

CONNIE XIAO

DATA SCIENTIST

CONTACT

✉ co.xiao1996@gmail.com
☎ 917 (392) 2809
📍 New York, NY
in linkedin.com/in/cxiao96/
🌐 cxiao96

EDUCATION

Stony Brook University June 2019
Bachelor of Science Respiratory
Care Therapy
Alpha Eta Honor Society
Lambda Beta Honor Society

SKILLS

PYTHON

Pandas
NumPy
Scikit-learn
Keras
NLtk
BeautifulSoup
Selenium

MACHINE LEARNING

Generalized linear regression
Logistic Regression
Neural Networks
Natural Language Processing
Clustering
Dimensionality Reduction
Topic Modeling
Image Classification

VISUALIZATIONS

Matplotlib
Seaborn
Plotly
Tableau
Streamlit
PowerPoint

DATA MANAGEMENT

MongoDB
SQL
GCP

SUMMARY

Highly motivated and adaptive data scientist and medical practitioner focusing on problem solving and critical thinking through data-driven mechanisms.

EXPERIENCE

NYU Langone Health Medical Center New York, NY
Registered Respiratory Therapist Sept. 2019 to Current

- Analyzed and discovered existing data, additional data and recommended data to evaluate the respiratory status of patients, to develop a respiratory care plan, and to determine the appropriateness of prescribed therapy
- Educated patients and families about patient's conditions and appropriate disease management techniques
- Oversaw and operated ventilators in surgical intensive care units as well as high frequency oscillation and high frequency jet ventilation in neonatal and pediatric intensive care units
- Managed and charted electronic medical records on Epic Systems software

Metis New York, NY
Data Scientist June 2021 to Sept. 2021

Completed Metis's 14- week intensive accredited data science and machine learning bootcamp with a strong emphasis on Python programming, machine learning, statistical modeling, data engineering, data visualization, project design/flow, and communication deliverables. Project highlights include:

COVID-19 Dashboard

- Collected data and created visualizations for a better understanding on COVID-19 infections across countries
- Built a pipeline to store data onto SQLAlchemy, created an app using Streamlit and deployed analytics into a web using Heroku

Dog Breed Image Classifier

- Built a classification model using Scikit-learn and Keras to classify different images of dog breeds using convolutional neural networks
- Used transfer learning to optimize leveraging existing data for a more efficient way to train the model

Yelp Reviews using NLP Techniques

- Used NLP techniques such as NMF to topic model across 20,000+ restaurant reviews
- Investigated topics and sentiment on positive and negative reviews

Stroke Prediction

- Compared performances of multiple models (logistic regression, random forest, and XGBoost) to predicted whether a patient will have a stroke
- Utilized feature engineering to improve algorithms and handled class imbalance