TOM O'CONNELL

tomoccollege@gmail.com | (623) 888 9955 | New York, NY

EDUCATION

Arizona State University

- **Major:** Data Science, BS GPA: 3.6 / 4.00
- **Certificates:** Applied Business Data Analytics and Mathematical Concepts of Engineering
- **Coursework:** Foundations of Machine Learning, Statistical Modeling and Inference, Applied Linear Algebra

PROFESSIONAL EXPERIENCE

Oak Street Health, Machine Learning Engineer Apr 2024 – present

- I researched and developed ML models for identifying early onsets of medical conditions, progression of diseases, and various other medical-related insights of patients.
- I utilized the OpenAI API to create a LLM that can be queried to retrieve specific health related insights from patient's medical documents within OSH clinics.
- The LLM is used by 100's of doctors and practitioners to establish correlations between health defects and the progression of certain diseases in 170 clinics across 21 states.
- I improved performance by testing different OpenAI models and temperatures. My team and I also improved the LLM BLEU scores from 0.54 to 0.76 by implementing unit tests. I mitigated hallucination and toxicity, and improved RAG retrieval tests by utilizing methods such as chunking, prompt permutation, reranking, and zero shot learning.
- I utilized Azure Databricks, leveraging libraries such as PyTorch, PySpark, Pandas, and NLTK for data processing and natural language understanding. I utilized also DeepEval to measure unit testing metrics.

Intel, Software Developer Aug 2022 – Apr 2024

- I worked on the development team for a LLM model called 'ChipChat'. The LLM reduces downed machinery time by 9.3% by helping technicians diagnose problems with tools throughout the FAB Labs. The LLM utilizes OpenAI API, Scapy, Beautiful Soup, and NLTK. Currently being tested with over 100 technicians within Intel FAB labs in Arizona.
- I worked on the development team in the development of numerous Intel websites, including a scheduling app used by 3000+ technicians and 200+ managers in FAB labs.
- The app allows managers to assign tasks to their technicians in an agile environment, improving production and communications across the FAB Labs. The front-end was built using React.js.
- I worked in a dynamic developer team environment utilizing Azure and Git to effectively streamline sprint planning, coordination and version control during the development process.

Hindman Auctions, IT Analyst Apr 2021 – Aug 2022

- I worked as a Software Developer with an external development team in the production of 'Gavelizer', an application built for Hindman Auctions to effectively catalogue their inventory. I developed full stack using React.js, ExecJS and Ruby. App is used by 270 Hindman employees daily.
- I also directly worked with Cisco Systems in the implementation of Cisco Meraki, a cloud-based application which allows the management of multiple systems, servers, and security devices within Hindman.
- I used RubyOnRails, Administrate, Carrierwave, Sidekiq, Docker and Linux.

PROJECTS

Machine Learning and Credit Default (Thesis)

- My team and I used the loan applicant data provided by the Home Credit Group to create a ML model that could identify the applicants who are most likely to default using both Supervised and Unsupervised techniques.
- Using over 287 features of applicants, we tested four different algorithms: Logistic Regression, XGBoost, Neural Networks, and Random Forest. We evaluated our classification models using the following metrics: Accuracy, Precision, Recall, F1 Score, and AUC Score. Our XGBoost model performed the best of the models with 92.2% accuracy and an AUC score of 0.8412.
- Used TensorFlow, Scikit, Pandas, NumPy, Seaborn, and Sklearn.

SKILLS

Languages: Python, C/C++, JavaScript, R, C#
Libraries: Spark, PyTorch, NLTK, Pandas, TensorFlow, React.js, Matplotlib
Database Management: SQL (MySQL, MariaDB), Cisco Meraki, Dbeaver
Models/Algorithms: LLMs, XGBoost, Random Forest, Artificial Neural Networks