FISHER MORITZBURKE

DATA SCIENTIST

fmoritzburke@gmail.com
snisher.github.io
(415) 342-1513
San Francisco, CA
fisher-moritzburke-536b01162
snisher

SKILLS

PROGRAMMING LANGUAGES: Python, Julia, C++
SOFTWARE LIBRARIES: Scikit-learn, TensorFlow, pandas, NumPy, Flux , PyTorch Geometric
DATABASES: PostgreSQL, MongoDB

EDUCATION

University of California at Santa Cruz

BS Neuroscience 2019
Minor Computer Science
Related Coursework: machine learning, machine learning & data mining, probability theory, database systems, vector calculus, data structures

EXPERIENCE

Metis Mar. 2020 to June 2020 · San Francisco, CA
Data Scientist

Completed an immersive 12 week full-time data science bootcamp focused on developing skills in Python, Machine Learning, Statistical Modelling and Data Visualization. Designed, implemented and presented five end-to-end data science projects demonstrating concepts such as data acquisition, supervised and unsupervised learning, classification, NLP and deep learning.

Predicting Which Genes are Associated with Diseases

- Trained a graph neural network using PyTorch Geometric to predict whether a gene is associated with a disease, formulated as link prediction problem on a heterogeneous graph (built with NetworkX) of 519 diseases and 7,294 genes. Achieved F1 score of 0.71.
- Processed clinical descriptions of diseases and genes into feature vectors using sklearn's TF-IDF vectorizer after removing stopwords.
- Disease-gene associations can help focus research into new drugs and therapies, reducing costly experimental validations.

ADHD Diagnosis Prediction from Functional Brain Connectivity Matrices

- Trained a soft-voting ensemble classifier (logistic regression, SVM with RBF kernel, and naive Bayes) to predict whether someone will be diagnosed with ADHD based on brain connectivity. Achieved an AUC ROC score of 0.774.
- Network statistics (degree, clustering coefficient, and closeness) were calculated using NetworkX and used as features along with connection strengths.
- Feature importance was determined using SHAP, providing physiological insights about which brain connections are associated with ADHD.

Presidential Speech Analysis

- Presidential speeches were processed with spaCy and topics were analyzed using Gensim's LDA implementation, and CorEx, which allows for topic seeds for a specific topic of interest.
- Stored data in MongoDB on an AWS EC2 instance.

Buck Institute for Research on Aging June 2017 to Aug. 2017 · Novato, CA
Research Intern

- Developed skills with core laboratory techniques such as Western Blot, Polymerase Chain Reaction and Immuno-Cyto Chemistry
- Managed 30+ laboratory experiments on Huntington’s disease bio-markers and effectively recorded and communicated the results