LEE, KI MIN

Klee604@outlook.com

(206) 886-9891

1617 East 37th Ave, Vancouver, BC

SKILLS

Languages: R, Python, and SQL (MySQL, Impala & SQLite3)
Data Visualization: R graphics and dashboard with Shiny, Power BI, Sisense (Periscope) and Tableau
Modelings: Statistical and Machine Learning methodologies
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 with Python and R. Detail-oriented and highly motivated individual committed to hard work and intensive analysis.

WORK HISTORY

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

- Provided visual context and dashboard that describes 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) .Created a prediction model for type II diabetes incident by each ethnic sub-group.
- 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

- Comparative study by simulating statistical models in Blang and Python

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
- DNA and plasmid validation for secondary screening for antibody production for partners (Pfizer)
- Designed and simulated protein conformation in 3D for mouse humanization to increase antibody production

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

- Mutant cell-cryopreservation development: 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 visual contexts 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 describe models for the passenger's behaviours
- Logistic regression & decision tree Airline choice model: 82% accuracy, Airport choice model: 81% accuracy

LendingClub Investment Analysis & Borrower's Default Prediction (Python libraries: Pandas, Numpy and Scikit-learn)

- Integrated three large datasets and created loan default prediction models using Naïve Bayes, Decision Tree, Random Forest, Neural Network and Gradient Boosting: The best model (Neural Network with resampling technique) showed 66% accuracy, 67% recall and 66% AUC scores.

Shuris Coffeeshop Business Management Database (MYSQL and Tableau)

- Constructed database (DB) for coffeeshop business in MySQL – Designed relational diagram. ETL (Extract, Transform, and Load) data and generated a fully functional DB with business rules with pre-implemented functions from SQL queries