ZHAOYUN MA

(412) 596-3997 • mazhaoyunyh@gmail.com • LinkedIn • GitHub • Columbia, SC

SUMMARY

Highly motivated data scientist with advanced research skills and ability to communicate complex concepts across business units. Seeking for opportunities to leverage data science, machine learning skills and research background. Strong technical, business, and problem-solving acumen. Content creator of "66 Days of data" on GitHub, Medium and LinkedIn. US Permanent Resident.

EDUCATION

Ph.D. in Mechanical Engineering, *South Carolina, SC* 2020
University of South Carolina, Overall GPA: 4.0/4.0,

Master of Science in Civil Engineering, *Pittsburgh, PA* 2016
University of Pittsburgh. Overall GPA: 3.9 /4.0

PROFESSIONAL EXPERIENCE

University of South Carolina, *South Carolina, SC*
Research Associate 01/2020 – present

- Spearheaded the DOE material evaluation project to detect material degradation and prevent nuclear structural failure
- Incorporated a dispersion curve based ultrasonic method for material/coating thickness evaluation and achieved micrometer scale of 3 µm (compared to traditional ultrasonic methods ~0.5 mm)

Research Assistant 01/2017 – 12/2020

- Led the NASA advanced composite project at UofSC to evaluate and further develop ultrasonic method reliability on composite structures and cocreate a Nondestructive Evaluation Handbook for future researchers and engineers. Journal Paper
- Developed a fully automated laser inspection system combined with effective algorithms for damage detection purpose
- Invented a network imaging method for quantification of complex structural damage US patent, Journal Paper
- Mentored 3 high school students in Summer Program for Research Interns 2017-2019 to develop their research interests
- Taught the mechanical engineering Lab courses to help students with sensor installation and measurement

Advanced Skills: Research, Mechanical Engineering, Signal Processing, Image Processing, Data Science, Statistics, Machine Learning, Data Visualization

PROJECTS

Sentiment Analysis for Web App Deployment (NLP, RNN) [GitHub]

- Deployed an LSTM RNN model using PyTorch with test accuracy of 87% to a Web App though Amazon SageMaker
- Trained the model using LSTM model with hidden dimension set to 200 and epochs 10
- Extracted features using text preprocessing and Bag of Words

Dog Breed Classifier (Computer vision, CNN) [GitHub]

- Trained and deployed a dog breed classifier with test accuracy 87% though Amazon EC2
- Implemented a human face detector using OpenCV and a dog face detector using VGG16

Customer Segmentation Report [GitHub] [Medium]

- Predicted customer conversion probability with 0.8 AUROC score using XGBoost
- Improved model using Bayesian Optimization and GridSearchCV
- Identified customer segments that describes the core customer base using PCA and K-Means

Predict Bike Sharing Pattern [GitHub]

- Built a neural network model to predict daily bike rental ridership using 24 nodes, 5000 epochs and 0.2 learning rate
- Predicted the results with a high rate of success, MSE loss is 0.056 and the final validation MSE loss is 0.160

Boston Airbnb Price Estimator [Medium]

- Created a heatmap that shows top 15 numerical factors that have highest influence on the price
- Trained data using a linear regression model to predict the listing price with 68% accuracy, which is reasonable since there are so many factors that affect the price

SKILLS

Programming:	Python, SQL, MATLAB, R, JavaScript, C++
Modules & Libraries:	Pandas, Numpy, Matplotlib, Scipy, Pytorch, TensorFlow, Keras
ML/AI, Deep Learning:	Predictive Modeling, Convolution Neural Network, Recurrent Neural Network, Generative Adversarial Network, Computer Vision, Natural Language Processing, Recommendation Engines, Model Deployment
Data Analytics:	Data Mining and Visualization, Data and Quantitative Analysis

ADDITIONAL INFORMATION

Social Impact: Mentored 3 high school student and research interns; Research Reviewer of Discover at UofSC Program
Certifications: Machine Learning Engineer Nanodegree (Udacity, 2021); Data Scientist Nanodegree (Udacity, 2020); Deep Learning Nanodegree (Udacity, 2021), R Programming (Coursera)
Languages: Fluent in English, Chinese