

ISOM 3360 Group Project

2020 Spring

Li Chi Ho (Cloud), 2051 0659
Jiwoo Han (Liz), 2038 8795
Li Ka Ho (Jim), 2042 8363
Yao, Yen-Lun(Nora), 2056 1488
Leung Cheuk Hang (Matthew), 2042 4379

Business Problem:

Recruiting new employees and providing professional training are costly to a company. Therefore, minimizing employee's turnover rate is significantly important. We plan to use various features to predict the attrition of valuable employees. This prediction can provide insights to help human resources managers in the recruiting process and managing current employees.

Data Source:

The dataset is obtained from "[IBM HR Analytics Employee Attrition & Performance. Kaggle](#)".

Data Description:

The dataset has 1470 examples and 35 features with no missing value. The dataset is imbalanced, 1233 examples of "No Attrition" and 237 examples of "Attrition". In terms of data types, nine features are string and twenty-six features are integer.

The target variable is "Attrition", a binary variable. There are mainly two groups of independent variables -- employee's personal information and company-based information. Employee's personal information includes "Age", "Education Level", "Education Field", and "Distance From Home". Company-based information includes "Frequency of Business Travel", "Department", "Standard Hours", "Years At Company", "Stock Option", "Job Satisfaction", "Performance Rating", and "Years Since Last Promotion".