AbderRahman N. Sobh

187 Acalanes Dr, Sunnyvale, CA 94086, Phone: 217-979-4127
abbysobh@gmail.com

Work Experience

Software Developer, University of Illinois at Urbana-Champaign (UIUC), (2012-)

Data Science Platform (Python, Spark)
- Designed machine learning and feature selection pipelines.
- Built machine learning tool for NLP, Text Features.
- Optimized distributed performance using clusters.

Online Database Platform (PHP, HTML, MySQL)
- Created webpages allowing users to record data to a MySQL db.
- Provided graphical views of data from the MySQL database.
    Source: (https://github.com/NanoBioNode/ISU_website_PublicVers)

Simulation Workflows (Python, TCL, Fortran, Rappture for XML)
- Combined applications within a virtual container for a single experience.
- Created a set of GUIs to replace text-based programs.
- Optimized memory handling and run-time on millions of entries.
- Applied parallel threading to serial coding procedures.
- Connected workflow to remote cluster venues for parallel processing.
- Visualized simulation result data using VTK toolkit.

Discrete Dipole Scattering Simulation Workflow ("nanoDDSCAT+")
    Source: (https://github.com/NanoBioNode/nanoDDSCATplus)
    App: (https://nanohub.org/tools/ddaplus/)

Molecular Dynamics Workflow ("Functionalization Workbench")
    Source: (https://github.com/NanoBioNode/InorganicBuilderPlugin)
    App: (https://nanohub.org/tools/partsub/)

Technical Skills

Python, Spark, TCL,
PHP, MySQL, SQL, NoSQL,
R, SAS, Matlab, Mathematica,
Tableau,
Microsoft VBA/Excel,
Fortran90, C, C++
HTML, JavaScript, CSS, XML
OSX. Linux. Windows

Personal Experience

Exposure to Data Science
- Member of Data Scientist Engineers Project Group Meetup (2016-)
- Member of New York Open Statistical Programming Meetup (2012-13)

Kaggle Data Science Competitor
(Python/Pandas, GraphLab, scikit-learn, XGBoost, AWS)
- Cleaned data sets, normalized data, predicted missing values.
- Used NLP, clustering to investigate text features.
- Performed feature selection, created models to predict new data values.
- Used Amazon Web Services for parallel data processing workflows.
- Received public recognition for open-sourced script (aka a "Kaggle Kernel").
    Source: (https://github.com/itg-abby/KaggleScripts/)

Auto Level-Creator for "In the Groove 2" Video Game (Python)
- Analyzed music/time series and mapped relevant features to a timing signature.
    Source: (https://github.com/itg-abby/StepGen)

Education

BS in Statistics, May 2012, University of Illinois, Urbana-Champaign, IL
- Dean's Honor List (2008), (2012)

Relevant Coursework

- Advanced Data Analysis
- Statistical Data Management
- Applied Regression and Design
- Biostatistics
- Computer Engineering
- Applied Linear Algebra