John Corio

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

Experience

Data Scientist Mar 2022 – Dec 2023
IBM *New York, NY*

- Developed machine learning assets using Python, C++, SQL, open-source libraries (SKLearn, PyTorch, Pandas, Apache Spark, etc.), and proprietary software to address individual business use cases and increase client adoption, contributing to the most successful software sales quarter in IBM Financial Services Market history.
- Established project requirements with client, performed data pre-processing, and lead implementation of regression models in Scala, Python, and Java using Apache Spark MLLib, transitioning a major banking client's overdraft system that mediates millions of commercial and personal transactions per day to AI.
- Co-authored NLP assets that perform document sectioning, entity extraction, and section classification on PDF and Word documents using BERT algorithm in Python Jupyter notebooks, via open-source text processing libraries and proprietary machine learning libraries.

Game Developer, Volunteer Mar 2025 –
Toontown Corporate Clash *Remote*

- Volunteer position with a fan-run MMORPG using Python's Panda3D, Astron, and Git to implement live features serving thousands of players worldwide.
- Designing, implementing, and refactoring code for systems involving gameplay mechanics, account management, networking, distributed systems, and more.

Data Analyst Jun 2021 – Feb 2022
ImageCare Centers *Newton, NJ*

- Drafted first designs of a proprietary SQL database for analytics, migrated large datasets from a third-party software vendor, used Python to analyze data informing on key KPIs.

Technical Skills

Languages: Python, SQL, C#, C++, Scala, R
Libraries: Apache Spark, scikit-learn, PyTorch, Tensorflow, CUDA, Pandas, NumPy, Matplotlib, C++ STL
Developer Tools: Git, Cloud Computing, Jupyter Notebooks, CMake, Google Colab, VSCode, Linux, Bash
Certifications: JLPT N2 145/180

Projects

Album Art Generator | *Python, PyTorch, OpenCV, Matplotlib, Pillow, SpotiPy*

- Implemented 3 different generative adversarial neural networks based on state-of-the-art computer vision research publications to generate album covers displaying aesthetics of various genres.
- Implemented a data engineering pipeline using Pillow and OpenCV to produce a dataset of 150,000 images.
- Developed Python scripts to rapidly query, extract, and categorize images and desired metadata from JSON objects stored on the Spotify developer API.

Yelp Review Classifier | *Python, scikit-learn, Matplotlib, Pandas, NLTK*

- Achieved top 10 percent in class on testing dataset accuracy for an NLP-based classification model of the emotional modality of Yelp reviews in Python using Jupyter notebooks.
- Created an automated framework for training and optimizing hyperparameters, evaluating and comparing support vector machines and deep learning models, and reviewing validation set accuracy results.

Raytracing Renderer | *C++, STL*

- Implemented a raytracing renderer that supports spheres and quad shapes, texturing of objects using noise and UV image mappings, materials, light-emitting surfaces, bounding volume hierarchical optimization using AABB bounding boxes, and object rotations and translations via instancing.
- Camera placement configurable via rotation and translation, output images configurable via adjustments to per-pixel sampling, ray bounce recursion depth, FOV, disk defocus, and aspect ratio.

Education

University of Michigan Ann Arbor, MI
Bachelor of Science in Data Science, Minor in Mathematics *Aug 2017 – May 2021*