

NORBERT MARKOWICZ

DEVELOPER / DATA ENGINEER

www.norb.dev

norbert.markowicz@gmail.com

919-691-1771

SKILLS

Programming

Python (advanced)
Golang (proficient)
SQL (proficient)
Javascript (basic)
Scala (basic)
Java (basic)

Technology/Frameworks

Airflow
Akka
Oozie
Spark/PySpark
HBase
Hive
Flask
Tornado

Cloud

Blob/ADLS
Databricks on Azure
Eventhub

DevOps

Docker
Kubernetes/Openshift
Jenkins
Git/SVN
Unix scripting

EDUCATION

The Cooper Union, New York, NY
M.Eng in Mechanical Engineering
2009 - 2012
B.Eng in Mechanical Engineering
2004 - 2008

EXPERIENCE

Credit Suisse, New York, NY / Raleigh, NC

Innovation & Core Engineering / Platform & Data Engineering Lead
2019 - Present

- Developed core components of cloud-based data sourcing, processing, validation & profiling platform. The multi-tenant platform was meta-data driven with Apache Airflow used for dynamic DAG generation, run orchestration & pipeline execution.
- Mentored a team of 4 platform data engineers of varying skill levels
- Developed custom Airflow operators to support a broad scope of functional requirements
- Created Python modules that allowed implementation of custom processing, profiling & validation logic through a standardized framework for each entity in Databricks clusters
- Rolled out Spark-streaming functionality to support real-time market data captured in Azure Eventhub to flow through into data lake as Delta tables
- Simplified several platform components through refactoring into more isolated modules to achieve greater separation of responsibility and better testability
- Implemented a caching mechanism for serialized DAGs accessible to all worker nodes, reducing object generation times and eliminating unnecessary network API calls
- Through self-initiative, wrote a highly-regarded, cross-platform CLI in Go to improve user experience & expedite the addition of workloads to run on the platform
- Expanded code unit test coverage and functional tests to expedite deployments into production environments

Projections Platform / Software Engineer

2016 - 2019

- Architected and developed end-user libraries in Python & R used by 320+ users across divisions for critical business scenario planning
- Created a model development framework to separate model, model parameters & execution run context to allow for standardization and readability of over 230 models
- Heavily focused on unit, functional and end-to-end tests for automated regression testing and expedient deployments
- Created extensions on the platform to measure usage, traceability and automated library updates
- Initiated and developed core components of Model-as-a-Service layer in Go to allow for rapid production of models on Openshift

PROFESSIONAL DEVELOPMENT

General Assembly 2015

Data Science Fundamentals

Credit Suisse 2016-Present

Firm-wide Hackathons

MEET-UPS

Data Driven NYC

Golang NYC

NYC Apache Airflow

RTP Python

Client Predictive Analytics / Software Engineer

2014 - 2016

- Core developer of a graphed based client relationship network framework, backed by HBase, from internally & externally sourced communication data
- Re-engineered sourcing workloads using Hive & Impala to reduce pipeline times of daily core presentation data tables
- Created API layer Scala + Akka framework to streaming chunked data to reduce UI rendering times by 40%
- Optimized row key index to expand the nature of use cases the platform was able to support

Finance IT / Systems Analyst

2010 - 2014

- Analyzed & designed front-to-back dataflows to support new business & function requirements
- Built queries and views which were later productized into Oracle DB stored procedures
- Performed end-to-end testing, automating steps in Python in the process
- Identified and worked to resolve issues relating to poor upstream data and flaws in logic
- Created architecture and data flow diagrams to support the development process
- Supported users groups in debugging data breaks and issues with system pipelines