Ericka Smith

(425) 308-7120

smith.ericka.b@gmail.com
erickabsmith.github.io
linkedin.com/in/erickabsmith

SUMMARY

Data scientist experienced with statistical modeling, data exploration, communication, coding, & machine learning.
Expertise in: Python, R, Shiny, H2O, AWS, SQL, HTML.
Toolkit: GitHub, Excel, EC2, Jupyter Notebook, PyCharm, PuTTY, GIS, AWS

WORK EXPERIENCE

Data Scientist   *Hartford Insurance Group*   *04/2022 – current*
Associate Data Scientist   *05/2021 – 03/2022*
- Build monitoring analyses for model validation.
- Create & evaluate advanced machine learning model solutions used in the implementation of a novel sales strategy.
- Collaborate closely with interdisciplinary team to produce analytical solutions & deliver modeling solutions.

Research Assistant   *Cooperative Institute for Marine Resource Studies*   *06/2020 – 12/2020*
- Leveraged data cleaning skills & discipline knowledge for analysis of large, messy datasets.
- Researched & implemented clustering algorithm to solve a data grouping problem.

Intern   *National Oceanic & Atmospheric Administration*   *08/2020 – 11/2020*
- Automated data management processes in R, resulting in a dramatic decrease in process time & a reduction in errors.

Technician   *Quantum Spatial Inc.*   *03/2019 – 09/2019*
Intern   *10/2018 – 03/2019*
- Used R, Python, Excel, Shiny, GIS for machine learning modeling & predictive analytics of large spatial datasets.

EDUCATION

M.S. Statistics   **Oregon State University**   **03/2021**
Teaching Assistant: Methods of Data Analysis (Graduate Level), Intro to Statistical Methods
Coursework: Probability, Computing, & Simulation; Data Visualization; Comm& Line Analysis; Python;
                            Sampling; Survey Methods; Generalized Regression Models; Time Series; Unix/Linux

B.S. Math/Biology   **Western Washington University**   **06/2015**
Employment: Database Assistant using Excel & Access; STEM Tutor teaching R & Excel
Coursework: Computer Programming & Data Structures; Differential Equations; Linear Algebra

PROJECTS

Classification Using Satellite Imagery & Random Forests
12/2020 – 03/2021
- Built simulations in R training many random forests, compared accuracy with different hyperparameters.

A Comparison of Generalized Regression Models for Interpretation of COVID-19 Case Counts
11/2020 – 12/2020
- Analyzed, fit, & evaluated negative binomial, quasi-Poisson, & Poisson regression models.

Bootstrapping the Pacific Crest Trail in Parallel
10/2020 – 12/2020
- Tested use of socket method of parallel processing workflow (R, Amazon Elastic Compute Cloud, AWS) to efficiently manage bootstrap resampling & simulations.

Interactive & Explorative Web-Based Tool for Communication of Large Ensemble Models
01/2020 – 07/2020
- Collaborated on Shiny app that consolidates & analyzes multiple high-volume datasets.
- Honorable Mention (2nd place) nationwide from the American Statistical Association.
- Designed visualizations to explain ensemble models to technical & non-technical audiences.

An Interactive Shiny App for Data Visualization
May. 2020 – June 2020
- Independently cleaned and compiled data into an accessible and interactive Shiny app that uses visualization to elucidate trends in time series data.