

# Mike Kutilek

Data Scientist

makutilek@gmail.com

412-638-9880

Pittsburgh

www.mikekutilek.com

github.com/mikekutilek

Data Engineer proficient in statistical analysis eager to take the plunge into the world of Data Science

## WORK EXPERIENCE

### Data Engineer Highmark Inc.

02/2019 – Present

Insurance company in Pittsburgh, PA

#### Achievements/Tasks

- Creating data pipelines with PySpark to move member data from legacy source systems to Hadoop environments
- Working closely with a team of Data Engineers to add routes to a Data API

### Software Developer Management Science Associates

05/2016 – 02/2019

Pittsburgh

Private data company focused on providing Big Data solutions

#### Achievements/Tasks

- Developing data driven tools within a monitoring portal to better track and notify the status of the system
- Utilizing APIs to inspect analytical patterns within system data and write the data out to a custom API
- Pulling subsets of data from Hive/Redshift tables to generate reports/graphs using SQL and Matplotlib
- Deploying code and kicking off long running jobs within Hadoop environments
- Collaborating in an Agile/Scrum environment with a focus on better communication within dev and qa teams
- Implementing and maintaining code and deployment standards for source and version control

## EDUCATION

### AI Programming with Python Nanodegree Udacity

05/2019 – 08/2019

#### Projects

- Neural Network Image Classifier

### BS in Computer Science University of Pittsburgh

08/2014 – 12/2016

#### Courses

- Software Engineering
- Artificial Intelligence
- Calculus
- Database Mgmt Systems
- Operating Systems
- Linear Algebra

## SKILLS

Python Numpy Pandas Git SQL

Matplotlib BeautifulSoup Scikit-learn

CRISP-DM Linear Regression Logistic Regression

PyTorch Nodejs Angularjs Javascript

HTML/CSS Spark Hive AWS Oracle

MongoDB JIRA Linux Selenium

Agile/Scrum Slack Communication Research

Teamwork

## PERSONAL PROJECTS

### Personal Website (06/2018)

- Portfolio Website showcasing a series of Pandas driven statistical analysis projects backed by a MongoDB data warehouse with a front end built in NodeJS

### Neural Network Image Classifier (05/2019 – 08/2019)

- Part of Udacity course listed below. Built using a Logistic Regression model & PyTorch

### MLB WAR Model (06/2018 – Present)

- Simple linear regression model aimed at Team & Player WAR ratings to better predict future team success and player performance.

## ACHIEVEMENTS

### Slack Initiative (07/2018 – 02/2019)

Directed the effort to use Slack across our project at MSA

### Standardized Deployment Process (07/2018 – 02/2019)

Reduced excess work for scrum teams and better met audit requirements

## INTERESTS

Photography Reading Running Data Science

Machine Learning