**LUCHAO QI**  +1 (443) 839-9129 |  luchaoqi.email@gmail.com

 <https://www.linkedin.com/in/luchaoqi/>

 Baltimore, MD  <https://luchaoqi.com/>

 WORK EXPERIENCE

**Ekohealth** Oakland, CA / Remote

Data Scientist Intern09/2020 – 12/2020

* Assisted in developing the product EKO-CORE, an FDA-cleared digital stethoscope, saving around $900 monthly cost for hemodialysis patients
* Led team of three engineers and designed machine learning algorithms detecting vascular stenosis in fistula with 73.68% accuracy
* Productionalized customer-facing analysis pipeline in Python using AWS (S3, EC2, SageMaker) to enable doctors to more accurately track clinical data, leading to a 13% reduction in errors
* Prepared regulatory submissions for FDA; successfully secured $295,881 in SBIR funding from the National Institutes of Health (NIH)

**The Johns Hopkins Data Science Lab** Baltimore, MD

Research Assistant08/2019 – 05/2020

* Led research project to analyze associations between demographic patterns, physical activity, and body mass index (BMI)
* Extracted 10K+ time-series SAS data from CDC in R and manipulated data using dplyr and tidyverse
* Designed CNN-based neural network using Keras for BMI prediction with 25.45 mean squared error (MSE)
* Reduced data dimensionality using principal component analysis (PCA); improved prediction by 23% training a generalized linear model (GLM)
* Hosted R Shiny website performing cluster analysis; visualized clustering results using ggplot2 and plotly

Demo: <https://github.com/LuchaoQi/Shiny_clustering>

 PROJECTS

**Reinforcement Learning: OpenAI Gym** 02/2020 – 05/2020

AI that Learns to Play Super Mario Bros Using Deep Q-Network (DQN) Demo: <https://github.com/LuchaoQi/Reinforcement_Learning>

* Developed DQN-based convolutional neural network (CNN) model as an AI agent using TensorFlow
* Accelerated network training by 30% training model parallelly with JAX
* Achieved 2X faster than average of human players with the trained agent completing tasks successfully

**Natural Language Processing: Amazon Reviews** 09/2019 – 12/2019

Use of Machine Learning to Detect Fake Amazon Fine Food Reviews

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

* Processed Amazon Food Review data using Pandas, NumPy, and dfply
* Tokenized unstructured text of user reviews using NLTK; converted text to vector using bag-of-words models with scikit-learn
* Predicted customer ratings using logistic regression with 0.94 AUC
* Improved negative reviews detection by 3% using random forest

 SKILLS

**Programming Languages**

Python, R (R Shiny), SQL / MySQL, HTML, CSS, Shell scripting

**Data Visualization**

Tableau, Matplotlib, Seaborn, ggplot2, plotly

**Packages**

Pandas, NumPy, SciPy, NLTK, scikit-learn, dplyr, tidyverse, Selenium, beautifulsoup4

**Machine Learning & Deep Learning**

GLM, Random Forest, SVM, PCA, CNN, LSTM

**Frameworks & Platforms**

PyTorch, TensorFlow, Keras, Hadoop, AWS, Google Analytics, Kubernetes

 EDUCATION

**Johns Hopkins University**

Baltimore, MD 05/2020

Biomedical Data Science M.S.E.

GPA: 3.6

* JHU Fellowships (Tuition Scholarship with 20% acceptance rate)

**Northeastern University**

Liaoning, CN 05/2018

Biomedical Engineering B.S.

GPA: 3.8

* Dean’s List (2014 – 2018)