

# John Corio

973-896-6625 | corioj@umich.edu | linkedin.com/in/john-corio-63b128196 | github.com/corioj

## EDUCATION

---

**University of Michigan: College of Literature, Arts, and Sciences**

*Bachelor of Sciences in Data Science, Minor in Mathematics*

Ann Arbor, MI

Aug. 2017 – May 2021

## EXPERIENCE

---

**Client Engineering Data Scientist**

*IBM*

Mar. 2022 – Present

*New York, NY*

- Increasing IBM software sales by engaging in pre-sales data science related proofs of concept and MVPs to rapidly solve clients' business problems and create further business opportunities for IBM
- Developing models using Python (SKLearn, Pandas, NumPy, Matplotlib) in Jupyter notebooks as well as engineering proprietary software solutions to address individual business use cases
- Serving as subject matter expert on ML, AI, statistical models and data science for client engagements, and how they can be applied to solve individual clients' business problems

**Data Quality Assurance Analyst**

*ImageCare Centers*

Dec. 2020 – Feb. 2022

*Newton, NJ*

- Drafted first designs of a proprietary SQL database specifically for analytics, migrated large datasets from a third-party software vendor, and then ran quality assurance via SQL queries
- Provided the groundwork for a BI analytics platform to inform on key KPIs for an estimated 70 million dollar company

## PROJECTS

---

**Album Art Generation via Neural Networks** | *PyTorch, Pillow, OpenCV, Matplotlib* Nov. 2020 – Dec. 2020

- Generated album covers from random noise tensors designed to fit the aesthetics of albums from given genres such as rock, pop, and others
- Designed a script to query the Spotify API using the SpotiPy library, extract desired data from JSON objects, and export cleaned image dataset to a preprocessing routine using Pillow and OpenCV
- Implemented 3 different generative adversarial neural networks based on computer vision research publications and individual design, then evaluated them based on image generation clarity

**Sentiment Classifier of Yelp Reviews** | *SKLearn, Matplotlib, Pandas, NLTK*

Sep. 2020

- Achieved top 10 percent in class on testing dataset accuracy for an NLP-based classification model of Yelp reviews in Python-based Jupyter notebooks
- Processed natural language data using NLTK, Pandas, and Matplotlib
- Performed hyperparameter optimization and evaluated various modelling choices using SKLearn

**Asset Return Predictor** | *R, ggplot*

Sep. 2020

- Placed second out of forty five teams on final project test dataset performance for predicting the forward returns of stock market assets
- Preprocessed time series finance data using various R libraries and self-made functions, then fitted and evaluated the holdout dataset performance of a variety of models such as KNN, Regression, Principal Components Regression, Random Forests, Bagging, and Boosting

**Traveling Salesperson Problem** | *C++, STL*

Dec. 2019

- Created custom coordinate data structures from input data and designed a map to display the information
- Implemented optimal and fast solutions for the Traveling Salesperson utilizing branch-and-bound algorithms, minimum spanning trees, and Prim's algorithm

## TECHNICAL SKILLS

---

**Languages:** Python, C++, SQL, R, MongoDB, JavaScript, HTML/CSS

**Libraries:** Pandas, NumPy, Matplotlib, SKLearn, PyTorch, C++ STL