

Kenneth Lee

kennethleeje.com | +1 (518) 960 8390 | kennethleeje8@gmail.com

EDUCATION

Rensselaer Polytechnic Institute (RPI), New York

August 2019 – May 2023

Dual B.S. Information Technology & Web Science and Business Analytics

RELEVANT SKILLS

Languages: Python | SQL | HTML/CSS | JavaScript | R | C++ | Java | Swift | Latex

Technologies: Google Cloud | Microsoft SQL Server | Git | Apache Spark | MongoDB | Angular.js | Node.js

Skills: ETL Pipelining | NLP | API Configuration | Data Warehousing | Data Scraping

RELEVANT COURSES

Managing IT Resources | Data Resource Management | Data Structures | Web Science Systems Development
| Foundations of Computer Science | Econometrics | Computer Organization | Statistical Methods

PROFESSIONAL EXPERIENCES

Software Engineering Intern, IOI Properties Group Berhad

October 2021 - December 2021

- Developed a document extraction program using **RPA** software, **UiPath**, that collects 8 types of monthly reports to be saved locally for further processing
- Streamlined the document extraction process using **Task Scheduler**, saving 16 hours of document extraction for the client's sales department every month with no downtime
- Wrote custom Javascript code for functions in **UiPath** that populates document search details

Data Science Intern, Snapcart Global

August 2021 - October 2021

- Implemented a **Python** script as part of an existing **OCR** (Optical Character Recognition) process developed in-house to replace outsourced OCR service
- Worked in an **Agile** team to construct a model that matches receipt items to their corresponding price, quantity, and discount
- Utilized **Google Cloud Platform** to store over 4400 data points on individual retail items
- Used **scikit-learn Classification Report** to optimize the balance of accuracy, precision and recall of the model, improving the accuracy of the model from **83% to 94%**

PROJECTS

NYS Health Department Project

August 2022 - November 2022

Source Code: github.com/KennethLeeJE8/health-data-NY

- Extracted raw CSV data from NYS Health Department Restaurant Inspection database and utilized Microsoft **SQL** Server to build a **Data Warehouse** with over 21000 data points
- Created **Template Queries** and **Cubes** to abstract the data retrieval process

IOI Properties Data Scraping

June 2022 – August 2022

Source Code: github.com/KennethLeeJE8/datascraping-lowyat

- Built an **ETL Pipeline** to scrape over 2000 online forums comments to gain public consensus on IOI Properties Group's brand
- Filtered and formatted text data by keywords into Microsoft Excel with Python's **xlrd**

Amazon Sentiment Analysis

May 2021 - August 2021

Source Code: kaggle.com/code/kennethleeje8/amazon-sentiment-analysis

- Conducted Sentiment Analysis on Amazon reviews from Kaggle dataset using NLP tools such as **Bag-Of-Words** and **TF-IDF**
- Leveraged **GridSearchCV** to tune the hyper-parameters of the **SVC** and **Linear SVC** models used

Other Projects are detailed on my personal website: kennethleeje.com/portfolios/