The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2802– 2815

An Infographic Approach for Employee Attrition for Corporate Company Based on Hard and Soft Voting Classifiers.

***Mamillapali Dianasaroj**
School of Computer Science and Engineering, Sandip University Nashik
*dianasarojcc@gmail.com
Pravin Gundalwar
School of Computer Science and Engineering, Sandip University Nashik
pravin.gundalwar@sandipuniversity.edu.in

Abstract- Attrition, also referred to as employee turnover or staff churn, is a significant problem for businesses operating in a variety of different industries. If organizations can anticipate and comprehend the variables that contribute to attrition, they will be in a better position to be able to take preventative action and reduce the negative repercussions of this phenomenon. In recent years, machine learning algorithms have matured into sophisticated tools that can analyze vast quantities of data and identify tendencies that could lead to attrition. One of the most important applications of these techniques is in the healthcare industry. For this research, a dataset was obtained from a corporate company in Nashik City. Additionally, to evaluate machine learning algorithms, Logistic Regression, Decision Trees, KNN, Support Vector Classifiers (SVC), and Ensemble Learning Hard and Soft Voting Classifiers were proposed. The ensemble soft model and the DT soft model have the best accuracy, precision, recall, and F1-score values at 0.909, whereas the SVC and LR soft models have a sensitivity of 0.1 and 1.0, respectively. The F1 score obtained by the SVC also has the worst accuracy, precision, and recall (0.680), as well as the lowest score overall.

Keywords- Infographic, Employee Attrition, Machine learning, Label encoder, Ensemble Learning, Voting Classifiers.

Introduction

Infographics are graphical representations of information, facts, and knowledge that are generated to express the subject matter in a manner that is simple and succinct. The term "infographics" is a shortened combination of the words "information" and "graphics." The human visual system can be helped in spotting patterns and trends by the use of graphics, which in turn can improve the quality of thought. Data visualization, information design, information architecture, and statistical graphics are all related fields. Due to its recent development as a tool for mass communication, infographics make fewer presumptions about their audience's prior knowledge than other visualization formats. Information graphics have come a long way from their humble beginnings in isotypes.

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2802– 2815
The term "attrition" is used to describe the rate at which employees quit a company. Several variables, including those already mentioned as well as others like remuneration, organizational culture, & market conditions, might affect employee retention. The rate of turnover varies greatly between different types of businesses. Nasik, in the Indian state of Maharashtra, is a major center for a variety of different types of production and industry, including the engineering, automobile, pharmaceutical, and IT sectors. Numerous city residents find gainful employment in these sectors. In general, macroeconomic conditions, employment market dynamics, and sector-specific trends can all impact employee turnover rates. When the economy is doing well and there is a lot of demand for talented individuals, turnover tends to go up because people have more possibilities to try new things. When times are tough, though, workers may be less likely to leave their positions in search of greener pastures. If you're looking for recent statistics and insights about workforce trends in Nashik, it's a good idea to get in touch with local government authorities, industry groups, and human resources research organizations. Furthermore, Nashik-based businesses might keep tabs on internal turnover by tracking employee departures and providing more precise figures.

There are many different reasons why employee turnover should be considered a big problem for any firm. These include the requirement to retrain including rehiring former employees, the possibility of losing key intellectual property, or delays in the completion of projects that are already in progress[1]. It has a substantial impact on organizational management, and it also contributes to the improvement of organizational culture when employee turnover is decreased, especially for organizational leaders. Instead of reacting after the fact to staff turnover as was the case in the past, it is now possible to take preventative action by anticipating it with artificial intelligence (AI)[2][3].

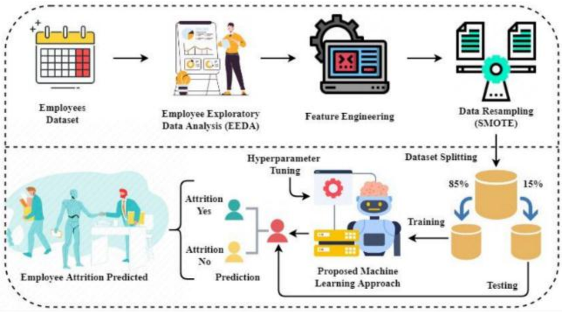

Figure 1 Employee attrition using machine learning

To make decisions and choose long-term trends, data mining analyzes data to look for patterns and trends[4]. The most recent and active area of research is data mining, and methods from this field have also been applied to classification, clustering, and prediction. There are many sectors, academic disciplines, and research fields that can benefit from different machine-learning techniques[5] [6]. Currently, businesses across a variety of industries routinely employ machine learning techniques. Only a few of them are in the retail, healthcare, finance, software, and insurance industries. Businesses that use machine learning also combine data processing with skills like pattern recognition, computer science, and statistics, along with other essential tools. Due to its ability to help organizations have a

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2802– 2815
comprehensive understanding of their employees and customers so they can make wise decisions that will benefit them, data mining is becoming more and more important in human resource management. For a company to satisfy the requirements of its business, the company must invest a significant amount of time & resources into the training of each employee. When an employee leaves a company, not only does the company lose one of its essential workers, yet it also loses the capital that it invested in recruiting, choosing, and instructing that worker for the role in the issue. On the other hand, the company must continue to make significant investments in the hiring, training, and development of new employees to fill its open roles[7]. Because of these considerations, every business makes it a priority to reduce employee turnover by cultivating more gratifying working conditions and implementing employee-friendly regulations[8][9]. The majority of companies may stand to gain from learning more about the levels of employee satisfaction and gaining information that could help them in managing attrition rates through the research study that is currently being conducted. Machine learning techniques are applied in this research to the issue of employee attrition caused by two factors. First of all, employee attrition issues have not been addressed by machine learning techniques recently[10]. Second, the attrition problem for employees is outperformed by machine learning approaches. With a predictive model, the current study effort has been performed in the way described below, from data collecting to determining the cause of staff turnover[11][12]. Predictive analytic methods have been applied here to data from human resources[13].

In corporate firms, especially those situated in Nashik City, machine learning techniques can be used to anticipate staff churn. Voting classifiers, both hard and soft, are ensemble learning techniques that aggregate the predictions of various separate classifiers to provide a final prediction. The following is how you can utilise them to forecast employee attrition:assembling a dataset Gather historical information on the attrition of employees, taking into account factors like age, gender, job level, department, performance ratings, compensation, commute distance, work-life balance, etc. Based on previous data, assign each employee an attrition status (attrited or not attrited).Preprocessing of Data To comprehend the distribution of the characteristics, find missing values, outliers, and correlations, use exploratory data analysis.Use methods like label encoding or one-hot encoding to encode categorical information. The dataset should be divided into training and test sets. Feature Picking Use feature selection strategies like correlation analysis, mutual information, and tree-based feature importance to choose the features that are most pertinent for attrition prediction. Model Training: For ensemble learning, select a group of classifiers from the logistic regression, decision trees, random forests, gradient boosting, or support vector machine families.On the training set of data, train each classifier.The Hard Vote Using the trained classifiers, build an ensemble model.Let each classifier offer a prediction for every case. The ultimate prediction is determined by which classifiers get the most votes overall.Soft Voting Classifier: Assign each classifier's prediction a weight based on its level of confidence rather than only taking the majority vote into account.To arrive at the final prediction, combine the weighted guesses.Model Evaluation: Assess the effectiveness of the

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2802– 2815
ensemble models using suitable evaluation metrics, such as accuracy, precision, recall, or F1 score.To evaluate the models' capacity for generalisation, use the testing dataset.

Literature Review

Muzafary 2021 et al. [14] studies have demonstrated that a mediator function is played by intrinsic task drive in the link between natural advantages for creativity and information sharing. These analyses also demonstrated that the intrinsic incentives for innovation rise when people share information. The present study contributes significantly to the existing literature on employee creativity by examining the mediating role of information sharing and intrinsic motivation in the link between intrinsic rewards over creativity as well as employee creativity on the one hand, and the mediating role of inbuilt inspiration in the relationship between intrinsic benefits for creativity and employee imagination on the other.

Pea-Assounga 2021 et al. [15] have started to realize the relevance of making use of Internet banking services and the function that these services play in the banking industry. The major goal of this study was to determine if staff innovation moderates the connection between Internet banking and performance at a selection of banks in the nation of the Congo. The data were analyzed by making use of a model of structure equation in conjunction with partial least squares, where a sample size of 350 was taken into consideration at all stages of the procedure. The findings of the study suggested that conducting financial transactions via the Internet had a positive impact on the levels of productivity & inventiveness displayed by workers. They also found that staff efficiency or employee creativity is linked in part to employees' innovativeness. This function bridges the gap between the two ideas. The theoretical model was developed with the help of three underlying theories: the Work Demands-Resources (JD-R) Model, the Absorptive Capacity Theory (ACT), and the Resource-Based View Theory. These three hypotheses serve as the backbone of the model. This research improves upon prior studies on the subject of online banking's effect on productivity in the workplace. The report also suggested directions for future research and made some recommendations.

Maharjan 2021 et al.[16] explored the possibility of applying machine learning as a solution to the problem of staff turnover. We have constructed models for machine learning by using methods that are supervised or classification-based, such as Logistic Regression or Support Vector Machine (SVM). These techniques were utilized in our model development. The models were trained using data from a public dataset on kaggle.com (https://kaggle.com) containing information about IBM HR employees. The models were then fine-tuned to boost their overall performance when the outcomes were assessed. Precision, recall, the area under the curve (AUC), and the receiver operating characteristic (ROC) slope were only a few of the metrics utilized to evaluate the model's efficacy. As compared to the Logistic Regression model's 0.67 reliability, 0.65 sensitivity, 0.70 specificities, 0.35 type I error, or 0.35 type II error, and 0.96 area under the curve (AUC), the SVM model achieved 0.93 accuracies, 0.98 sensitivity, 0.88 specificities, 0.12 type I error, and 0.01 type II error. The accuracy of the

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2802– 2815
Logistic Regression models was 0.0. The accuracy of the Logistic Regression model was calculated to be 0.67.

Al-Darraji 2021 et al.[17] Considered a well-known problem that requires the management to make suitable judgments to keep highly skilled staff members on the payroll. It's noteworthy to note that machine learning is frequently utilized as a method that's successful in anticipating problems like these, so it's something to keep in mind. To improve the accuracy of employee attrition prediction, the study that was suggested makes use of the deep learning method in conjunction with a few preprocessing methods. Many different factors can contribute to employee turnover. The purpose of examining these variables is to demonstrate how they interact with one another and to identify which ones are the most important. The IBM analytics imbalanced dataset, which consists of 35 characteristics for a total of 1470 workers, was utilized in the assessment of our research. To get results that were more accurate to the real world, we produced a version that was more balanced based on the original. In the end, we resort to cross-validation so that we can accurately evaluate our efforts. Our study has been put to the test in a variety of different ways to demonstrate its usefulness. When employing a synthetic dataset, the accuracy of prediction is approximately 94%, whereas when utilizing the real dataset, the accuracy of prediction is approximately 91%.

Korichi 2021 et al.[18] by using statistical procedures (such as sum, max, etc.) to turn into non-time-series ones. Such techniques lead to information loss and less accurate predictions. In this research, we introduce a real-time approach to predicting employee turnover by using the longitudinal nature of the data. allowing the models to generalize to different behaviors and giving a more precise assessment of the employee departure risk.

Table-1 Literature summary

Author/Year	Dataset	Method/Model	Parameters	References
Shete/2021	UCI	SVM, KNN, Decision tree	---	[19]
G. Klop/2021	pandemic and non-pandemic dataset	Decision tree, Random forest, AdaBoost	Accuracy=99%	[20]

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2802– 2815

Mansor/2021	IBM"s fictional dataset created	ANN, DT, SVM	Accuracy=88.87%	[21]
Wijnen/2019	Conduent Human Resource Dataset	XGBoost SVM, LR, DNN	Accuracy=93.2%,Precision=93.90%,Recall=73.2%	[22]
Khamar/2019	IBM Watson dataset	RF, gradient booster	---	[23]

Proposed Methodology

A. Data collection

Data from the Nasik city region's Corporate company and Corporate Company should be gathered; it includes columns for Manager, Section, EmpID, Business Unit, Division, and so on. Score Type, Top 2% in Advocacy, Top 2% in Advocacy Mean Score, Top 2% in ESI, Top 2% in Loyalty, Top 2% in Loyalty Mean Score, Top 2% in Final, and Top 2% in Final Mean Score. 14 columns and 244 samples are contained. Data is modified using the cut-copy method five times due to the low number of samples; the final sample is 1220 with 14 columns.

B. Perform EDA

Exploratory Data Analysis is the abbreviation. Before using more sophisticated methods or modeling, it is a common practice in data science and statistics to review and summarise data sets. To acquire insights into the data and comprehend its underlying patterns, relationships, and distributions, EDA uses descriptive statistics, visualizations, and numerous analytical approaches.

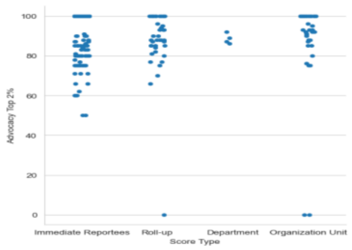

Figure 2 Scatter plot of Advocacy top 2% and score type

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2802– 2815

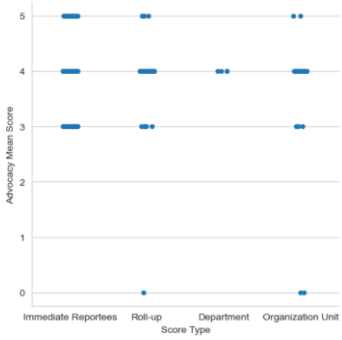

Figure 3 Scatter plot of Advocacy mean score and score type

Figures 2 and 3 display scatter plots for data frame variables such as Immediate Reporters, Rollup, Department, and Organization Unit.

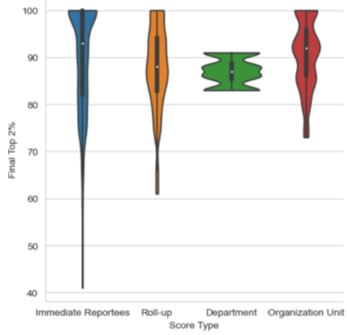

Figure 4 shows a cat plot of the final top 2% vs score type

Figure 4 displays the range of values for immediate reporters (about 41 to 100), roll-up score final accuracy (above 60 to 100), department (above 80 to above 90), and organization unit (above 70 to above 100).

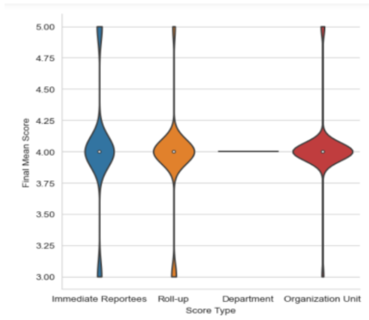

Figure 5 shows the cat plot final mean score vs score type

The final mean accuracy cat plot with score type is shown in Figure 5. The final mean accuracy ranges from 3 to 5 for the immediate reports score type, 3 to 5 for the roll-up mean accuracy, 4 for the department, and 3 to 5 for the organization unit accuracy.

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2802– 2815

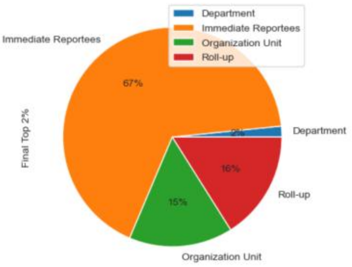

Figure 6 shows a piechart of final top 2%

Figure 6 explains the score types in a pie chart. 67% have scores for direct reports, 16% for roll-up, 15% for organization units, and 2% for departments.

C. Data preprocessing

- Type conversion

Type conversion, commonly referred to as typecasting, is the process of translating data between different formats. Implicit conversion refers to the automatic transformation of the used data types by scripting languages like JavaScript and compiled languages like C++. Another possibility is that the source code itself will need to be converted.

- Null value & Missing value and fill in missing value using zero.

The value "null" is assigned to a column in a relational database when the information contained in that column cannot be determined or is not there. For data types that deal with characters, dates, or times, a null is not the same as an empty string or a value of zero, and vice versa for data types that deal with numbers. Missing values may occasionally be substituted for known quantities. Since [NA] values in the context of skip logic can either signify zero or a response from a prior iteration, this is feasible.

- Label encoder

Label encoding is a frequent encoding strategy used when working with categories. This technique assigns a unique number to each label based on where it appears in the alphabet.

- Drop unnecessary columns

The drop() method, which by default doesn't change the original DataFrame but instead returns a new one after eliminating the columns you specify, can be used to remove certain columns from a DataFrame. A preexisting DataFrame object's columns can be removed by using the inplace=True argument.

D. Data Splitting

It is referred to as "splitting" when data is divided into numerous subgroups. Usually, when data is split in half, one half is utilized to train the model, and the other half is used to test it. Data splitting is an important stage in the data science process for developing models from raw data. 90% of the data were used for training and 10% for testing in this study's 90:10 ratio.

E. Machine learning Models

Machine learning techniques can be employed in corporate businesses, particularly in Nashik City, to predict workforce turnover. Hard and soft voting classifiers are ensemble learning approaches that combine the results of various independent classifiers to get a final prediction. The way you can use them to predict staff attrition is as follows.The implementation of the machine learning & voting classifier-based strategy used both hard and soft voting techniques. In addition to using hard and soft voting classifiers, the model also used decision trees, KNN, SVM, and logistic regression.

- Logistic regression

The logistic regression technique is one of the most fundamental and popular methods for resolving classification problems. Even though the fundamental methodology of "Logistic Regression" is nearly identical to that of "Linear Regression," it is given the name "Logistic Regression." This type of classification uses the Logit function, from whence the word "Logistic" is derived.

- Decision tree

A non-parametric supervision learning technique known as a decision tree can be used to perform both classification and regression operations. You can utilize decision trees. It is structured like a hierarchical tree, having internal nodes, leaf nodes, roots, and branches in addition to the other nodes.

- KNN

The k-nearest neighbors algorithm, often known as k-NN or simply KNN, is a non-parametric, supervised learning classifier that makes use of a data point's proximity to another to create predictions or classifications about how that point is grouped.

- SVM

When data points aren't naturally divided along a linear axis, SVM translates them to a high-dimensional feature space for classification. The information is transformed so that the determined boundary between the categories can be represented graphically as a hyperplane.

- Hard and soft voting classifiers

There are two ways to settle disputes involving categorization: hard voting and soft voting. A vote is viewed as hard if the forecast receives the most votes, whereas a vote is viewed as soft if the prediction receives the highest overall probability.

I. RESULT AND DISCUSSION

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2802– 2815
The first source of information to collect is the Corporate Company, which has columns for Manager, Section, EmpID, Business Unit, Division, and other terms. Score, followed by a preprocessing step that removes extraneous columns and uses Label Encoder, Null value, and Missing value to fill in missing values. when data is used to train the suggested models in a 90:10 training: testing ratio. Listed below are metrics that are used for evaluation.

$$Accuracy = \frac{(TP+TN)}{(TP+FP+TN+FN)} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F-score = \frac{2}{\frac{1}{precision}+\frac{1}{recall}} \quad (4)$$

$$Sensitivity = \frac{TP}{TP+FN} = recall \quad (5)$$

$$Specificity = \frac{TN}{TN+FP} \quad (6)$$

Table. 2 Performance evaluation of models

Model	Accuracy	Precision	Recall	F1-score	Specificity	Sensitivity
LR	0.704	0.704	0.704	0.704	0.0	1.0
DT	0.909	0.909	0.909	0.909	1.0	1.0
KNN	0.868	0.868	0.868	0.868	1.0	1.0
SVC	0.680	0.680	0.680	0.680	0.0	0.1
Ensemble soft	0.909	0.909	0.909	0.909	1.0	1.0
Ensemble Hard	0.885	0.885	0.885	0.885	0.2	1.0

The ensemble soft model and the DT soft model have the highest accuracy, precision, recall, and F1-score value of 0.909, while the sensitivity of the SVC and LR soft models are 0.1 and 1.0, respectively. And the SVC's F1-score has the lowest accuracy, precision, recall, and score of 0.680.

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2802– 2815

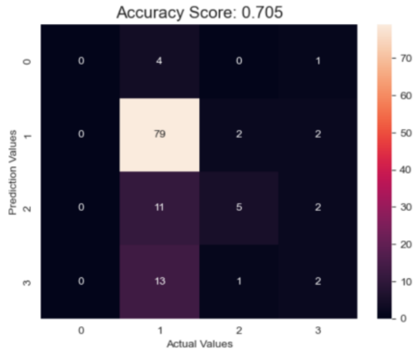

Figure 7. Confusion matrix for Logistic regression

Figure 7 depicts the logistic regression confusion matrix, with an accuracy score of 0.70 for the logistic regression of actual and predictive values.

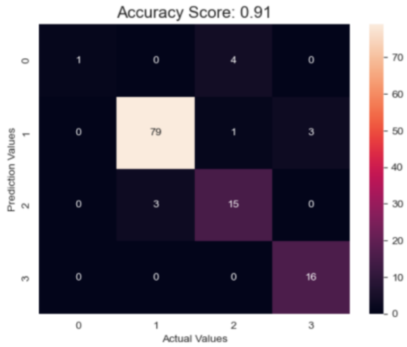

Figure 8. Confusion matrix for Decision Tree

Figure 8 depicts the confusion matrix of logistic regression, with a decision tree's accuracy score of 0.70 for both real and predicted values.

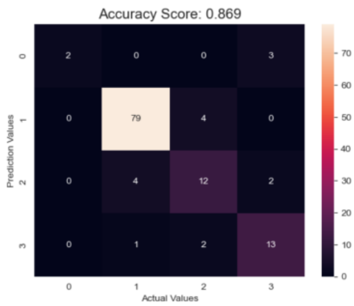

Figure 9. Confusion matrix for KNN

Figure 9 illustrates the confusion matrix of KNN, with an accuracy score of 0.70 for the KNN of actual and predicted values.

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2802– 2815

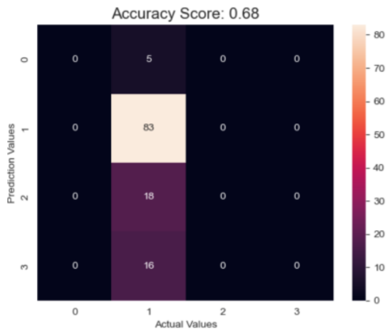

Figure 10. Confusion matrix for Support vector classifier (SVC)

Figure 10 depicts the confusion matrix for SVC, which has an accuracy score of 0.70 for both the actual and predicted values.

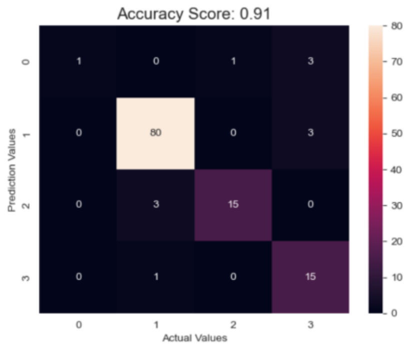

Figure 11. Confusion matrix for Ensemble Soft

The Ensemble Soft confusion matrix is shown in picture 11. where Ensemble soft of actual and predicted values has an accuracy score of 0.70.

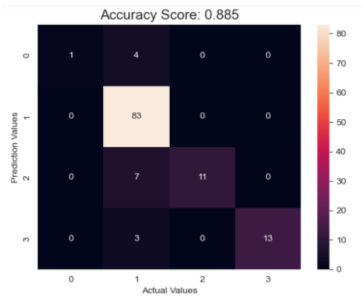

Figure 12. Confusion matrix for Ensemble Hard

The confusion matrix for Ensemble Hard is depicted in Figure 12 with an accuracy score of 0.70 for both actual and predicted values.

Conclusion

Attrition, often known as employee turnover or staff churn, is a significant problem for businesses in a variety of different industries. If organizations can anticipate and comprehend the variables that contribute to attrition, they will be in a better position to be able to take preventative action and reduce the negative repercussions of this phenomenon. In recent years, machine learning algorithms have matured into sophisticated tools that can analyze vast quantities of data and identify tendencies that could lead to attrition. One of the most important applications of these techniques is in the healthcare industry. In this study, a dataset from the Corporate company Nasik City was utilized, and various machine learning techniques, including logistic regression, decision trees, kernel neural networks (KNN), support vector classifiers (SVC), and hard and soft ensemble learning, were proposed for evaluation. The ensemble soft model and the DT soft model have the best accuracy, precision, recall, and F1-score values at 0.909, whereas the SVC and LR soft models have a sensitivity of 0.1 and 1.0, respectively. In addition, the SVC's F1 score has the worst accuracy, precision, and recall (0.680), as well as the lowest score overall.

References

[1] W. Sen, Z. Xiaomei, and D. Lin, "Impact of Job Demands on Employee Learning: The Moderating Role of Human–Machine Cooperation Relationship," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–11, 2022, doi: 10.1155/2022/7406716.

[2] B. Saeidipour, P. Akbari, and Z. Alizadeh, "The Role of Employee Demographic Characteristics and Selection of Human Resource Management Methods on Organizational Tensions and Turnover Intention," *Educ. Res. Int.*, vol. 2022, 2022, doi: 10.1155/2022/5365860.

[3] A. Mehbodniya *et al.*, "Data Mining in Employee Healthcare Detection Using Intelligence Techniques for aIndustry Development," *J. Healthc. Eng.*, vol. 2022, 2022, doi: 10.1155/2022/6462657.

[4] N. Mduma, "Data Balancing Techniques for Predicting Student Dropout Using Machine Learning," *Data*, vol. 8, no. 3, p. 49, 2023, doi: 10.3390/data8030049.

[5] M. T. Alharbi and M. M. Almutiq, "Prediction of Dental Implants Using Machine Learning Algorithms," *J. Healthc. Eng.*, vol. 2022, 2022, doi: 10.1155/2022/7307675.

[6] S. Soner, A. A. Hussain, and R. Khatri, "Predictive Deep Learning approach of employee attrition for imbalance datasets using SVMSMOTE algorithm with Bias Initializer," pp. 1–19, 2022.

[7] M. Atef, D. S. Elzanfaly, and S. Ouf, "Early Prediction of Employee Turnover Using Machine Learning Algorithms 135 Original Scientific Paper," *Int. J. Electr. Comput. Eng. Syst.*, pp. 135–144, 2022.

[8] Y. Qian and J. Yin, "Study on Employee Performance Evaluation Based on Adaptive Feature Selection Fuzzy Algorithm," *Mob. Inf. Syst.*, vol. 2022, 2022, doi: 10.1155/2022/6985508.

[9] C. Onyutha, "Multiple Statistical Model Ensemble Predictions of Residual Chlorine in Drinking Water: Applications of Various Deep Learning and Machine Learning Algorithms," *J. Environ. Public Health*, vol. 2022, 2022, doi: 10.1155/2022/7104752.

[10] G. Pratibha and D. N. Hegde, "Analytical Survey on Prediction of Employee Attrition Non parametric tuning Algorithms," vol. 10, no. 6, pp. 818–826, 2022.

[11] S. Aggarwal, M. Singh, S. Chauhan, M. Sharma, and D. Jain, "Employee Attrition Prediction Using Machine Learning Comparative Study," *Smart Innov. Syst. Technol.*,

vol. 265, no. 9, pp. 453–466, 2022, doi: 10.1007/978-981-16-6482-3_45.

[12] F. Fallucchi, M. Coladangelo, R. Giuliano, and E. W. De Luca, “Predicting employee attrition using machine learning techniques,” *Computers*, vol. 9, no. 4, pp. 1–17, 2020, doi: 10.3390/computers9040086.

[13] R. Yedida, R. Reddy, R. Vahi, R. Jana, A. GV, and D. Kulkarni, “Employee Attrition Prediction,” 2018.

[14] S. S. Muzafary, M. N. Wahdat, M. Hussain, B. Mdletshe, S. A. Tilwani, and R. Khattak, “Intrinsic Rewards for Creativity and Employee Creativity to the Mediation Role of Knowledge Sharing and Intrinsic Motivation,” *Educ. Res. Int.*, vol. 2021, 2021, doi: 10.1155/2021/6464124.

[15] J. B. B. Pea-Assounga and H. Yao, “The Mediating Role of Employee Innovativeness on the Nexus between Internet Banking and Employee Performance: Evidence from the Republic of Congo,” *Math. Probl. Eng.*, vol. 2021, 2021, doi: 10.1155/2021/6610237.

[16] R. Maharjan, “Employee Churn Prediction using Logistic Regression and Employee Churn Prediction using Logistic Regression and Support Vector Machine Support Vector Machine,” pp. 52–54, 2021.

[17] S. Al-Darraji, D. G. Honi, F. Fallucchi, A. I. Abdulsada, R. Giuliano, and H. A. Abdulmalik, “Employee attrition prediction using deep neural networks,” *Computers*, vol. 10, no. 11, pp. 1–11, 2021, doi: 10.3390/computers10110141.

[18] A. R. Korichi, H. Kheddouci, and D. J. West, “Predicting employee attrition with a more effective use of historical events,” *ESANN 2021 Proc. - 29th Eur. Symp. Artif. Neural Networks, Comput. Intell. Mach. Learn.*, pp. 53–58, 2021, doi: 10.14428/esann/2021.ES2021-110.

[19] M. B. Shete, “Prediction of Employee Attrition Using Machine Learning Approach,” *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 9, no. 8, pp. 2479–2483, 2021, doi: 10.22214/ijraset.2021.37796.

[20] G. Klop, “Predicting Employee Turnover and Reducing Turnover Costs Using Machine Learning Techniques,” *Ind. Eng. Innov. Sci.*, p. 98, 2021.

[21] N. Mansor, N. S. Sani, and M. Aliff, “Machine Learning for Predicting Employee Attrition,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 11, pp. 435–445, 2021, doi: 10.14569/IJACSA.2021.0121149.

[22] C. Wijnen, “Predicting employee attrition,” no. May, pp. 1–62, 2019.

[23] N. R. Khamar, R. No, K. G. Upadhyay, and R. No, “EMPLOYEE ATTRITION PREDICTION USING MACHINE LEARNING AND SENTIMENT ANALYSIS DEPARTMENT OF INFORMATION TECHNOLOGY FACULTY OF ENGINEERING AND TECHNOLOGY SRM INSTITUTE OF SCIENCE AND TECHNOLOGY KATTANKULATHUR-603203,” no. May, 2019.