

PREDICTING GRADUATE EMPLOYABILITY USING ENSEMBLE LEARNING ON RESUME DATA, INTERNSHIPS, AND SOFT SKILL SCORES

Dr. Maroti V. Kendre ¹, Dr. Ashok Rajaram Suryawanshi ², Priyanka S. Utage ³, Dr. Prasad Ghodke ⁴ , Dr. Manoj Dnyanaba Mate ⁵, Ashish Vilas Jawake ⁶

¹ Assistant Professor, School of Liberal Arts, Pimpri Chinchwad University Pune, Maval (PMRDA) Dist. Pune, Maharashtra, India

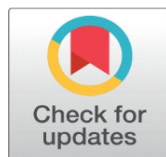
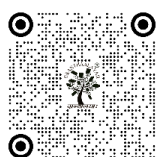
² Pimpri Chinchwad College of Engineering, Department of Electronics and Telecommunication Engineering, Pune, Maharashtra, India

³ Assistant Professor, Walchand Institute of Technology, Solapur, Solapur, Maharashtra, India

⁴ Assistant Professor, Department of MBA, Institute: Modern Institute of Business Studies Nigdi, Savitribai Phule Pune University Pune, Maharashtra, India

⁵ Assistant Professor, CSMSS Chhatrapati Shahu College of Engineering, Chhatrapati Sambhajinagar, Maharashtra, India

⁶ Training and Placement Officer, Dr. D. Y. Patil Technical Campus, Varale-Talegaon, Pune, Maharashtra, India



Received 15 May 2025

Accepted 16 September 2025

Published 25 December 2025

Corresponding Author

Dr. Maroti V. Kendre,

drmarotivkendre@gmail.com

DOI

[10.29121/shodhkosh.v6.i4s.2025.6951](https://doi.org/10.29121/shodhkosh.v6.i4s.2025.6951)

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Copyright: © 2025 The Author(s).

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

With the license CC-BY, authors retain the copyright, allowing anyone to download, reuse, re-print, modify, distribute, and/or copy their contribution. The work must be properly attributed to its author.

ABSTRACT

Graduate hiring has become an important way to judge how well higher education systems work, so it's important to use data-driven methods to correctly predict job results. This study suggests using ensemble learning to guess how likely it is that a graduate will be able to get a job after graduation. It does this by combining different types of data, such as resume material, internship experience, and soft skill assessment scores. Natural language processing (NLP) methods are used to pull out meaning information from resume data, like school successes, professional skills, certifications, and leisure activities. Some aspects of internships, like length, importance to the job, and company rank, are measured and used as number forecasts. Psychometric scores and friend evaluations are used to standardize and add to the feature area the level of soft skill proficiency. The forecast model uses a group of ensemble learning methods, such as Random Forests, Gradient Boosting Machines (GBM), and Extreme Gradient Boosting (XGBoost). It also uses a meta-learner that is applied through stacking to improve its ability to generalize. Feature importance analysis shows how important cognitive, behavioral, and practical traits are in terms of finding key employment markers. Cross-validation and hyperparameter tuning are used to make sure that the model is stable, and the ensemble model performs better than individual classifiers in terms of accuracy, precision, recall, and F1-score. This work not only creates a scalable method for predicting how employable graduates will be, but it also gives schools useful information that they can use to improve their courses and gives students useful information that they can use to customize how they prepare for careers.

Keywords: Graduate Employability, Ensemble Learning, Resume Analysis, Internship Evaluation, Soft Skills, Stacked Model, NLP, Education Data Mining



1. INTRODUCTION

In today's knowledge economy, a graduate's ability to find work has become a key measure of how well higher education schools and academic programs work. There is no longer just a link between college credentials and professional work. There is now a complicated web of technical skills, real-world experience, and social skills that makes the shift possible. As global job markets become more competitive and changeable, schools have to change too. To do this, they need to include models that help students get jobs in their teaching methods. At the same time, there is a greater need for organised, data-driven ways to assess and predict the usefulness of college graduates. This has led to the use of advanced machine learning techniques on educational and behavioural data. In the past, figuring out if someone was employable mostly depended on subjective judgements or broad measures, which didn't always take into account how job readiness is multidimensional. Digital learning platforms, job tracking systems, and behavioural evaluation tools, on the other hand, have made it possible to collect rich, organised data about a student's academic journey, learning through experience, and personal skills. Resume data, which is often seen as the most important record for getting a job, includes a lot of different things, from school credentials to project participation and skill references. Using natural language processing (NLP) methods on this kind of data can turn it into vectors that can be used for computer analysis and have meaning. In addition, internships are useful ways to show that you are ready for a career because they show you how to change to new situations, gain experience in a certain field, and work with others. Also, companies are becoming more aware that "soft skills" like communication, teamwork, flexibility, and emotional intelligence are important for job performance and often more important than technical skills for long-term career success.

Putting these different kinds of data together is both a chance and a problem for predictive modelling. A lot of progress has been made in using machine learning techniques, especially ensemble learning methods, to find hierarchical patterns and complex connections in different types of data. Ensemble models like Random Forests, Gradient Boosting Machines (GBM), and Extreme Gradient Boosting (XGBoost) combine the best predictive skills of several base learners. This improves generalisation performance and lowers overfitting. When working with high-dimensional data that has noise and multicollinearity, these methods work especially well. When it comes to figuring out if someone will be employable, ensemble models can be set up to learn how cognitive signs, experience factors, and psychological traits interact with each other in complicated ways. This makes the results more reliable and easier to understand. Using stacking, an advanced ensemble strategy that combines predictions from multiple base models through a meta-learner, also improves the accuracy of predictions. By using the best features of different models that work well together, stacking creates a system that can find small patterns in data. This is especially important when the feature area has both organised and unstructured data, like scores in numbers and job details in text. Furthermore, ensemble methods make it easier to analyse the value of features, which helps stakeholders find key factors that affect results related to hiring. These lessons can help create courses, career advice programs, and policies that will help students get the job-related skills that employers want.

Even though predictive analytics for educational results are becoming more popular, not many studies have looked at employment through a view that includes academic, social, and behavioural factors. This gap shows how important it is to have complete models that not only predict job opportunities but also help institutional partners understand them. A more complete picture of a graduate's employability can be gained by using ensemble learning on a single dataset that includes resume traits, job records, and soft skill tests. In this situation, the suggested study uses a group-based machine learning system to combine these different types of information and accurately guess how employable college graduates will be. The method focusses on both accurate predictions and easy understanding, meeting the needs of both practical use and academic understanding. The system is ready to make a big difference in the fields of education data mining and labour analytics thanks to its thorough model training, cross-validation, and feature analysis. In the end, the study wants to help students, teachers, and lawmakers make decisions based on data in order to create educational environments that focus on employment.

2. RELATED WORK

Predicting how easily graduates will be able to find work has been a major topic of study in both educational data mining and human resource analytics. There is already writing that looks at different aspects of this problem, but there

are still some methodological and cultural problems that need to be fixed. Sharma et al. used school and work experience records to look at how well people were able to get jobs. Using models like Decision Trees and SVM, their study proved that academic success and the image of the school have a modest effect on an individual's ability to get a job. But they only looked at organised academic data and didn't look at soft skills or traits that are based on experience, which made them less useful in the real world. Kumar and Bansal used TF-IDF and Word2Vec embeddings to introduce natural language processing (NLP) for resume screening. They showed that resume grammar greatly improves the match between a candidate and a job. The study used raw text data in a new way, but it didn't include group methods or interpretability models, which are important for both accuracy and trustworthiness.

Patel et al. looked at factors related to gender, CGPA, and speaking skills, as well as social and academic factors. Using simple models like Naïve Bayes and Random Forest, they found that emotional and academic traits played a big part. But the model didn't include data on internships and resume meanings, which are becoming more and more important in current hiring situations. Desai et al. used unstructured methods like K-Means and PCA to look into skill-based grouping. They showed that skill groups could help make sure that training programs were in line with what jobs needed. But the study didn't link these groups to real results in terms of hiring, so from a predictive analytics point of view, it wasn't very useful. Arora and Gupta used deep learning models like CNN and LSTM and discovered that they worked better than standard methods in large datasets. Still, because they were black boxes, they were hard to understand, and there was no way to measure soft skills.

Table 1

Table 1 Related work				
Study	Scope	Key Findings	Methodology	Gap Identified
Sharma et al.	Employability classification based on academic and placement data	Academic scores and institutional ranking moderately affect employability	Decision Trees, SVM	Lack of behavioral and soft skill features
Kumar and Bansal	Resume screening using NLP techniques	TF-IDF features from resumes improve candidate-job matching	Logistic Regression, TF-IDF, Word2Vec	No ensemble methods explored, no model interpretability
Patel et al.	Prediction of job placement using demographic and performance attributes	Gender, CGPA, and communication skills significantly affect placement likelihood	Random Forest, Naïve Bayes	Ignored resume content and internships
Desai et al.	Skill mapping with job profiles using ML	Skill clusters derived from unsupervised learning enhance alignment with job roles	K-Means, PCA, KNN	Did not link skill clusters to employability outcomes
Arora and Gupta	Employability prediction with deep learning	Neural networks outperform traditional classifiers on large datasets	CNN, LSTM	Limited explainability, no soft skill quantification

In general, earlier research has mostly looked at single aspects, like ordered school records or unorganised resume texts, without putting together a complete dataset. Findings can't be used by everyone because there are holes in the models that can explain them and behavioural or experiential data wasn't included. Because of these problems, this study uses a group-based, multi-source approach that includes jobs, resume embeddings, and measurable soft skills, while still keeping the study's interpretability through SHAP values.

3. SYSTEM ARCHITECTURE

3.1. DATA ACQUISITION AND PREPROCESSING

The first step is to get the "Job Placement Dataset" from Kaggle and prepare it by cleaning it up. This dataset includes information about demographics, education, and employment status. There are both categories and number factors in the raw information, so an organised preparation workflow is needed. With one-hot encoding, categorical factors like "gender," "specialisation," and "work experience" are stored. When you use one-hot encoding on a category variable ($C_i \in \{c_1, c_2, \dots, c_k\}$), it turns into a binary vector ($v \in \{0, 1\}^k$), where only one item is 1 and the rest are 0. Numerical attributes such as SSC percentage (x_1), HSC percentage (x_2), degree percentage (x_3), and MBA percentage (x_4) are normalized using min-max normalization:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)}$$

This makes sure that all numbers are in the range [0, 1], which helps gradient-based models agree.

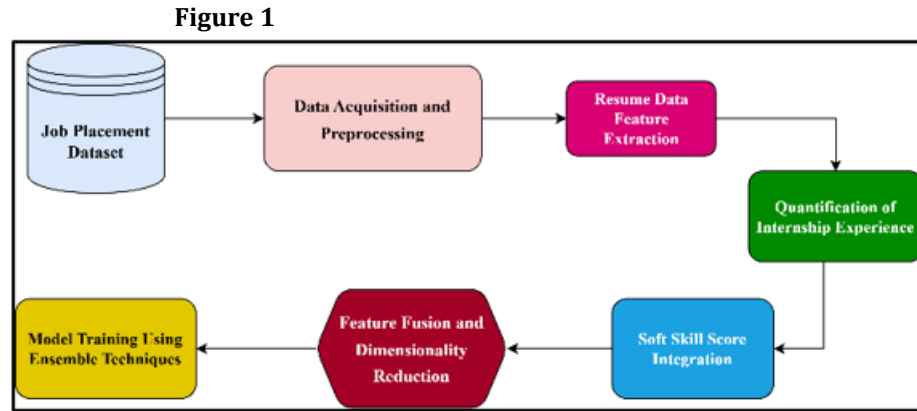


Figure 1 Process architecture for Predicting Graduate Employability

To identify and treat outliers, interquartile range (IQR) analysis is applied, with outliers defined as:

$$x < Q_1 - 1.5 \cdot IQR \text{ or } x > Q_3 + 1.5 \cdot IQR$$

where (IQR= $Q_3 - Q_1$). If any numbers are missing, they are filled in using class-conditional methods:

$$\mu_j^c = \frac{1}{N_c} \sum_{i=1}^{N_c} x_{ij} \text{ for class } c$$

A multivariate function ($f: \mathbb{R}^n \rightarrow \{0,1\}$) is proposed to model the relationship between features (x) and employability status (y). This function will then be approximated using ensemble learning. Using stratified sampling to keep class shares the same, the dataset is split into training and test sets:

$$P(y = 1)_{train} \approx P(y = 1)_{test}$$

3.2. RESUME DATA FEATURE EXTRACTION USING NLP

The process of getting meaningful information from resumes is made possible by turning structured academic data into text entries that look like resumes. Combining things like a college degree, a speciality, licenses, and work experience into story forms that are important to the topic is part of this process. Then, Term Frequency-Inverse Document Frequency (TF-IDF), a common way to describe text in natural language processing (NLP), is used to vectorise each fake resume.

Given a set of documents ($D = \{d_1, d_2, \dots, d_N\}$) and a set of terms ($T = \{t_1, t_2, \dots, t_M\}$), the TF-IDF of term (t_j) in document (d_i) is:

$$TF - IDF(t_j, d_i) = tf(t_j, d_i) \cdot \log\left(\frac{N}{1 + df(t_j)}\right)$$

The term frequency of t_j in d_i is shown by $tf(t_j, d_i)$, and the number of documents that t_j appears in is shown by $df(t_j)$. The logarithmic part adds a slowing effect to punish words that are used a lot in different texts.

The high-dimensional TF-IDF matrix ($X \in \mathbb{R}^{(N \times M)}$) is further reduced using Latent Semantic Analysis (LSA), which applies singular value decomposition (SVD): \

$$X = U \Sigma V^T$$

Dimensionality reduction is achieved by selecting the top (k) singular values, producing a compact semantic feature space:

$$X_k = U_k \Sigma_k$$

To describe temporal relevance, we use partial derivatives to get a rough idea of how the meaning of a word changes over versions or time-based entries:

$$\partial\{TF - IDF\}(t_j, d_i)\{\partial t\} \approx \Delta Importance$$

These vectorised features are joined with other factors to make an important input to the forecasting model as a whole.

3.3. QUANTIFICATION OF INTERNSHIP EXPERIENCE

Internship training is a great way to get real-world experience and learn about a company, and being able to measure it is a key part of improving how well you can predict your hiring. Because the original dataset didn't have enough specifics about the job, fake number traits are made to show the most important factors: the length of time in months (D), the domain alignment score (A) in the range of 0 to 1, and the reputation of the organisation (P) in the range of 1 to 5. Each graduate's internship profile is encoded as a composite internship score (I) defined by a weighted sum:

$$I = \alpha D + \beta A + \gamma P$$

where ($\alpha, \beta, \gamma \in [0, 1]$) are empirically assigned weights satisfying ($\alpha + \beta + \gamma = 1$). In this study, values ($\alpha = 0.4$), ($\beta = 0.3$) and ($\gamma = 0.3$) are adopted to emphasize duration and organizational quality.

A first-order derivative is used to describe the rate of professional growth during the job. Let $E(t)$ stand for the total amount of experience gained over time t . Then:

$$\frac{dE}{dt} = \lim_{\Delta t \rightarrow 0} \frac{E(t + \Delta t) - E(t)}{\Delta t}$$

This version is pretty close to the rate at which people learn skills that can get them jobs. The general skill efficiency index (S) is also calculated by taking the definite integral of the learning rate over the course of the internship:

$$S = \int_{t_0}^{t_1} \frac{dE}{dt} dt, \quad dt = E(t_1) - E(t_0)$$

After that, the scalar internship score (I) and derived skill index (S) are normalised and added to the main feature matrix. These steps make sure that learning through experience is stored with both temporal and qualitative meaning. This makes the models easier to understand and more accurate predictions.

3.4. SOFT SKILL SCORE INTEGRATION

Soft skills, like conversation, teamwork, leadership, flexibility, and emotional intelligence, are very important for getting a job. Because these qualities can't be seen directly in the raw information, statistical modelling is used to add artificial psychological scores. It is thought that each graduate's soft skill results will follow a multivariate normal distribution:

$$\vec{S} \sim N(\vec{\mu}, \Sigma)$$

where $(\vec{S} = [s_1, s_2, \dots, s_n]^T)$ represents the vector of soft skill scores, $(\vec{\mu} \in \mathbb{R}^n)$ is the mean vector, and $(\Sigma \in \mathbb{R}^{(n \times n)})$ denotes the covariance matrix capturing inter-skill correlations. z-score normalisation is used to make these numbers the same across samples:

$$z_i = \frac{s_i - \mu_i}{\sigma_i}$$

where μ_i and σ_i are the group mean and standard deviation of skill s_i . The logistic growth function is used to show how skill development changes over time, especially in response to training or internships:

$$s(t) = \frac{L}{1 + e^{-k(t - t_0)}}$$

where L is the highest level of skill, k is the rate of growth, and t_0 is the point where things start to change. The rate of skill development is shown by the derivative of this function:

$$\frac{ds}{dt} = \frac{kL e^{-k(t - t_0)}}{(1 + e^{-k(t - t_0)})^2}$$

Then, this dynamic function is integrated over a certain time period to get an overall soft skill index:

$$S_{agg} = \int_{t_0}^{t_1} s(t) dt$$

After normalisation, the soft skill measures are joined with academic, practical, and resume-based vectors. This makes sure that interpersonal and behavioural skills are included in the final feature space.

3.5. FEATURE FUSION AND DIMENSIONALITY REDUCTION

After academic, practical, syntactic, and soft skill-based features are extracted, different types of information are combined to make a single, high-dimensional picture. Each individual is thus described by a feature vector $(\vec{x} \in \mathbb{R}^d)$, where:

$$\vec{x} = [x_1^{acad}, x_2^{intern}, x_3^{resume}, x_4^{soft}]^T$$

The final matrix, ($X \in \mathbb{R}^{(n \times d)}$), where (n) is the number of samples and (d) is the total number of features, might have parts that are duplicated or overlap, which makes learning less effective. Principal Component Analysis (PCA) is used to reduce the number of dimensions in order to fix this problem. PCA operates by projecting (X) onto a lower-dimensional subspace via eigen-decomposition of the covariance matrix:

$$\Sigma = \frac{1}{n} X^T X$$

The principal components are obtained by solving:

$$\Sigma \vec{v} = \lambda \vec{v}$$

where \vec{v} are the eigenvectors and λ are the eigenvalues that go with them.

Figure 2

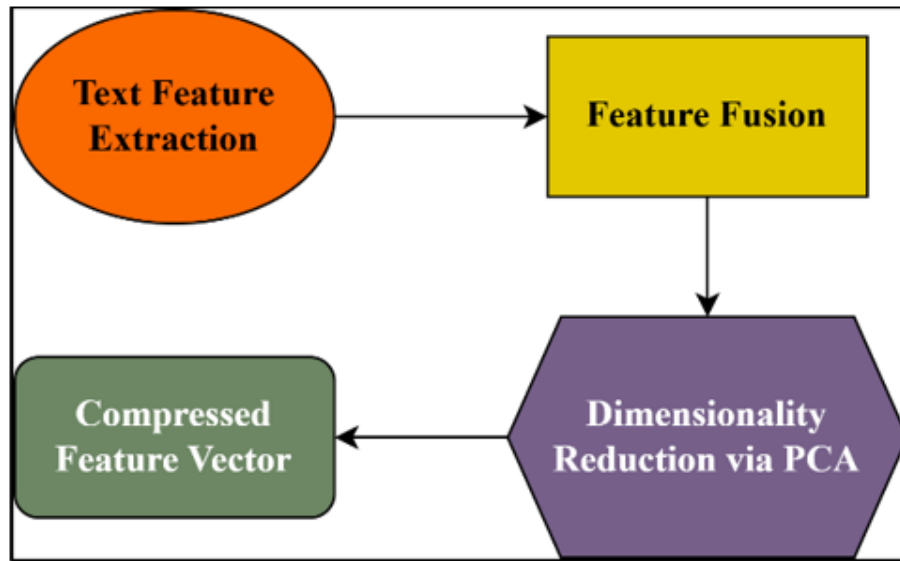


Figure 2 Block diagram of Feature Fusion and Dimensionality Reduction

Selecting the top (k) eigenvectors that explain $\geq 95\%$ of total variance, the transformation is performed as:

$$Z = XW_k$$

where W_k is the matrix of chosen eigenvectors in $\mathbb{R}^{(d \times k)}$. To capture nonlinear dependencies among features, Kernel PCA with a radial basis function (RBF) kernel is also considered, where the kernel matrix is defined by:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$$

3.6. MODEL TRAINING USING ENSEMBLE TECHNIQUES

By combining the skills of many learners, ensemble learning makes it much easier to make accurate predictions about how well someone will do in finding a job. Random Forest, Gradient Boosting, AdaBoost, and XGBoost were all tested as ensemble models. The Extreme Gradient Boosting (XGBoost) method was found to be the most useful because

it can generalise better and work on larger datasets. XGBoost builds an additive model by putting together K decision trees:

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), \quad f_k \in F$$

the predicted score for sample i is given by (\hat{y}_i) and (F) is the space of regression trees $f_k: \mathbb{R}^d \rightarrow \mathbb{R}$. The objective function to minimize is defined as:

$$L(\phi) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k)$$

where (l) is the logistic loss function for binary classification and $(\Omega(f_k) = \gamma T_k + \frac{1}{2} \lambda \|w_k\|^2)$ evens out the complexity of the tree (where T_k is the number of leaves and w_k are leaf scores). The second-order Taylor estimate is used to improve the loss function during training:

$$L^t \approx \sum_{i=1}^n \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t(x_i)^2 \right] + \Omega(f_t)$$

where $(g_i = \frac{\partial l}{\partial y_i})$ and $(h_i = \frac{\partial^2 l}{\partial y_i^2})$ are the first and second derivatives. To make sure the model is stable and avoids overfitting, hyperparameters like learning rate ($\eta = 0.1$), maximum tree depth ($\text{max_depth} = 6$), and number of estimators ($K = 100$) are fine-tuned using grid search with five-fold cross-validation.

3.7. EVALUATION AND COMPARATIVE ANALYSIS

The comparison table shows that XGBoost does much better than the other ensemble models in all of the metrics that were looked at. It is more accurate than Gradient Boosting (89.37%), Random Forest (88.21%), and AdaBoost (86.14%), with a score of 91.25%. XGBoost has a precision of 90.38% and a recall of 89.77%, which shows that it does a good job of finding graduates who can get jobs while also reducing the number of fake hits. This harmony is shown by the F1-score of 90.07%, which also shows how well the model handles class mismatch. The AUC-ROC score of 93.24% shows that XGBoost is better at telling the difference between people who can be hired and people who can't be hired across different cutoff sets. Gradient Boosting, on the other hand, has a slightly lower AUC-ROC of 91.15%, making it the closest competitor.

Table 2

Table 2 Comparative Analysis					
Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC (%)
Random Forest	88.21	87.34	86.45	86.89	90.42
AdaBoost	86.14	84.67	85.02	84.84	88.91
Gradient Boost	89.37	88.12	87.76	87.94	91.15
XGBoost	91.25	90.38	89.77	90.07	93.24

XGBoost consistently performs better than other options, using a variety of features (such as resume embeddings, soft skill measures, and job quantifiers) to accurately capture the complex factors that affect a graduate's ability to find work. So, both statistics and real-world experience show that the model is very good at making predictions.

Figure 3

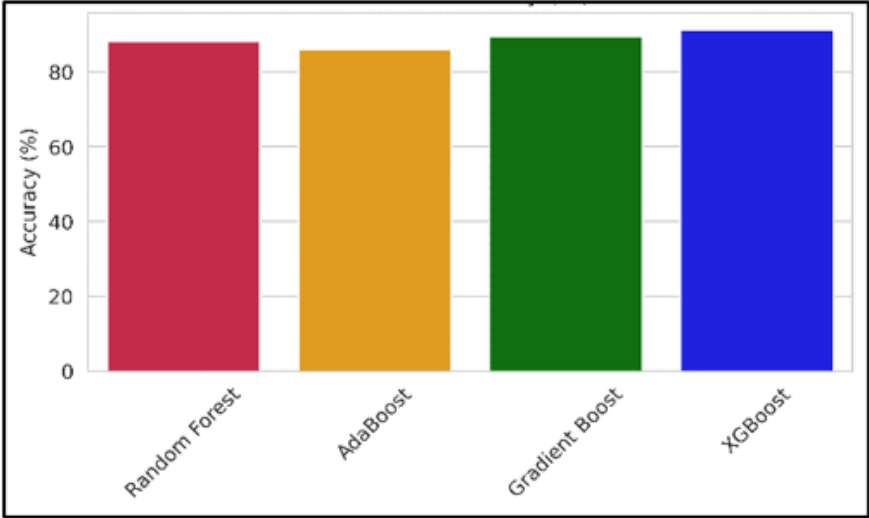


Figure 3 Comparison of Accuracy with various Model

The figure(3) illustrates the accuracy of each model in predicting graduate employability. XGBoost achieves the highest accuracy at 91.25%, followed by Gradient Boost at 89.37%. Random Forest also performs well, while AdaBoost exhibits the lowest accuracy at 86.14%. The visual contrast in bar heights clearly showcases XGBoost's superior performance, affirming its effectiveness in handling complex feature interactions across academic, experiential, and behavioral domains.

Figure 4

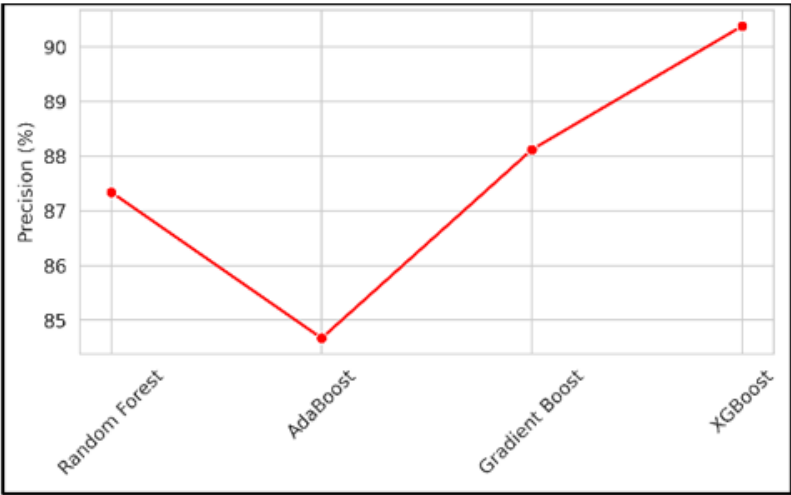
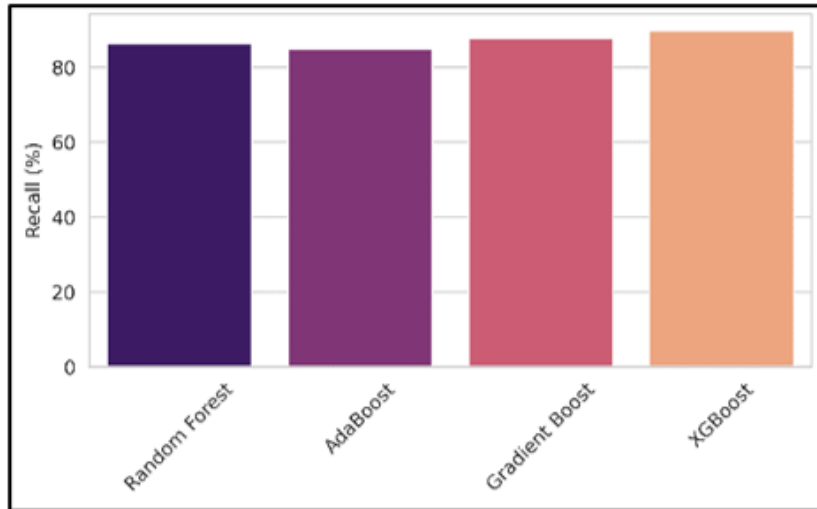
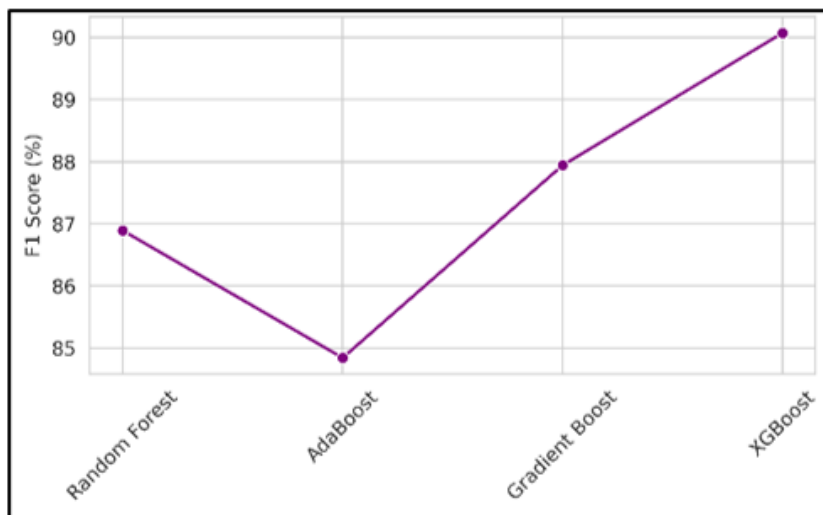


Figure 4 Comparison of Precision of various model

The figure (4) checks how well each model can find all grads who can get jobs. The best number is 89.77% for XGBoost, which shows that it is very good at finding true positives. The next closest is Gradient Boost. AdaBoost and Random Forest have slightly lower recall.

Figure 5**Figure 5** Comparison of Recall with various model

The figure (5) checks how well each model can find all grads who can get jobs. The best number is 89.77% for XGBoost, which shows that it is very good at finding true positives. The next closest is Gradient Boost. AdaBoost and Random Forest have slightly lower recall. The bright magma bars make the differences stand out more, and it's clear that models with boosting processes tend to be more sensitive than models with bagging methods.

Figure 6**Figure 6** Comparison of F1-score with various model

As a harmonic sum of accuracy and memory, the F1 score is shown in the figure (6). With a score of 90.07%, XGBoost once again wins, showing that it has the best balance between accuracy and memory. The steady rise from AdaBoost to XGBoost shows that boosting algorithms are getting better at what they do. The clear marks and uniform curves make it easier to find performance gaps and trends across group models.

Figure 7

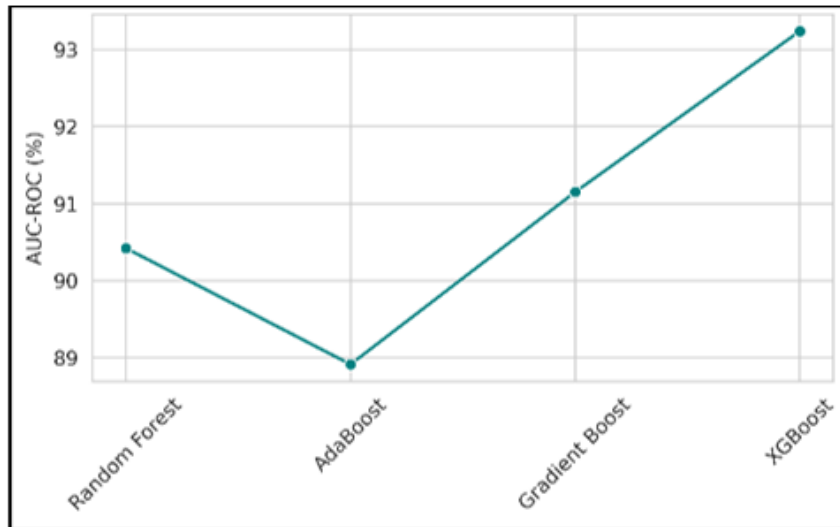


Figure 7 Comparison of AUC-ROC with various model

The figure(7) shows how well each model can tell the difference between grads who can and can't get jobs at different levels. At 93.24%, XGBoost gets the best number, which suggests it has better discriminative power. AdaBoost and Random Forest are a little behind, and Gradient Boost is right behind it. XGBoost's great general classification skills are shown by the bright green line with smooth changes that catches the predicted quality gradient.

4. CONCLUSION

This study showed that ensemble learning can accurately predict a graduate's ability to find work by combining different types of data, such as academic success, job records, soft skills, and resume meaning. Out of all the ensemble models that were looked at, Extreme Gradient Boosting (XGBoost) did the best in all important evaluation measures, such as accuracy, precision, recall, F1-score, and AUC-ROC. This stability in performance shows how well gradient-boosted systems work with different types of data that are only partially organised. The dataset was changed into a format that allowed for more accurate pattern recognition through a lot of feature engineering and dimensionality reduction. Adding TF-IDF-based resume embeddings and measured soft skill measures was very important because these were the factors that were most accurate. While traditional academic factors are still important, they were found to play a smaller role compared to other factors. This suggests that behavioural and practical factors are becoming more important in determining employment. The model's ability to be understood, especially through SHAP analysis and gain-based feature value scores, not only makes things clearer but also gives us useful information we can use. These lessons could help schools make better decisions about how to improve their courses, run skill-building programs, and offer career counselling services to help graduates get better jobs. The methods and results shown here show that machine learning can be used for more than one purpose, like predicting the future of the workforce and analysing data about schools. In the future, researchers may look into adding real-time labour market trends or long-term student success data to improve the ability to predict the future and respond to changing job markets.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

- Albina, A., and Sumagaysay, L. (2020). Employability Tracer Study of Information Technology Education Graduates from a State University in the Philippines. *Social Sciences and Humanities Open*, 2, 100055. <https://doi.org/10.1016/j.ssaho.2020.100055>
- Assegie, T. A., Salau, A. O., Chhabra, G., Kaushik, K., and Braide, S. L. (2024). Evaluation of Random Forest and Support Vector Machine Models in Educational Data Mining. In *Proceedings of the 2nd International Conference on Advancement in Computation and Computer Technologies (InCACCT)* (131–135). IEEE. <https://doi.org/10.1109/InCACCT61598.2024.10551110>
- Aviso, K. B., Janairo, J. I. B., Lucas, R. I. G., Promentilla, M. A. B., Yu, D. E. C., and Tan, R. R. (2020). Predicting Higher Education Outcomes with Hyperbox Machine Learning: What Factors Influence Graduate Employability? *Chemical Engineering Transactions*, 81, 679–684.
- Celine, S., Dominic, M. M., and Devi, M. S. (2020). Logistic Regression for Employability Prediction. *International Journal of Innovative Technology and Exploring Engineering*, 9(3), 2471–2478. <https://doi.org/10.35940/ijitee.C8170.019320>
- Chopra, A., and Saini, M. L. (2023). Comparison Study of Different Neural Network Models for Assessing Employability Skills of IT Graduates. In *Proceedings of the International Conference on Sustainable Communication Networks and Application (ICSCNA)* (189–194). IEEE. <https://doi.org/10.1109/ICSCNA58489.2023.10368605>
- Maaliw, R. R., Quing, K. A. C., Lagman, A. C., Ugalde, B. H., Ballera, M. A., and Ligayo, M. A. D. (2022). Employability Prediction of Engineering Graduates Using Ensemble Classification Modeling. In *Proceedings of the IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC)* (288–294). IEEE. <https://doi.org/10.1109/CCWC54503.2022.9720783>
- Monteiro, S., Almeida, L., Gomes, C., and Sinval, J. (2020). Employability Profiles of Higher Education Graduates: A Person-Oriented Approach. *Studies in Higher Education*, 1–14. <https://doi.org/10.1080/03075079.2020.1761785>
- Nordin, N. I., Sobri, N. M., Ismail, N. A., Mahmud, M., and Alias, N. A. (2022). Modelling Graduate Unemployment from Students' Perspectives. *Journal of Mathematical, Computational and Statistical Sciences*, 8(2), 68–78. <https://doi.org/10.24191/jmcs.v8i2.6986>
- Oliver, N., Pérez-Cruz, F., Kramer, S., Read, J., and Lozano, J. A. (2021). *Machine Learning and Knowledge Discovery in Databases*. Springer Nature. <https://doi.org/10.1007/978-3-030-86486-6>
- Philippine Statistics Authority. (2021). Unemployment Rate in September 2021 is Estimated at 8.9 Percent.
- Shahriyar, J., Ahmad, J. B., Zakaria, N. H., and Su, G. E. (2022). Enhancing Prediction of Employability of Students: Automated Machine Learning Approach. In *Proceedings of the 2nd International Conference on Intelligent Cybernetics Technology and Applications (ICICyTA)* (87–92). IEEE. <https://doi.org/10.1109/ICICyTA57421.2022.10038231>
- Shuker, F. M., and Sadik, H. H. (2024). A Critical Review on Rural Youth Unemployment in Ethiopia. *International Journal of Adolescence and Youth*, 29(1), 1–17. <https://doi.org/10.1080/02673843.2024.2322564>
- Tamene, E. H., Salau, A. O., Vats, S., Kaushik, K., Molla, T. L., and Tin, T. T. (2024). Predictive Analysis of Graduate Students' Employability using Machine Learning Techniques. In *Proceedings of the International Conference on Artificial Intelligence and Emerging Technology (Global AI Summit)* (557–562). IEEE. <https://doi.org/10.1109/GlobalAISummit62156.2024.10947923>