

# Leon Zhang

*Machine Learning Engineer*

Phone (US): (206) 487-6780 | E-Mail: leon.zhang@duke.edu | Address: Durham, NC ; Vancouver, BC | [Website](#) | [LinkedIn](#) | [GitHub](#)

## Summary of Qualifications

---

**Programming Languages:** Python · Java · SQL · R · MATLAB · C++ · JavaScript · HTML · CSS

**Software Tools:** ML (Sklearn, PyTorch, TensorFlow, Keras, Hugging Face) · Cloud (AWS, Azure, GCP) · Git · Docker · ETL (Spark, MapReduce, Snowflake, RapidMiner) · CI/CD · Flask · Visualization (Tableau, Grafana)

**Skills:** Machine Learning · Deep Learning · Natural Language Processing · Statistical Modeling · Data Engineering · Database Management Systems (RDBMS, NoSQL) · Data Structures · Algorithms · Data Visualization · A/B Testing

**Certifications:** AWS Certified Solution Architect – Associate [\[Credentials\]](#)

## Education

---

**Duke University | Durham, NC**

**Aug. 2020 – Apr. 2022**

Master of Science, Data Science (MIDS)

Overall GPA: 3.83/4.00

- Duke Interdisciplinary Product Management Club: Co-Founder, [\[SQL Workshop Speaker\]](#)

**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

## Professional Experiences

---

**Data Science Intern | Windstream | Durham, NC (Remote)**

**May. 2021 – Aug. 2021**

- Improved chatbot customer services with transformer-based sentiment analysis models using PyTorch.
- Established REST API endpoints of deep learning models for cross-functional usage including chatbot and KPI dashboards.
- Accelerated customer digital adoption by 30% through communicating insights and building dashboards on user activities, remedy tickets, IVR routing, and customer satisfaction using SQL, Python, statistical modeling, and Grafana.

**ML Software Programmer | Duke Health System | Durham, NC**

**Jan. 2021 – Present**

- Assist doctors in finding targeted cancer treatment through implementing clustering algorithms using Python and R.
- Formulated a new algorithm that produces symptom-patient block clusters after effective communication with stakeholders on the research goal.

**Research Assistant | University of Washington | Seattle, WA**

**Jan. 2019 – Jun. 2020**

- Devised a variational autoencoder with the research team to explore chemical reaction pathways and predict intermediate chemical species using TensorFlow, NumPy, and Pandas.
- Implemented rotational techniques with quaternion coordinate system to visualize molecule interactions in Python and C++.

## Projects & Competitions

---

**Reinforcement Learning for Algorithmic Stock Trading** [\[Link\]](#)

**Sep. 2021**

- Develop trading strategies from deep reinforcement learning algorithms using PyTorch and StableBaseline3.
- Designed a simulated trading environment that provides reinforcement learning algorithms with feedback using OpenAI Gym.

**2020 Duke Datathon – 1<sup>st</sup> Place** [\[Link\]](#)

**Oct. 2020**

- Collaborated in a team of four in presenting insights of COVID-19 economic impacts and proposing relieves using time-series and regression models in R and Python.
- Built a comprehensive metric using PCA that reflects the economic condition of a country over time to perform modeling.

**Movie Recommendation Web Application** [\[Link\]](#)

**Aug. 2020**

- Launched a full-stack web app on GCP to provide movie recommendations using Python, JavaScript, Flask, HTML, and CSS.
- Automated code testing, deployment, and scaling pipelines through CI/CD practices.