

Determinates of Employee Voluntary Turnover and Forecasting in R&D Departments: A Case Study

Xiaojuan Zhu¹, Rapinder Sawhney¹, Girish Upreti¹

¹Department of Industrial and Systems Engineering, The University of Tennessee, Knoxville, TN 37996, USA

Correspondence: Girish Upreti, Department of Industrial and Systems Engineering, The University of Tennessee, Knoxville, TN 37996, USA

Received: June 2, 2016 Accepted: June 16, 2016 Online Published: July 25, 2016

doi:10.11114/set.v3i1.1635

URL: <http://dx.doi.org/10.11114/set.v3i1.1635>

Abstract

Researchers and human resource departments have focused on employee turnover for decades. This study examines employee voluntary turnover factors using logistic regression and forecasts employee tenure using a decision tree for four research and development departments in a large U.S organization. Company job title, gender, ethnicity, age and years of service significantly affect employee voluntary turnover behavior determined by logistic regression. The findings assist managers and human resource departments in specific employee retention strategies to reduce R&D departments' voluntary turnover rate. The decision tree method built a five-level depth tree model with 17 nodes. This model has the lowest AIC value and the best performance in the validation dataset. Age at hire, jobtitle, division, and race are statistically significant factors to predict employee tenure. The most important variable is age at hire located in the decision tree's first, third, and fourth nodes. Classification rules assist managers and human resource departments in quickly predicting employee tenure and in making hiring decisions.

Keywords: employee turnover, logistic regression, decision tree, R&D department

1. Introduction

Researchers and human resource managers have focused on employee turnover for decades because it has negatively affected organizations' performance (Shaw, 2011). Employee retention is one of the main challenges in organizations, especially for those with a long lead time to hire a new employee. In this study, besides the hiring and training time, the background security check takes a long time (>6 months) when filling positions. Our study's subjects were 731 terminated employees in four research and development departments across more than a ten-year study window. This study focuses on employee voluntary turnover, i.e., employees who voluntarily quit their job. Among the reasons for termination, voluntary turnover is one of the major ones, accounting for 26%. Compared to retirement and layoff, voluntary turnover is harder for companies to control. Because voluntary turnover is expensive for companies (Selden & Moynihan, 2000), they do not want their employees to voluntarily leave (Allen, Bryant, & Vardaman, 2010). Therefore, this study's objective is to examine what determines voluntary turnover for R&D department employees and to forecast employee tenure, thus assisting in hiring decisions.

Employee turnover represents a major loss of intellectual property to an organization. The total cost associated with employee turnover is estimated to range from 100-300% of a departing employee's annual salary (Moody, 2000). Both private firms and governmental organizations spend billions of dollars every year to manage employee turnover (Leonard, 2001). Furthermore, employee turnover can significantly reduce the production system's reliability (Ahmad Niknam & Sawhney, 2014; Sawhney, Subburaman, Sonntag, Rao Venkateswara Rao, & Capizzi, 2010). For organizations to execute their tasks efficiently and effectively with the highest quality, they must ensure that the right people are available at the right places and at the right times (Khoong, 1996).

Khan determined the factors affecting employee turnover in an organization of a developing country. Job security, financial crises and retirement benefits have significant effect on employee turnover (Khan, 2014). Similarly, in Ethiopia, Yohannes Melaku analyzed turnover data collected from surveys from an Ethiopian religious organization. Dissatisfaction with pay structure, mismanagement, unfair reward and promotion systems were some of the factors for employee turnover (YOHANNES, 2014). Jamie Liu interviewed a group of people who volunteered for retirement. The results show that 70% of respondents agreed that having chances to develop themselves and grow is more important

than a wage raise. However, only 52% of respondents would prefer getting recognition instead of a wage raise. In addition, women with children have a significant tendency to remain in their jobs, whereas their male counterparts are more likely to changing jobs (Lee Liu, 2014).

Many researchers have attempted to identify turnover factors to prevent and reduce turnover. (Bluedorn, 1982) reported that turnover is related to an individual's, routine, age, length of service, and perception of environmental opportunities. (Balfour & Neff, 1993) suggested that caseworkers with more education, less experience, and less stake in an organization are more likely to turnover. According to (Griffeth, Hom, & Gaertner, 2000), pay and pay-related variables modestly affect turnover. They also examined the relationship among pay, a person's performance, and turnover. They concluded that when high performers are inadequately rewarded, they quit. In our dataset, company job titles are highly and positively correlated with employees' salaries, sensitive information that cannot be released. We can hypothesize the following:

Hypothesis 1: The company job title is a significant predictor of voluntary turnover.

Besides job titles, we think employees' college majors are also a factor causing variation of employees' salaries. Thus, we hypothesize the following:

Hypothesis 2: The College major (branch of study) is a significant predictor of voluntary turnover.

(Griffeth, et al., 2000) study shows that employees' demographic characteristics such as gender and education level are also highly correlated to employee turnover. They also found that men are less likely to turnover than women. Thus, we hypothesize the following:

Hypothesis 3: Employee education level is a significant predictor employee voluntary turnover.

Hypothesis 4: Female employees are more likely to voluntary turnover than male employees.

Many researchers have examined the relationship between employees' age and turnover. Younger employees are more likely to move from one job to another (Burke, 1994), particularly those less than 35 years old (McShane & Von Glinow, 2003). Thus, we hypothesize the following:

Hypothesis 5: Age is a significant factor in employee voluntary turnover.

Employees are more likely to quit during the first five years before they have a strong commitment to the organization. (Bluedorn, 1982) found that turnover is related to length of service. Thus, we hypothesize the following:

Hypothesis 6: Length of service is a significant factor in employee voluntary turnover.

Furthermore, some races may react differently to other races. For example, (Thaden, Jacobs-Priebe, & Evans, 2010) found that employee turnover is significantly related to race. Ethnic minorities' language and accent barriers may cause social differences with the ethnic majority. Thus, we hypothesize the following:

Hypothesis 7: Ethnicity is a significant predictor of employee voluntary turnover.

Many researchers have used logistic regression to build models and to identify turnover's attributes. For example, (Balfour & Neff, 1993) applied logistic regression to build an employee turnover model. (Wright & Cropanzano, 1998) used correlation and logistic regression to examine whether emotional exhaustion is a predictor of turnover. (Morrow, McElroy, Laczniak, & Fenton, 1999) applied logistic regression to determine whether employees' absence and performance record can be employee-turnover indicators. (Nagadevara, Srinivasan, & Valk, 2008) applied several statistical methods including logistic regression to predict employee turnover.

To predict employee turnover behavior, many other statistical techniques (such as regression, neural network (NN), and data mining) have also been used. For example, (Ng, Cram, & Jenkins, 1991) used a proportional hazards regression (PHR) to develop a turnover prediction model. (Beng Ang, Tee Goh, & Chye Koh, 1994) developed a turnover prediction model for accountants in a Singapore organization. (Jenkins & Paul Thomlinson, 1992) used multiple regression to explore turnover intention. (Chang & Xi, 2009) combined Taguchi's method and Nearest Neighbor Classification Rules to select feature subsets and analyze factors to find the best predictor of employer turnover. (Alao & Adeyemo, 2013) used a decision tree to classify employees based on various types of attrition using employees' records. However, all these modeling methods have attempted to forecast employees' turnover behaviors rather than providing a forecast model or rules that forecast employee tenure and that also assist in hiring employees who are less likely to voluntarily quit.

This study has five parts. Part one is the introduction including the objective and literature review. Part two provides a synopsis of tools for data preparation and discusses forecasting methods. Part three presents results. Part four discusses the results followed by conclusion.

2. Methodology

2.1 Data and Preparation

In 2011 a large U.S. organization provided the human resource data for this study. The dataset contained 731 observations of terminated employee records from October 2000 to June 2011, including metrics such as ID, division (department), job title, termination reason, last hired date, termination date, years of service (YCS), gender, race, age at hiring, age at termination, highest education degree, and branch of study. Eight variables are created from the datasets for analysis listed under results.

In addition, two variables were created from the dataset for analysis. One variable was major, including 15 levels. The first three letters of the branch of study are used for identification; for example, Eng refers to engineering, and Che refers to chemistry. The other was a binary target variable Y for logistic regression. Y is equal to 1 if an employee voluntarily quits; otherwise, y is equal to 0 if an employee leaves the organization for other reasons.

2.2 Logistic regression

The hypotheses were tested using logistic regression to identify voluntary turnover's significant predictors. Logistic regression, or logit regression, is a popular model for binary data (Agresti, 1996). It is used to predict a binary response from a binary predictor, to predict a categorical variable's outcome based on one or more predictor variables, and to estimate a qualitative response model's parameters. Unlike linear regression, logistic regression can handle both categorical and numeric variables. The logistic regression is expressed in Equation 1:

$$\text{logit}[\pi(x)] = \log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \alpha + \beta x \quad (1)$$

where x represents the explanatory variables; $\pi(x)$ denotes the target variable's probability at value x ; $\pi(x)/(1 - \pi(x))$ is the odds of x ; and α, β are the estimated coefficients. All the predictors were tested using 0.05 as a p value criteria. The other variables, such as department, were also included in the model to eliminate the other variables' variance. Three statistics criteria were used to evaluate the logistic regression model: Akaike's information (AIC), Schwartz's Bayesian criterion (SBC), and the Hosmer-Lemeshow Goodness of Fit Test. AIC and SBC are both information criteria using likelihood value. Usually, the best model comes with the lowest AIC or SBC values. The Hosmer-Lemeshow test is a chi-square test to determine whether the logistic regression model is correctly specified (Hosmer Jr, Lemeshow, & Sturdivant, 2013). The high p-value ($P > 0.05$) indicates that this model passes the test.

2.3 Decision tree

A decision tree, also called a classification tree, is one method of classification via an intermediate tree-like structure in data mining (Hand, Mannila, & Smyth, 2001). A decision tree's purpose is to develop a classification rule from the data based on attributes, or explanatory variables. Similar to logistic regression, a decision tree can handle categorical or numeric variables. Creating a decision tree involves splitting the significant variable until only one unique classification is on each branch of the tree. The decision tree was used for data compression and prediction.

In this study, YCS was used as a target variable generating a classification rule for managers and human resource departments to forecast how many years an employee will work in the organization. It also provides managers a tool for making accurate hiring decisions. The turnover dataset was divided into two parts: training (60%) used to build a decision tree and validation (40%) used to validate the model performance. The decision tree model was built using AIC values as model selection criteria. The decision tree was analyzed using SAS Enterprise Miner 7.

3. Results

3.1 Hypothesis Test Results

All variables proposed in the hypothesis were tested with other control variables using the logistic regression model- Jobtitle, Division, Gender, Degree, Major, Race, Age, and Length of service. The parameter estimations are shown in Table 1 and Table 2. The variables included *Div* (division), *jobtitle* (job title), *race*, *gender*, *degree* (highest education degree), *major* (branch of study), *age* (age at termination), and YCS (years of company service). Five variables were statistically significant ($P < 0.05$) in the logistic regression model: *jobtitle*, *gender*, *race*, *age* and YCS. This finding indicates that employees voluntarily quitting the organization are statistically affected by their company job title, gender, race, age, and years of company service.

Table 1. Logistic Regression Parameter Estimates

Parameter	Level	DF	Estimate	Std. Error	Wald Chi-sq.	P-value
Intercept		1	2.441	0.803	9.247	0.002
Dir	BS	1	-0.106	0.595	0.032	0.858
Dir	CS	1	0.677	0.57	1.409	0.235
Dir	ES1	1	-0.756	0.552	1.877	0.171
Jobclus	1	1	-0.837	0.45	3.454	0.063
Jobclus	2	1	-1.564	0.358	19.108	<.001
Jobclus	3	1	-1.039	0.348	8.902	0.003
Gender	F	1	0.611	0.22	7.714	0.006
Majorclus	1	1	-3.866	0.623	38.468	<.001
Majorclus	2	1	-2.119	0.601	12.426	0.001
Majorclus	3	1	-2.255	0.571	15.574	<.001

The results of the hypotheses tests are discussed below.

Hypothesis 1: The employees' job title is a significant predictor of voluntary turnover.

The company job title is a pivotal factor in voluntary turnover since it is highly associated with employee salary ($\chi^2=37.2$, $df=14$, $P<0.05$). We have no evidence to reject the null hypothesis. Compared to the reference group (Managers), four job title groups are statistically significantly different than the manager group. This company's job titles are Administrator 1, Administrator 2, Administrator 3, Staff Level 1, and Technician; their coefficient estimates are -6.7 -2.9, -3.6, -1.8 and -3.0, respectively. These results show that compared with managers, employees in these positions are less likely to quit: Administrator 1 has 99.9% ($1-\exp(-6.7)$) less odds; Administrator 2 has 94.5% less odds; Administrator 3 has 97.3% less odds; Staff 1 has 83.4% less odds; and Technician has 95% less odds. The other job titles are not statistically significant, indicating they have the same turnover odds as managers.

Hypothesis 2: The college major (branch of study) is a significant predictor of voluntary turnover.

According to Table 1, employees' major has 9.7 of Wald χ^2 with 14 degrees of freedom, and P value is 0.78. We have evidence to reject the null hypothesis. Employees' college major (branch of study) does not significantly affect employee turnover behaviors in R&D departments.

Hypothesis 3: Employee education level is a significant predictor of employee voluntary turnover.

In this study, an education degree is not significantly correlated with employee voluntary turnover ($\chi^2=2.25$, $df=5$, $P=0.81$). We have evidence to reject the null hypothesis, thus showing that employees with a higher education degree have voluntary turnover behavior similar to those without a higher education degree.

Hypothesis 4: Female employees are more likely to voluntarily turnover than male employees.

We do not have evidence to reject the null hypothesis. Gender is also significant ($\chi^2=10.7$, $df=1$, $P<0.05$), indicating that female and male employees have different voluntary quit behaviors, as the parameter estimation for female is 1.29. Also, the odds of female voluntary turnover are 3.6 times that of male employees, indicating female employees are more likely to voluntarily quit.

Hypothesis 5: Age is a significant factor in employee voluntary turnover.

We do not have evidence to reject this null hypothesis. Employees' age significantly affects voluntary turnover ($\chi^2=38.8$, $df=1$, $P<0.05$). The coefficient estimate is -0.13 and the odds ratio is 0.87, indicating that the turnover odds decrease 13% when an employee's age increases one year older. As a result, younger employees tend to be more likely to quit than older employees.

Table 2. Logistic regression parameter estimates

Parameter	Levels	DF	Estimate	Standard Error	Wald χ^2	Pr > χ^2
Intercept		1	10.605	2.039	27.048	<.0001
Job title	Administrator1	1	-5.710	1.589	12.912	0.001
Job title	Administrator2	1	-2.157	1.166	3.425	0.064
Job title	Administrator3	1	-2.892	1.155	6.265	0.012
Job title	Administrator4	1	-0.042	1.089	0.002	0.969
Job title	Staff 1	1	-1.788	0.966	3.428	0.064
Job title	Staff 2	1	-0.688	0.803	0.733	0.392
Job title	Staff 3	1	0.065	0.667	0.009	0.923
Job title	Staff 4	1	0.877	0.688	1.622	0.203
Job title	Staff Member1	1	-1.479	1.216	1.479	0.224
Job title	Staff Member2	1	0.935	0.983	0.903	0.342
Job title	Staff Member3	1	-0.683	0.793	0.741	0.389
Job title	Staff Member4	1	0.907	0.862	1.108	0.293
Job title	Technician	1	-2.652	0.992	7.146	0.008
Job title	Others	1	-11.141	820.7	0.000	0.989
Gender	F	1	1.293	0.395	10.720	0.001
Race	Asian	1	-1.489	0.632	5.545	0.019
Race	Black or African American	1	-3.917	0.733	28.595	<.0001
Race	Hispanic/Latino	1	-1.495	1.068	1.959	0.162
Race	Native American/Alaskan	1	12.905	1679.0	0.000	0.994
Age		1	-0.129	0.021	38.812	<.0001
YCS		1	-0.149	0.022	45.618	<.0001

Hypothesis 6: Length of service is a significant factor in employee voluntary turnover.

We do not have evidence to reject this null hypothesis. Employees' years of service significantly affect voluntary turnover ($\chi^2=45.6$, $df=1$, $P<0.05$). The coefficient estimate is -0.15 and the odds ratio is 0.86, indicating that the turnover odds decrease 14% with each additional year of service. Thus, employees with fewer years of service tend to be more likely to quit.

Hypothesis 7: Ethnicity is a significant predictor of employee voluntary turnover.

We do not have evidence to reject this null hypothesis. Ethnicity has a statistical impact on employee voluntary turnover behavior ($\chi^2=30.6$, $df=4$, $P<0.05$). Asian, Black or African American, and Hispanic/Latino are significantly different statistically from white employees, the reference group, in terms of turnover because these three groups have negative coefficient estimates. Black or African American employees have the lowest coefficient estimates (-3.7) and odds ratio ($0.024 = \exp(-3.7)$). Asian and Hispanic/Latino employees have -1.3 and -1.5 of coefficients and 0.27 and 0.22 of odds ratio, respectively. These three groups have lower turnover probability than white employees. However, Native American employees are not significantly different statistically from white employees ($P>0.05$), showing that Native American employees have the same turnover probability as white employees.

3.2 Decision Tree Analysis Results

This study's decision tree model is a five-level depth tree with 17 nodes and has the lowest AIC value (Figure 1).

Four variables are statistically significant in the model: *ageh* (age at hire), *jobtitle* (job title), *div* (division), and *race*. The most important variable is *ageh* located in the decision tree's first, third, and fourth nodes, indicating that an employee's YCS is mainly determined by age when hired. Two other variables—job title and division—also significantly affect an employee's YCS. Based on the model, DIV1 is significantly different from the other three departments (DIV 2, DIV3, and DIV4). The employees with the following job titles are significantly different from those with other job titles: Staff 2, Staff Member 1, Administrator 2, or Administrator 1. Compared to other ethnicities, white employees have more YCS if they are younger than 28 when hired.

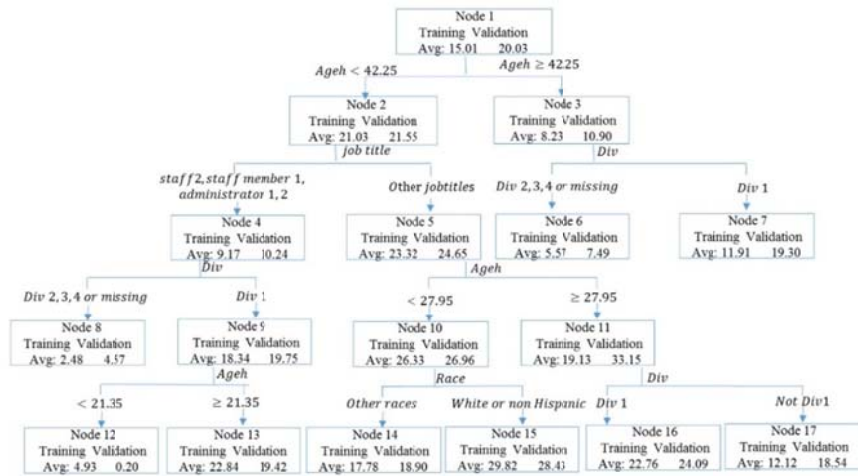


Figure 1. R&D Employee YCS Decision Tree Model

The decision tree’s rules are the following:

- if $ageh \geq 42.25$ and $div = DIV1$, then $E(YCS) = 11.9$;
- if $ageh \geq 42.25$ and $div \neq DIV1$, then $E(YCS) = 5.6$;
- if $21.35 \leq ageh < 42.25$, $jobtitle = Staff\ 2, Staff\ Member\ 1, Administrator\ 2, or\ Administrator\ 1$, and $div = DIV1$, then $E(YCS) = 22.9$;
- if $ageh < 42.25$, $jobtitle = Staff\ 2, Staff\ Member\ 1, Administrator\ 2, or\ Administrator\ 1$, $div = DIV1$, and $ageh < 21.35$, then $E(YCS) = 4.9$;
- if $ageh < 42.25$, $jobtitle = Staff\ 2, Staff\ Member\ 1, Administrator\ 2, or\ Administrator\ 1$, and $div \neq DIV1$, then $E(YCS) = 2.5$;
- if $27.95 \leq ageh < 42.25$, $jobtitle \neq Staff\ 2, Staff\ Member\ 1, Administrator\ 2, or\ Administrator\ 1$, and $div = DIV2, DIV3 or\ DIV4$ then $E(YCS) = 12.1$;
- if $27.95 \leq ageh < 42.25$, $jobtitle \neq Staff\ 2, Staff\ Member\ 1, Administrator\ 2, or\ Administrator\ 1$, and $div \neq DIV2, DIV3 or\ DIV4$ then $E(YCS) = 22.8$;
- if $ageh < 42.25$, $jobtitle \neq Staff\ 2, Staff\ Member\ 1, Administrator\ 2, or\ Administrator\ 1$, $ageh < 27.95$, and $race = white$ or $white/non-Hispanic$ origin, then $E(YCS) = 29.8$;
- if $ageh < 42.25$, $jobtitle \neq Staff\ 2, Staff\ Member\ 1, Administrator\ 2, or\ Administrator\ 1$, $ageh < 27.95$, and $race \neq white$ or $white/non-Hispanic$ origin, then $E(YCS) = 17.8$;

where, $E(YCS)$ denotes expected or average YCS. The predicted YCS’s range is from 2.5 to 29.8 years. The model’s performance is validated by predicting $E(YCS)$ in the validation dataset. As shown in Figure 2, the predicted values are close to the actual values in each leaf, indicating the decision tree model has strong forecasting ability. After the decision rule is determined, predicting YCS for a new employee is easily calculated by following the rules. This ease of calculation is one of the decision tree method’s advantages.

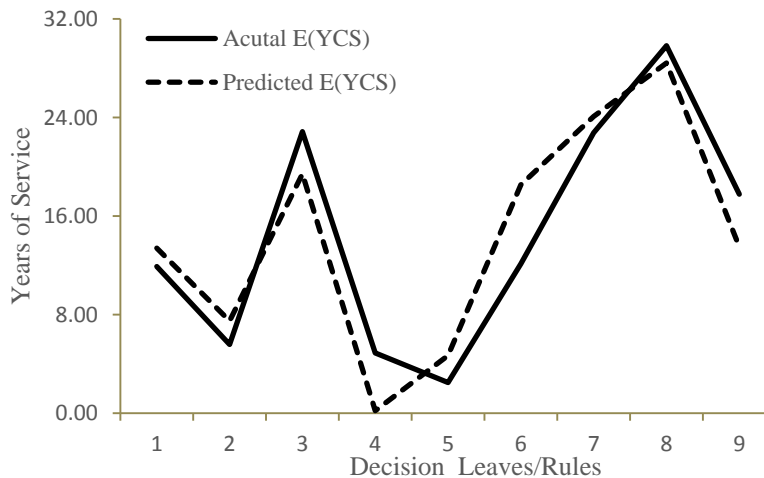


Figure 2. Actual vs. Predicted E(YCS) for validation dataset

4. Discussions

This study proposes investigation of factors contributing to R&D employee voluntary turnover and to support employee hiring and retention strategies. The variables proposed in the hypothesis were tested with other control variables using the logistic regression model to determine the probability of an employee's voluntary turnover. Five variables identified are statistically significant ($P < 0.05$): job title, gender, ethnicity, age, and years of service. However, an employee's division, education level, and major are not associated with employee turnover.

Although we found that employee job title is a significant predictor of employee turnover, the comparison results show that employees with lower salaries are more likely to stay. This finding is different from the previous study. Compared to the other staff levels, level 1 staff are also more likely to stay because they do not have much work experience. Also, their average salary is 20% higher than the job market's average. On the other hand, the other staff levels have the same probability to quit as the reference group (managers). According to (Tuji, 2013), when employees find no opportunities for advancement within the system, they will not remain in the work situation. The promotion competition is intense among R&D staff members with the same education degree and experience. However, they easily find higher paying positions in other companies or become members of faculty or research staff in organizations when they have several years of work experience. Therefore, R&D staff are more likely to quit when they are not in staff level 1.

Furthermore, female employees are more likely to quit than male employees. According to the work/family life-balance theory (Clark, 2000), as economic pressure increases, an employee may have to do the work of more than one person. Especially when an organization downsizes or restructures, the same amount of work has to be done by fewer employees. Thus, employees spend more time on their work life and have less time for their personal life (Smith, 2009). Female employees serve a more important role in the family than male employees. Forced to make a choice between family and work, female employees are more likely to quit.

Many studies have proved that age has a negative relationship with turnover (Rhodes, 1983). Researchers believe younger workers leave for two main reasons: lack of training opportunities and lack of mentors in the workplace (Paul, 2012). Although many managers are promoted because of their strong capability and achievement, they are not necessarily good coaches or team leaders and are unable to help employees improve their performance (Smith, 2009). Furthermore, R&D employees work more independently than employees in other departments. For example, sometimes one employee is responsible for one project or part of a project. Compared to experienced staff, younger staff feel frustrated and stressed when they meet difficulties or barriers and cannot receive adequate advice. As a result, younger employees are more likely to voluntarily leave.

The number of years an employee has worked in the organization is another significant factor in turnover. Many studies have found that length of service is highly related to organization commitment (Jena, 2015; Kelarijani, Heidarian, Jamshidi, & Khorshidi, 2014; Popoola, 2006). The longer the employee works, the more that employee feels attached to and involved in the organization. As a result, employees with longer length of service are more likely to stay. On the other hand, employees with few years of service have not formed a strong commitment to the organization. Thus, they tend to switch to a new job to increase their work experience, receive higher pay, or be in a higher position.

Ethnicity is another significant factor affecting employee turnover. According to our analysis, Black or African American employees have the lowest probabilities of quitting compared to the other ethnic groups. Asian and Hispanic

employees are also less likely to quit than the white employees in the study. However, Native American/Alaskan employees have the same turnover probability as white employees.

Employee turnover does not correlate with majors as shown in Table 1. Although some majors like computer science or engineering have more job openings, they do not have a higher turnover probability in R&D departments because employees have 20% higher salaries than the job market's average levels, according to our study.

Education degree is not a significant factor in voluntary turnover. Most employees (80%) in the four R&D departments are highly educated with at least a bachelor's degree. However, technicians and administrators are more likely to have a high school or an associate's degree. As discussed above, they are all less likely to quit. Therefore, determining which employees with which degree are more likely to quit is difficult.

The decision tree model predicts the average employee's working years in the R&D department under all types of employee turnover. The significant variables—age when an employee was hired, job title, division, and ethnicity—are easily obtainable. Based on the decision rules, managers can quickly determine whether an interviewee will have a long or short tenure. Based on the rules shown in section 4.2, a new employee's years of service can be predicted. If a new employee is applied to rules (2), (4), and (5), that employee's predicted YCS will be fewer than 10 years. If a new employee is applied to rules (1), (6), and (9), the predicted YCS will be 10 to 20 years. If a new employee is applied to rules (3), (7), and (8), the predicted YCS will be above 20 years.

5. Conclusions and Recommendations

In this study, employee turnover's significant attributes and classification rules are identified based on employee records from four R&D departments. Eight variables in the hypothesis are tested in the logistic regression model to find the variables' significance. This logistic regression determines the probability of an employee's voluntary turnover and identifies five significant factors affecting voluntary turnover: job title, gender, ethnicity, age, and years of service. These results can assist managers and human resource departments in developing employee-retention strategies to reduce R&D departments' voluntary turnover rate. As an example, employee with higher titles are more likely to quit. So the organization needs to increase benefits for higher ranks professionals.

The decision tree generates nine rules to predict the average length of YCS of an employee. The models' results show that combining models is suitable for forecasting employee turnover. Applications of the models can be used with hiring strategies. For example, when the data and related variables are accessible, a decision tree can generate a decision rule, such as hiring or not hiring an applicant. Results from the decision tree indicate that the age of employees when hired is the most important variable.

Several statistical software packages are available to conduct logistic regression and decision tree models, such as SAS, SPSS. If human resource departments have limited budgets, they can also use statistical software R, which is an open source. These two models are created based on the termination dataset with limited variables. If more data and variables can be identified, the models can be further improved. Furthermore, this study will be improved by an interview or a survey from the current active employees in the future study.

References

- Agresti, A. (1996). *An Introduction to Categorical Data Analysis* (Vol. 135): Wiley New York.
- Ahmad, N. S., & Sawhney, R. (2014). A model for reliability analysis of multi-state manufacturing systems. *International Journal of Quality & Reliability Management*, 31(8), 938-949. <http://dx.doi.org/10.1108/IJQRM-05-2012-0065>
- Alao, D., & Adeyemo, A. (2013). Analyzing Employee Attrition Using Decision Tree Algorithms *Computing, Information Systems, Development Informatics and Allied Research Journal*, 4(1), 17-28.
- Allen, D. G., Bryant, P. C., & Vardaman, J. M. (2010). Retaining Talent: Replacing Misconceptions with Evidence-based Strategies. *The Academy of Management Perspectives*, 24(2), 48-64. <http://dx.doi.org/10.5465/AMP.2010.51827775>
- Balfour, D. L., & Neff, D. M. (1993). Predicting and Managing Turnover in Human Service Agencies: A Case Study of an Organization in Crisis. *Public Personnel Management*, 22(3), 473-486. <http://dx.doi.org/10.1177/009102609302200310>
- Beng, A. K., Tee, G. C., & Chye, K. H. (1994). An Employee Turnover Prediction Model: A Study of Accountants in Singapore. *Asian Review of Accounting*, 2(1), 121-138. <http://dx.doi.org/10.1108/eb060649>
- Bluedorn, A. C. (1982). A Unified Model of Turnover from Organizations. *Human relations*, 35(2), 135-153. <http://dx.doi.org/10.1177/001872678203500204>
- Burke, R. J. (1994). Generation X: Measures, Sex and Age Differences. *Psychological Reports*, 74(2), 555-562.

<http://dx.doi.org/10.2466/pr0.1994.74.2.555>

- Chang, H. Y., & Xi, L. (2009). *Employee Turnover: A Novel Prediction Solution with Effective Feature Selection*. Paper presented at the WSEAS International Conference. Proceedings. Mathematics and Computers in Science and Engineering.
- Clark, S. C. (2000). Work/family border theory: A new theory of work/family balance. *Human relations*, 53(6), 747-770. <http://dx.doi.org/10.1177/0018726700536001>
- Griffeth, R. W., Hom, P. W., & Gaertner, S. (2000). A meta-analysis of antecedents and correlates of employee turnover: Update, moderator tests, and research implications for the next millennium. *Journal of management*, 26(3), 463-488. <http://dx.doi.org/10.1177/014920630002600305>
- Hand, D. J., Mannila, H., & Smyth, P. (2001). *Principles of Data Mining*: MIT press.
- Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Logistic Regression Models for Multinomial and Ordinal Outcomes. *Applied Logistic Regression, Third Edition*, 269-311. <http://dx.doi.org/10.1002/9781118548387.ch8>
- Jena, R. (2015). An assessment of factors affecting organizational commitment among shift workers in India. *Management: Journal of Contemporary Management Issues*, 20(1), 59-77.
- Jenkins, M., & Paul, T. R. (1992). Organisational Commitment and Job Satisfaction as Predictors of Employee Turnover Intentions. *Management Research News*, 15(10), 18-22. <http://dx.doi.org/10.1108/eb028263>
- Kelarijani, S. E. J., Heidarian, A. R., Jamshidi, R., & Khorshidi, M. (2014). Length of service and commitment of nurses in hospitals of Social Security Organization (SSO) in Tehran. *Caspian journal of internal medicine*, 5(2), 94.
- Khan, S. A. (2014). The factors affecting employee turnover in an organization: The case of Overseas Pakistanis foundation. *African Journal of Business Management*, 8(1), 25. <http://dx.doi.org/10.5897/AJBM2013.2508>
- Khoong, C. (1996). An Integrated System Framework and Analysis Methodology for Manpower Planning. *International Journal of Manpower*, 17(1), 26-46. <http://dx.doi.org/10.1108/01437729610110602>
- Lee, L. J. (2014). Main causes of voluntary employee turnover a study of factors and their relationship with expectations and preferences.
- Leonard, B. (2001). Turnover at the Top. *HR Magazine*, 46-52.
- McShane, S. L., & Von Glinow, M. A. Y. (2003). *Organizational Behavior: Emerging Realities for the Workplace Revolution*: McGraw-Hill/Irwin New York, NY.
- Moody, R. W. (2000). GOING, GOING, GONE By Reporting to Management on the Causes and Aevastating Effects of Employee Turnover, Internal Auditors May Help to Curb the Costs. *Internal Auditor*, 57(3), 36-41.
- Morrow, P. C., McElroy, J. C., Laczniak, K. S., & Fenton, J. B. (1999). Using Absenteeism and Performance to Predict Employee Turnover: Early Detection through Company Records. *Journal of Vocational Behavior*, 55(3), 358-374. <http://dx.doi.org/10.1006/jvbe.1999.1687>
- Nagadevara, V., Srinivasan, V., & Valk, R. (2008). Establishing a Link between Employee Turnover and Withdrawal Behaviours: Application of Data Mining Techniques [Article]. *Research & Practice in Human Resource Management*, 16(2), 81-99.
- Ng, S. H., Cram, F., & Jenkins, L. (1991). A Proportional Hazards Regression Analysis of Employee Turnover among Nurses in New Zealand. *Human Relations*, 44(12), 1313-1330. <http://dx.doi.org/10.1177/001872679104401205>
- Paul, A. M. (2012). This Is The Biggest Reason Talented Young Employees Quit Their Jobs. *Careers*, 2016, from http://mobile.businessinsider.com/why-young-employees-quit-their-jobs-2012-9?comments_page=2
- Popoola, S. (2006). Personal factors affecting organizational commitment of records management personnel in Nigerian State Universities. *Ife Psychologia*, 14(1), 183-197. <http://dx.doi.org/10.4314/ifep.v14i1.23702>
- Rhodes, S. R. (1983). Age-related differences in work attitudes and behavior: A review and conceptual analysis. *Psychological bulletin*, 93(2), 328. <http://dx.doi.org/10.1037/0033-2909.93.2.328>
- Sawhney, R., Subburaman, K., Sonntag, C., Rao, V., Rao, P., & Capizzi, C. (2010). A modified FMEA approach to enhance reliability of lean systems. *International Journal of Quality & Reliability Management*, 27(7), 832-855. <http://dx.doi.org/10.1108/02656711011062417>
- Selden, S. C., & Moynihan, D. P. (2000). A Model of Voluntary Turnover in State Government. *Review of Public Personnel Administration*, 20(2), 63-74. <http://dx.doi.org/10.1177/0734371X0002000206>
- Shaw, J. D. (2011). Turnover rates and organizational performance Review, critique, and research agenda.

Organizational Psychology Review, 1(3), 187-213. <http://dx.doi.org/10.1177/2041386610382152>

Smith, J. L. (2009). 12 Reasons Employees Leave Organizations. *InterBusiness Issues*. Retrieved 3-30, 2016

Thaden, E., Jacobs-Priebe, L., & Evans, S. (2010). Understanding Attrition and Predicting Employment Durations of Former Staff in a Public Social Service Organization. *Journal of Social Work*, 10(4), 407-435. <http://dx.doi.org/10.1177/1468017310369606>

Tuji, A. (2013). *An Assessment of the Causes of Employee Turnover in Oromia Public Service Organizations*. Master, Addis Ababa University. http://www.academia.edu/7912727/An_Assessment_of_the_Causes_of_Employee_Turnover.

Wright, T. A., & Cropanzano, R. (1998). Emotional Exhaustion as a Predictor of Job Performance and Voluntary Turnover. *Journal of Applied Psychology*, 83(3), 486. <http://dx.doi.org/10.1037/0021-9010.83.3.486>

Yohannes, M. (2014). *Factors Affecting Employee Turnover And Its Impact On Ethiopian Evangelical Church Mekane Yesus*. MA, Addis Ababa University, Addis Ababa, Ethiopia.



This work is licensed under a [Creative Commons Attribution 3.0 License](https://creativecommons.org/licenses/by/3.0/).