Tony DiRubbo

+1 (315) 876-6023 ● afdirubb@syr.edu ● linkedin.com/in/tonydirubbo ● adirubbo13.github.io

Education

Syracuse University, School of Information Studies, Syracuse, NY   *May 2026*
M.S. Applied Data Science *Secondary Core:* Artificial Intelligence – 4.0 GPA
Relevant Coursework: Software Engineering for AI Applications | Deep Learning | Applied Machine Learning
Natural Language Processing | Scripting for Data Science | Big Data Management and Concepts

St. John Fisher University, Rochester, NY, *Summa Cum Laude*   *May 2024*
B.S. Media Management | Minor in Data Analytics – 4.0 GPA
Relevant Coursework: Foundations of Computer Science | Web Design | Database Management and Security

Applicable Technical Skills/Certifications

Python: TensorFlow, PyTorch, Scikit-learn, NumPy, Pandas, Matplotlib, NLTK, Flask, spaCy, CUDA
Software Development: Java, C, C++, Swift, Git, Docker, Microsoft Azure, MongoDB
Microsoft Excel: Pivot Tables, VLOOKUP, Statistical and Solver Add-On Packages, 3D Mapping, VBA
SQL Relational Databases: MySQL, SQLite, MS SQL Server, PostgreSQL, Microsoft Access, pyodbc
Web Development: JavaScript, HTML + CSS, WordPress, Google Analytics

Work Experience

Global Master Data / Project Analyst Co-op, CooperVision, Rochester NY   *March 2025 – August 2025*
- Utilized Java, Visual Basic, and Oracle to develop applications to assist with the data pipeline
- Constructed Python scripts to perform large-scale data changes and updates
- Reviewed and signed-off on data entry/changes to ensure data governance between regional systems
- Pulled data for inquiry and analysis for finance, sustainability, and customer service teams using SQL

Graduate Student Researcher, Nexis Technology Lab, Syracuse NY   *January 2025 – Present*
- Developing adaptive AI resources to enhance student learning and engagement
- Collaborating with a team of five graduate students, meeting weekly to coordinate research efforts
- Leading semester-based projects focused on consistent AI themes across multiple terms
- Presenting research progress and outcomes at end-of-semester poster symposiums

Project Experience

Open-Source Model Distillation *Natural Language Processing*   *Spring 2025*
- Developed a paper on how to utilize a large language model to produce smaller, specific, models
- Trained a model from Meta's Llama 3.1 and Google's BERT to identify spinal cord diseases
- PyTorch was utilized to establish the model's framework, with CUDA allocation for efficient processing
- Presented findings at a keynote presentation, walking through the overall process of distillation

Machine Learning to Predict Diabetes *Applied Machine Learning*   *Spring 2025*
- Built ML pipelines to predict Type II diabetes using non-clinical data from the CDC BRFSS survey
- Engineered and encoded 28+ features; applied SMOTE and PCA to handle high dimensionality
- Compared classification models (Logistic Regression, Random Forest, XGBoost, Neural Networks)
- Applied cross-validation, threshold tuning, focal loss, and synthetic sampling to improve performance

Leadership Experience

Boys Cross Country Coach, Skaneateles Central School District, Skaneateles, NY   *August 2024 – Present*
- Assists with the development and leading of practices and coordination of bus travel
- Practices are planned on a weekly basis, with the week's competition in mind
- Contributes to runners achieving faster 5km times and higher race placements