Jie Dong

515-525-5372 | jie.dong.00@gmail.com
Lowell, MA
LinkedIn | GitHub | Portfolio

SUMMARY

I'm a data scientist with background in mathematics and statistics. My Interests are astronomy, martial arts, data visualization. In my previous roll, I worked collaboratively on a team because the projects had different timelines and many moving pieces. I am confident that my passion for technology and training in mathematics will serve me to excel as a data scientist.

SKILLS

Category 1: Mathematical Modeling, Experimental Design, Probability Theory
Category 2: Imputation, EDA, Transformations, Feature Engineering, Dimensionality Reduction, Regression, Classification, Data Structures, Supervised and Unsupervised Learning.

PROJECTS

Boston Marathon | *GitHub*: Boston Marathon EDA with Stats
- Here I implemented some concepts form EDA and statistics. I set up a null hypothesis test to see how significant the difference in performance between males and females is.
- Tech: Jupyter Notebook, Pandas, NumPy, Seaborn

UFC Fight Prediction | *GitHub*: Supervised UFC Fight Model
- I used supervised learning algorithms to build a model that can predict the winner of a given UFC bout. Random forest classifier worked great on this data set. During cross validation, the model performed better than the current UFC model.
- Tech: Python, Seaborn, scikit-learn, Plotly

Anomaly Detection | *GitHub*: Unsupervised Anomaly Detection
- In this project I used ideas from unsupervised learning and anomaly detection to see if I can predict credit card fraud. I compared several unsupervised models. The idea here is algorithms that use non-linear dimensionality reduction will capture hidden correlations not captured by algorithms based on linear algebra.
- Tech: Python, TensorFlow, Keras, scikit-learn

EXPERIENCE

THINKFUL — Remote
Data Science Apprenticeship — 2019 -present
- Data visualization, model preparation, statistics and probability with Python, SQL for data scientists, experimental design, regression, classification, unsupervised learning.

EDUCATION

UNIVERSITY OF MASSACHUSETTS LOWELL — Lowell, MA
BS, Mathematics — Fall 2018
- For my senior seminar I used Heat Equations to model conditions for spontaneous combustion resulting from mass storage of organic materials.

MIDDLESEX COMMUNITY COLLEGE — Bedford, MA
AS, Physical Science — Spring 2016
- Related coursework: Linear Algebra, Differential Equations, Multivariable Calculus