NICHOLAS TEDESCO

nicholas.r.tedesco@gmail.com | 267-772-0490 | github.com/nrtedesco | linkedin.com/in/nicholas-r-tedesco

TECHNICAL SKILLS

Programming Languages: [Proficient] – Python, SQL (SQL Server), C / C++, Bash, R. [Familiar] – Java, JavaScript, Golang.
AI & Machine Learning: PyTorch, TensorFlow, Scikit-learn, LightGBM, Hugging Face Transformers, Deep Learning, NLP (BERTopic), LLMs & Generative AI, RL (Gymnasium), Anomaly Detection, MLflow, Optuna, SHAP.
Distributed & Parallel: PySpark, OpenMP, MPI, CUDA.
Data & Analytics: Databricks, Snowflake, Pandas, NumPy, Matplotlib, Power BI, Tidyverse.
Cloud & Dev Tools: Azure, Git / GitHub, Conda, Docker, CI/CD, Pytest, Linux, Vim, YAML.

WORK EXPERIENCE

Humana – MarketPoint Remote
Data Scientist II 09/2024 – Present

- Eliminated $3M in annual vendor fees by designing and deploying an end-to-end machine learning system, taken from prototype to production, to rank Medicare Advantage brokers for contact and replace external contracting dependencies with a custom automated ML and ETL pipeline.
  [Tech] – Python (Scikit-learn, LightGBM, SHAP, MLflow, Pytest) | Databricks | Snowflake | SQL Server | Git / Azure DevOps
- Developed a leader-facing dashboard to surface trends in the Medicare Advantage market by applying transformer-based NLP (neural topic modeling) to 80,000+ call transcripts between Humana sales managers and external brokers.
  [Tech] – Python (BERTopic, Transformers, PySpark) | Power BI | Databricks
- Evaluated the impact of a company-wide AI-based marketing intervention on Medicare Advantage sales by leading an A/B test (clustered randomized controlled trial) involving 60+ Humana sales managers.
  [Tech] – R (lme4) | Python (Scikit-learn) | Power BI | SQL Server

University of Michigan School of Medicine Remote
Data Scientist 03/2023 – 09/2024

- Built an ETL pipeline and regression framework to analyze the All of Us dataset from the National Institutes of Health (200K+ patients), resulting in a publication highlighting specific social determinants of health as predictors for liver-related disease.
  [Tech] – R (Tidyverse, BigRQuery, Tidymodels) | Google Cloud Platform | Python (Matplotlib, Seaborn)
- Developed a predictive modeling pipeline using penalized regression (LASSO) to identify speech-related predictors of hepatic encephalopathy (HE), contributing to a novel scoring system evaluating HE risk.
  [Tech] – R (Caret, REDCapR, ggplot2) | Git / GitHub

PROJECTS

Reinforcement Learning: implemented Deep Deterministic Policy Gradient (DDPG) from scratch in PyTorch to solve the continuous Lunar Lander environment (Gymnasium), achieving an average reward of +258.25 over 100 evaluation episodes.

High Performance Computing: parallelized the list ranking algorithm (Wyllie, Helman-JaJa) in C++ using OpenMP and CUDA, achieving a maximum speedup of 13X over the naïve sequential implementation on a 24-core compute cluster.

BidCentral: developed and containerized a full-stack application to simulate an online auction platform. Product included components for database (PostgreSQL), backend API to query database (Python Flask), and frontend user interface (Flask + JavaScript).

EDUCATION

Georgia Institute of Technology Atlanta, GA
MS in Computer Science – Machine Learning (GPA: 4.0) 08/2023 – 08/2026 (Expected)

University of Pittsburgh Pittsburgh, PA
MS in Biostatistics – Health Data Science (GPA: 4.0) 08/2021 – 12/2022
BS in Biochemistry, BA in Writing (GPA: 3.87) 08/2017 – 05/2021

CERTS: TensorFlow Developer, Coursera Stanford ML, Coursera Statistics with Python.
AWARDS: Humana STAR Award, Delta Omega Honorary Society, Brackenridge Research Fellowship, National Merit Scholar.