

# Sohaib Syed

Chicago, IL • [Website](#) • 773-876-7254 • sohaibsyed6@gmail.com • [LinkedIn](#)

---

## EDUCATION

ILLINOIS INSTITUTE OF TECHNOLOGY, CHICAGO, IL

May 2024

- Bachelor of Science, Computer Science, **3.91/4.0**
- Master of Applied Science, Data Science, **3.85/4.0**

Relevant Courses: Machine Learning, Applied Statistics, Data Mining, Data Prep & Analysis, Big Data Tech

---

## EXPERIENCE

*Machine Learning Intern*

May 2022 - August 2022

**Motorola Solutions Inc., Applied Technology** (Python, PyTorch, Bash)

Schaumburg, IL

- Reduced 1000+ lines of code: extracted stand-alone Wav2Vec2.0 model from Hugging Face and reduced fine-tuned model delivery time by 1-mo, which led to increased productivity to complete 5+ experiments vs. 1 to improve Wav2Vec2.0 for Voice Activity Detection
  - Experimented with building 2 additional model versions, including a CNN and MLP layer, with PyTorch
  - Produced score files to classify every 20ms as speech or non-speech to simplify model evaluation
  - Documented detailed steps from model extraction to model evaluation to reduce time to reproduce results by 40%
- 

## PROJECTS

**Late Tax Payment Classifier for Cook County Treasurer's Office** (Python, PySpark, Jupyter Notebooks)

- Collaborated with 3 subject-matter experts for 8 weeks with a team of 5, developed classifier with 75% recall to determine which properties will miss payment date, increased predictive capabilities 100%
- Provided ability to allow CCTO to send additional warnings in advance to targeted individuals potentially saving 90% on mailing expenses, additional potential of collecting 100% of taxes on time
- Presented weekly to CCTO to report progress leading to on-time delivery of product

**Credit Card Fraud Detection Analysis** (R, RStudio)

- Rebalanced training set using SMOTE in R to achieve a 65/35 split between positive and negative class from the original 99.9/.1
- Led a team of 4 by delegating tasks, creating schedules and meetings, and producing deliverables all by deadline
- Improved accuracy of the logistic model and decision tree model by 12% and 6% respectively
- Analyzed residuals and correlation to reduce feature size for simpler and more efficient models

**Employee Database** (SQL, Java, DBeaver, PostgreSQL)

- Implemented a SQL script to create a database with 7 tables used to store employee data
  - Programmed a Java application to communicate with a PostgreSQL database to keep and update employee information
  - Configured database to allow admin, manager, and employee users with different information privileges for improved security
  - Managed a team of 3 by organizing meetings and delegating tasks to meet project deadline
- 

## SKILLS AND INVOLVEMENT

Programming Languages: Python, R, SQL, Java

Technologies: PyTorch, TensorFlow, PySpark, Jupyter Notebooks, Git, DBeaver, PostgreSQL, Linux/Bash, AWS: S3, EC2, EMR

Volunteer: Weekly volunteer at Sacred Learning Center, SLC Tech Group