DANIEL TSHIANI

Machine Learning Engineer

DTshiani@Gmail.com
https://DGT-International.com
https://Github.com/DGT-INT
www.Linkedin.com/in/Daniel-Tshiani/

EDUCATION

American University, Washington, DC
Master of Science: Data Science, 2025
Overall GPA: 3.9

Loyola University Maryland, Baltimore, MD
Bachelor of Arts: Quantitative Economics, 2022
Minor: Mathematics; Major GPA: 3.5

Montgomery College, Rockville, MD
Associate of Arts: General Studies, 2020
Overall GPA: 3.8

PROGRAMMING LANGUAGES

- Python
- R
- STATA
- SQL
- HTML
- CSS

ML FRAMEWORK

Supervised Learning: Linear & Logistic Regression, Ridge & Lasso Regression, Decision Trees, Random Forest, Support Vector Machines (SVM)

Unsupervised Learning: Clustering (K-means, Hierarchical), Principal Component Analysis (PCA), Independent Component Analysis (ICA), Canonical Correlation Analysis (CCA), Independent Vector Analysis (IVA)

Deep Learning: Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Generative Adversarial Networks (GANs), Graph Neural Networks (GNNs)

OTHER SKILLS

Data Visualization
- Tableau
- PowerBI

Tools
- Github
- Excel

Languages
- English
- French

PROFESSIONAL SUMMARY

Machine Learning Engineer with expertise in developing and deploying end-to-end ML solutions using Python, TensorFlow, Keras, PyTorch, and SQL. Experienced in automating data workflows, building predictive models, and applying MLOps principles for scalable, reliable systems.

FEATURED PROJECTS

MLS Soccer Player Recruitment App | American University | 2025

- Reduced scouting costs by ~30% and expanded talent coverage by ~50% by developing a full-stack application that integrates API data and empowers clubs to make data-driven roster and long-term value decisions.
- Built customizable KPIs and advanced search functionality, enabling scouts to filter and compare players by performance metrics and positional needs.
- Implemented and refined a predictive salary model, improving accuracy (R^2) from 0.32 to 0.86 by advancing from a linear regression baseline to tuned advanced algorithms, which enhanced the identification of undervalued players and supported data-driven recruitment.

Tech Stack: R-Shiny, tidyverse, caret, ggplot2, API Integration, Random Forest, XGBoost

Credit Card Approval Prediction App | American University | 2025

- Automated a credit card approval pipeline, reducing review time by 60% while standardizing decisions and ensuring compliance with financial regulations.
- Implemented Random Forest models with feature importance analysis to eliminate demographic bias, achieving a 25% reduction in disparate impact score
- Improved credit approval predictions by addressing class imbalance with SMOTE, which increased predictive accuracy by 35%, contributing to a 20% reduction in projected financial losses and an 18% higher precision for qualified applicants.

Tech Stack & Methods: R, tidyverse, caret, Random Forest, SMOTE, F1 Score

PROFESSIONAL EXPERIENCE

Professor Assistant

American University | Jan 2024 – Present | Washington, DC

- Supported over 20 students by clarifying AI algorithms and guiding thesis development through feedback on research design and methodology.

Quantitative Research Assistant

American Institute of Research | July 2024 – July 2025 | Washington, DC

- Automated large-scale data processing workflows in R and STATA, improving accuracy, efficiency, and quality assurance in education program evaluations.
- Designed and analyzed surveys for the PASCOFI project, applying statistical methods to evaluate cash transfer programs that informed evidence-based policy decisions.
- Contributed to a randomized controlled trial on multilingual teacher development, conducting data cleaning, statistical modeling, and impact evaluation to provide actionable insights for partners, including Dalberg and IDRC

Research Assistant

KNG Health Consulting, LLC | April 2023 – July 2024 | North Bethesda, MD

- Built county-level geospatial analyses and visualizations to evaluate Medicare beneficiaries' reliance on Hospital Outpatient Departments, delivering insights to the American Hospital Association.
- Conducted statistical analysis of CPT code utilization trends (e.g., 92523) using Medicare claims data, providing evidence-based recommendations on healthcare access and cost patterns.

WORK IN PROGRESS

Football Player Tracking & Performance Analysis System (Expected Completion: Dec 2025)

Developing a computer vision and machine learning pipeline to analyze player movements, team assignments, and performance metrics from video footage. The system uses YOLOv8 and custom-trained detectors for player, referee, and ball detection, KMeans pixel clustering for team assignment, and optical flow with perspective transformation to measure player movement, detect team formations and pressing patterns, providing coaches actionable insights to anticipate and break down the opposition's press.