mgalarny@eng.ucsd.edu
(858) 999-7481
La Jolla, CA

Michael Galarnyk

github.com/mGalarnyk
linkedin.com/in/michaelgalarnyk
mgalarnyk.github.io

EDUCATION

University of California, San Diego

M.S. Computer Science (Data Science) 2017

B.S. NanoEngineering, Minor Biology 2015

SKILLS

Programming:	Python (pandas, sklearn, PySpark), SQL, MATLAB, Bash Shell Scripting, JavaScript, HTML, CSS, C++
Database Management:	Amazon AWS (EC2), MSSQL, PostgreSQL, MySQL
Operating Systems	Linux (Ubuntu, Red Hat), Windows, Mac
Other Technologies:	Hadoop, HBase, Spark (PySpark), Tableau, IPython, Jupyter, Git, LaTeX, SolidWorks, AutoCAD

EMPLOYMENT

DUV Systems Engineering Intern (Data Analysis) at Cymer Summer 2015

- Designed MATLAB GUI tools to automate SQL queries (MSSQL) and to automatically generate data reports.
- Reduced average SQL query time by ~27% for Cymer GUI tools.
- Presented to CEO and earned my coworkers recognition at the Cymer All-Hands Meeting

Researcher at UCSD NanoBioElectronics Lab 2013-2015

- 9 coauthored publications in peer reviewed journals, 120+ citations (h-index: 7), 1 publication featured on cover
- Gathering and analyzing data using MATLAB and Python (NumPy, pandas)
- Graphic Design/SolidWorks Modeling for schematic illustrations in peer-reviewed journals and news organizations (BBC, Nanowerk)

RELEVANT COURSEWORK SAMPLE

Python for Data Analysis
Probability and Statistics using Python
Data Management Systems (DBMS)
Tableau for Data Visualization
Machine Learning
Data Analysis using Hadoop and Spark
Python for Informatics

TECHNICAL WORK SAMPLE

Twitter API for Sentiment Analysis

- Created a Python based API to collect steaming data of live tweets
- Sentiment analysis using TF-IDF

ETL & Analysis of Ebola Data

- Scraped and cleaned Ebola data from the World Health Organization
- Python (NumPy, pandas, matplotlib) based analysis on the data

Relational Database for Sales Data

- Designed and optimized database schema using E/R diagrams (Cube Schema)
- Created database on Amazon AWS instance (EC2)
- Reduced query and maintenance cost by using optimized relational algebra, indices, and materialized views.

Relational Database for Cats Data

- Designed and optimized database schema using E/R diagrams (Star Schema)
- Created database on Amazon AWS instance (EC2)
- Reduced query and maintenance cost by using optimized relational algebra, indices, and materialized views.

PCA, K-Means on Animals with Attributes Dataset

- Bash scripting, PCA and K-means using sklearn, hierarchical clustering (scipy), showed the position of each animal using matplotlib
- Code: http://mgalarnyk.github.io/Clustering.html

Pokemon Battle Game

- HTML, CSS, Javacript, and JQuery
- Game: http://mgalarnyk.github.io/web_development/pokemon/index.html

Publications: http://mgalarnyk.github.io/publication_list.html