Joshua (Yixuan) Ji

Seattle, WA · ji.yix@northeastern.edu · (206) 291 – 0050 · Github · LinkedIn

EDUCATION

Northeastern University **May 2021 – May 2023**

Master of Science – Data Analytics Engineering, GPA: 3.9

Coursework: Machine learning (Spark, PyTorch), Database Management for Analytics (MySQL, MongoDB), Data Analytics Foundations, Deterministic Operations Research, Data Visualization(R), Data Mining (Python)

University of Washington **Sep 2016 – Dec 2020**

Bachelor of Arts – Geography with Data Science

SKILLS

Data: Python (Beautiful Soup, Spark, PyTorch, MapRedeuce), SQL, NoSQL, R (ggplot2, dplyr, shiny), Neo4j (Cypher)
Software: Java, JavaScript (Leaflet), Kotlin, HTML, XML, CSS, D3.js
Management System: MySQL, PostgreSQL, MongoDB, Oracle DB, ArcGIS, MATLAB

GROUP PROJECTS

Northeastern University | Crack Detection Study **Nov 2021 – Dec 2021**

- Developed an end-to-end **machine learning model** with ResNet-50 to achieve 98.45% accuracy on crack detections
- Transformed over 40,000 crack images into Tensors in **Python(Pytorch, Numpy),** used the data to train different neural networks and cross-validated models for image processing and pattern recognition tasks
- Generated graphs on **RStudio (ggplot2, dplyr)** to visualize model accuracy and parameters tuning

Northeastern University | DBMS Use-Case Study **July 2021 – Aug 2021**

- Developed a database model that simulates the performance of online discussion platforms such as Reddit & Baidu to analyze stock market (Gamestop) fluctuations
- Designed EER & UML diagrams and implemented data models on **MySQL** and **MongoDB (NoSQL)** servers
- Generated data-driven insights through crawling over 1,000+ user profiles with **Python (BeautifulSoup)** and performed queries over 5,000+ data points in **SQL**, and **Terminal (Bash, Git)**

EXPERIENCES

University of Washington HGIS Lab | Data Research Assistant **Jan 2020 – Aug 2020**

- Built an interactive COVID-19 dashboard to **visualize** infection rates worldwide, worked alongside a research team to facilitate the University on raising awareness of the pandemic (hgis.uw.edu/virus/)
- Created data frames and performed regression models on Jupyter Notebook by utilizing **Python (Pandas, SciPy, scikit-learn)** to analyze COVID spreading trends
- Collected and transformed 50,000,000+ COVID data points into .csv files, managed data in both **SQL** database & Excel
- Performed **SQL** queries on datasets with 600,000+ lines to extract information tables used for data visualization tasks

OpptIn | Android App Developer Intern **July 2019 – Sep 2019**

- Built a testing framework for Android devices that included the initial version of APP via **Android Studio** which secured users' data and personal information
- Implemented blockchain algorithm in **Java** for location validation, used **XML** and **Kotlin** to design user interfaces

- Led team meetings to plan, execute, and manage progress using **GitHub** to ensure timely deliveries