Applications in Predictive Analytics:

Predicting Employee Attrition

Tamara Williams

Northwestern University

Applications in Predictive Analytics:

Predicting Employee Attrition

In November of 2016 3.1 million people, or 2.1% of the workforce, quit their jobs (U.S. Dept. of Labor, 2016). During the first 11 months of 2016, roughly 32.5 million people voluntarily left their employers. The estimated the cost of replacing an employee ranges from $3,372 for an hourly retail employee to $260,000 for a senior-level executive (Boushey & Glynn, 2012). Employee attrition, or voluntary separation, is clearly a concern to business and impacts company finances. Research into employee satisfaction is not new, but the rise of analytics permits researchers to explore ways to predict attrition, rather than merely define what causes it.

Hewlett-Packard has assigned an attrition prediction to every employee (Siegel, 2016). Wal-Mart, Credit Suisse, and Box Inc. also have prediction initiatives around employee attrition (Silverman & Waller, 2015). Google has similar metrics (Morrison, 2009). Specifics on how each of these companies are modeling their attrition predictions do not appear in a literature search. However, we can gain insight into the probable methods used by looking at published research on the topic.

Methods Used

Research approaches for predicting attrition differ. Saradhi and Palshikar (2011) draw an analogy between employee and customer retention. Their work evaluates using customer retention prediction techniques to predict employee retention. Their findings show machine learning models can predict employee attrition. Mossholder, Settoon, and Henagan (2005) use employees' relationships to model attrition risk. They could predict attrition based on aspects of at-work relationships. Hom and Xiao added family connections to work relationships in their model. They conclude the strength of aggregate social ties are predictors of employee retention.

Data Collection Methods

Internal corporate data. Saradhi and Palshikar (2011) elected to work with a subset of available employee work history data. Employee selection for inclusion in the data set was based on an employee value model (EVM). The EVM weights employees based on their assigned projects and the number of months billed. The study population is those employees on important projects with a high number of billed months.

Surveys. Mossholder et al. (2005) used a combination of self-reported survey ratings (1–5 scale), social relationship and demographic questions to create their data set. Hom and Xiao (2011) similarly used employee survey data for their analysis.

Data Analysis Methods

Supervised Learning Models (Machine Learning). Sarahdhi and Palshikar (2011) used supervised learning to perform classification on the employee data. They found support vector machines (SVM), random forests and logistic regression were the most accurate for predicting customer churn. Their research showed that SVM gave the best predictions of employee churn for their data set. The article lacked details which attributes which effected model performance.

Cox Regression Analysis. Mossholder et al. (2005) used the Cox regression method to examine the effect of multiple variables, in this case, types of relationships, on the outcome of employee attrition. They found coworker support, a sense of obligation to coworkers, and size workplace network can predict employee turnover. Hom and Xiao (2011) also used Cox regression. Their study focused on the effect of community, family and workplace ties in predicting attrition. They conclude strong ties in all areas bolster job satisfaction and are predictors of employee retention.

Contribution to Management

There are three primary ways predicting employee attrition can assist management in making choices which have positive business value. First, by identifying high-value employees at risk of quitting, management can attempt to "re-recruit" them to back into the company. Second, by identifying poor performers at risk of flight, management can make an informed decision on whether they choose to invest in that employee. Management may elect to spend little time addressing the poor performance, may encourage the employee to leave, or may outright sever them. Finally, to identify the high-value employees, you must be able to define who they are and the characteristics that make them the most valuable. Roberts (2015) found that the definitions for good and poor performance, along with key performance indicators (KPIs) can be combined into predictive metrics regarding performance. She observes that predictive performance metrics can thus, inform the recruiting function. Ability to predict performance allows for better screening of key success attributes.

Knowing who is a flight risk has one other potential benefit to management. If the at-risk employee is key to the business, succession planning can be initiated to develop a mitigation strategy, should efforts to retain the employee fail. The ability to minimize disruption to the team, caused by the departure of a key employee, has real, dollar value to an organization. For management, retaining your best employees, mitigating the investment in poor performers, and being able to improve the recruiting pipeline based on known success predictors, all have a positive impact on the bottom line.

References

Boursey, H., Glynn, S.J. (2012). *There are Significant Business Costs to Replacing Employees*, Center for American Progress. Retrieved from https://www.americanprogress.org/wp-content/uploads/2012/11/CostofTurnover.pdf.

Hom, Peter W., & Xiao, Zhixing (2011). Embedding social networks: How guanxi ties reinforce Chinese employees' retention, *Organizational Behavior and Human Decision Processes*, Volume 116 (issue 2), pp 188-202, ISSN 0749-5978, doi: 10.1016/j.obhdp.2011.06.001.

Morrison, S. (2009, May 19). Google searches for staffing answers. *The Wall Street Journal.* Retrieved from http://www.wsj.com.

Mossholder, K. W., Settoon, R. P., & Henagan, S. C. (2005, August). A Relational Perspective On Turnover: Examining Structural, Attitudinal, And Behavioral Predictors. *Academy Of Management Journal*, Volume 48 (Issue 4), pp. 607-618. doi:10.5465/AMJ.2005.17843941.

Roberts, G. (2015, May 21) Reducing the costs of employee churn with predictive analytics. *Predictive Analytics Times*. Retrieved from http://www.predictiveanalyticsworld.com /patimes/reducing-the-costs-of-employee-churn-with-predictive-analytics-0521151/5398/.

Saradhi, V., & Palshikar, G. (2011, March), Employee churn prediction. *Expert Systems with Applications*, Volume 38 (Issue 3), pp. 1999–2006. doi:10.1016/j.eswa.2010.07.134.

Siegel, E. (2016), "With Power Comes Responsibility Hewlett-Packard, Target, the Cops, and the NSA Deduce Your Secrets*",* pp 46 – 101, *Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die, Revised and Updated*, Hoboken, New Jersey, John Wiley & Sons, Inc.

Silverman, R. E., & Waller, N. (2015, March 13). The algorithm that tells the boss who might quit. *The Wall Street Journal*. Retrieved from http://www.wsj.com

United States Department of Labor, Bureau of Labor Statistic (2017, January 10*). Job Openings and Labor Turnover Summary*. Retrieved from https://www.bls.gov/news.release/jolts.nr0.htm