ARIA
APPLIED RESEARCH IN ACTION

Multi-Regional Salary Prediction Model

Empowering strategic decisions with comprehensive insights into compensation trends worldwide
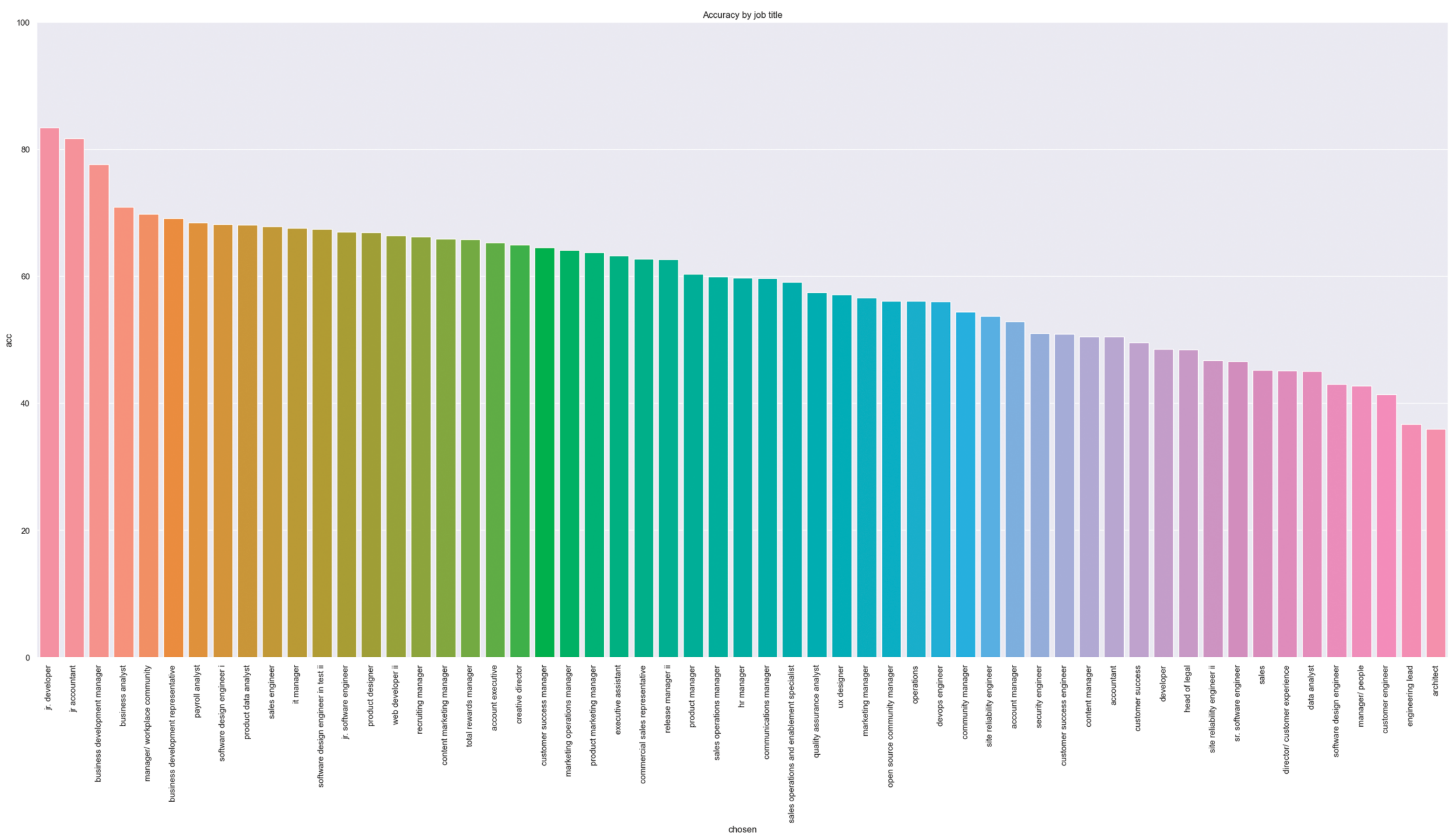
Sarah Rozaidi

Mariano Consens
ACADEMIC SUPERVISOR

Dave Kim
INDUSTRY SUPERVISOR

Accuracy by job title

PROJECT SUMMARY

Predicting salaries is complex due to factors like cost of living, job demand, and education levels, which vary globally. This complexity intensifies as companies hire globally. Developing a machine learning model that considers these factors offers valuable insights for quarterly budgeting and recruitment. It aids in understanding regional job market trends and compensation practices, enabling more informed decisions. The goal is to create a model providing salary ranges based on key indicators. A broad 37,000+ global dataset was collected and cleaned, then subject to an extensive feature engineering and selection process wherein features with significant correlation and importance were integrated. Three supervised machine learning techniques were applied to the resulting dataset, K-Nearest Neighbour, Linear Regression and Random Forest. The outputs of the models were evaluated, and the best performer (Random Forest with an accuracy of 90.32% and a mean absolute error of $6,861) was selected for further tuning and optimizing before deploying to customers for testing. We anticipate that the collection of feedback and additional data while testing will provide a valuable avenue for continual refinement and enhancement of the model.

Agentnoon

Computer Science
UNIVERSITY OF TORONTO

Master of Science in
Applied Computing