LEE, K. M.

Klee604@outlook.com

(672) 999-9891

1617 East 37th Ave, Vancouver, BC

SKILLS

Languages: R, Python, MS Excel and SQL (MySQL, Impala & SQLite3)
Data Visualization: R graphics and dashboard with Shiny, MS Power BI, MS Visio, Sisense (Periscope), Google Analytics and Tableau
Modelings: Machine Learning Algorithms
Cloud Computing: Apache Hadoop & Spark
Website Development: Wix

EDUCATION

Seattle University
Master of Science: Business Analytics
Seattle, WA, 09/2020

University of British Columbia
Bachelor of Science: Cell Biology
Vancouver, BC, 05/2014

WEBSITE

https://kilee722.github.io/

https://github.com/kilee722

https://www.linkedin.com/in/ki-min-lee-1b5b39191/

CERTIFICATIONS

IBM Professional Data Science
Cloudera Big Data Analysis with SQL

PROFESSONAL SUMMARY

Data Analytics Specialist with experience in big data analysis and data visualization. Strong mathematical and statistical skills with good programming skills in Python and R. Detail-oriented and highly motivated individual committed to hard work

WORK HISTORY

Research Assistant - Health Analytics and Visualization, Remote
Seattle University – School of Nursing, Seattle WA 08/2020 - Present

- Provided visualization and dashboard that describe lifestyle and diet habit influencing type II diabetes by each ethic sub-group in California from California Health Issue Survey data (103482 observations with 328 features in the year 2013-2017) and created a prediction model for type II diabetes incident by each ethnic sub-group.
- Found early age of onset (30s) of diabetes in Filipino and Hispanic by 4-7% higher than average. Positive correlation between BMI and onset age of diabetes. Large difference in diabetes incidents by gender in Japanese (41.7% higher in male than female in Japanese vs. 7.8% lower in male than female on average)

Research Assistant
UBC – Department of Statistics, Vancouver BC 05/2018 - 01/2019

- Simulated deep learning models in Blang (developed by supervisor) and in Python and compared performances between two programming languages

Research Scientist
AbCellera Biologics, Vancouver BC 05/2016 - 03/2017

- Tested/modified PCR experiment protocol: Increased cDNA production by 10%, reduced experiment time by introducing Gibson Assembly method
- Designed and simulated protein conformation in 3D for mouse humanization to increase antibody production

Research Technician
UBC-Department of Botany, Lee Lab, Vancouver BC 08/2014 - 04/2016

- Developed Mutant cell-cryopreservation for laboratory samples: Increased post cryopreservation survival rate at 62%
- Created a database and conducted statistical analysis for the mutant cell library using MS Excel, provided reports with visualizations and regression analysis for research publication

PROJECTS

Passengers' Airport and Airline Choice (Python libraries: pandas, numpy and scikit-learn)

- Analyzed passengers' airport and airline choice between ICN and GMP airports and developed prediction models for the passengers' selection
- Constructed prediction models: Logistic regression & decision tree Airline choice model: 82% accuracy, Airport choice model: 81% accuracy
- Identified core factors (airfare & destination) for airline business to increase revenue and provided recommendations to increase traffic for airports

LendingClub Investment Analysis & Borrower's Default Prediction (Python libraries: pandas, numpy and scikit-learn)

- Analyzed a large online social lending dataset and created borrower's default prediction models using machine learning algorithms
- Applied random sampling technique to mitigate data imbalance issue, created over 20 predictive models including random forest and neural network, and achieved 66% accuracy and AUC, and 67% recall
- Provided prediction models and interpretations on important features in loan status definitions for investors to make informed decision for investment