

eamonn.tweedy@gmail.com | mobile: (919) 280-1782  
Professional website: e-tweedy.github.io  
linkedin.com/in/eamonn-tweedy | github.com/e-tweedy |

- **Python libraries:** Pandas, NumPy, Scikit-learn, SciPy, PyTorch, XGBoost, Statsmodels, Matplotlib, Seaborn, Plotly, NLTK, Hugging Face Diffusers and Transformers, Streamlit, Gradio
- **Machine learning and statistical models:** linear models, KNN, SVM, decision trees, clustering, ensemble models, MLP, CNN, RNN, GAN, LLMs, encoders and decoders, transformers, diffusion, ARIMA, (Neural) Prophet.
- **Data science principles:** dimensionality reduction e.g. PCA, LDA, manifold learning; data cleaning and feature engineering; and data exploration and visualization
- **Math expertise:** Calculus and differential equations, linear algebra, abstract algebra, statistics and probability, graph theory and discrete math, geometry, and topology.
- Proficiency in SQL (PostgreSQL, MySQL), LaTeX, Git, Microsoft Office Suite

- **Built BikeSaferPA**, a machine learning model which predicts the severity of bike crashes in Pennsylvania.
  - Procured and cleaned PENNDOT cyclist crash data from 2020-2021, created data visualizations which reveal prevalence of crash factors and their influence on severity, and designed a feature engineering pipeline.
  - Selected salient input features using a baseline logistic regression model, and then selected BikeSaferPA - a gradient boosted decision tree model - using a cross-validation process and randomized search hyperparameter optimization.
  - Explained BikeSaferPA's predictions and its feature importances using a SHAP value analysis, and articulated concrete recommendations for improving cyclist crash outcomes in Pennsylvania based on my findings.
  - Designed a BikeSaferPA web app, allowing users to visualize the data and experiment with the model.

View my project on GitHub: <https://github.com/e-tweedy/BikeSaferPA>

- Trained a **U-Net model brain tumor region segmentation** on MRI image data.
  - Trained a 3-dimensional U-Net, a convolutional neural network architecture, on the BraTS 2020 challenge dataset using significant data augmentation to avoid overfitting.
  - Implemented test-time augmentation for more robust model predictions; on the holdout test set, the model achieves mean Dice scores of 88%, 80%, and 74% on the whole tumor, tumor core, and enhancing tumor regions, respectively.

- Built a **RoBERTa language model for extractive question answering** by fine-tuning a base model on v2 of SQuAD (Stanford question answering dataset).
  - The model achieves approximately 80% exact-answer-match accuracy on the evaluation dataset.
  - Designed a web app which demonstrates both standard Q&A functionality and Wikipedia-assisted Q&A functionality.

View the fine-tuning project on GitHub: <https://github.com/e-tweedy/roberta-qa-squad2>

Widener University | 2019-2023  
Widener University | 2014-2019  
Rice University | 2011-2014

- Planned, coordinated, and executed individual and collaborative research projects in math. I have authored or co-authored seven peer-reviewed academic articles published in national and international math journals and presented my research at invited seminars and national conferences.  
View my Google scholar profile: [🔗](#)
- 12 years of experience developing and teaching advanced math courses to undergraduate and graduate students, earning outstanding teaching evaluations from students and praise from colleagues.
- Chaired a University-wide faculty Committee on Technology and Instructional Resources at Widener (2019-2022).

Univ of California Los Angeles | 2006-2011 | Cum. GPA 3.848  
North Carolina State Univ | 2002-2006 | Cum. GPA 3.929