weaknesses. Equations 1-5 show the calculations for all.

$$\text{Accuracy} = (TP+TN)/(TP+FP+TN+FN) \tag{1}$$

$$\text{Precision} = (TP)/(TP+FP) \tag{2}$$

$$\text{Recall} = TP / (TP + FN) \tag{3}$$

$$\text{Specificity} = (TN)/ (TN)(FP) \tag{4}$$

$$\text{F1 score} = 2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall}) \tag{5}$$

3 RESULTS AND DISCUSSION

Before implementing the machine learning models, the dataset is split into (80/20, 70/30, and 60/40) for the training and testing sets. The model performance analysis across the three data splits (80/20, 70/30, and 60/40) yielded several noteworthy observations. These observations provide insights into how each model behaves with varying training and testing data proportions, highlighting the robustness and generalisation capabilities of the algorithms used. In these experiments, early stopping is used. Table 2 presents the experimental results for each experimental approach on the student employability dataset.

Below are the detailed observations based on the accuracy results obtained from each data split. Firstly, the K-Nearest Neighbors (KNN) model demonstrated a noticeable variance in performance across different splits, with the highest accuracy observed in the 80/20 split (0.72). The performance drops slightly as the training data decreases, indicating that KNN benefits from having more training data. Moreover, the Support Vector Machine (SVM) models exhibited stable performance across the splits. The linear kernel SVM achieved consistent results, with the highest accuracy in the 80/20 split (0.73). The RBF and polynomial kernels also performed similarly across the 70/30 and 60/40 splits, indicating their robustness regardless of the data split proportions. The XGBoost model, on the other hand, showed a significant drop in accuracy with larger test sets, particularly in the 70/30 (0.62) and 60/40 (0.65) splits. This suggests that XGBoost may have been overfitting the training data and did not generalise as well with less training data, highlighting the importance of tuning and validation.

The AdaBoost model also demonstrated strong and consistent performance, particularly excelling in the 80/20 split (0.75). This indicates that AdaBoost effectively leverages weak learners to boost performance, maintaining high accuracy across different data splits. LightGBM's performance was highest in the 80/20 split (0.72) and the 60/40 split (0.72) but dropped in the 70/30 split (0.62). This variability suggests sensitivity to the amount of training data, highlighting the importance of parameter tuning for optimal performance. CatBoost maintained stable performance across all splits, with accuracies ranging from 0.70 to 0.74. This consistency indicates that CatBoost handles different data sizes without significant performance degradation, proving its robustness and reliability. The Neural Network model exhibited high accuracy across all splits, with the highest being in the 80/20 split (0.86). This indicates robust generalisation capabilities and practical learning from the data, making it a strong performer irrespective of the data split. Furthermore, stacking achieved the highest accuracy among all models, particularly in the 80/20 split (0.91). This model's superior performance highlights the effectiveness of combining multiple base learners to capture diverse patterns in the data, resulting in enhanced predictive performance. Lastly, bagging was consistently performed across all splits, with accuracies ranging from 0.84 to 0.85. This stability indicates that bagging effectively reduces variance and improves model robustness, making it a reliable choice for varying data sizes. The boosting model also demonstrated high accuracy, with the best performance in the 80/20 split (0.87). The consistent results across other splits indicate its effectiveness in sequentially building strong learners from weak ones, reinforcing its value in ensemble learning approaches. The results of this comprehensive evaluation indicate that different models exhibit varying sensitivity levels to the amount of training data. Stacking emerged as the top-performing model, demonstrating the highest accuracy and robustness across different data splits. Bagging and boosting methods also showed strong and consistent performance, underscoring their effectiveness in enhancing model stability and generalizability. These findings provide valuable insights into the selection and application of machine learning models for different data scenarios, ensuring reliable and accurate predictions.

Table 2. The result from different splits.

Model	**80/20**	**70/30**	**60/40**
	Accuracy	**Accuracy**	**Accuracy**
KNN	0.72	0.66	0.67
SVM Linear	0.71	0.69	0.71
SVM RBF	0.71	0.69	0.71
SVM Polynomial	0.71	0.69	0.71
XGBoost	0.70	0.62	0.65
AdaBoost	0.75	0.70	0.73
LightGBM	0.72	0.62	0.72
CatBoost	0.74	0.70	0.71
Neural Network	0.86	0.84	0.85
Bagging	0.85	0.84	0.85
Boosting	0.87	0.86	0.84
Stacking	**0.91**	**0.87**	**0.87**

3.1 Confusion Matrix

In this study, the evaluation method chosen is a binary classification confusion matrix. Figure 4 shows an example of the confusion matrix of ensemble methods and Neural network algorithms.

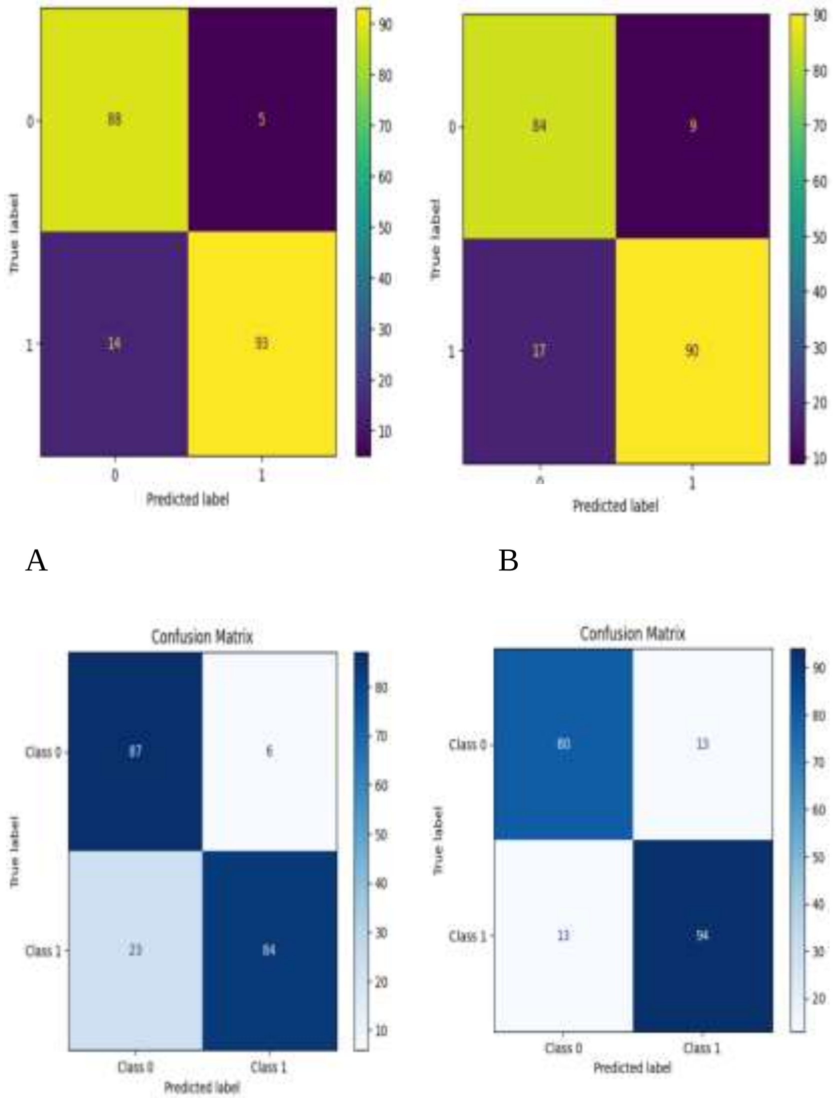

Figure 4. Confusion Matric (A: stacking, B: for boosting, C; Bagging and D; NN).

Overall, the ensemble methods outperformed the other models regarding precision, recall, and F1 weighted average scores, as shown in Table 3. This indicates that these models are better at accurately identifying positive and negative cases, leading to a more balanced performance. Neural Networks also demonstrate strong performance across all metrics. This suggests its ability to capture complex patterns in the data effectively. Traditional machine learning models like KNN, SVM, and XGBoost perform poorly. While they achieve reasonable recall, their precision is notably lower, indicating a higher rate of false positives. AdaBoost, LightGBM, and CatBoost show moderate performance. They exhibit a balance between precision and recall, but the ensemble and neural network models surpass their overall performance.

Table 3. Results of confusion matrix analysis (Weighted Avg Scores).

Model	Precision	Recall	F1-score
KNN	0.67	0.72	0.67
SVM Linear	0.51	0.71	0.59
SVM RBF	0.51	0.71	0.59
SVM Polynomial	0.51	0.71	0.59
XGBoost	0.58	0.62	0.59
AdaBoost	0.73	0.74	0.68
LightGBM	0.61	0.64	0.62
CatBoost	0.69	0.72	0.65
Neural Network	0.87	0.87	0.87
Bagging	0.87	0.85	0.85
Boosting	0.87	0.87	0.87
Stacking	**0.91**	**0.91**	**0.91**

A comparative analysis synthesised insights from current and past studies, delineating the strengths and limitations of various ML models in employability prediction. Key findings underscored the efficacy of SVM and ensemble methods in achieving high prediction accuracies while acknowledging challenges such as overfitting and dataset heterogeneity. Practical insights derived from the study suggested optimising feature selection techniques and exploring hybrid model architectures to further enhance prediction accuracy and applicability across diverse employment contexts.

Table 4. Comparative performance of employability prediction models.

Classification Technique	Present Study	OPT-BAG Model (2023)	Employability Prediction of IT Graduates (2022)	Predicting Student Employability through Internship Context (2022)	Predicting Students Employability using SVM (2020)
KNN	0.72	-	-	0.58	0.79
XGB	0.7	0.82	-	0.74	-
AdaBoost	0.75	0.65	-	0.56	-
LightGBM	0.72	0.83	-	0.77	-
CatBoost	0.74	-	-	0.75	-
Neural Network	0.86	0.9	1	-	-
DT	-	-	0.92	0.66	0.56
NB	-	-	0.98	0.44	0.61
LR	-	-	0.97	0.61	0.65
RF	-	0.9	-	-	0.64
LDA	-	-	-	0.6	-
ANN	-	-	-	-	-
OPT	-	-	-	-	-
-BAG	-	-	-	-	-
SVM	0.71	-	0.98	0.47	0.91
Stacking	0.91	-	-	-	-
Bagging	0.85	-	-	-	-
Boosting	0.87	0.91	-	-	-

This study comprehensively evaluated machine learning models to forecast employability outcomes across varying data splits—80/20, 70/30, and 60/40. K-Nearest Neighbors (KNN) demonstrated consistent performance with an accuracy of 0.72 in the 80/20 split, highlighting its reliability in capturing underlying patterns in employability data, particularly on larger datasets. Support Vector Machines (SVMs), employing Linear, RBF, and Polynomial kernels, exhibited stable performance with accuracies around 0.71 across all splits, showcasing their resilience in handling diverse data complexities. However, XGBoost, initially promising with an accuracy of 0.70 in the 80/20 split, showed susceptibility to overfitting as accuracy declined with larger test sets, suggesting a need for regularisation techniques. In contrast, AdaBoost consistently achieved high accuracies of 0.75 across all splits, effectively leveraging ensemble learning to enhance predictive outcomes.

Moreover, advanced models such as LightGBM (0.72), CatBoost (0.74), and neural networks (0.86) demonstrated robust stability and generalisation capabilities across different dataset sizes. Ensemble methods like Stacking (0.91), Bagging (0.85), and Boosting (0.87) outperformed individual models, emphasising their effectiveness in complex employability prediction tasks. As shown in Table 4, previous studies by Haque et al. (2024) highlighted the superior performance of Artificial Neural Networks (ANN) and Support Vector Machines (SVM), achieving accuracies of 80% and 79%, respectively, underscoring their capability to capture complex employability patterns. Vo et al. (2023) explored the OPT-BAG model, achieving a leading accuracy of 0.91, surpassing traditional models like Random Forests and Decision Trees. ElSharkawy et al. (2022) demonstrated near-perfect accuracies for Decision Trees and SVMs (1.0 and 0.98, respectively), particularly in predicting employability outcomes for Information Technology graduates. Additionally, studies by Saidani et al. (2022) and Casuat et al. (2020) underscored the effectiveness of Support Vector Machines (SVM) across various prediction contexts, consistently outperforming traditional classifiers like KNN and Logistic Regression (LR). These findings collectively highlight the diverse strengths of machine learning models in employability prediction and emphasise the importance of model selection and ensemble techniques in optimising predictive accuracy.

Given these findings, it is evident that practical implications significantly extend to the application of ensemble methods in employability forecasting. The comparative analysis reveals diverse performance among various machine learning models. K-Nearest Neighbors (KNN), Support Vector Machines (SVMs), and XGBoost show initial promise but are prone to overfitting, particularly with larger datasets. In contrast, AdaBoost consistently demonstrates robust performance by effectively utilising weak learners. Previous research highlights the efficacy of advanced models, including Artificial Neural Networks (ANN), SVMs, and ensemble techniques like Random Forest (RF) and Logistic Regression (LR). Notably, the OPT-BAG model's exceptional accuracy of 0.91 underscores the potential of tailored ensemble methods for optimising employability predictions.

Practical implications emphasise the importance of implementing ensemble methods, such as stacking, bagging, and boosting, to enhance predictive accuracy in employability forecasting. This strategic approach enables organisations and educational institutions to identify high-potential students more effectively, informing career development and recruitment decisions. Understanding each model's strengths and weaknesses is crucial for selecting the most suitable approach for specific predictive tasks. This advancement improves the precision of employability predictions and guides proactive strategies for shaping educational programs and student career paths. Future research should continue exploring innovative methodologies and datasets to refine and validate predictive models, advancing our understanding of machine learning applications in student career outcomes.

ACKNOWLEDGEMENTS

This research received no specific grant from public, commercial, or not-for-profit funding agencies.

REFERENCES

Baker, I., & Fitzpatrick, D. (2022). Student Experiences in Pre-COVID Virtual Internships: Integration, Barriers, Motivation Challenges, Supportive Supervisors, and Intern Growth. American *Journal of Distance Education, 36*(2), 90–102. https://doi.org/10.1080/08923647.2022.2034399

Casuat, C. D., Festijo, E. D., & Alon, A. S. (2020). Predicting students' employability using support vector machine: a smote-optimized machine learning system. *International Journal of Emerging Trends in Engineering Research, 8*(5), 2101-2106. https://doi.org/10.30534/ijeter/2020/102852020

Del Rio Rajanti, S. (2024). Global perspective in career development: The influence of international experiences on post-graduation employment (Bachelor's thesis, Tampere University of Applied Sciences). https://urn.fi/URN:NBN:fi:amk-2024051712981

ElSharkawy, G., Helmy, Y., & Yehia, E. (2022). Employability prediction of information technology graduates using machine learning algorithms. *International Journal of Advanced Computer Science and Applications, 13*(10), 359-367. https://urn.fi/URN:NBN:fi:amk-2024051712981

Grillo, J. L. (2023). Students' Perceptions of the Effects of Internships on Confidence and Post-Graduation Outcomes in *Higher Education (*Doctoral dissertation, St. John's University*)*. https://scholar.stjohns.edu/theses_dissertations/553

Hassock, L., & Hill, C. (2022). Employability and employment: The role of higher education in a rapidly changing world. In *Higher education and job employability* (pp. 155-178). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-05716-8_8

Haque, R., Quek, A., Ting, C. Y., Goh, H. N., & Hasan, M. R. (2024). Classification Techniques Using Machine Learning for Graduate Student Employability Predictions. *International Journal on Advanced Science, Engineering & Information Technology, 14*(1), 45-56. https://doi.org/10.18517/ijaseit.14.1.19549

Herbert, I. P., Rothwell, A. T., Glover, J. L., & Lambert, S. A. (2020). Graduate employability, employment prospects and work-readiness in the changing field of professional work. *The International Journal of Management Education, 18*(2), 100378. https://doi.org/10.1016/j.ijme.2020.100378

Kim, Y.-A., Kim, K.-A., & Tzokas, N. (2022). Entrepreneurial universities and the effect of the types of vocational education and internships on graduates' employability. *Studies in Higher Education, 47*(5), 1000–1009. https://doi.org/10.1080/03075079.2022.2055324

Margaryan, S., Saniter, N., Schumann, M., & Siedler, T. (2022). Do Internships Pay Off? *Journal of Human Resources*, 57 (4), 1242-1275. https://doi.org/10.3368/jhr.57.4.0418-9460R2

Ministry of Higher Education (2021). Graduate Tracer Study Microdata. https://great.mohe.gov.my/penerbitan/LAPORAN%20KAJIAN%20PENGESANAN%20GRADUAN%20TVET%202020.pdf.

Nisha, S. M., & Rajasekaran, V. (2018). Employability skills: A review. *IUP Journal of Soft Skills, 12*(1), 29–37. https://ssrn.com/abstract=3251255

Oberman, W., Hunt, I., Taylor, R. K., & Morrisette, S. (2021). Internships and occupational self-efficacy: Impact and gender differences. *Journal of Education for Business, 96*(7), 424–434. https://doi.org/10.1080/08832323.2020.1848768

Perusso, A., & Baaken, T. (2020). Assessing the authenticity of cases, internships and problem-based learning as managerial learning experiences: Concepts, methods and lessons for practice. *The International Journal of Management Education, 18*(3), 100425. https://doi.org/10.1016/j.ijme.2020.100425

Rogers, S. E., Miller, C. D., Flinchbaugh, C., Giddarie, M., & Barker, B. (2021). All internships are not created equal: Job design, satisfaction, and vocational development in paid and unpaid internships. *Human Resource Management Review, 31*(1), 100723. https://doi.org/10.1016/j.hrmr.2019.100723

Saidani, O., Menzli, L. J., Ksibi, A., Alturki, N., & Alluhaidan, A. S. (2022). Predicting student employability through the internship context using gradient boosting models. IEEE Access, 10, 46472-46489. https://doi.org/10.1109/ACCESS.2022.3170421

Tamrat, W. (2023). The shifting landscape of graduate employment in Ethiopia: changes, challenges and responses. *Policy Reviews in Higher Education, 7*(2), 211-228. https://doi.org/10.1080/23322969.2023.2196552

Vo, M. T., Nguyen, T., & Le, T. (2023). OPT-BAG Model for Predicting Student Employability. *Computers, Materials & Continua*, *76*(2), 1555-1568. https://doi.org/10.32604/cmc.2023.039334

Webb, M. E., & Paretti, M. (2022, October). Exploring Professional identity development research on displaced higher education students. In *2022 IEEE Frontiers in Education Conference (FIE)* (pp. 1-9). https://doi.org/10.1109/FIE56618.2022.9962604