Joe Bui

Education

Master's in Data Science | Eastern University | GPA: 4.0   Jan. 2021 – Nov. 2021

Bachelor's in Biochemistry | University of Waterloo | GPA last two years: 3.7   Sep. 2016 – Apr. 2020

Projects

Tasty.co Recipe Insights | GitHub   Dec. 2021

- Created a Tableau Dashboard providing insights on different cuisines and recipes from Tasty.co involving the most common ingredients used, nutritional values, and popularity

- Cleaned data and performed an exploratory data analysis using SQL (PostgreSQL)

- Web scraped recipes and cuisine data from Tasty.co using the Python library, Selenium

Graphical Search Method | GitHub | URL: lung-cancer-dmpymw37jq-ue.a.run.app   Nov. 2021

- Deployed a Plotly Dash Application on the Google Cloud Platform, where users may graphically search for scientific journal articles on a plot that is filtered by similarity search, date, publication type, etc., and can be selected and saved into a csv file

- Applied SBERT (NLP) and cosine similarity for similarity search of keywords to journal articles

- Web scraped 200k+ journal articles from PubMed using the Python library, BeautifulSoup4, and performed data cleaning/wrangling using Jupyter Notebooks (Pandas, NumPy)

Predicting Appliance Energy Usage in a Low Energy Building | GitHub   Jul. 2021

- Applied time-series regression models to predict future appliance energy usage within a low energy building

- Performed time-series data cleaning, exploratory data analyses and feature engineering (time periods, mean encoding and lags) using Python (Plotly, Pandas, NumPy)

Professional Development and Academic Services

Biochem Student Researcher | Dr. Guillemette's Lab | University of Waterloo   Sep. 2019 – Apr. 2020

- Determined efficacy of inhibitors for the enzyme, Nitric Oxide Synthase, by analyzing enzyme kinetic assay data for research and development in treating brain-degenerative diseases including Alzheimer's and Epilepsy

Co-Founder | Cultured (Start-up) | University of Waterloo, Velocity   Jan. 2017 – Mar. 2018

- Interviewed professors and graduate students on the workflow of research, the publishing and funding system, and the tools and skills required for replicability to investigate the problem of irreproducibility in academic research

- Designed an electronic lab notebook to improve organization and workflow efficiency of experimental procedures while preventing biases and poor habits in recording data

- *2nd Place Winner for Velocity Problem Pitch Competition W18:* Awarded $2500 to R&D by pitching the problem of irreproducibility in academic research, revolving the history, scope, and impact at an industrial level

- *1st Place Winner for St. Paul's Big Ideas Challenge S17:* Awarded $3000 and invited to the social entrepreneurship incubator, St. Paul's Greenhouse Fellows, by presenting the idea of sustainably synthesizing lab grown meat to replace the expensive costs of raising livestock for the contribution towards the sustainable UN world goal of Life on Land

---

Personal Information

- Kitchener, ON
- 226-507-2648
- dsjoebui@gmail.com
- linkedin.com/in/dsjoebui
- github.com/dsjoebui
- dsjoebui.github.io/Portfolio

Languages and Aptitudes

Python
(Pandas, NumPy, Scikit-learn, Selenium, BeautifulSoup4)

R
(Tidyverse, Statistical Analysis, Visualization, Modeling)

Data Visualization
(Tableau, Excel, Qlik, Seaborn, Matplotlib, ggplot, Plotly Dash)

SQL
(PostgreSQL, Queries, Nested Subqueries, Joins, Procedures, Triggers)

Deployment/Pipeline
(Git, Github, Heroku, AWS S3, Google Cloud)

Machine Learning
(Regression, Classification, Clustering, Time Series, NLP, TensorFlow)

Volunteer Work

UWaterloo Biochem Executive
GRSM Physiotherapy Assistant
A&K Pharmacy Assistant
Epilepsy Fundraiser Manager
(epilepsy SCO)

Hobbies and Interests

Volleyball
Reading
Cooking
5am club