

Evaluating Impact of Skill Oriented Courses on Women Empowerment Using Machine Learning

Blessy Paul P^{1*} and Cini Kurian²

¹ Research Scholar, School of Computer Sciences, MG University, Kottayam, 686560, Kerala, India.

² Principal, Al-Ameen College, Edathala, Kochi, 683564, Kerala, India.

Abstract. A nation can not advance fully without the advancement of its women. The conventional way of life for a woman restricted to her family has changed in modern-times. The level of education and employment that women have acquired has an impact on women's empowerment. On this occasion, it should be investigated whether vocational education is necessary for women employment and whether all women who receive education find employment and are satisfied with their jobs. This research examines the contribution of vocational training on women's empowerment by gathering data from working women. The machine learning algorithms used here to assess the affects of skill-oriented courses on employability include Artificial Neural Networks (ANN), Support Vector Machines (SVM), Naive Bayes, Random Forest, and Decision Tree. The ANN algorithm was used to perform a more accurate evaluation.

Keywords: Women Empowerment, Employability, Skill Oriented courses, Machine Learning Techniques, Accuracy.

1 Introduction

India's economy has one of the fastest growth rates in the world. With the Eleventh Five Year Plan, skill development has been a key initiative of the Indian government. India is the country with the second-largest population, the second-largest labour force, and the highest percentage of people who are working age. 49 percent of the population is made up of women. Yet, Indian women's status is not admirable. Women that are empowered are independent, self-assured, and capable of handling difficult situations. The economic development of every country depends heavily on the advancement of women. In India, a society where men predominate, women frequently encounter several barriers. According to Manjushree[1], women and men must work together to achieve the goals of the nation's peaceful development, which calls for the empowerment of women. In order to empower women, education and skills acquisition will have an essential impact.

Gaining new skills is essential for success since it increases output, employability, and financial options. It is now seen as a significant and essential weapon for women's empowerment. According to shetty[2], the majority of obstacles to obtaining skills and

* Corresponding author: blessypaul91@gmail.com

useful job are faced by Indian women. In order to build a trained workforce and economic prosperity, significant effort is required. In line with the conventional educational model, practical training now takes precedence over theoretical studies. Employment prospects were particularly limited for women because they had less opportunities for practical education. But, as practical topics began to appear in college studies, new employment prospects developed. This study assesses the value of skill development programmes and training for the workforce and how it influences women's empowerment.

2 Literature Review

Student performance evaluation and employability have been done using several machine learning methods. Combining various methodologies, like Bayesian technique and the decision tree techniques [3], can be used to make certain predictions about students' employability. Some researchers used [4], WEKA tool and R programming for student performance evaluation. In [5], the researcher described his technique and contrasted it with the decision tree, Naive Bayes, and neural network classification algorithms. According to Vijaya [6], increasing women's employability will result in higher output and income. Their ability to be more inventive is a result of their confidence. In traditional method, the major factors considered for employability evaluation are university exam scores and performance outcomes [7]. But, this paper handles the user attributes of employed women. Because they are more productive and require less training than graduates without experience, those with expertise are in high demand during the hiring process. Since it is thought to be the primary factor in producing this workforce, HEIs must go through regular reviews to offer future graduates with the necessary skills [8]. Many academics have tried to improve graduate attributes and curriculum to assist employability in higher education in recent years by using machine learning [9]. By utilizing the three different classification techniques DT, ANN, and SVM, the writers in [10], [11] sought to determine the most important determinants determining graduate employment. The characteristics used in this study are demographic characteristics, extracurricular and co-curricular activities, university characteristics, and internships. Data for the study were gathered from institutional databases. In a different study [12], hard skills, demographic factors, and extracurricular activity traits were utilized to forecast college students' employability for part-time job with local firms. The data for these analyses were gathered via student surveys. In [13], the authors used graduates survey data and institutional databases to identify the most predictive characteristics through relevant skills, personal qualities, and demographic profiles to ascertain why learners are most likely to obtain employment. The authors then used and contrasted three methods: SVM, RF, and DT. The authors in [14] used a hyper box-based machine learning algorithm with 78% correctness to explore the effects of several institution attributes on graduate employability. This research gives more perfect solution to evaluate the employability of women who completed the skill oriented training. At the same time it compares the difference in job opportunities for the women who doesn't have skill training. The above algorithm is used in this research, but according to the difference in the way the algorithm is implemented and the specificity of the data attribute, 95.78% accuracy has been obtained.

3 Data Collection and Pre-processing

Women who are employed from all around Kerala were surveyed to gather the necessary data for this study. A significant amount of data was provided by alumni

associations from all colleges, particularly women’s colleges in Kerala. The women have either completed their undergraduate or graduate degrees. The sample data includes females in the age ranges of 25 and older, 26 to 40, and over 40. 2610 females contributed the data who are working in different job field. Based on the institution databases, Anita [15] projected the students’ employability based on their technical talents. By referring Figure 1, it is clear that, useful piece of information to determine one’s freedom for employment, self-dependence and education are marital status and age group.

Are you financially self-sufficient and independent		
Age group	Married or Unmarried?	
26-40	Married	1.
	Un-Married	1.
Above 40	Married	1.
	Un-Married	1.
Below 25	Married	1.
	Un-Married	0.

Fig. 1. Relationship between age group, marital status and self-dependence

This study found a connection between employment and both general education courses and skill-based courses. Those who have attended skill-oriented courses are more likely to find employment, according to the employment statistics. The direct perspectives of working women indicate that vocational training is essential for professional advancement. The importance of skill development in empowering women is underscored by the fact that women who received vocational training surpass other women in terms of earnings, independence, and standard of life. This pertinent data demonstrated the significance of skill development for women’s empowerment. Google form was the main platform used for data collection. Table 1 includes the selected features and their definition. As a total, 2610 employment data has been applied to various machine learning model. Here, each attribute has a range of values, because the numerical representation is more convenient to calculate. Three incomplete rows were avoided and outliers removed. No fabrication or falsification applied. The categorical data where encoded and applied feature scaling for the accuracy improvement. Table 1 includes all the attribute values based on the feedback. The real data helped to reach a correct evaluation of women employment.

Table contains different attributes and corresponding values within a range. The possible values are depend upon particular criteria. The data collected through online platform, Google form. The major questions asked through google questionnaire are based on course studies, presence of skill component in course, employability and effect of skill components in employability. The possible values are either zero or one or within a range.

Table 1. Employment attributes.

Attribute	Attribute Definition
Age group	Below 25 – 1, 26-40 - 2, Above 40 -3
Married or Unmarried?	Unmarried – 0 Married - 1
Are you graduate or post-graduate?	Graduate – 1 Post-graduate -2
Course type	Normal course-BA, B. Com, B. Sc etc or corresponding PG -1 Professional course-B.Tech, BCA etc or corresponding PG -2 Skill oriented course-B.Voc, M.Voc etc -3 Certificate course -4
Did the course you studied include vocational training components?	Yes -1 No -0
Are you employed or not ?	Employed -1 Unemployed - 0
Has the course you studied affected your employability?	Yes -1 No -0
Do you think that getting vocational training can improve your career?	Yes -1 No -0
Status of income	Good - 3 Average - 2 Below Average - 1
Are you satisfied in your Job?	Yes -1 No -0
Are you the financial support of your family?	Yes -1 No -0
Are you financially self-sufficient and independent?	Yes -1 No -0
Women in your community can be employable after acquiring vocational skills?	Yes -1 No -0
Has vocational education/Skill oriented courses helped in improving your lifestyle as a woman?	Yes -1 No -0

4 Methodology

The writers in [16] developed a structure using several machine learning techniques, to predict candidate selection by utilising various statistical data on feature selection, including job skills, characteristics, and professional experience. Since it is challenging for a machine to interpret raw information and provide the desired results, data preparation is a crucial step in the creation of a machine learning model [17]. Here the pre-processed data has been applied to five machine learning methodologies Artificial Neural Networks (ANN), RF, DT, SVM, and Naive Bayes.

4.1 ANN

It is the algorithm which gave highest accuracy with 95.78%. Like the neurons present in the biological nervous system, artificial neurons are computational units. The majority of the ANN model is composed of the input, hidden, and output layers. Sequential

modelling and dense layers were used in the ANN implementation. The network has 14 variables in the input layer, 3 hidden layers with 6 nodes, and an output layer with 1 node, as indicated by the notation 14/6/1. The chosen characteristics are supplied as input. Relu and Adam are the activation and optimizer functions, respectively. The initial weight is 0.5, and there are 6 layers. The loss function used in this study, known as binary cross entropy, shows the discrepancy between output that was anticipated and what was actually produced. ANN achieved the greatest result over 100 epochs when batch size was 10 and verbose was 1.

4.2 DT

It is a clear and easy to understand supervised machine learning method. Here, the DT was trained using the following ideal parameters: criterion selected="gini", number of random state="100", maximum depth="3", and minimum samples leaf="5" Here, after utilising 10-fold cross-validation, a final accuracy of 95.6 percent was attained.

4.3 Random Forest

It is used as a supervised learning approach in this research. Regression and classification are also possible uses for it. A data set with x features selects a random feature called y first if there are x features in the data set. Following that, the algorithm tries to combine trees using the voting process and predicted result. 95.65% of predictions were accurate using Random Forest Classifier with n estimators=100.

4.4 SVM

To classify the objects in the provided dataset, SVM uses predefined classifications. In order to improve performance accuracy, it assigns one or more classes to transactions to categorise them. [18]. Here, it is employed a linear SVC (Linear Support Vector Classification). The linear kernel type provided an accuracy of 90.67%.

4.5 Naïve Bayes

A variation of the Naive Bayes method is called Gaussian Naive Bayes (Gaussian NB). It is a chance based ML technique related to Bayes theorem that is used for a variety of classification tasks. It strongly presumes that predictors should be independent of one another. The maximum likelihood method is applied to estimate the parameters. The least accuracy was of this algorithm, which is 85.44%.

5 Implementation

In India, the majority of women workers are currently unskilled. They can be inspired to acquire life skills that would enable them to land well-paying employment with better living conditions and the self-assurance to support their families. They will become more capable and self-sufficient as a result. The current situation needs to be assessed in order to promote the skill development of women. It is necessary to address the current research' shortcomings. The majority of research employs hard skill levels, demographic data, and characteristics of extracurricular activities as the attributes for model training. This research contains direct feed-backs from women that help to find more accurate

solution. As per the collected data, more than half percent women studied courses which include vocational training components. As per Figure 2, 70% of the women agreed that the course they have studied, affected their employability. 90% women believe that getting vocational training can improve their career. This opinion is from their direct experience at workplace. According to the data gathered, more than 50% of women do not earn a good salary. The majority of women concur that taking skill-oriented courses and receiving training will enhance their performance at workplace and result in higher pay. According to the ANN algorithm, women who have completed skill-oriented courses will have a higher priority when applying for jobs. Such successful women are more independent, according to the analysis. The empowerment of women is led by it.

Do you think that getting vocational training can improve your career?

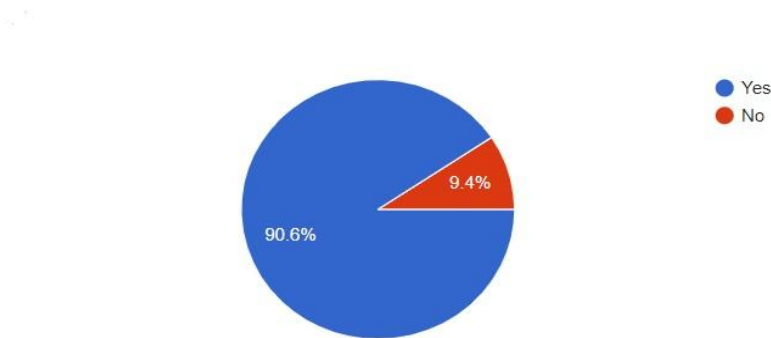


Fig.2. Feedback on importance of skill training on career development

Some of the negative results that can be understand by implementing the algorithm. Even though more than 50 percent of women have jobs, they are not financially stable. Algorithms point to lack of skill training as the reason. In the data collected, 58% women are the financial support of the family. Out of this 45% women are not satisfied with their job. The decision tree algorithm found that the reason for this was stress caused by work, low income and unstable work. More self-satisfying work is the foundation of women's empowerment.

Do you agree that women in your community/neighbourhood can be employable after acquiring vocational skills?

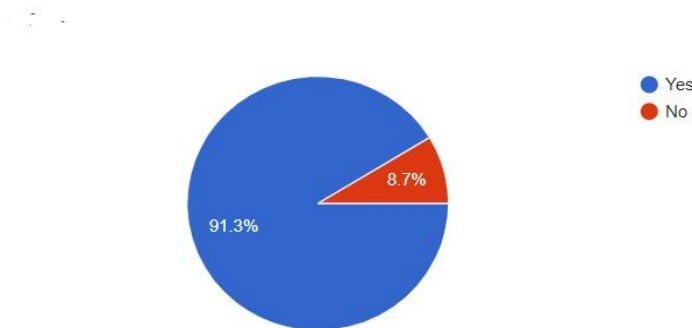


Fig. 3. Effect of skill training in getting job to others

Has vocational education/Skill oriented courses helped you in improving your lifestyle as a woman?

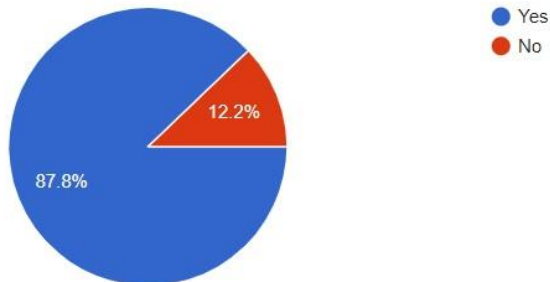


Fig. 4. Importance of skill oriented courses on life style and self-esteem

Figure 3 and Figure 4 illustrates that, women empowerment can be ensured by providing more skill oriented courses to women. All women agrees that if they get skill training, they can improve in career.

6 Results

This research implements five machine learning algorithms Artificial Neural Networks(ANN), Random Forest, DT, SVM, and Naive Bayes with accuracy 95.78%, 95.65%, 95.65%,90.67%, and 85.44% respectively. The major methodologies for finding accuracy measures are precision (rate of exact accuracy), recall (true positive rate) and f1-score (model accuracy). Ratio of correct positive occurrences to all instances that were positively predicted can be defined as precision [19]. Table 2 shows the accuracy measures of various algorithms used in this research. In this research the maximum accuracy achieved by the neural network model. Thus it indicates that vocational components in graduation or post-graduation courses are positively affecting the employability.

Table 2. Confusion matrices and Performance measures of ANN, DT, Rndom Forest, SVM And Naive Bayes.

model	TP	TN	FP	FN	Accuracy	Precision	Recall	F1 Score	ROC
ANN	448	52	11	11	95.78	97.60	97.60	97.79	0.95
DT	676	73	12	22	95.65	98.25	96.84	97.60	0.889
RF	676	73	12	22	95.65	98.25	96.84	97.60	0.889
SVM	667	43	42	31	90.67	94.07	95.55	95.23	0.801
NB	623	46	39	75	85.44	94.10	89.25	91.65	0.789

Figure 5 shows the skill component and employability. It is the diagram drawn by python matplotlib library file. Figure shows that, most of the employed woman studies the course which contains skill component. It means the skill training helped them to achieve good career. Figure 6 shows the final result of the research, which indicates the course type and relevance of the course in employability. From this figure, it is clear that normal courses have less employability and skill oriented courses have increased employability rate.

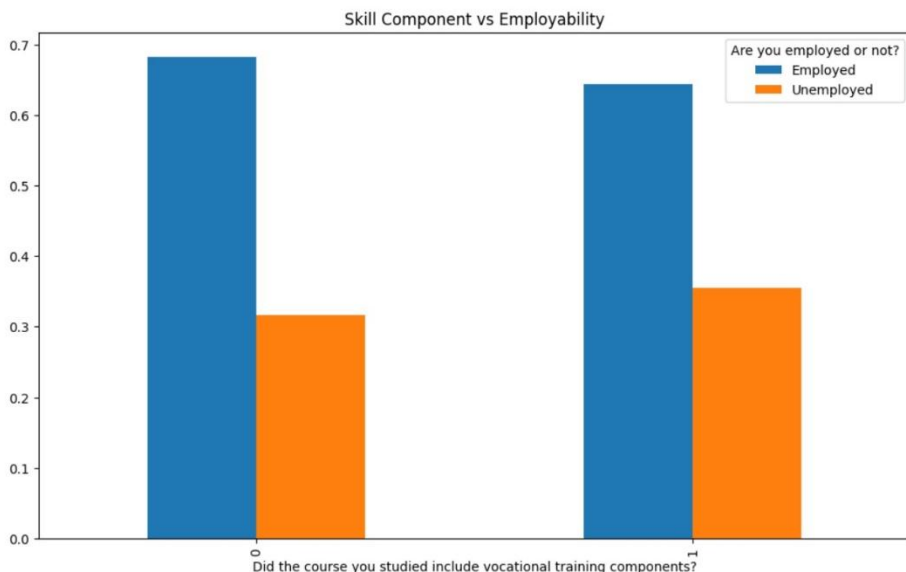


Fig. 5. Skill component and Employability.

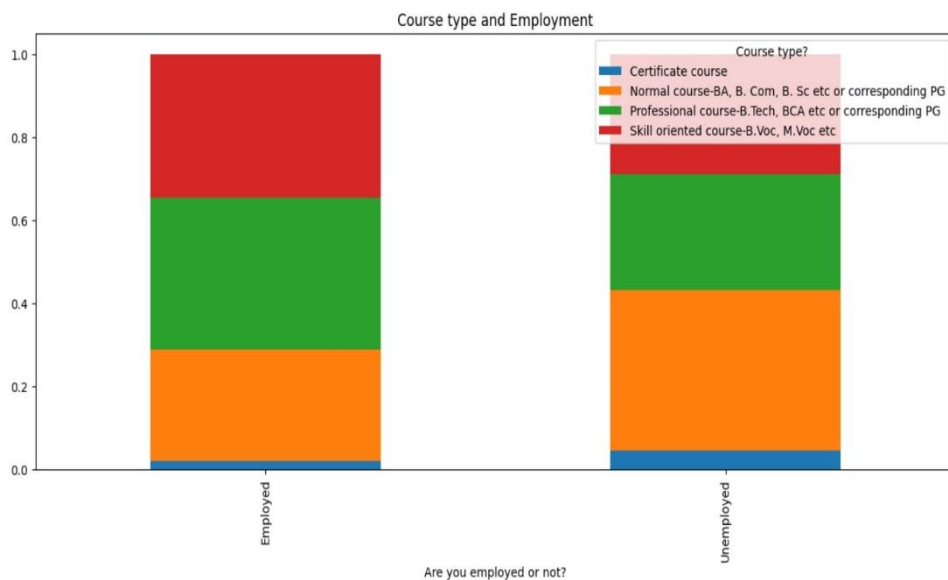


Fig. 6. Course type and Employability.

7 Conclusion

The implemented models are effective and helpful for higher education sectors to build qualified long-term systematic strategies for college students who are well packed with content knowledge, skill set, and meet the requirements of the industry. If government can identify the drawbacks of current education system, they can develop the new education policy with advanced vocational training. As per accuracy measurement, ANN achieved 95.78% accuracy. ANN is the best algorithm having the ROC index 0.95.

The foundation for India's development is an independent populace. Women who are educated, have stable jobs, and can get through tough times play a significant role in this advancement. Women face a lot of challenges, including unemployment, despite India making progress in terms of literacy. It is necessary to assess the current scenario in order to find solutions for their unemployment issues and the discrepancy between their skill sets and the demands of the labour market. This research implements five machine learning algorithms. This study makes it easy to comprehend how much age and marital status influence career opportunities. According to this study, there is a relationship between employment and both regular courses and skill-oriented courses, and individuals who have taken skill-oriented courses are more likely to get job. Vocational training is crucial for career growth, according to the direct opinions of working women. The significance for skill training in women's empowerment is highlighted by the fact that women who got vocational training outperform other women in terms of income, self-reliance, and living status. This research proved the importance of skill training in women empowerment.

References

1. Manjushree, D.L., Varsha, T.L., Arvind, W.K., Laxman, D.N.: Performance analysis of the impact of technical skills on employability. *International Journal of Performability Engineering* 17(4), 371 (2021)
2. Shetty, S., Hans, V.: Education for skill development and women empowerment. *EPR International Journal of Economic and Business Review* 7 (2019)
3. Mavani, U., Lobo, V.B., Pednekar, A., Pereira, N.C., Mishra, R., Ansari, N.: Naïve bayes classification on student placement data: A comparative study of data mining tools. In: *Information and Communication Technology for Sustainable Development: Proceedings of ICT4SD 2018*, pp. 363–372 (2020). Springer
4. Daud, A., Aljohani, N.R., Abbasi, R.A., Lytras, M.D., Abbas, F., Alowibdi, J.S.: Predicting student performance using advanced learning analytics. In: *Proceedings of the 26th International Conference on World Wide Web Companion*, pp. 415–421 (2017)
5. Ashok, M., Apoorva, A.: Data mining approach for predicting student and institution's placement percentage. In: *2016 International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS)*, pp. 336–340 (2016). IEEE
6. Anuradha, R., Reddy, G.: Empowerment of rural women through vocational education and training. *Conflux Journal of Education* 1(2), 12–18 (2013)
7. Zeineddine, H., Braendle, U., Farah, A.: Enhancing prediction of student success: Automated machine learning approach. *Computers & Electrical Engineering* 89, 106903 (2021)

8. Oswald-Egg, M.E., Renold, U.: No experience, no employment: The effect of vocational education and training work experience on labour market outcomes after higher education. *Economics of Education Review* **80**, 102065 (2021)
9. Saidani, O., Menzli, L.J., Ksibi, A., Alturki, N., Alluhaidan, A.S.: Predicting student employability through the internship context using gradient boosting models. *IEEE Access* **10**, 46472–46489 (2022)
10. ElSharkawy, G., Helmy, Y., Yehia, E.: Employability prediction of information technology graduates using machine learning algorithms. *International Journal of Advanced Computer Science and Applications* **13**(10) (2022)
11. Othman, Z., Shan, S.W., Yusoff, I., Kee, C.P.: Classification techniques for predicting graduate employability. *International Journal on Advanced Science, Engineering and Information Technology* **8**(4-2), 1712–1720 (2018)
12. Dubey, A., Mani, M.: Using machine learning to predict high school student employability—a case study. In: 2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA), pp. 604–605 (2019). IEEE
13. Casuat, C.D., Festijo, E.D.: Identifying the most predictive attributes among employability signals of undergraduate students. In: 2020 16th IEEE International Colloquium on Signal Processing & Its Applications (CSPA), pp. 203–206 (2020). IEEE
14. Aviso, K.B., Janairo, J.I.B., Lucas, R.I.G., Promentilla, M.A.B., Yu, D.E.C., Tan, R.R.: Predicting higher education outcomes with hyperbox machine learning: What factors influence graduate employability? *Chemical Engineering Transactions* **81**, 679–684 (2020)
15. Bai, A., Hira, S.: An intelligent hybrid deep belief network model for predicting students employability. *Soft Computing* **25**(14), 9241–9254 (2021)
16. Reddy, D.J.M., Regella, S., Seelam, S.R.: Recruitment prediction using machine learning. In: 2020 5th International Conference on Computing, Communication and Security (ICCCS), IEEE, pp. 1–4 (2020).
17. Rahman, S.R., Islam, M.A., Akash, P.P., Parvin, M., Moon, N.N., Nur, F.N.: Effects of co-curricular activities on student’s academic performance by machine learning. *Current Research in Behavioral Sciences* **2**, 100057 (2021)
18. Casuat, C.D., Festijo, E.D., Alon, A.S.: Predicting students’ employability using support vector machine: a smote-optimized machine learning system. *International Journal* **8**(5) (2020)
19. Thakar, P., Mehta, A., et al.: Role of secondary attributes to boost the prediction accuracy of students employability via data mining. arXiv:1708.02940