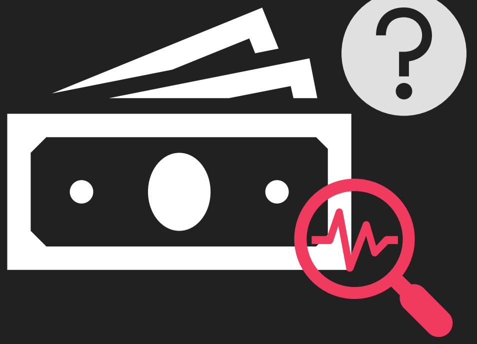
Salary prediction based on job descriptions

Problem Statement

- Although salary is a big factor for people to decide whether to take the job or not, it is not specified on the job description in the United States. Job seekers often times invest their time and effort to get low paying jobs and it turns out to be a disappointment at the end of the process. It would be helpful to build a model that would predict the salary based on the job description.

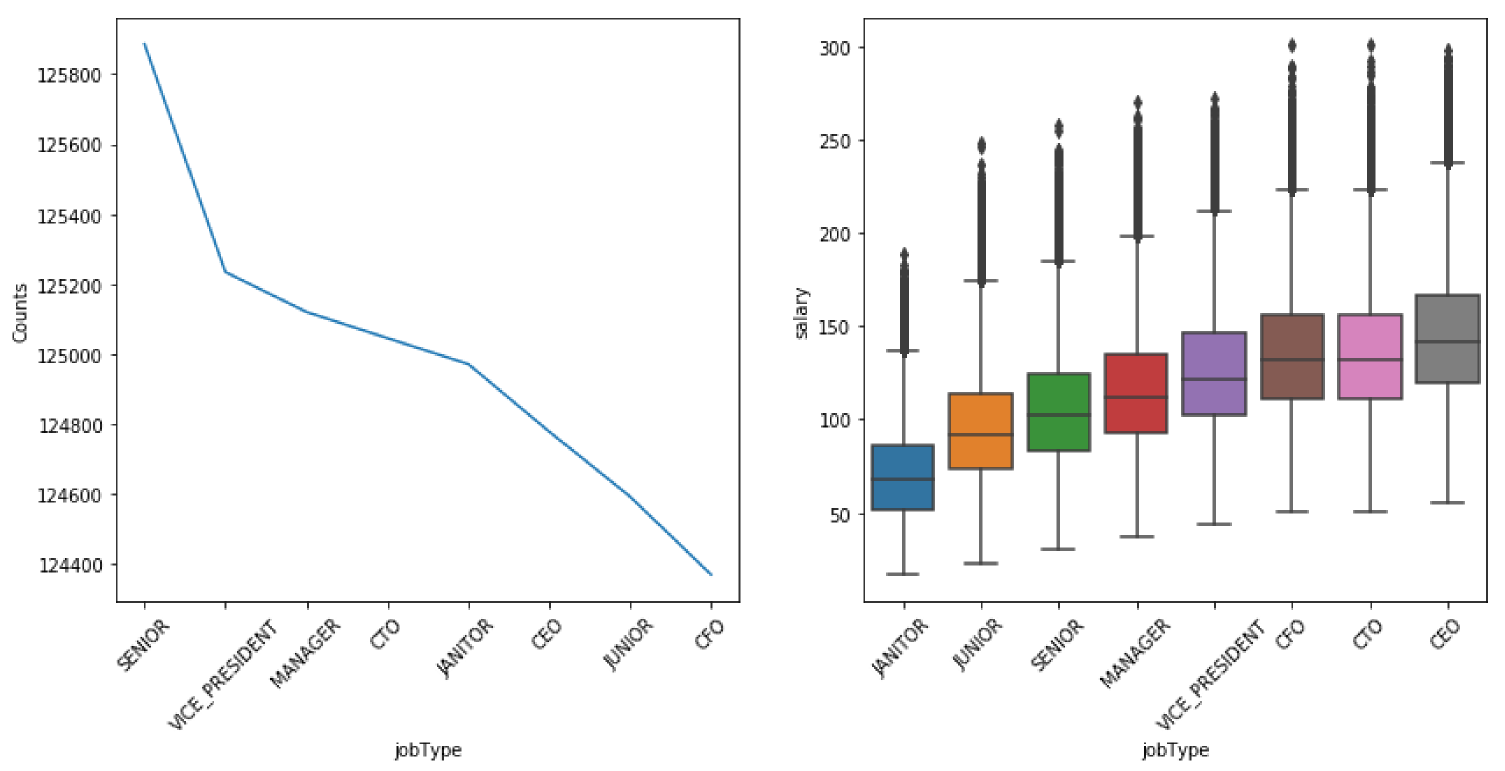


Exploratory Data Analysis (EDA)

Plot on the left is distribution of the feature and the one on the right is the dependence of target variable on the feature

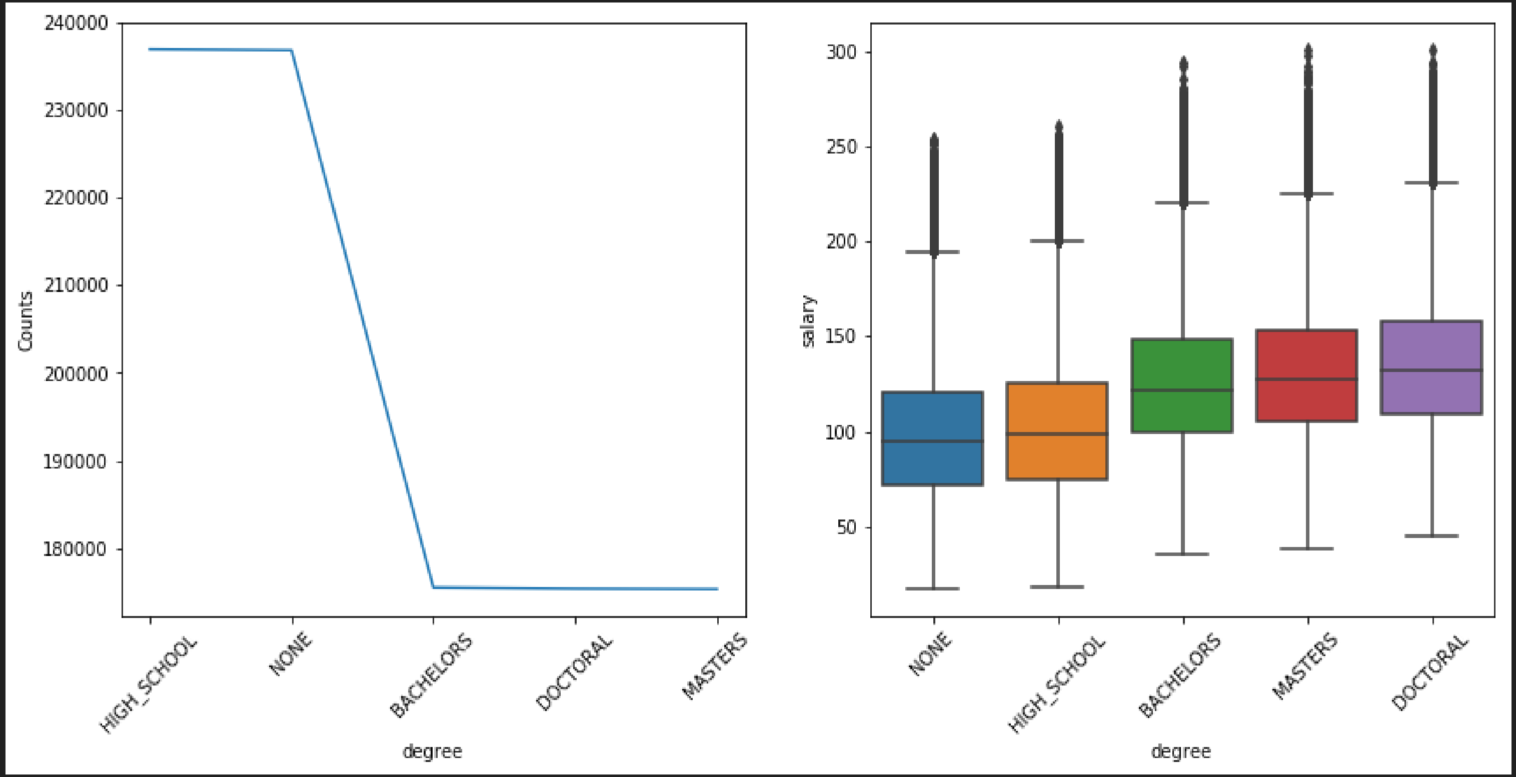
Job type and salary

Higher position tends to receive higher salary



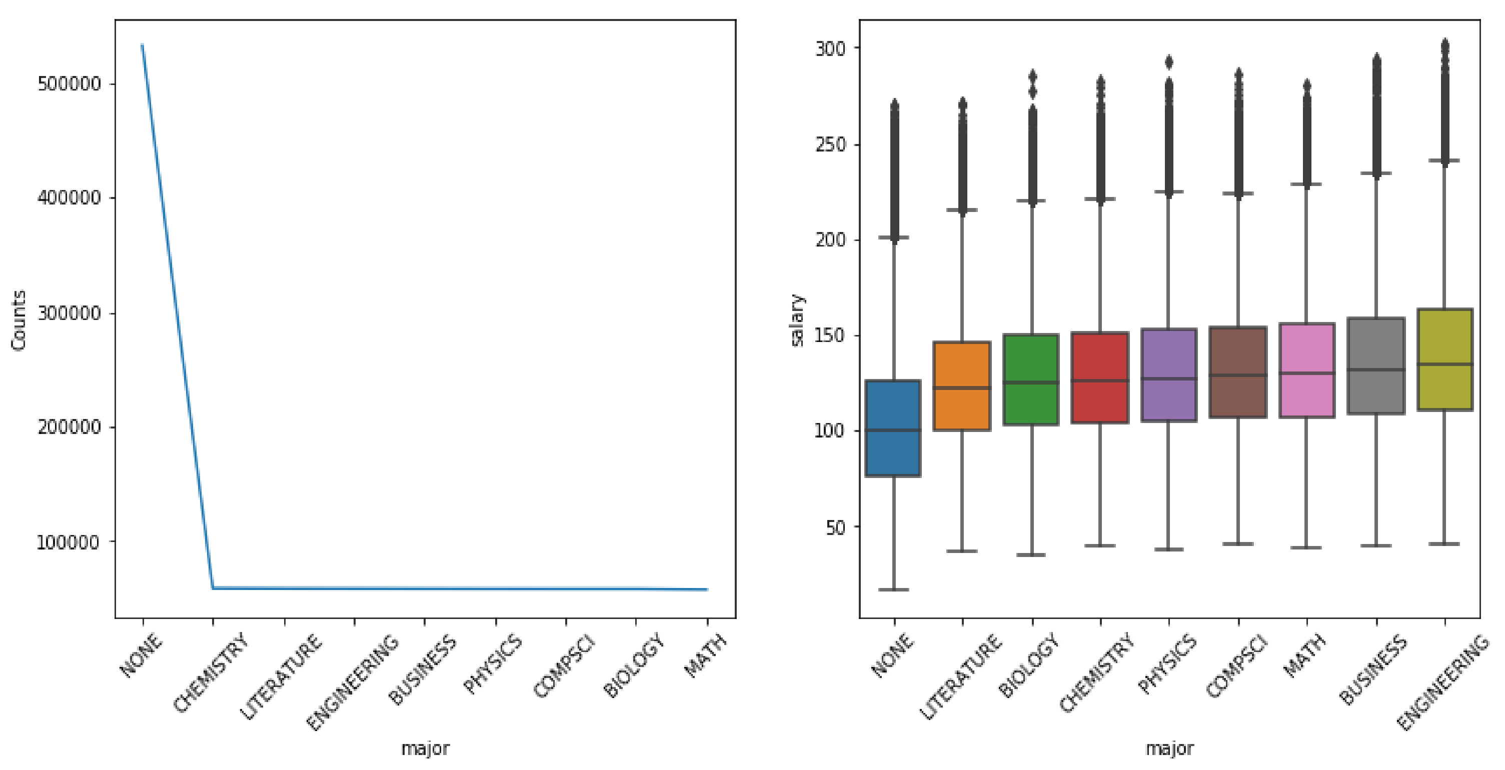
Degree and salary

More advanced degree tends to receive higher salary



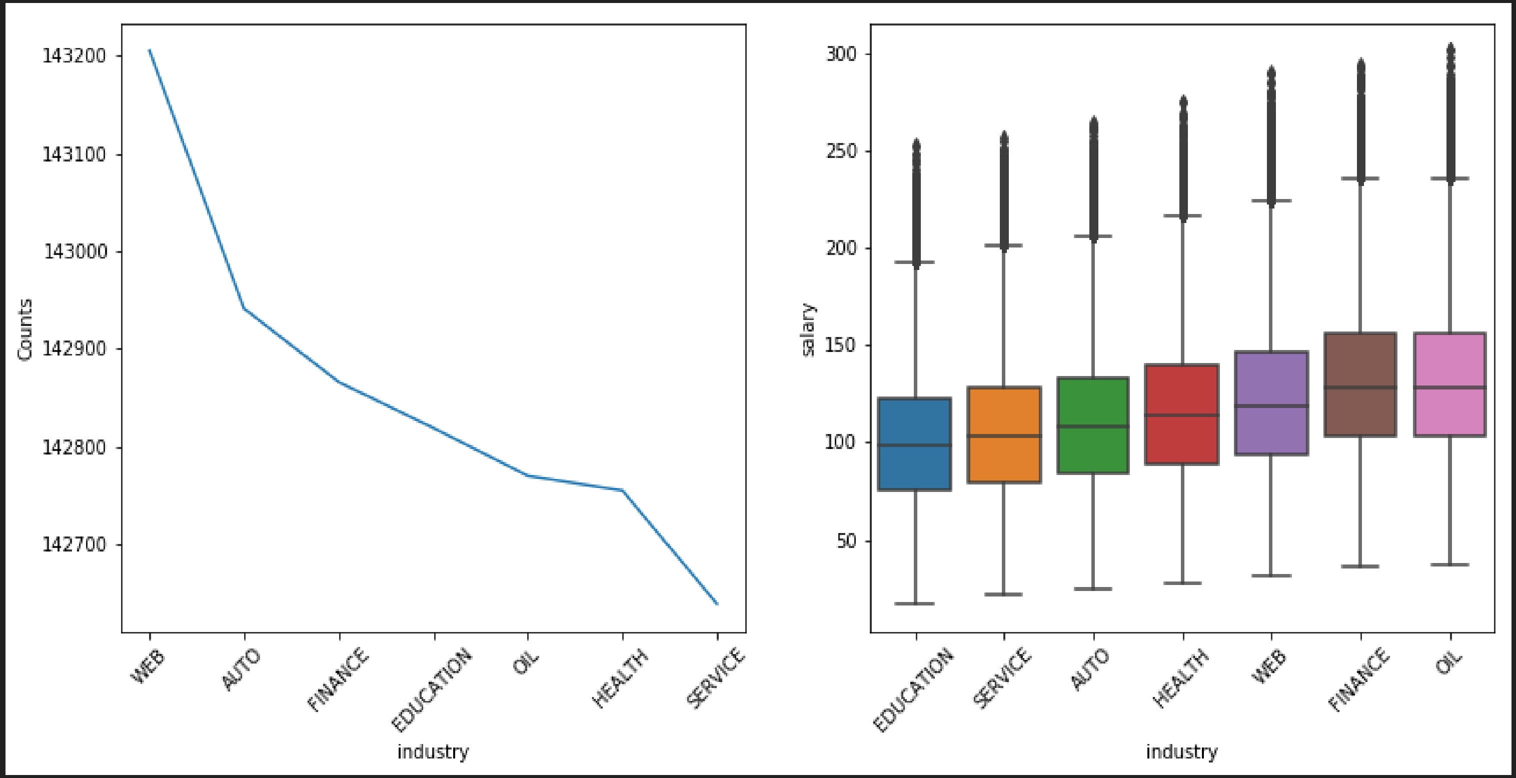
Major and salary

People who studied engineering or business tend to receive higher salary than other majors



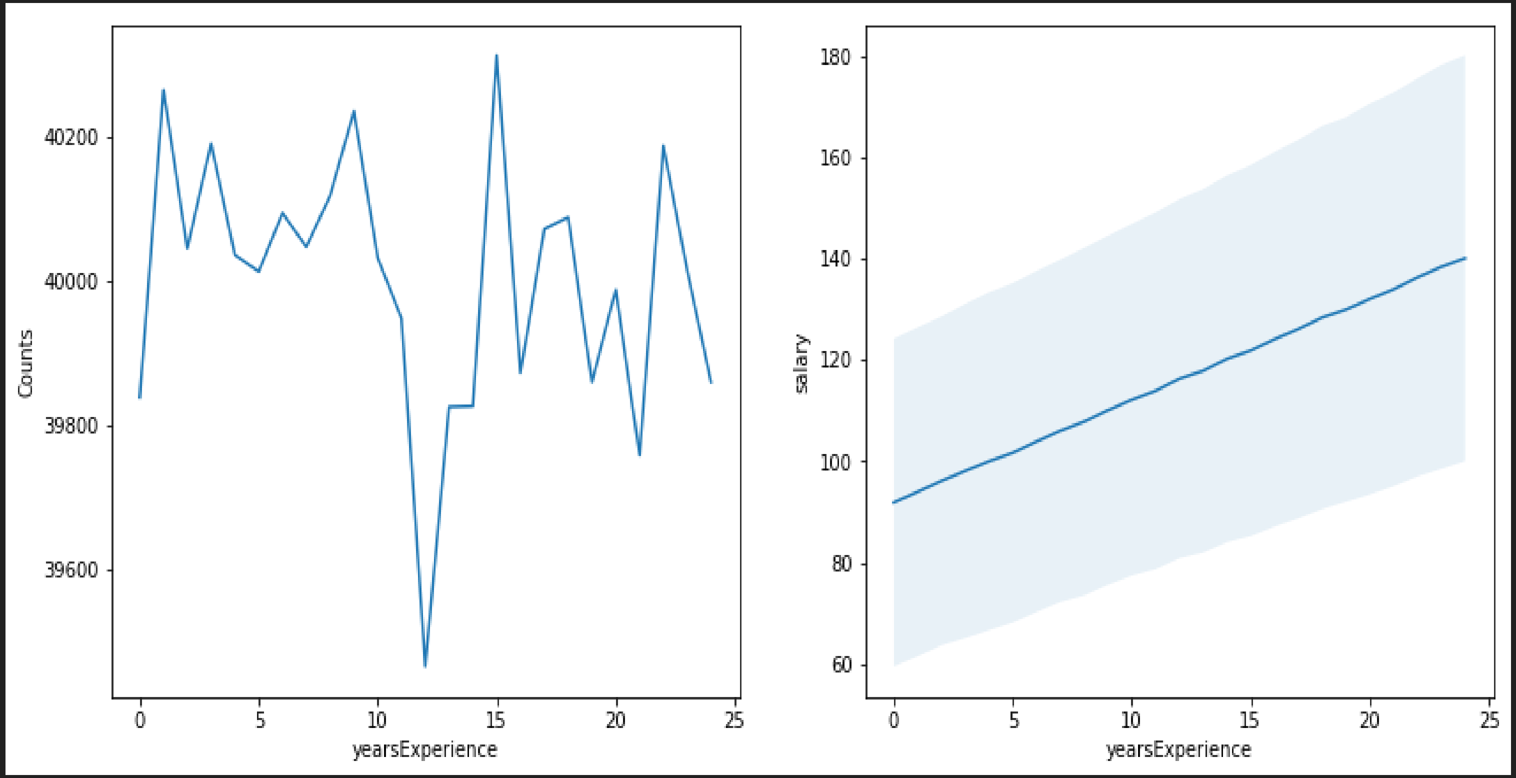
Industry and salary

Oil and finance industry tend to pay higher salary



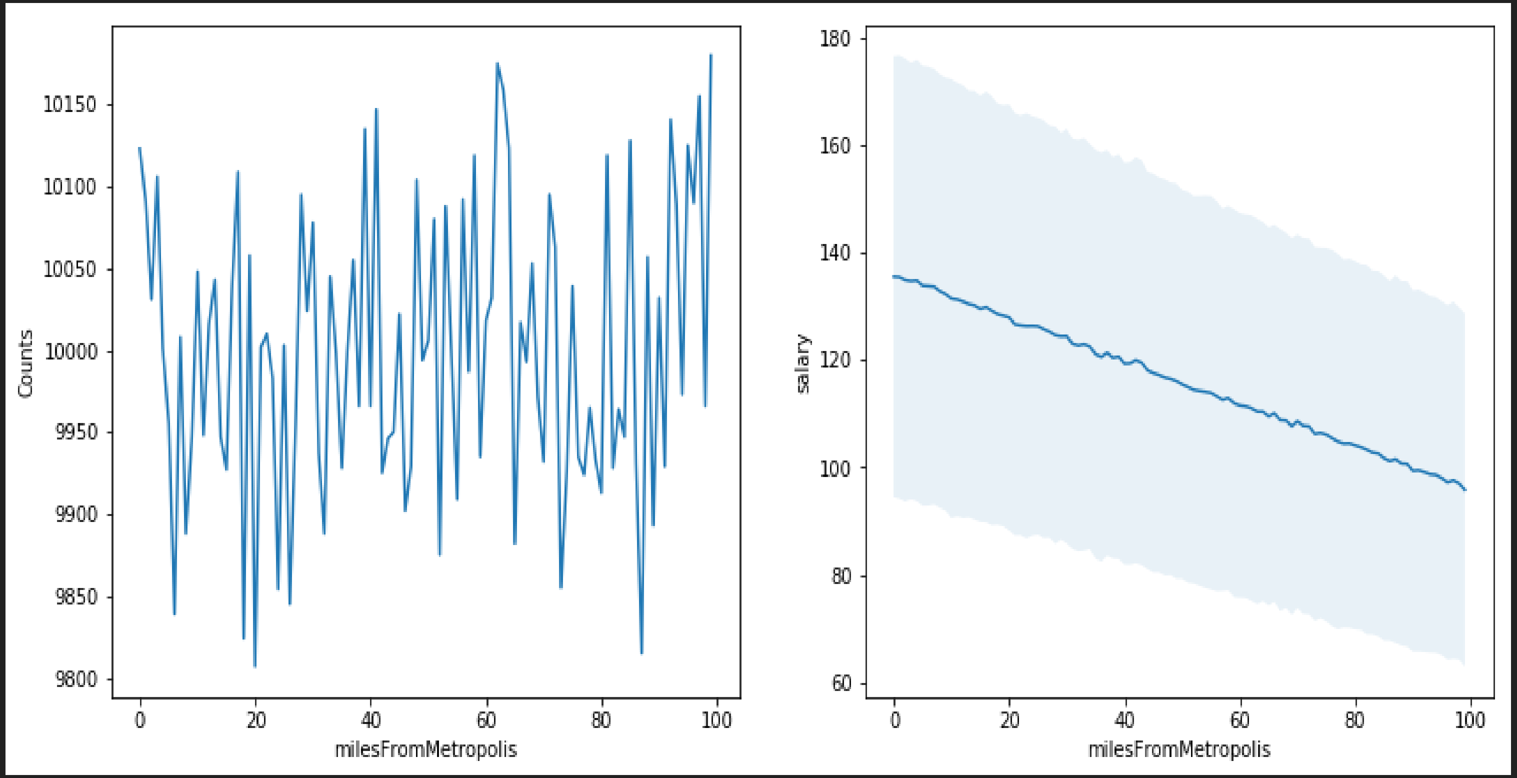
Years of experience and salary

There is a positive correlation between years of experience and salary



Miles from metropolis and salary

Further from the metropolis tends to receive less salary



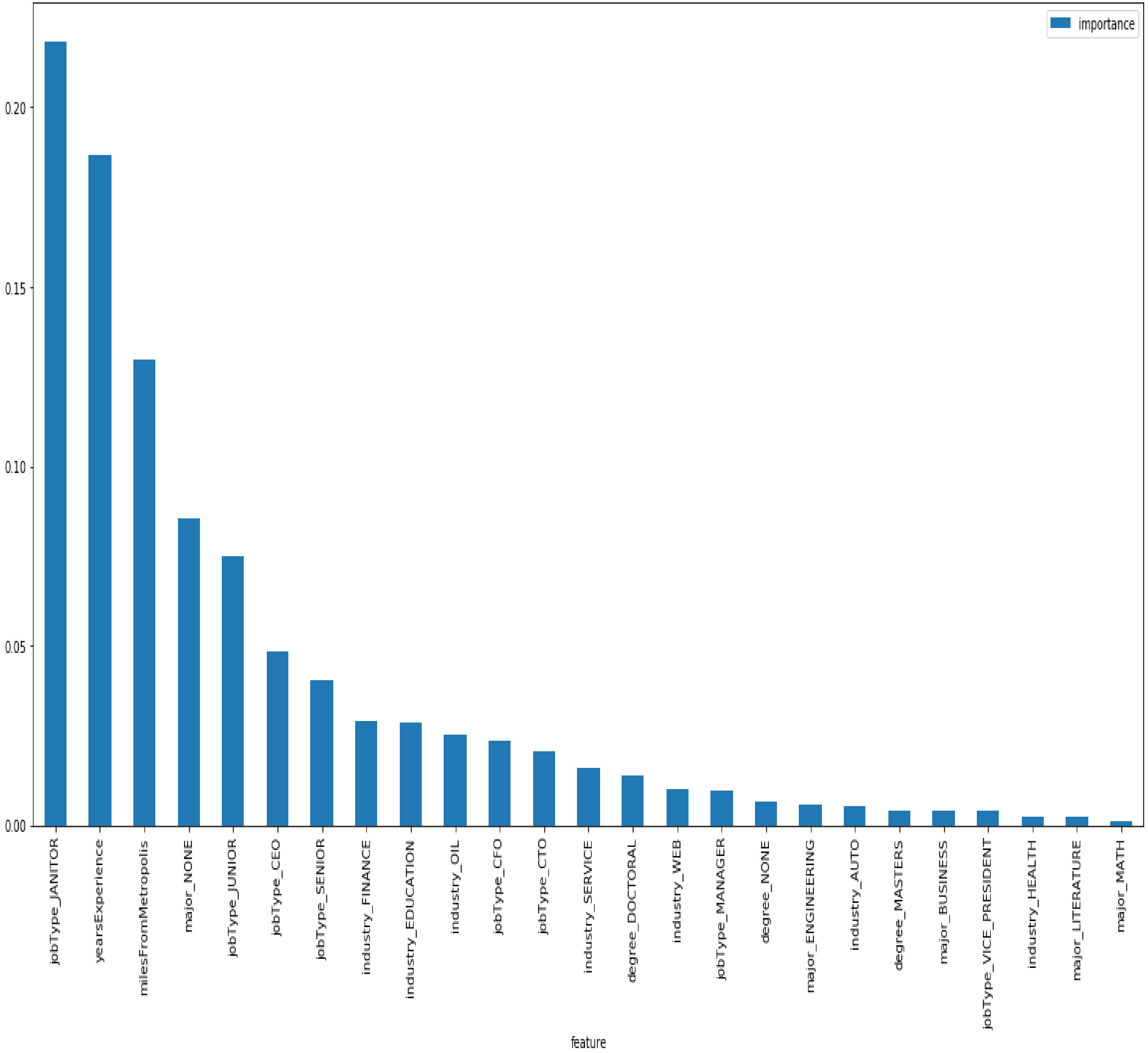
In-Depth Analysis Using Machine learning

Supervised Learning - Regression

Model Performances

	Linear Regression	Scaled linear regression	Gradient boosting regression	Random forest regression
Average mean squared error	384.49	384.49	357.20	367.76
Standard deviation	1.40	1.40	0.81	1.30
Runtime	7.89	28.16	1764.50	953.19

Gradient boosting has the longest run time but the best performance



Feature Importance

Top 3 features

- JOB TYPE (JANITOR)
- YEARS OF EXPERIENCE
- MILES FROM METROPOLIS

Thank You

Jaesuk Kim

jaesukkim4402@gmail.com

Please see Jupyter Notebook for further analysis