Luchao Qi

Baltimore, MD
☎ *+1 (443) 839 9129*
✉ *luchaoqi@luchaoqi.com*
Personal Website: `https://luchaoqi.com/`
Github: `https://github.com/LuchaoQi`

Education

2018 – 2020
Aug   May

Master of Science in Engineering *Biomedical Engineering - Data Science*
Johns Hopkins University, Baltimore, MD, GPA: 3.61.
Thesis Topic: *Associations between Body Mass Index (BMI) and Accelerometer Time Series Data: National Health and Nutritional Examination Survey (NHANES) 2005-2006*
Advisor: Professor Brian Caffo, Professor Ciprian M. Crainiceanu, Dr. Jiawei Bai

2014 – 2018
Aug   May

Bachelor of Science *Biomedical Engineering*
Northeastern University, Shenyang, Liaoning, CHINA, GPA: 3.82.
Thesis Topic: *Calcium fluorescence response of human breast cancer cells by 50-MHz ultrasound microbeam stimulation*
Advisors: Professor Kwok Ho Lam, Professor Ming Qian, Professor Kun Yu

Professional Experience

2020 – Present
Sep

Data Scientist Intern *Eko,* Berkeley, CA.

- Help build the product Eko-core, an FDA-cleared digital stethoscope attachment device, saving monthly cost for patients with arteriovenous fistula (AVF)

- Spearhead a project to build the prototype of an audio-based dialysis fistula assessment algorithm detecting stenosis, which help secure a $295,881 in SBIR funding from the National Institutes of Health (NIH)

- Maintain the database loaded to Amazon S3 bucket using AWS

- Implement Fast Fourier transform (FFT) on audio signals for feature engineering based on frequency domain

- Constructed a random forest model (acc: 73.68%, AUC: 0.85) detecting stenosis caused by AV fistula

2019 – 2020
Aug   May

Research Assistant *The Johns Hopkins Data Science Lab,* Baltimore, MD.

- Spearheaded a project focusing on association analysis between lifestyle patterns, physical activity, and body mass index (BMI)

- Migrated data in SAS transport file format from external databases (National Health and Nutrition Examination Survey) using R and processed data using dplyr and tidyverse

- Decreased the data dimensionality using principal component analysis (PCA) and predicted user BMI with 46.07 mean squared error by training a generalized linear model (GLM)

- Achieved a 13% error reduction rate utilizing random forest and nested ANOVA (F-test) on principal component groupings

2019 – 2019

Mar   July

Data Analyst Intern *Johns Hopkins Bloomberg School of Public Health,* Baltimore, MD.

- Executed and managed research project on survival analysis of accelerometer time-series data

- Created a convolutional neural network (CNN) using Keras to predict the 5-year mortality with 71% accuracy

- Improved the accuracy to 86.45% by implementing a regularized logistic regression model using principal component scores

- Hosted R Shiny website comparing machine learning algorithms (PCA, k-means, UMAP, and t-SNE) & visualized clustering results using ggplot2 and plotly

Projects

2020 – 2020

Mar   May

Reinforcement Learning: Super Mario Bros.

AI that Learns to Play Super Mario Bros Using Deep Q-Network (DQN) in TensorFlow.

Demo: `https://github.com/LuchaoQi/Reinforcement_Learning`

- Built a reinforcement learning environment using OpenAI Gym; emulated Nintendo Entertainment System using Nes-Py in Python

- Designed a convolutional neural network (CNN) model with 5 hidden layers as an agent in TensorFlow

- Trained the agent using deep Q-learning and reduced training time by 20% using Adam optimizer

- Achieved 2X faster than averaged human players for trained agent to complete tasks successfully

2019 – 2019

Sep   Dec

Natural Language Processing: Amazon Rating Prediction.

Use of Machine Learning to Detect Fake or Abusive Amazon Product Reviews.

Demo: `https://www.kaggle.com/luchaoqi/amazon-review-rating-prediction`

- Processed Amazon Food Review data using pandas, NumPy and dfply in Python

- Tokenized unstructured text of user reviews using NLTK for feature construction

- Converted text to vector using bag-of-words model (unigram/bigram) with scikit-learn

- Predicted customer ratings using logistic regression with 0.94 AUC

- Improved bad review detection by 3% finding abusive entities (sellers & reviewers) via random forest

Publications

[1] **L. Qi**, A. Leroux, S. Marudheri, C. Crainiceanu, J. Bai, and B. Caffo. Associations between body mass index (BMI) and accelerometer time series data: National health and nutrition examination survey (NHANES) 2005-2006. *Sensors | Special*

Issues: Data Analytics and Applications of Wearable Sensors in Healthcare, 2020. Manuscript submitted.

[2] **L. Qi**, Q. Zhang, Y. Tan, K. H. Lam, H. Zheng, and M. Qian. Non-contact high-frequency ultrasound microbeam stimulation: A novel finding and potential causes of cell responses. *IEEE Transactions on Biomedical Engineering*, 67(4):1074–1082, 2020.

[3] **L. Qi**, Q. Zhang, K. H. Lam, R. Guo, R. Chen, J. Huang, R. Meng, Z. Wang, H. Zheng, and M. Qian. Calcium fluorescence response of human breast cancer cells by 50-mhz ultrasound microbeam stimulation. In *2017 IEEE International Ultrasonics Symposium (IUS)*, pages 1–3, 2017.

Software

R Packages

MRIPCA: Principal component analysis (PCA) on MRI data
https://github.com/LuchaoQi/MRIPCA.

MRIcloudT1volumetrics: T1 volumetric analysis of MRIcloud output
https://github.com/bcaffo/MRIcloudT1volumetrics.

R Shiny Web Applications

Clustering analysis using K-means, PCA, T-sne, and Umap
https://github.com/LuchaoQi/Shiny_clustering.

BMI Calculator
https://luchao-qi.shinyapps.io/BMI_Calculator/.

Skills

Programming: Python, R (Shiny), SQL, Shell scripting.

Packages: Pandas, NumPy, SciPy, NLTK, scikit-learn, Tidyverse.

Frameworks & Platforms: PyTorch, TensorFlow, Keras, Hadoop, AWS.

Machine Learning & Deep Learning: GLM, Random Forest, SVM, PCA, CNN, LSTM.

Honors and Awards

2019–2020 **Whiting School Graduate Fellowships (Tuition Scholarship)**.
2014–2018 **Dean's List**.