Leon Zhang

Machine Learning Engineer & Data Scientist

Phone (US): (206) 487-6780 | E-Mail: lz198@duke.edu | Address: Durham, NC ; Vancouver, BC | Website | Linkedin | GitHub

Summary of Qualifications

Programming Languages: Python, Java, R, SQL, MATLAB, C++, JavaScript, HTML, CSS.

Software Tools and Skills: ML (Sklearn, PyTorch, TensorFlow, Keras, Hugging Face), Git (GitHub, GitLab), Containers (Docker), Cloud Computing (AWS, Azure, GCP), Spread Sheets (Excel, Google Sheet), Continuous Integration & Deployment (CI/CD), Flask Application, Data Visualization Tools (Tableau, Grafana), Data Warehouse and Orchestration (Snowflake, RapidMiner).

Relevant Coursework: Machine Learning, Deep Learning, Statistical Modeling, Natural Language Processing, Data Engineering Systems, Database Management Systems, Data Visualization, A/B Testing, Linear Algebra, Data Structures, Algorithms, Probability.

Certifications: AWS Certified Solution Architect – Associate [Credentials]

Professional Experiences

Data Science Intern, Windstream – Durham, NC **May. 2021 – Aug. 2021**

• Facilitated proactive live customer services through sentiment analysis by building, training, and packaging NLP transformer models for chatbot systems in PyTorch.

• Established access to topic classification and sentiment analysis models across different platforms through model deployment by programming REST API endpoints.

• Accelerated customer digital adoption by providing data-driven suggestions and building insight dashboards on user activities, remedy tickets, system message errors, and customer satisfaction using SQL, Python, RapidMiner, and Grafana.

• Enabled A/B testing and customer personas through building census demographic lookup of the customer base.

ML Software Programmer, Duke Health System – Durham, NC **Jan. 2021 – Present**

• Assist doctors in finding effective treatment solutions through modifying existing Bi-Clustering algorithms in R and Python, which match patient demographics with cancer symptoms.

• Researched and tested Bi-clustering algorithm applications in symptom patient study to help professor propose R01 grant.

Research Assistant, University of Washington – Seattle, WA **Jan. 2019 – Jun. 2020**

• Devised a deep learning model - variational autoencoder with the research team to explore chemical reaction pathways and predict intermediate chemical species using TensorFlow.

• Built a training dataset by computing all possible chemical species from reaction pathways using Python libraries such as NumPy, Pandas, etc.

• Implemented molecular rotational techniques with quaternion coordinate system in Python and C++ to visualize molecule movement in space and study interactions with different interfaces.

Education

Duke University, Durham, NC **Aug. 2020 – Apr. 2022**

Master of Science, Data Science (MIDS) Overall GPA: 3.78/4.00

University of Washington, Seattle, WA **Sep. 2016 – Jun. 2020**

Bachelor of Science, Chemical Engineering Overall GPA: 3.55/4.00 Computer Science GPA: 3.76/4.00

Projects & Competitions

2020 Duke Datathon – 1st Place [Link] **Oct. 2020**

• Collaborated in a team of four and achieved 1st place in presenting insights of COVID-19 economic impacts across the world and suggesting relieves using regression modeling and time series forecasting in R and Python.

• Designed a comprehensive metric using PCA that reflects the economic condition of a country over time by aggregating multiple economic indicators to perform modeling.

Movie Recommendation Web Application [Link] **Aug. 2020**

• Designed a visually appealing, scalable web application to provide movie recommendations using Flask, Python, JavaScript, HTML, and CSS.

• Adapted continuous integration and deployment for automated code test and production using cloud services from GCP.