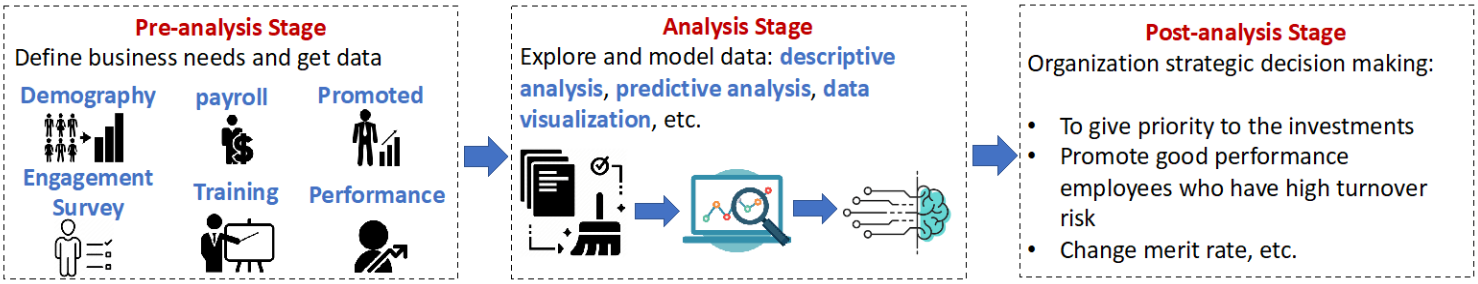
KENNESAW STATE UNIVERSITY
ANALYTICS AND DATA SCIENCE INSTITUTE

datascience.kennesaw.edu

Using HR Analytics to Support Managerial Decisions: A Case Study



BACKGROUND

More and more organizations are becoming employee-centric in the 21st century. The employment and workforce industry are increasingly relevant because the value of human capital is directly linked to an organization's profitability. Human Resource (HR) Analytics enables HR managers to make strategic contributions and support managerial decisions. However there are several challenges with integrating analytics into HR: the data is significant, messy, and imbalanced, it is hard to harness both structured and unstructured data, most HR managers lack data mining skills, and there is a lack of related empirical research that gives detailed analytics guidelines. The contribution of this study is that we develop a framework to support an industrial aluminum company to make the decisions and to improve strategy execution. The framework includes descriptive analysis, predictive analysis, and entity sentiment analysis. We analyzed an industrial aluminum company's HR data as a case study and found some actionable issues using descriptive analysis. Then we employed machine learning algorithms to predict employee turnover rate and identify risk factors. Finally, we applied the entity sentiment analysis on the unstructured data collected from employees' engagement survey.

APPROACH

We defined our HR analytics framework as a three-stage framework. The first stage is the pre-analysis stage, which aims to define business needs and collect necessary data. The second stage is the analysis stage. The two main analysis methods are descriptive analysis and predictive analysis. In the descriptive analysis, the main task is understanding the organization's HR problems and using statistical metrics or understandable visualizations to show the insights behind the data. There are three main methods for predictive analysis: statistics-based analysis, machine learning-based analysis, and deep learning-based analysis. Post-analysis is the third stage in our framework. In this stage, the managerial strategies are developed using the analytical results from the second stage. An organization could make decisions, for instance, promoting high-performing employees who have been identified as high-turnover risk.

Since the dataset is imbalanced, we also applied three balancing strategies to solve the imbalanced problem that could also improve the model's performance. The three balanced strategies are random oversampling, SMOTE, and ADASYN. We tested five machine learning models: K-nearest Neighbors (KNN), Logistic Regression (LG), Random

Forest (RF), Gradient Boosting Tree (GRB), and Decision Tree (DT). For the unstructured data, which we collected from employees' engagement survey, we utilized Google Cloud NLP API to address the entity sentiment analysis.

Results

Table 1 below shows the results of the turnover prediction models' performance. From the results, we see that gradient boosting and random forest have the best performance with random oversampling. When we rank the features, we found that the most critical factors that affect attrition are Actual vs. Target pay, Tenure, and age, as Figure 1 shows.

Balanced Strategy	Models	AUC	Recall	Precision	Accuracy
Random Over-sampling	Logistic Regression	0.60	0.58	0.24	0.62
	KNN	0.53	0.43	0.18	0.61
	Decision Tree	0.54	0.21	0.24	0.77
	Random Forest	0.62	0.31	**0.43**	**0.83**
	Gradient Boosting	**0.68**	**0.63**	0.28	0.73
SMOTE	Logistic Regression	0.58	0.54	0.26	0.61
	KNN	0.55	0.48	0.22	0.64
	Decision Tree	0.61	0.38	0.24	0.75
	Random Forest	0.56	0.18	0.21	0.60
	Gradient Boosting	0.61	0.32	0.25	0.69
ADASYN	Logistic Regression	0.62	0.56	0.23	0.66
	KNN	0.55	0.51	0.18	0.60
	Decision Tree	0.59	0.33	0.25	0.76
	Random Forest	0.60	0.25	0.38	**0.83**
	Gradient Boosting	0.61	0.31	0.25	0.81

Table 1: Models' performance for turnover prediction.

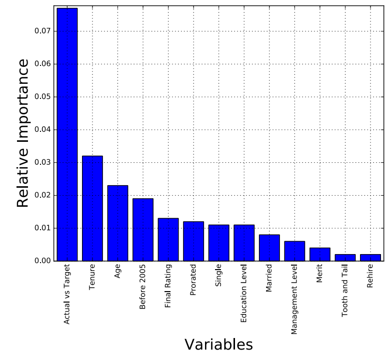


Figure 1: Factors ranking

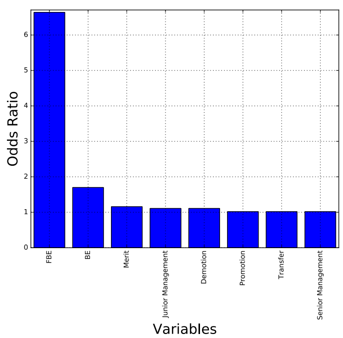


Figure 2: Most Important Factors Affecting Turnover

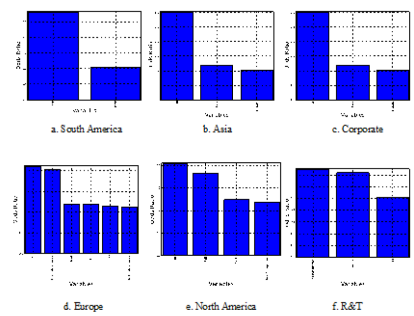


Figure 3: Factors Affecting Turnover by Region

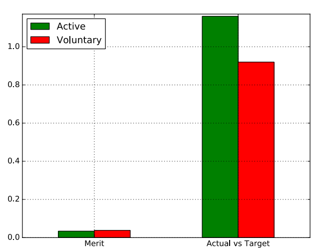


Figure 4: Average Merit and Actual vs Target Pay by Termination

Region	Low performers	Average performers	High performers
Overall	41.98%	21.25%	18.29%
Asia	52%	23.43%	23.17%
Corporate	40%	25.34%	21.18%
Europe	37.50%	19.84%	23.23%
North America	29.28%	21.58%	17.90%
R&T	22.22%	14.46%	7.14%
South America	50%	18.57%	9.38%

Table 2: Voluntary Ratio of Different Regions by Performers Type

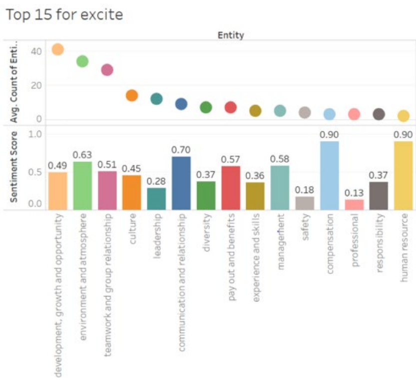


Figure 5: Entity employee like most

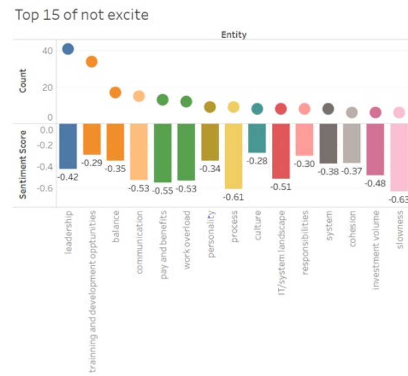


Figure 6: Entity employee like least

Performance rating is the most critical factor that affects turnover. The experiment results also show that higher merit, higher voluntary departure risk, which also gives an alert to the company about payroll management. We also performed segment models by regions since the company has different branches all over the world. Not surprising, the different locations have different factors that contribute to voluntary departure, as Figure 3 shows. Table 2 shows the different voluntary departure ratio by regions. We mainly answered four questions via descriptive and predictive analysis.

The first question is, "What suggestions could we make for monitoring high-risk employees?" We make several suggestions that could support the company in making managerial decisions. First, the outputs of logistic regression show that employees who have been transferred or promoted are more likely to leave, especially those with lower merit payouts. Second, there exist critical years of departure, which is around three to six years of tenure. Organizations should pay more attention by doing a better job in the employee care with employees who have three to six years of tenure. Third, make sure there is no experience of a mismatch between job bands (position and personal). Fourth, we found that employees who are leaving voluntarily are consistently performing higher (constant merit increases) or lower (constant merit decrease). Finally, we suggest companies pay more attention to a problem that there are no discrepancies among incentives: Higher merit scores but lower actual vs. target payouts, the employee most likely to voluntary departure as Figure 5 shows.

The second question we answered is, "Does the company retain high-performers?" We mainly use descriptive analysis to solve it. In general, the answer is "Yes". All the high performers in different regions have a lower voluntary attrition rate compared with average and low performers, but the top performers in Europe have higher voluntary attrition rates than average performers. The third question is, "Does previous year merit and payout drive higher performance for the following year?" With the descriptive analysis, we found the answer is: In general, high merit drives high performance, but not for Actual vs. target payouts. Figure 5 shows the top 15 entities which most excited employees in the company using text mining. Employees had the most positive attitudes towards development, growth and opportunity, environment and atmosphere, teamwork and group relationships. Figure 6 shows the top 15 entities which least excited employees. With these analyses, we found that employees had the most negative attitudes towards leadership, manager; training, development and opportunity; work and life balance.

CITATION FOR FULL ARTICLES

Liyuan Liu, Sanjoosh Akkineni, Paul Story and Clay Davis, Accepted by 2020 ACMSE.