Thomas Christo

551-256-1552 | christot091702@gmail.com | github.com/christot20 | linkedin.com/in/thomas-christo-aa2031222

Experience

Freewheel   *New York, NY*
Software Engineer (Data) Intern   *May 2023 - Aug 2023*

- Prepared ad hoc queries on raw data tables **>5TB** in size located in **AWS S3** for master table validations and data analysis with **Pyspark**, **Pandas**, and **SQL**.
- Designed an **ETL pipeline** by developing **Python Wheel** and **Databricks Workflows** to streamline process of creating and saving **Delta Tables** into **Databricks Unity Catalog** with updates **>70 million rows per day (>10 GB)** in **<20 minutes**.
- Optimized **SQL** queries for **<1 second execution time** allowing for quick dashboard updates.
- Completed task of developing an **Anomaly Detection Dashboard** within **Databricks** for Samba viewership/device data complete with **daily alerting and updates** saving engineers development time from debugging and data validations by **15%**.

Brookhaven National Laboratory   *Upton, NY*
Research Intern   *Jun 2022 - Aug 2022*

- Automated stellar model running process on a SLURM cluster and data collection of models using **Bash**, **Sed**, and **Python** which **decreased model runtime and data extraction** by **55%**.
- Developed tools that organized hundreds of stellar model outputs, sped up computation/analysis by **70%**, and established data visualizations for the subsequent paper using **NumPy**, **Pandas**, **SciPy**, **Matplotlib**, and **Excel**.
- Collaborated with team and saved stellar modelling tools codebase on NNDC **Gitlab** server.
- Wrote **>1500 word report** and poster presentation of research process and work accomplishments using **LaTeX** and **Microsoft Office**.

Projects

R.A.M. Trading
Personal Project

- **Python** stock trading bot to compare the success of **3** different trading strategies that traded for **>270 days**.
- **LSTM** neural network trained by **3 years of S&P 500 stock data** implemented with **Keras**, **NumPy**, **Pandas**, and **SKlearn**.
- Gathered and analyzed stock data from algorithmic value trading method with **Pandas** and **IEX Cloud API** for **<50 second** computations.
- Accounts' data saved using a **MYSQL** database and explored **>8000 transactions** in **MYSQL Workbench**.
- **Streamlit** dashboard generated to monitor activity and status of each trading account by **1 minute** intervals.

Indeed Scraper
Personal Project

- **Python web-scraping** bot utilizing **BeautifulSoup** and **Selenium** to search Indeed.com and compare different Data jobs by salary ranges, skills, and requirements in **6** large US cities.
- Able to parse through **500** job listings in **<30 minutes** without captcha detection.
- Required **NLTK** and **Regex** usage to clean raw data and **Matplotlib**/**Seaborn** for data visualizations.
- Stored **>1000** entries of gathered job data and performed data exploration with a **SQLite3 database** and **SQL** scripts.
- **Tableau** dashboard made for further visualization of generated metrics for **>250** job entries.

Education

Rutgers University   *New Brunswick, NJ*
B.S. Computer Science   *Sept 2022 - Current*

- GPA: 3.946
- **Courses:** Mathematical Theory of Probability, Discrete Mathematics, Software Engineering, Data Structures, Computer Architecture, Bayesian Data Analysis, Regression Methods, Data 101

Union County College   *Cranford, NJ*
A.S Computer Engineering   *Sep 2020 - May 2022*

- GPA: 4.0
- Phi Theta Kappa Member
- NASA Community College Aerospace Scholars