John Corio

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

Experience

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

- Developed machine learning assets using Python, Scala, Java, and SQL, open-source libraries (scikit-learn, Tensorflow, Pandas, Apache Spark, etc.), and proprietary software frameworks to address individual business use cases and increase client adoption, helping to achieve the most successful quarter for software sales in IBM Financial Services Market history.
- Gathered project requirements with client, performed data exploration and processing, and lead implementation of regression models in Scala, Python, and Java using Apache Spark MLLib to transition a major banking client's overdraft system that mediates millions of commercial and personal transactions per day from rule-based architecture to AI.
- Co-authored NLP assets for a banking client that performs document sectioning, entity extraction, and section classification on PDF and Word documents using deep-learning and rule-based algorithms in Python Jupyter notebooks, via open-source text processing libraries and proprietary machine learning libraries.

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

- Drafted first designs of a proprietary SQL database specifically for analytics, migrated large datasets from a third-party software vendor, and wrote SQL test queries to verify data quality
- Created SQL queries, tables, and views to pull and evaluate data used in analytics informing on key KPIs.

Japanese Language Student Jul 2023 – May 2024
ISI Language School *Tokyo, JP*

- Studied Japanese in upper level language classes at ISI Takadanobaba. Currently possess JLPT N2 certification.

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.
- Prepared a data processing, labeling, and splitting routine using Pillow and OpenCV to produce a cleaned dataset of over 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.

Tsuneflip | *C#, Unity Engine, Ableton Live 11*

- 3D arcade-style game jam submission made in one week, under theme 'inside out.'
- Designed and implemented grid and game management systems, a basic level editor, tile and grid system, data exchange systems between menus and levels, game control systems, input handling, movement.
- Currently working on an update including individual enemies, a globally connected leaderboard, greater variety in levels and worlds, configurable player preferences, Japanese language support, power-ups, and more.

Education

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

Technical Skills

Languages: Python, C#, C++, Scala, Java, SQL, R, JavaScript
Libraries: Apache Spark, scikit-learn, PyTorch, Tensorflow, CUDA, Pandas, NumPy, Matplotlib, C++ STL
Developer Tools: Git, Xcode, Visual Studio, VS Code, Jupyter Notebooks, CMake, Google Colab, IntelliJ