Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 3844 – 3857
Received 05 March 2021; Accepted 01 April 2021.

Talent Flow Employee Analysis based Turnover Prediction on Survival Analysis

Sumathi K[1], Balakrishnan D[2], Naveen V[3], Hariharan P[4], Rahul Iniyan M[5]
[1,2,3,4,5]*Department of Computer Science and Engineering*
M. Kumarasamy College of Engineering,Thalavapalayam,Karur-639113.
[1]*sumathik.cse@mkce.ac.in* , [2]*balabalu.2001@gmail.com* , [3]*sullannaveen08@gmail.com* ,
[4]*hariharanpalpandian@gmail.com*, [5]*rahuliniyan07@gmail.com*

ABSTRACT

When an employee-departures form a company, it can result in a significant loss. Therefore, we are predicting employee turnover with the help of human resource management. There are some previous researches are happened but they are mainly focused on centered turnover of an employee. They are ignored past events of company's turnover behavior and also the statistical data for each job. For that we are using an algorithm called CoxRF an employee centered turnover prediction, in this we are combining the statistical results of survival analyzing with the help of assembly learning to reduce the problem called conservative supervised binary classification for an event centered perspective. To help with the structure of survival data from censored data, we coined the terms "event-person" and "time event". We interrelate CoxRF to a number of baseline approaches using an original dataset of china's largest technological social network. The result shows that it is a good turnover interpreter. The following discoveries have been made: i) Employee turnover varies by industry, with the IT sector having a slightly higher rate than the government sector; ii) Gender plays major role if it is a women after marriage some are relieving from work and other factors also; iii) The person with good academic records can work more efficient than the other with low; iv) GDP plays an vital role in turnovers of an company and an employee in the current situation which have been neglected in previous studies; v) And the final one is for salary hike they are relieving is one of the problem for that we are losing a good employee.

Keywords

Survival analysis, Cox relative hazards model, Random forest algorithm, CoxRF turnover prediction algorithm.

Introduction

Employee turnover costs the company a lot of money and time, so the person in charge of human resources spends a lot of time and money to avoid it. As a result, it's critical to investigate and anticipate employee turnover. Voluntary and automatic turnover are the two types of turnover. Involuntary turnover can be caused by layoffs at a company and is difficult to prevent predictability as a result, we concentrate primarily on forecasting employee voluntary turnover. There are two types of datasets that have been used in previous studies.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 3844 – 3857
Received 05 March 2021; Accepted 01 April 2021.

The data of the employee has been collected from the human resource department of the company is used in some studies [1]-[9], the simulation dataset of an employee from the IBM Watson analytics[10] is representative and common. Employees in the current firm, which has a short time span, have a wealth of information about their working contents. Studies [3, [5,] and [11] have also used data from online specialised networks. LinkedIn, Maimai, and Viadeo are examples of websites with extensive social information about employees and a longer time frame. Although a few studies have looked at the real time of turnover [3], [5], [11], binary classification issues [12]-[14], [7] is the majority of issues which mainly focus on turnover of employee happens or not. There are four different kinds of research when it comes to algorithms. The first is a test of machine learning algorithms [4], [6], [7], [12], and [14]. They studied which method performed the best in these experiments by comparing different methods. Furthermore, different optimal learning models for datasets of various sizes [8] and features [13] were developed by researchers.

Some studies [15] looked at the differences between different companies which implements different model of prediction to every model. Survival analysis algorithm is coming under the second category of [11], [5], [3], that predict turnover of employee is the likelihood for relieving a post at a specific duration. Semi-Markov algorithms [2] and the algorithms which blend the individuality of the social network of employee [14], [13], [17] fall into the third and fourth categories, respectively. The term "survival analysis" refers to an approach called statistical analysis that takes into account with both the time and the event before it occurs. In sociology, it's known as Event History Analysis, while in engineering, it's known as Failure Analysis and in biostatistics, it's known as Survival Analysis [17]. The following traits can be found in survival data: The data takes into the account until both the time and the event occur. There are a lot of censored data in the data, and the survival time distribution is difficult to detect. These features match the records of work experience in our database. There is no precise turnover time only the occurrence of turnover events is present as many as in our dataset. Avoiding of these records would cause the result in loss of a great deal of useful information and also make serious imbalance in our data examples, and it would make the prediction incorrect. For this we are going with CoxRF. It is an innovative employee turnover prediction method, that is coming under the model called survival analyses from that we are taking out the perspective of events. To start we have to reframe the issue, we have to focus on the employees but we are focusing on the turnover events. From the observation in this model every the single employee will take unique decisions in unique situations and in the businesses. It's worth to find our task(model) is to determine if the turnover crisis will occur. After that we have to find the likelihood of the crisis in a certain point of period and use that feature to determine the probability score us the property to determine if the employees can maintain control over turnover at that time. High level of F1-score and AUC for prediction in our CoxRF in compare with the different baseline ml algorithms with more accuracy we proceed with this approach. We also do a thorough analysis of the data and come up with the fascinating things: i) Sex(gender) play an major contribution in the turnover of employee behaviour, in which the female candidates having a greater rate of turnover compare to male candidates; ii) GDP growth often play a significant role in employee turnovers, which has been overlooked from the older data-driven analyses; iii)

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 3844 – 3857
Received 05 March 2021; Accepted 01 April 2021.

Candidates turnover behaviour differs by industry likewise the government has less turnover rate but the IT sector having more. iv) The person with good academic records can work more efficient than the other with low; The following is the outline for the rest of this paper: The second section contains a short description of the data. In Section III, we go over our strategy in greater depth. The results of the experiments are summarised in Section IV. Section V incorporates related survival analysis works. This paper is summarised in Section VI.

Data Description

Dispersed crawlers are to collect a huge amount of chartered communal data from the India's largest working social media platform. Employee's private information, like hometown, gender, educational qualification, salary, and work experience was added to our dataset. Our dataset had up to 460 thousand work experiences from different employees and 243 thousand educational experiences of 290 thousand employees. Data of the user is primarily classified into three stages. The first one is demographic information, like sex, date of birth, native, constellation, and so on; The second one is current job status, like name of company, working location, work area, and job role; the final one is social platform information, like views, likes earned, comments, impact in public, how he maintain his private information, and so on. The starting and ending of a particular work, work role, work title, work objective, beginning time, updating time, and the label, all of these details are available. The educational background information includes when he start and when he finish education, school details, achievements, awards, degree and objective about his life and finally the additional learning experience.

The data set has all the necessary details of the employee from his birth to his present life and also from his social media platform we can take data so that we can judge him accurately.

Methodology

Problem Findings

From this step, we have to find the calculation of income of the employee by using event-centred perspective and which will be focus mainly on whether an turnover of the employee function will happen or not. There is a way to solve the problem which was a direct way to find the turnover from behaviour of candidates only by revenue of his previous crisis. It is a "biased" type of problem definition, if an employee leaves from his job then he is not related to the history of his previous income, but it also depend upon, his ongoing position individuality. Furthermore, temporary factors will not to be ignored, as the candidates probability income changes at a time. As the result, the best way to discover or solve the issue is to compare the turnover past historical characteristics, the ongoing work, and the time point at which the turnover forecast is to be made.

We incorporate the appropriate prediction time from the survival function and to find the candidates time-variant and time-invariant data from every post(job) with the aid of the

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 3844 – 3857
Received 05 March 2021; Accepted 01 April 2021.

above considerations to make turnover predictions based on the above considerations. The following are examples of preparations that can be used to prepare the problem:

Definition 1 (Prediction of Turnover): The turnover calculation problem uses survival analysing and algorithms of machine learning to predict the person will quiet his ongoing work at a time(t) given his turnover events on past, on working job information, information in social networks, and a specific time t..

Survival Analyses

Survival analysis, is the statistical analysing procedure, investigates time-relative occurrence probability of income functions by taking into account both events and time. As a result, without a distribution of strong theory, censored data can be used fully. Survival analysis is now used in a variety of disciplines, including biostatistics, engineering, sociology, and others.

• Calculate the rate of survival and draw a survival curve using the method of Kaplan-Meier.

• Use the Log Rank or Breslow methods to see if there are any obvious dissimilarities in the group of survival curves.

• Take the proportional hazard model is to see the factor influences the likelihood of the functions occurrences.

Basic Concepts:

Basic concepts of survival analysis are listed below as:

Time: Calculate the time for the beginning of the event to the occurrence of an event or the end of an event.

Event(Observation): During observation some events may occur like death, turnover or failure.

Censor: An event can occur as a non-occurrence during an observation. The data from the right censor shows that the object being observed has left before the incident occurs, indicating that the observation has come to an end. The knowledge from the left censor shows that the entity is seen before the events.

Role of Survival function: As compared to t, T is a higher individual survival rate. All experimental species are alive at time 0, according to the assumption S (0) = 1, finally survival feature is:

$S(t) = (T > t)P$ (1)

T - Death Time(period)

t - Specific time.

Hazard function: As compared to survival time, the likelihood of occurrences of event t is greater at time t. If t is greater than the individual survival time, the likelihood that it will not greater in survive(fig 1).

$$P\{ X \in (t, t + dt) | X > t \}$$

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 3844 – 3857
Received 05 March 2021; Accepted 01 April 2021.

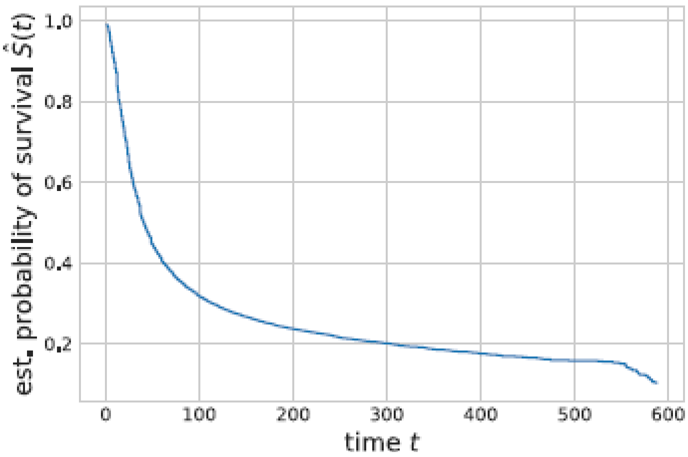

Fig. 1. Kaplan-Meier survival curve

Cox Proportional Hazards Model (CoxPH): When there are covariates in the survival data other than event and time, the Cox proportional hazards model, is also called the Cox model. It is used to measure the effects of covariates on time period. In additional, this Cox model is used to estimate the likelihood of survival from the given point of period(time). Two theories underpin this method:

• **Proportional Hazard hypothesis:** Cox's regression is a method for looking at many variables at once. Another name for it is proportional hazards regression analysis.

• **Logarithmic Linear hypothesis:** The logarithmic hazard ratio has a linear relationship with the covariates.

For individual i,

Yi –time

Ci- event occurs

andXi = $\{Xi1, Xi2, . . . , Xip\}$represents p individual covariates.

Baseline Methods

SVM: An SVM algorithm is a classification or categorization algorithm to assign to data(feature vector) from one class ot the another(higher dimensional space). It involves discovering hyper planes which cleanly segregate data into classes. Once idea hyper planes are discovered new data planes are easily classified.

Logistic regression: Logistic regression comes under the supervised learning classification algorithm which is used for the prediction of the value which will be targeted. The nature of target or dependent variable is dichotomous, which means there would be only two possible classes.

Naive Bayes: It uses Bayes theorem for object classification. Naive Bayes classifiers are based on the assumption of solid, or naive, independence between data point attributes. Spam filters, text interpretation, and medical diagnosis are all popular applications of naive Bayes classifiers. Since they are easy to implement, these classifiers are commonly used in machine learning. Simple Bayes or freedom Bayes are other names for naive Bayes.

Decision Tree(Graph): It is an intuitive graphical approach that uses statistical probability analysis and is like a tree oriented decisional graph with additionally probabilistic outcomes. It comes under the prediction model in machine learning that relates a mapping procedure between two things (object attributes and object values).In this the object attribute's judgement condition is represented by every node from the tree, and the thing(object) that meets the node condition is represented by its branch. The prediction outcome to which the object belongs is represented by the tree's leaf node.

XGBoost: eXtreme Gradient Boosting is the improvement from the GBDT-based boosting algorithms. This Gradient boosting process is the correcting of all the poor learner's residuals by inserting the fresh one. Then finally, more(multiple) learners are combined to the closing forecast, which has a higher accuracy than a particular one. Random Forest algorithm: It is an ensemble learning technique that combines several result trees. Training set of each tree is taken random. It and reversibly from the training samples. In training, all features are taken at random without being proportioned.

CoxRF: A Turnover Prediction Model

We can easily see that the threat rate λ(t) has a value from [0,∞] and S(t) Survival Rate has a value range of [0, 1]. When choosing covariates that if Xij(0 j p) on things varies over period, that if the influencing factor of Xij(0 j p) is equal to Xij, and also equal to the time, then j(0 j p) not equal to constant and a function on time period t, rendering the model of cox model useless. We choose features whose effect on events does not shift over time, based on the Cox model's strong hypotheses. Then, along with time and events, adding that feature to as basic covarriates on to shape the survival information. The first step is to calculate the rate of survival for every individuals using the cox model at the end. New features are relearned from the dataset by using the algorithm called supervised classification.

Experiments

With the help of past research findings and , first step is to extracted appropriate features in this section. The presentation of CoxRF is then compared to baseline algorithms. Finally, Kaplan-Meier analysis used to see difference factors influence employee turnover behaviour.

Feature Extraction

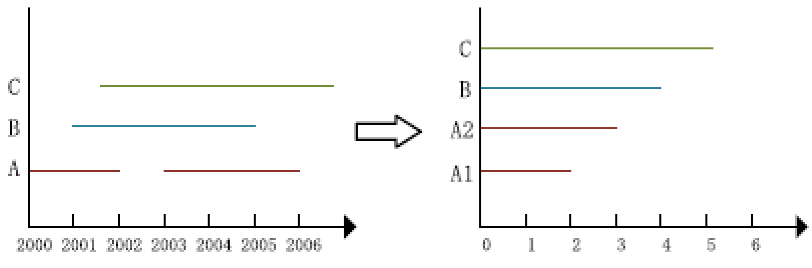

Fig. 2. Concepts of event-person and year-event

Personal characteristics of employees , managerial factors, environmental factors and structure factors are all factors that influence employee voluntary turnover,

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 3844 – 3857
Received 05 March 2021; Accepted 01 April 2021.

according to previous studies [16], [18]. Personal characteristics of employees include gender, age, marital status, education, and so on. Gender is a factor (male employees high turnover rate) [19], education ranking (high school high turnover rate) [5], and marital status (low turnover rate of married employees) [19] are some of the most important factors in the first-category, according to some studies. Working hours in the sector (high turnover rate of long working hours) [4], promotion (high promotion times of low turnover rate) [19], role (high positionsofhigh turnover rate) [19], are some of the organisational variables, and work performance [5] (fig 2). The macroeconomic climate, work opportunities outside of the institution, and labour market conditions, among other things, are all examples of external environmental factors. Voluntary turnover is heavily influenced by the external climate, according to some studies [12], [15], and [20]. Take, for example, the research [20]. For example, the study [20] found a link between the increase and decrease of company stock prices and employee mobility between the companies. Other research [15] found that economy is good, employees are more like to leave without being forced, while if the economy is bad, employees are more likely to leave passively. Social networks of an employee are referred by the structural factors. The centrality, in-degree, out-degree, and other characteristics of the employee network, for example [13].

Human resources data’s is highly confidential and rarely shared between companies. As a result, Some conventional statistical methods [5] relied on employee data collected over a short period of time within an organisation. As a result, conducting long-term career assessments of employees has become difficult, and the majority studies are based on employee self and organisational factors. The gender, highest level of education, and school type are associated with the high level of education are the characteristics we extracted for employee personal factors.

The current job's start time, end time (The system time user changed the record is used to fill in the missing values), sector classification, title rank, turnover numbers, and working time are all extractable characteristics for organisational variables. Since anyone can have several turnovers, the number of turnovers was taken into account. This means that the item will appear several times during the inspection phase, according to survival analysis. Repetitive incidents can be dealt with using current case history analysis [17]. Methods for repetitive counting and characterising data using relative time were suggested."Event-person" and "Time-event" are the concepts introduced in this paper, which are based on the same idea. Over the monitoring span from 2000 to 2006, employee A's event was observed twice, employee B's event was observed only once, and C's event was not observed, meaning that occurrences of censor.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 3844 – 3857
Received 05 March 2021; Accepted 01 April 2021.

TABLE 1. summarises the four feature types we use, totalling 24 features.

TABLE IFEATURE DESCRIPTION

Category	Name	Description
Employee personal features (3)	max_degree	Highest education, such as post doctor, doctor, master, undergraduate, etc.
	max_sch_type	The highest educational level corresponds to the level of the school, such as 985, 211, etc.
	gender	Gender of employee
Organizational features (7)	start_year	Year of job start
	end_year	Year of job end
	industry_type	Industry categories, such as education, IT, services, etc.
	position_level	Position levels, such as senior, intermediate, junior, etc.
	has_turnover_num	Number of turnovers
	has_timelength	The time that has been worked (including the current working hours)
	timelength	Length of working time
External environment features(1)	GDP	Annual growth rate of Chinese GDP
Platform information (13)	interactions	Interactions number
	dongtai	Posts number
	guandian	Standpoints number
	zhuanlan	Column number
	dianping	Comment number
	likes	Liked number
	views	Viewed number
	recent_feeds	Recently received feeds number
	influence	Influence number
	inf_defeat	Percentage of influence over others
	info_ratio	Percentage of information perfection
	imp_tag_num	Impression tag number
	pro_tag_num	Profession tag number

To begin, we divide employee A's two events into two categories: A1 and A2. Then, rather than concentrating on the calendar year, transform absolute time to relative time, or in other words, use the numbers 0-6 to show off the period 2000 to 2006, with the duration of work time as primary consideration. A1 and A2 files have been processed as "case-person," which means that the event has been divided into parts depending on who is involved. Similarly, a "year-event" is a relative time period ranging from 0 to 6, which is part of the "time-event" definition, which divides time by the event.

Only a certain studies have examined environmental considerations; for example, study [20] focused solely on movement of people between fixed companies. Because of data's integrity and accessibility, we chose the GDP of annual growth rate as a function of the environmental considerations. Additionally, since data's are come from a technical social network, activity platform data from employees can be used as functions. Since the social platform cannot include data on time-varying turnover behaviours, the fourth category of function, structural features, is not included in this article.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 3844 – 3857
Received 05 March 2021; Accepted 01 April 2021.

Method Comparison

Following data cleaning, there are 287,229 job records. Label is set to land consider user to eventually quit the job based on begin and conclusion times of a experience work. The label is set 0 that is non-turnover and the user is in a work state if user only filled in the start time and left the end time blank. In a 7:3 ratio the data set is randomly divided into training and testing set. For assessing the model, we used Accuracy, Recall, F1-measure, and AUC as metrics. In TABLE 2 disorder matrix is given firstly. The calculations of other metrics are will be introduced.

Table 2. Confusion Matrix

Actual class	Predicted class		*Total*
	Turnover	*Non-turnover*	
Turnover	TP(True Positive)	FN(False Negative)	P
Non-turnover	FP(False Positive)	TN(True Negative)	N
Total	P	N	

This is primarily due to the DT model's poor generalisation ability compared to other ensemble methods, as well as its proclivity for over fitting.Furthermore, In terms of Accuracy, F1-measure, and AUC the CoxRF outperforms the competition.

Results Analysis

Gender: The feature important scores are given in Fig. 3, after by the calculated average value of score of Gini Index and normalizes them. We can see the gender and the profitable indication are two features to higher their scores up to 0.15, with gender being a category variable. Difference between groups is examined by the Kaplan-Meier technique to calculate survival rate and group the plot survival curves. The difference in genders are similar in shape of survival curves, as shown in Figure4.

However, if we look at the specifics of the curve, majority of time female survival rates are lower than male survival rates, implying that after working for the same number of years more females are like to leave a job. The female group of endurance curve are drop from 200-400 (month to event) (fig 3).

Industry: Our dataset also includes detailed information about the sectors in which employees work, they are 18 categories in total. This confirms our suspicion that job changers are mostly IT employees only, while government workers are more likely to stay in the government sectors only, the turnover rates of other companies reaches up to 20% (80% below survival probability) within 20 months after starting work, all show their similar drop patterns and their survival curves. It takes three years to achieve this level for the government sector. The big difference in decline rate after two years in survival curves of different industry groups. Within the first two years as for the education industry, the financial and cultural media industry is close the survival curve.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 3844 – 3857
Received 05 March 2021; Accepted 01 April 2021.

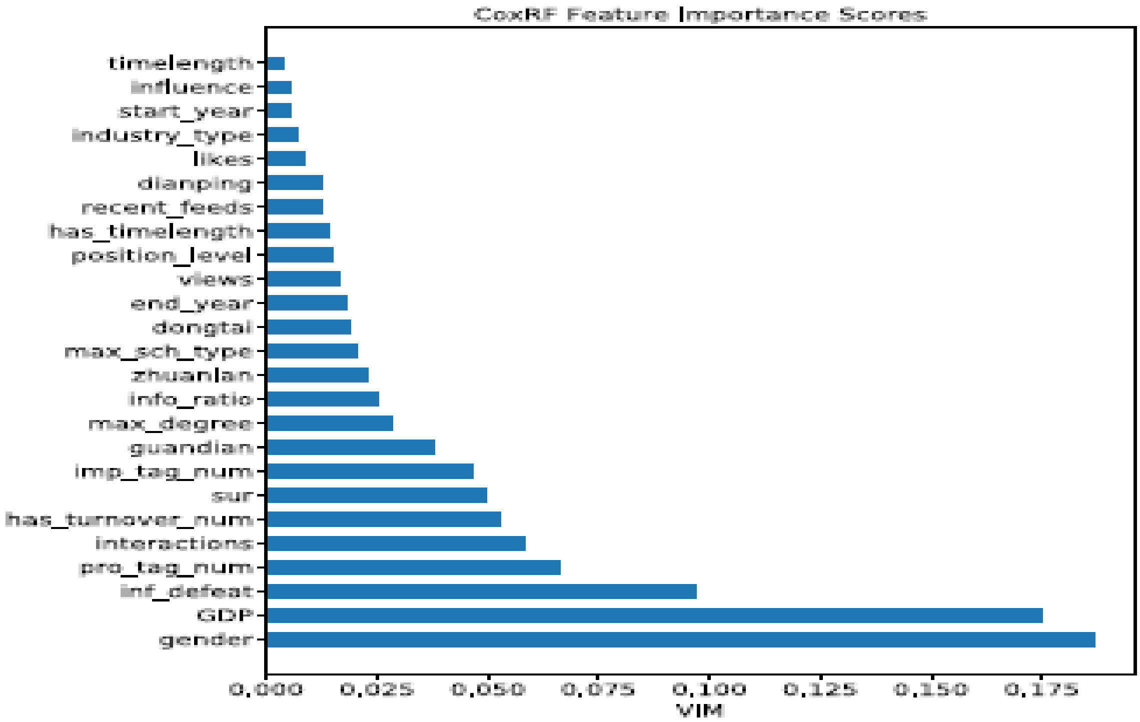

Fig 3. Feature Important Scores

Educational Background: We split three categories of employees based on their highest academic degree on the university level, i.e. project 985 university, project 211 university, and other, to investigate how educational background influences employee turnover behaviour.

Related work

We presented a CoxRF method to forecast employees changeover in this paper. The conventional survival analysis predict each person's rate of survival at a given point in period(time). The longest "time-event" determined by the time unit of "timed-event" in the data set determines the predictable time range, and the predictable time granularity. Survival analysis is a classic statistical model, as can be seen. Furthermore, in survival analysis the Cox model is a basic model, and linear model are not met when the assumptions are results ineffective. When evaluating the Concordance Index the researchers often use the survival analysis performance. Although it is different to AUC, and differently its values will be calculate. As a result, we cannot able to apply survival analysis directly to machine learning algorithms. The concept of hazard functions and survival includes the basic survival approach and as well as process the survival data on the censored data, it does not reply on hypothesis. Only time, events, covariates, etc are some data requirement organization. We conduct event-centred exploration because of these features. The Kaplan-Meier technique is for figure out the alive of people at any given time, and we can visualize each group's survival pattern. The correlation between characteristics and result are obtained by observing and different groups of survival curves are distinguishing between the dissimilarity or by the probable value can be obtained by directly testing, there is a small number of covariates and their categorization factors.

Conclusion

Based on the survival analysis, we suggested the CoxRF method for predicting employee turnover in this study. CoxRF combines random forest bagging ensemble learning with survival analysis. To calculate the survival rate the more events and passage are in this way.Meanwhile, our task has been transformed as a problem called standard supervised binary classifier and comparing with all other algorithms. We have also proposed the "time-event" and "event-person" concept for constructing data survival and maxi missing the usage of censored information. We discovered that sex(gender) had a significant impact on business turnover, in which the female candidates turnover being greater than male candidates turnover for the similar period of time(duration) worked. Furthermore, external factors influence turnover in decision-making. Furthermore, different industry groups have different turnover occurrence times, and the IT industry group's turnover rate is higher in comparison with government group's. After three to five years of employment, the turnover rate is higher than other schools and background of employee with good education. To studies of employee turnover behavior prediction, our findings adds a new viewpoints and different ideas.

References

[1] P. Ajit, “Prediction of employee turnover in organizations using machine learning algorithms,” *algorithms*, vol. 4, no. 5, p. C5, 2016.

[2] M. Fang, J. Su, J. Liu, Y. Long, R. He, and T. Wang, “A Model to Predict Employee Turnover Rate: Observing a Case Study of Chinese Enterprises,” *IEEE Systems, Man, and Cybernetics Magazine*, vol. 4, no. 4, pp. 38–48, 2018.

[3] Murugesan, M., Thilagamani, S. ,” Efficient anomaly detection in surveillance videos based on multi layer perception recurrent neural network”, Journal of Microprocessors and Microsystems, Volume 79, Issue November 2020, https://doi.org/10.1016/j.micpro.2020.103303

[4] J. Liu, Y. Long, M. Fang, R. He, T. Wang, and G. Chen, “Analyzing Employee Turnover Based on Job Skills.” ACM, 2018, pp. 16–21.

[5] Z. R. Mulla, K. Kelkar, M. Agarwal, S. Singh, and N. E. Sen, “Engineers’ voluntary turnover: application of survival analysis,” *The Indian Journal of Industrial Relations*, pp. 328–341, 2013.

[6] Thilagamani, S., Nandhakumar, C. .” Implementing green revolution for organic plant forming using KNN-classification technique”, International Journal of Advanced Science and Technology, Volume 29 , Isuue 7S, pp. 1707–1712

[7] Y. Yang, D.-C. Zhan, and Y. Jiang, “Which One Will be Next? An Analysis of Talent Demission,” 2018.

[8] Y. Zhao, M. K. Hryniewicki, F. Cheng, B. Fu, and X. Zhu, “Employee Turnover Prediction with Machine Learning: A Reliable Approach.” Springer, 2018, pp. 737–758.

[9] Thilagamani, S., Shanti, N.," Gaussian and gabor filter approach for object segmentation", Journal of Computing and Information Science in Engineering, 2014, 14(2), 021006, https://doi.org/10.1115/1.4026458

[10] I. W. Analytics, "SAMPLE DATA: HR Employee Attrition and Performance," 2015.

[11] J. Wang, Y. Zhang, C. Posse, and A. Bhasin, "Is it time for a career switch?" ACM, 2013, pp. 1377–1388.

[12] Rhagini, A., Thilagamani, S. ,"Women defence system for detecting interpersonal crimes", International Journal of Advanced Science and Technology, 2020, Volume 29,Issue7S, pp. 1669–1675.

[13] K. G. Bigsby, J. W. Ohlmann, and K. Zhao, "The turf is always greener: Predicting decommitments in college football recruiting using Twitter data," *Decision Support Systems*, vol. 116, pp. 1–12, 2019.

[14] J. Yuan, Q.-M. Zhang, J. Gao, L. Zhang, X.-S. Wan, X.-J. Yu, and T. Zhou, "Promotion and resignation in employee networks," *Physica A: Statistical Mechanics and its Applications*, vol. 444, pp. 442–447, 2016.

[15] A. C. C. de Jesus, M. E. G. Jnior, and W. C. Brando, "Exploiting linkedin to predict employee resignation likelihood." ACM, 2018, pp. 1764– 1771.

[16] K.Deepa, S.Thilagamani, "Segmentation Techniques for Overlapped Latent Fingerprint Matching", International Journal of Innovative Technology and Exploring Engineering (IJITEE), ISSN: 2278-3075, Volume-8 Issue-12, October 2019. DOI: 10.35940/ijitee.L2863.1081219

[17] P. D. Allison and G. Fan, *Event history and survival analysis*. Gezhi Press., 2017.

[18] Deepa, K., Kokila, M., Nandhini, A., Pavethra, A., Umadevi, M. "Rainfall prediction using CNN", International Journal of Advanced Science and Technology, 2020, 29(7 Special Issue), pp. 1623–1627. http://sersc.org/journals/index.php/IJAST/article/view/10849

[19] W. Wu, "Why Do Canadian Employees Quit? Results from Linked Employee-Employer Data," Ph.D. dissertation, University of Waterloo, 2012.

[20] Deepa. K , LekhaSree. R , Renuga Devi. B , Sadhana. V , Virgin Jenifer. S ,"Cervical Cancer Classification", International Journal of Emerging Trends in Engineering Research, 2020, 8(3), pp. 804–807 https://doi.org/10.30534/ijeter/2020/32832020

[21] Marchington, M.; Wilkinson, A.; Donnelly, R.; Kynighou, A. Human Resource Management at Work; Kogan Page Publishers: London, UK, 2016.

[22] Santhi, P., Mahalakshmi, G., Classification of magnetic resonance images using eight directions gray level co-occurrence matrix (8dglcm) based feature extraction, International Journal of Engineering and Advanced Technology, 2019, 8(4), pp. 839–846

[23] Deepak, K.D.; Guthrie, J.; Wright, P. Human Resource Management and Labor Productivity: Does Industry Matter? Acad. Manag. J. 2005, 48, 135–145.

[24] Santhi, P., Lavanya, S., Prediction of diabetes using neural networks, International Journal of Advanced Science and Technology, 2020, 29(7 Special Issue), pp. 1160–1168

[25] Santhi, P., Priyanka, T.,Smart India agricultural information reterival system, International Journal of Advanced Science and Technology, 2020, 29(7 Special Issue), pp. 1169–1175.

[26] Vijayakumar, P, Pandiaraja, P, Balamurugan, B & Karuppiah, M 2019, 'A Novel Performance enhancing Task Scheduling Algorithm for Cloud based E-Health Environment', International Journal of E-Health and Medical Communications , Vol 10,Issue 2,pp 102-117

[27] Nagadevara, V. Early Prediction of Employee Attrition in Software Companies-Application of Data Mining Techniques. Res. Pract. Hum. Resour. Manag. 2008, 16, 2020–2032.

[28] P. Pandiaraja, N Deepa 2019 ," A Novel Data Privacy-Preserving Protocol for Multi-data Users by using genetic algorithm" , Journal of Soft Computing , Springer , Volume 23 ,Issue 18, Pages 8539-8553

[29] Usha, P.; Balaji, N. Analysing Employee attrition using machine learning. Karpagam J. Comput. Sci. 2019, 13, 277–282.

[30] N Deepa , P. Pandiaraja, 2020 ," Hybrid Context Aware Recommendation System for E-Health Care by merkle hash tree from cloud using evolutionary algorithm" , Journal of Soft Computing , Springer , Volume 24 ,Issue 10, Pages 7149–7161

[31] Microsoft Docs: Team Data Science Process. Available online: https://docs.microsoft.com/it-it/azure/ machine-learning/team-data-science-process/ (accessed on 31 October 2020).

[32] IBM HR Analytics Employee. Available online: https://www.kaggle.com/pavansubhasht/ibm-hr-analytics- attrition-dataset (accessed on 31 October 2020).

[33] N Deepa , P. Pandiaraja, 2020 , " E health care data privacy preserving efficient file retrieval from the cloud service provider using attribute based file encryption ", Journal of Ambient Intelligence and Humanized Computing , Springer , https://doi.org/10.1007/s12652-020-01911-5

[34] Antecol, H.; Cobb-Clark, D. Racial harassment, job satisfaction, and intentions to remain in the military. J. Popul. Econ. 2009, 22, 713–738. [CrossRef]

[35] K Sumathi, P Pandiaraja 2019," Dynamic alternate buffer switching and congestion control in wireless multimedia sensor networks" , Journal of Peer-to-Peer Networking and Applications , Springer , Volume 13,Issue 6,Pages 2001-2010

[36] Theodossiou, I.; Zangelidis, A. Should I stay or should I go? The effect of gender, education and unemployment on labour market transitions. Labour Econ. 2009, 16, 566–577.

[37] Shankar, A., Pandiaraja, P., Sumathi, K., Stephan, T., Sharma, P. ," Privacy preserving E-voting cloud system based on ID based encryption " Journal of Peer-to-Peer Networking and Applications , Springer , https://doi.org/10.1007/s12083-020-00977-4

[38] Griffeth, R.W.; Hom, P.W.; Gaertner, S. A meta-analysis of antecedents and correlates of employee turnover: Update, moderator tests, and research implications for the next millennium. J. Manag. 2000, 26, 463–488.

[39] Cockburn, I.; Henderson, R.; Stern, S. The Impact of Artificial Intelligence on Innovation. In The Economics of Artificial Intelligence: An Agenda; University of Chicago Press: Chicago, IL, USA, 2019; pp. 115–146.

[40] Jarrahi, M. Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. Bus. Horiz. 2018, 61, 577–586.

[41] Sumathi, K., Naveena, K., Prashanth, P., Revanthkumar, S., Srikeerthanaa, A.P.,E-health based patient surveilling device, International Journal of Emerging Trends in Engineering Research, 2020, 8(3), pp. 792–796

[42] Paschek, D.; Luminosu, C.; Dra, A. Automated business process management-in times of digital transformation using machine learning or artificial intelligence. In MATEC Web of Conferences; EDP Sciences: Les Ulis, France, 2017; Volume 121.

[43] Varian, H. Artificial Intelligence, Economics, and Industrial Organization; National Bureau of Economic Research: Cambridge, MA, USA, 2018.

[44] Vardarlier, P.; Zafer, C. Use of Artificial Intelligence as Business Strategy in Recruitment Process and Social Perspective. In Digital Business Strategies in Blockchain Ecosystems; Springer: Berlin/Heidelberg, Germany, 2019; pp. 355–373.