

# Eric Keniuk

✉ [eric.keniuk@gmail.com](mailto:eric.keniuk@gmail.com) ☎ (224) 639-3648 [in https://www.linkedin.com/in/erickeniuk/](https://www.linkedin.com/in/erickeniuk/) [github https://erickeniuk.github.io/](https://erickeniuk.github.io/)

## SUMMARY

---

Master's in applied economics and bachelor's in mathematics. Expertise in financial technology, big data on the petabyte scale, and experience in USG. Highly proficient in both qualitative and quantitative analysis. Skillset includes: Hadoop, Linux admin & networking. Spark Scala, PySpark, Python, R, Java, C, SQL, SAS, Unix / Bash, shell scripting, Linux, Networks (Firewalls), certs, Chef, Git, Cloud & distributed computing, Docker, Agile principles (Scrum Teams), Machine Learning, Unsupervised Learning Techniques (Dimension Reduction, Data Clustering)

## EDUCATION

---

### Master of Arts in Applied Economics

University of Alabama • Tuscaloosa, AL • December 2019 • GPA: 4.0

### Bachelor of Arts in Mathematics and Economics

University of Alabama • Tuscaloosa, AL • May 2019 • GPA: 3.86

## EXPERIENCE

---

### T-Mobile

#### Data Engineer

March 2023 - Present, Bellevue, WA

- ETL pipeline builds within MS Azure standing up resource groups (ADF, data sinks, key vaults, storage accounts, etc)
- Development in Agile methodology within Python data migration framework to allow easy user configuration and deployment; Support for deployment and bug troubleshooting
- Databricks notebook development and workflow automation scheduling for readStream/writeStream data migration of Delta Tables to Snowflake including automated Voltage encryption translation

### Early Warning/Zelle

#### Sr. Data Scientist - Data Platform

May 2022 - March 2023, Seattle, WA (Remote)

- Lead developer (prototype DEV/QA) for synthetic identity fraud product in Spark Scala, version control (Git) w/ Artifactory deployment; led communications; regression, integration testing, automated ETL pipeline (Python, Unix)
- Lead maintainer of consumer behavior Spring APIs hitting endpoints to test PROD Java code; processes Solr, Hbase data and regression test internal code to cross-check for accuracy with PROD code; pipeline automation via Unix. Prototyping PROD Java code to retroactively score API requests (currently: real-time only) and PO for UI hosting internal Spring APIs for easier use by data scientists while working to externalize internal tool
- Lead for transition of 5 teams to new on-prem data center/cluster: Built entire virtual environment (Windows, Linux), established firewall openings, installed applications/software, configured Hadoop 3.x, resolved/refactored script errors, daily Linux, Hadoop, DevOps admin work, Chef and Cron automation; Began work on cloud environment transition

#### Data Scientist

September 2020 - May 2022, Chicago, IL

- Automation of daily processes— data ETLs, report aggregation, and model performance and data drift monitoring via Bash, Python, Spark Scala, R; RSA keys and SSH, employ DataRobot, and Universal Controller system automation
- Feature engineered +800 features via Spark Scala, Git for new payments ML model
- Model validation, performance & governance reporting via Hive, SAS, Python. Automated processing pipelines, cut lead times by more than 50%. Enhanced SpringBoot PROD API to run multiple instances at once and run more reliably.

## PROJECTS

---

### Synthetic Identity Fraud Attribute Product – Lead Development Internal Testing Code

- Attribute product to create input data for external modeling & rules frameworks for identifying fake identities

### Non-contributing Payments Machine Learning Fraud Model – Feature Engineering

- Feature eng. in Spark Scala of payments fraud ML model for targets with no contributed attributes data

### Real-Estate Investor Effects on Primary School Outcomes – Master's Thesis

- R language to extract & transform 3MM observations of home sales from 1952 to 2015, identify small, medium, and large investors by home ownership volume and model impacts on educational outcomes: <https://erickeniuk.github.io/>