Shivam Badkas

+1(647) - 636 - 8628 • shivambadkas@gmail.com • www.shivambadkas.com

EDUCATION

University of Toronto | Honours BSc. Double Major in Statistics and Global Health   June 2020

Relevant Courses: Immunology, Physiology, Global Health Policy, Methods of Data Analysis, Probability

Vice-President, Veg Club: Leader of a student group with 14 executive members and 1000+ general members

PROJECTS (2022)

Feature Selection for Diagnosis: Using tumour dataset; determined which predictors were significant using visualisations, recursive random forests, SelectKBest, Recursive Feature Elimination, then used PCA for feature extraction

Effect of Literacy on Covid-19 Death Ratio: Used pandas and scikit-learn to explore state-wise Covid-19 data and determine significance of education policy measured by literacy rate on ratios of cases to deaths using linear regression

Classification of Disease using sklearn and pandas: Used Decision Trees, Logistic Regression, K-NN, and Random Forests to classify presence of disease. Used train_test_split to separate training data and used ROC curve/AUC to evaluate models

Prediction of Medical Insurance Charges by Regression: Predicted medical insurance costs of families by multiple predictors using linear regression, L1 and L2 regularisation using ridge and lasso regression, and random forest regressor

Dimensionality Reduction using PCA and t-SNE: Standardised MNIST 28x28 hand-written digit dataset and used PCA to find number of principal components that explain variance, then used t-SNE to visualise how local structure is preserved

EXPERIENCE

Data Analyst   June 2020 - October 2020
GoalMogul| Toronto, Canada
- Created event tracking graphs, funnels, and models on Amplitude to visualise metadata from ~100 test clients
- Used visuals to observe trends in order to better understand user experience, diagnosed flaws of tracking framework, recommended events to track in the next sprint for deeper insights
- Coordinated task backlog with engineers across the stack in order to get the right analytics events tracked
- Managed product analytics dashboard on Asana to create and organise tables which held detailed descriptions of product user tracking events and to organise analytics objectives and goals

Research Assistant   June 2018 - September 2018
Stony Brook University, Shah Lab |New York, United States

Understudied for doctoral candidate for 4 months, assisted with live model training and data collection, was solely responsible for novel machine used to measure rat model gait, used software to complete analysis, wrote instructional documentation, presented to committee of lab professor and doctoral students on research findings

Key Accomplishments:
- Co-authored "Comprehensive investigation of forelimb kinematics during overground locomotion in non-injured rats"
- Took holistic ownership of the data collection efforts. This included proactively fixing procedural problems, and preparing a best-practices guide for future researchers
- Tested eleven live rat models on MotoRater to quantify gait cycles and determine baseline performance which allowed for comparison to post-treatment tests
- Analysed collected data in SIMI Motion Analysis to determine patterns of locomotion and interpret results of ~10 parameters which were present in the final paper

SKILLS

- **Data Science:** R, ggplot, dyplr, INLA, Python, pandas, numpy, matplotlib, sklearn, MySQL, BigQuery, HTML/CSS
- **Technologies:** AWS, Git, Excel, Amplitude, Unix, Jupyter, Tableau, PowerBI
- **Certifications:** Google Data Analytics Professional