PREDICTION OF EMPLOYEES TURNOVER

MOTIVATION

Reducing the cost of staff turnover is the challenging task for all companies. Python's package Scikit-learn, Speech Recognition and NILK helps management team lower risk of extra spending.

PACKAGE OVERVIEW

Scikit-learn ⬤

Explore Employee Churn Reasons
What factors are significant to employee turnover?
- Using IBM HR data on Kaggle as our dataset

01 Feature Selection

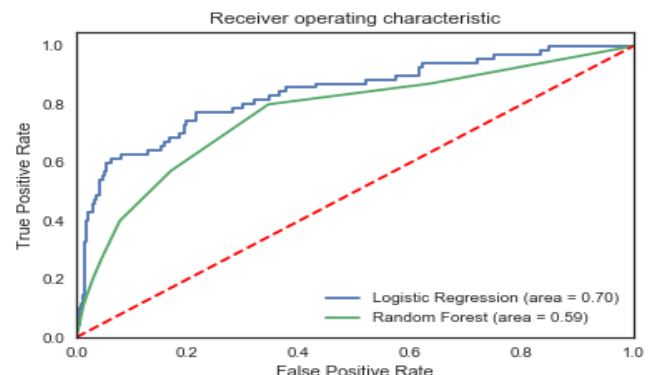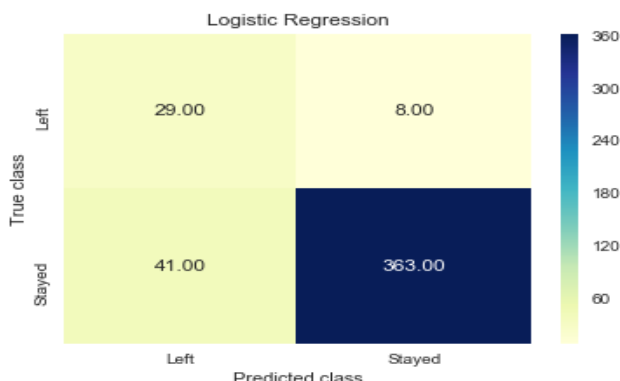Build up a correlation matrix of numerical variables to remove highly correlated variables

02 Model Selection

Separate training and testing data. Fit the data using Logistic regression, Random forest, SVC classifier. Select the model with highest accuracy and interpretability

03 Model Validation

Cross validate the prediction power for different data split. Plot ROC to check difference between models.
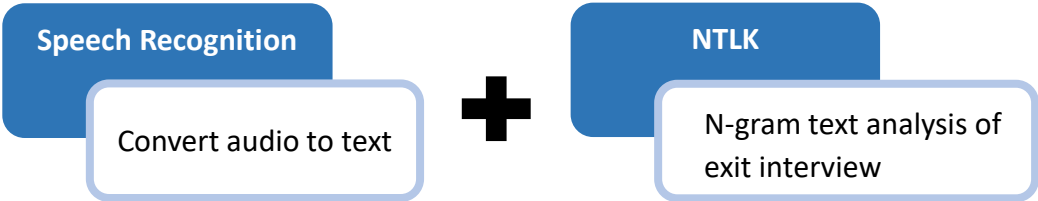➤ logistic regression is a better classification because it is relatively far away from the random guess line.

Logistic Regression confusion matrix and Receiver operating characteristic curve

Patrick Hoban (hoban014@umn.edu) | Kartikeya Tiwari (tiwar093@umn.edu) |
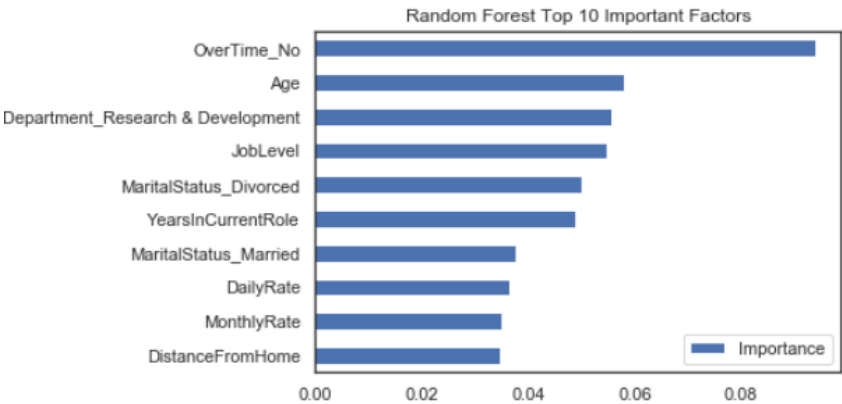Chia-Hsuan Chou (chou0176@umn.edu) | Sarah Wang (wang9379@umn.edu) | Chi-An Hsieh (hsieh142@umn.edu)

Direct Response from your Employee

Are there any other factors to put into consideration?
- exit interviews as sample set

Speech Recognition

Convert audio to text

+

NTLK

N-gram text analysis of exit interview

HERE'S OUR STORY

Result of HR dataset

Random Forest Top 10 Important Factors

OverTime_No
Age
Department_Research & Development
JobLevel
MaritalStatus_Divorced
YearsInCurrentRole
MaritalStatus_Married
DailyRate
MonthlyRate
DistanceFromHome

Importance: 0.00 — 0.02 — 0.04 — 0.06 — 0.08

Most important factors, but is there any omitted variable?

From another perspective with text analysis
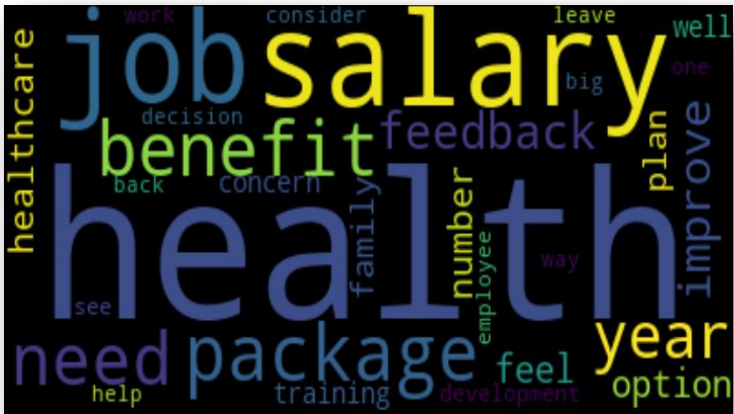
Health concern is the first reason for leaving job

📊 Conclude significant factors and primary reasons to reduce the cost of HR management

🙂 Know what your employees needs, design better retention plans and improve satisfaction

BUSINESS APPLICATIONS

01 **Predicting customer churn**

02 **Evaluating product performance**

03 **Exploring industry trends**

Download E-Edition

Patrick Hoban (hoban014@umn.edu) | Kartikeya Tiwari (tiwar093@umn.edu) |
Chia-Hsuan Chou (chou0176@umn.edu) | Sarah Wang (wang9379@umn.edu) | Chi-An Hsieh (hsieh142@umn.edu)