Hanieh Haeri

925-768-0299 | hhaeri0911@gmail.com | Alamo, CA94507
https://hhaeri.github.io/portfolio/ | https://www.linkedin.com/in/hanieh-haeri-9319b024/

PROFESSIONAL SUMMARY

Seasoned Data Scientist with a decade of expertise in modeling, statistical analysis, and machine learning. Demonstrated ability to lead impactful projects and deliver innovative solutions to complex business challenges. Excited to leverage advanced skills in data science and technology to generate actionable insights for clients.

EXPERIENCE

Freelance Data Scientist   **2023 - Present**
Alamo, CA

- Leading efforts to enhance the performance of a Generative model that uses a a novel Variational Autoencoder (VAE) architecture to provide interpretable deep learning models for medical imaging data (link)
- Leveraged Causal Machine Learning to identify customer groups for maximizing coupon campaign ROI. This approach yielded a 48% higher return on pricing and promotion investments compared to traditional experimentation methods like A/B testing. Deployed model on Databricks and AWS using PySpark to accommodate large volume of data (link)
- Trained a Semantic Segmentation model in TensorFlow to extract key geologic features from historic USGS maps (Dice Coefficient = 0.89, Precision = 0.85, Recall=0.93). The model provides a robust automated solution for large-scale geologic feature extraction from diverse maps, effectively reducing manual feature extraction efforts by 54% (link).
- Performed Natural Language Processing and sentiment analysis using the Yelp review dataset on the Hugging Face Hub. Built machine learning models in Scikit_Learn (bag of words, bigram and Naive Bayes models) to predict the rating of a review from its text (link).

Data Scientist / Engineer   **2014 - 2022**
Montgomery & Associates , Oakland, CA

- Spearheaded Python and SQL integration to streamline data processing tasks, achieving a 60% boost in efficiency through automation.
- Acquired, cleaned and transformed data from various sources. Identified and addressed data quality issues, missing values, and outliers.
- Developed custom data pipelines using Python to wrangle and preprocess large datasets for analysis.
- Utilized SQL queries and Python scripting to extract, transform, and aggregate data from various sources for analysis and visualization.
- Designed and implemented statistical and machine learning models to solve complex business problems.
- Evaluated and optimized model performance using appropriate techniques and metrics.
- Created visually compelling products, including dashboards and reports, to effectively communicate complex findings and insights to clients with varying technical backgrounds.
- Collaborate with cross-functional teams to translate data-driven insights into actionable recommendations.

SKILLS

Programming Languages: Python, MATLAB, Spark, SQL
Machine/Deep Learning :Scikit-Learn, PyTorch, TensorFlow, Keras
Computer Vision: OpenCV
Statistical Data Analysis, Hypothesis testing & Causal Inference: SciPy, Statsmodels, DoubleML, EconML
Data Wrangling & Visualization : SQL, Numpy, Pandas, Matplotlib, Tableau
Natural Language Processing : spaCy, NLTK, regex
Distributed Version Control: Git
Geospatial Data Analysis and Mapping: GIS, GeoPandas

EDUCATION

Ph.D., Engineering, University of California, Davis   **2011**
B.S., Engineering, Sharif University of Technology   **2004**

CERTIFICATIONS

- **Large Language Models Professional Certificate**, Databricks   **2024**
- **Creating Dashboards and Storytelling with Tableau**, UCDavis   **2024**
- **Interpreting Machine Learning Models,** Uplimit   **2023**
- **Deep Learning Specialization**, DeepLearning.AI   **2023**
- **Data Scientist Certification,** The Data Incubator   **2022**