A Workforce Management Turnover Model for Retail Human Resources

Mayank Gupta, Akhilesh Karumanchi, Matthew A. Lanham

Purdue University Krannert School of Management

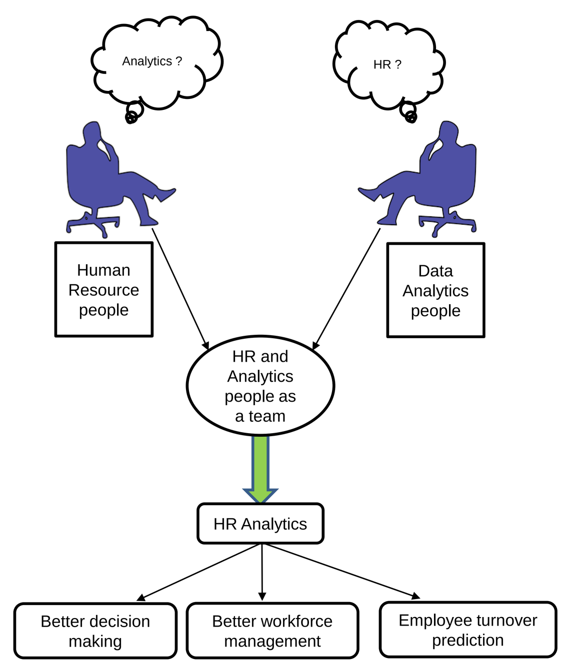
gupta363@purdue.edu; akarumanchi@purdue.edu; lanhamm@purdue.edu

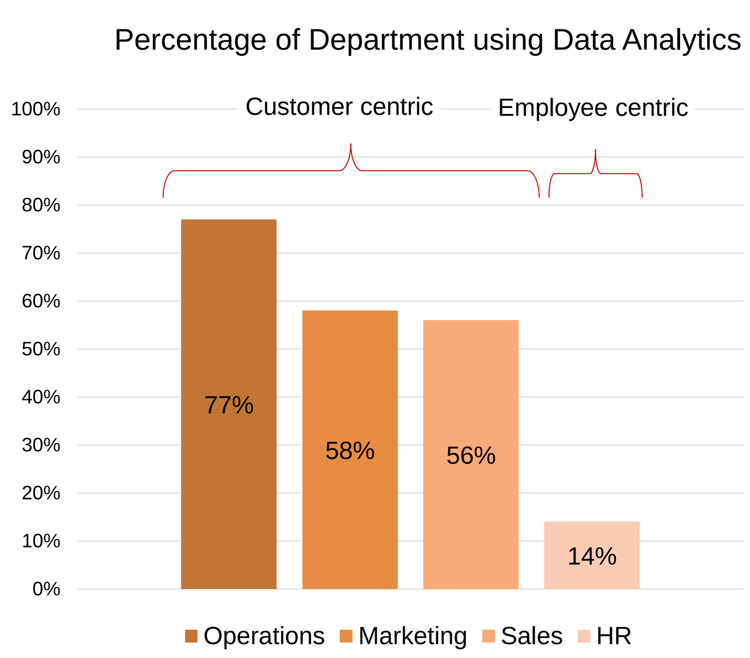
Abstract

This study builds and evaluates predictive models to estimate the probability that a team member will leave an employer within a future planning horizon. We then develop a theoretical decision model that provides guidance in how one might use the prediction outputs for future decision or policy making to better manage their workforce. This study proposes a framework to use analytics for human resource workforce management. The literature on analytics in human resources is scarce in comparison to other business verticals such as sales, marketing, and operations. Most research we have studied discuss employee turnover on theoretical grounds, rather than providing analytical decision-support solutions. Using data from a local retailer we develop a working framework that provides guidance to human resource professionals in how descriptive, predictive, and prescriptive analytics can work together to support workforce management decision and policy making.

Introduction

Retailers can achieve success when they retain and reward their best people. Employee turnover is costly if the employee who is leaving the company is a high performer. Nowadays, the big challenge for employers is to retain the best people by developing policies that engage their employees well, while having the ability to meet the expectations of their customers. Here comes the power of analytics which a lot of HR managers overlook as shown below.





Analytics in Retail

Most industry work on retail analytics comes from a supply chain and customer engagement (**CRM**) perspective, which focuses on the various stages of getting and presenting products directly to their customers. The key here is to **know your customers**. For example, retailers invest much time identifying consumer buying patterns which can provide assortment and pricing insights. Marketing departments will regularly analyze transaction log data, in-store checkout wait times, and store foot traffic to develop modified marketing strategies (Brust 2013). Many retailers will also have customer loyalty programs to help increase the transparency of these purchases.

Analytics in Human Resource

The remarkable amount of data collected and analyzed are also often used for many decisions and policies in retail. The interesting thing is that the advanced analytics seem to terminate there. We have found that retailers will invest much to understand their customers, but little to **know their employees**. This motivation led us to develop an analytical framework that HR departments could build upon to better understand and support their workforce decisions and policies within the firm. The study can be used directly by the HR decision makers using their own data to understand and model employee turnover.

Data

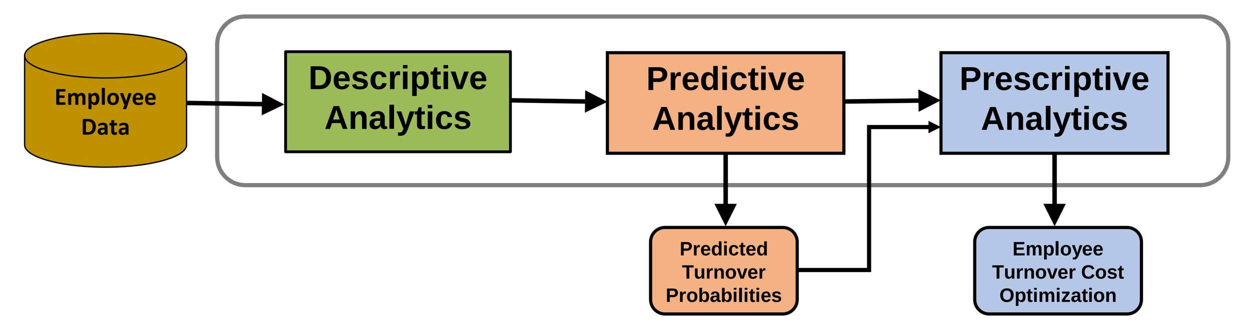
The data set used in this study was provided from a regional retailer faced with understanding their employee turnover problem. We were provided **500** records of employees that left the company and **500** records of employees that remained with the company over the previous six months. Features include information about store demographics, position titles, type of pay, pay amount, market index of pay per position, employee performance reviews over the previous two years, and store performance over the previous two year. Retailers can obtain whatever data is available about their employees, job functions, etc. to identify and estimate cause-and-effect relationships of the drivers of employee turnover.

Variable Modification

We began the analysis by deriving variables based off other features provided which made more sense for business interpretation. For Example, the wage level of that market level (*Pnormal*), was derived using its median and quartile values, dummy variables for some position titles, etc.

Methodological Framework

The methodological framework we propose includes using descriptive, predictive, and prescriptive analytics together.

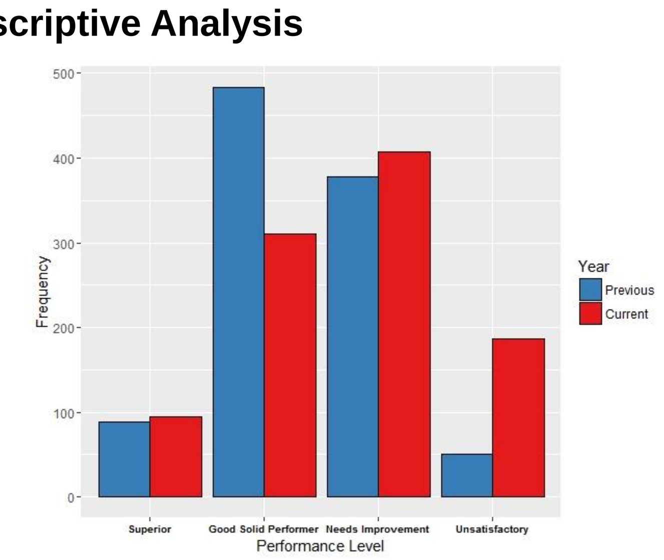


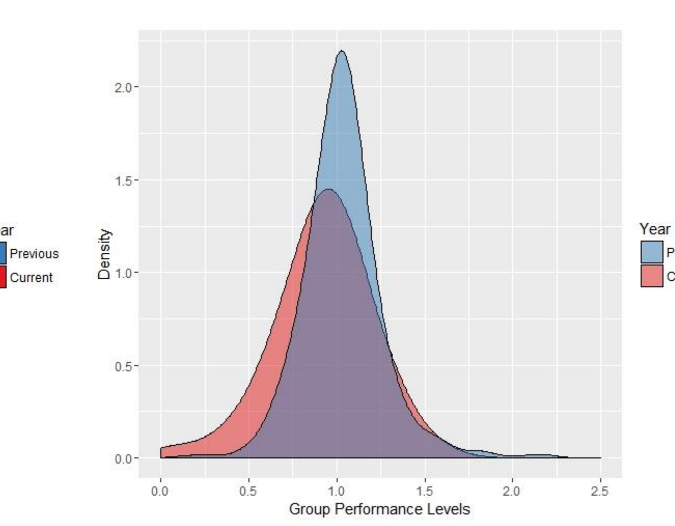
We built and evaluated **five different predictive models** to identify the drivers and their estimated effect on turnover. The predictive models were built using a **70/30% train/test partition**. The training set is used to build the model, while the test set allows us to gauge generalizability on future observations.

Once we identified a predictive model that best estimated the reasons for leaving, we formulated a decision model (prescriptive analytics) that incorporated those estimated effects from the predictive models. This model provides an HR professional an analytically-based means to decide **what actions to take** to help improve overall employee retention. Example decisions could be increasing employee pay, separating from poor performing employees whom are likely to leave anyway, providing educational incentives that can led to reduce turnover, etc. The problem lies in how to make these decisions company-wide while accounting for all business constraints. The decision model we develop provides a working prototype of this.

Results

Descriptive Analysis



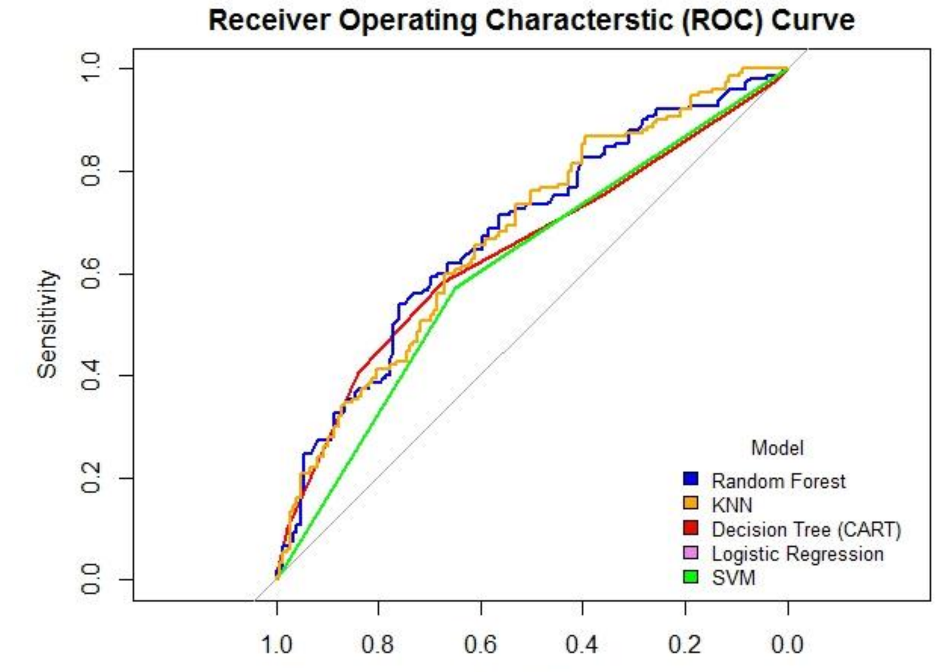


Predictive Analysis

- The Random Forests realized the greatest AUC value (0.68) in the test data. However, having a training AUC of 1 was a clear indicator that this model was overfitting to the training data. Therefore, we did not use this model.
- The next best model is the KNN model that we considered with AUC over test data set as 0.67.
- We choose to use the decision tree model because it provided clear interpretation for the HR professionals, allowed us to easily integrate the inputs into the decision model, and also had good predictive performance (AUC = 0.64).

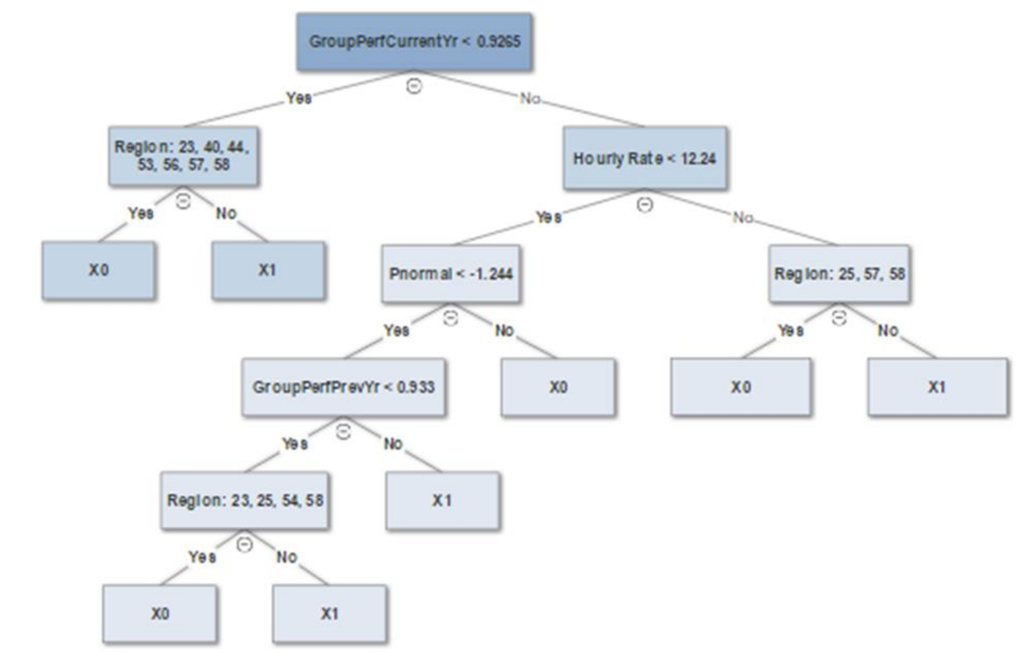
	Training				Testing			
Models	Accuracy	Sensitivity	Specificity	AUC	Accuracy	Sensitivity	Specificity	AUC
Random Forest	100.00%	100.00%	100.00%	1.00	64.88%	69.13%	60.67%	0.68
k-Nearest Neighbor	63.77%	61.43%	66.10%	0.70	60.54%	57.72%	63.33%	0.67
Decision Tree (CART)	65.34%	64.86%	65.81%	0.72	62.88%	67.79%	58.00%	0.64
Logistic Regression	65.91%	62.86%	68.95%	0.72	63.21%	63.76%	62.67%	0.63
Support Vector Machines	70.90%	76.29%	65.53%	0.71	60.87%	63.09%	58.67%	0.61

Receiver Operating Characterstic (ROC) Curve



Model	Significant Variables used in Model construction
Random Forest	Pnormal, GroupPerfCurrentYr, GroupPerfPrevYr, Hourly Rate, Region, Position Title
K-Nearest Neighbor	-
Decision Tree (CART)	Pnormal, GroupPerfCurrentYr, GroupPerfPrevYr, Hourly Rate, Region
Logistic Regression	Pnormal, GroupPerfCurrentYr, GroupPerfPrevYr, Pflag_below
Support Vector Machine	-

The decision tree model provides the HR professionals an understanding of the drivers impacting employee turnover.



Prescriptive Analysis

We use these decision tree splits and weights in our customized decision model below. The idea is as decisions are optimized, the probability of leaving will be adjusted.

<u>Decision variables</u>

x_{ij} = amount to increase wage of employee i in store j; $i = 1,..,N;\ j = 1,..,M$

y_{ij} = decision to fire employee i in store j, $y_{ij} \in \{0,1\};\ i = 1,..,N;\ j = 1,..,M$

<u>Parameters:</u>

N = total number store employees

B_{ij} = wage of store manager position i in store j; $i = 1,..,B;\ j = 1,..,M$

S_{ij} = wage of sales associate position i in store j; $i = 1,..,S;\ j = 1,..,M$

T_{ij} = wage of stockers position i in store j; $i = 1,..,T;\ j = 1,..,M$

τ_{ij} = latest performance review of employee type i in store j; $i = 1,2,3;\ j = 1,..,M$

φ_{ij} = hourly rate ($) of employee i in store j; $i = 1,..,N;\ j = 1,..,M$

ω_{ij} = wage index of employee i in store j; $i = 1,..,N;\ j = 1,..,M$

ρ_{ij} = estimated probability of turn in next 6 months of employee i in store j; $i = 1,..,N;\ j = 1,..,M$

ψ_{ij} = estimated class of turn in next 6 months of employee i in store j; $i = 1,..,N;\ j = 1,..,M;\ \psi_{ij} \in \{0,1\}$

K_{ij} = the average team performance of employee type i in store location j; $j = 1,..,M$ (specified by HR)

Z_j = the average team performance of stockers at store location j; $j = 1,..,M$

A = the next sixth month budget ($) for store managers

B = the next sixth month budget ($) for sales associates

Γ = the next sixth month budget ($) for stockers

<u>Based on how the decision variables change, these parameters will change</u>

φ_{ij}^* = new hourly rate ($) of employee i in store j; $i = 1,..,N;\ j = 1,..,M$

ω_{ij}^* = new wage index of employee i in store j; $i = 1,..,N;\ j = 1,..,M$

ρ_{ij}^* = new estimated probability of turn in next 6 months of employee i in store j; $i = 1,..,N;\ j = 1,..,M$

ψ_{ij}^* = new estimated class of turn in next 6 months of employee i in store j; $i = 1,..,N;\ j = 1,..,M;\ \psi_{ij} \in \{0,1\}$

<u>Objective function:</u>

$\max\left[\sum_j \sum_i \psi_{ij}^*\right]/N$ (maximize the percentage of expected non-turners to complete workload)

<u>Constraints:</u>

$\sum_i \tau_{ij}/S \geq K_{ij} \quad \forall j$ (average employee type performance should exceed some threshold)

$1040 * \sum_j \sum_i B_{ij} \leq A$ (budget for store managers must be satisfied)

$1040 * \sum_j \sum_i S_{ij} \leq B$ (budget for sales associates must be satisfied)

$1040 * \sum_j \sum_i T_{ij} \leq \Gamma$ (budget for stockers must be satisfied)

$\sum_j \sum_i \omega_{ij}^*/N \geq 0.95$ (average wage index of employee type i is at least 0.95; 1 would imply market avg.)

$x_{ij} \geq 0$ (hourly wages will can only increase)

$y_{ij} \in \{0,1\}$

…

other custom constraints

Conclusions

- The predictive models generated provide the HR professionals some guidance in understanding what will lead an employee to seek other employment. However, having limited features about team members will lead to limited insights.
- The point here is to show the HR professionals what could be done using analytics and then extend those insights from the predictive models to actions for decision-support or policy making
- Many times organizational units predict something, but having the ability to take those predictions and use them in a meaningful way by incorporating business knowledge and constraints is how an organizational unit can truly maximize the benefits of analytics.
- This study provides a working framework in the context of retail workforce management and employee turnover.

Acknowledgements

The poster is partially funded by Business Information and Analytics Center (BIAC), Purdue University. We really appreciate their support and the opportunity that they provided.