

# Sean Cooke

(516) 413-1876

[SeanCooke8892@gmail.com](mailto:SeanCooke8892@gmail.com)

[LinkedIn](#) | [GitHub](#) | [Personal Website](#)

## Skills

---

- Machine learning engineer with 5+ years of experience as a full-stack software engineer
- Experience working with
  - Programming Languages: Python (Jupyter, Pandas, scikit-learn, Matplotlib, pytest), JavaScript, PHP, Java
  - Databases: Snowflake, Oracle SQL, MySQL
  - Infrastructure: Linux, Bash, AWS, Git, Docker, Airflow, Apache, Tableau
  - Project Management: Agile Scrum, Kanban, Jira, Confluence
- Skilled in Machine Learning, Data Pipelines, Data Visualization, Big Data, and Cloud Computing

## Experience

---

### **athenahealth** *Watertown, MA*

#### *Data Analyst*

*March 2020-Present*

- [Coauthored published research](#) in top cardiology journal showing addition of Apple watch atrial fibrillation detection software did not increase AFib claims using Difference-in-Differences Estimation
- Prioritized rollout of feature that alerts doctors to wear PPE when seeing patients who may have been exposed to COVID by ranking 16k hospitals by number of COVID patients using Snowflake and Pandas
- Reduced Filecloud storage by 30% (17 TB/year) by showing Product Owners many old documents were being imported. Created probability density plot of 'document age at time of import' using Matplotlib

#### *Data Engineer*

*May 2018-March 2020*

- Reduced daily extract failures from 677 to 5 in 1 month through SQL optimization by leveraging indexes and limiting cross-joins. Helps prevent EHR outages for hospitals by decreasing ETL job failures
- Migrated revenue reporting machine learning model from DC/OS to AWS by using newly created AWS RDS, S3 instances; cost savings of \$120k/year

### **University of Rochester Medical Center** *Rochester, NY*

#### *Software Engineer*

*June 2014-April 2018*

- Architected, implemented, and supported [urattendance.urmc.rochester.edu](http://urattendance.urmc.rochester.edu), an attendance tracking website for 1k medical residents and fellows; allowing physicians more time with patients by reducing admin time by 2 hours per meeting
- Allowed high school students to apply for medical school summer programs online by replacing paper applications with mobile-friendly [STEP Application](#) and [SURF Application](#) using JavaScript and PHP
- 2017 Meliora Award nominee for years of outstanding service and organization contributions; highest honor awarded to staff members; 5 time Strong Star recipient for outstanding software engineering

## Independent Projects

---

### **titanic**

[www.github.com/SeanCooke/titanic](https://www.github.com/SeanCooke/titanic)

- Feature Engineering and XGBoost classifier achieves 86% accuracy on survival from Titanic Data Set

### **Tumor Classification**

[www.github.com/SeanCooke/tumor-classification](https://www.github.com/SeanCooke/tumor-classification)

- Random forest classifier achieves 96% accuracy on diagnosis from Breast Cancer (Wisconsin) Data Set

## Education

---

### **University of Rochester** *Rochester, NY*

*August 2015-May 2017*

- Master of Science in Computer Science; GPA 3.46

### **SUNY Geneseo** *Geneseo, NY*

*January 2011-May 2014*

- Bachelor of Arts in Mathematics, Minor in Computer Science; Dean's List Spring 2014