Tianyu (Frank) Zeng

Mobile : 206-512-4349

Email : tzeng6@wisc.edu

frankzengtianyu.rocks

Education

- **University of Wisconsin - Madison** — Madison, WI
  Bachelor of Science (BS) in Computer Science and Statistics; ***GPA:*** *3.43/4.0* — *Jan. 2016 – Dec. 2019*
- **Relevant Coursework:**
  Data Structure, Algorithm, Software Engineering, Data Science in R, Regression Analysis, Machine Learning

Experience

- **HelpCare Connect - Werner Lab** — Madison, WI
  Software Developer/Research Assistant — *Aug 2018 - Present*
  - Prototyping and implementing a Decision Tree model in Scikit-Learn for caregivers by predicting potential diseases based on symptoms.
  - Creating an API to wrap the machine learning model in Flask, enabling users to request predictions.
  - Developing and maintaining "Forum" features in PHP and collaborated with other developers to improve functionality of the existing website.
- **CUNA Mutual Group** — Madison, WI
  Data Science and Analytics Intern — *Jun 2018 - Aug 2018*
  - Implemented Gradient Boosting Model to predict the severity of automobile insurance claims and reduced MSE by 15%. Constructed 5x2 cross validation to evaluate the performance of different models.
  - Developed custom dashboards for visualizing results from LDA topic models and sentiment analysis in R shiny and adopted by peer stakeholders to drive potential products and business impact.
  - Researched and formulated approaches based on the sentiment analysis in order to provide insightful solutions for over 600 credit unions from a customer perspective.
  - Performed EDA and feature engineering on large data sets in Spark/Hive to facilitate the development of statistical models as well as improving model performance and flexibility.
- **USDA Forest Laboratory** — Madison, WI
  Undergraduate Research Assistant — *Apr 2017 - Aug 2017*
  - Implemented bilinear interpolation algorithm in Julia to enhance the image resolution for ML algorithms (neural nets) to recognize same wood images taken by different cameras.
  - Automated data-entry of over 10,000 wood specimens features in Python and decreased the documentation workload from one week to two days.

Skills

- **Programming languages:** Java, Python, R, SQL.
- **Machine Learning:** Regression, SVM, KNN, Decision Trees, Naive Bayes, Random Forest, etc.
- **Statistical Analysis:** A/B Testing, Hypothesis Testing.
- **Big Data Technologies:** Exposure to Spark/Hive, basic understanding of MapReduce.
- **Others:** Git, GitHub, Agile Development.

Projects

- **NYC Taxi Fare Prediction:** Trained a LightGBM model to predict taxi fare amounts given locations, time, passengers, etc. Currently ranked top 25%
- **Trump vs Stock:** Analyzed the sentiment of Trump's tweets and applied PCA to investigate potential effects on stock volatility and visualized the result on R shiny.