Shreya Patil

Mobile +1-443-453-8281 | kognole.shreya@gmail.com |Baltimore, MD

https://github.com/shrpatil/ https://www.linkedin.com/in/shreya-patil-470219105/

PROFESSIONAL SUMMARY

- Data Scientist with 3+ years of experience in Machine Learning, Deep Learning and High-Performance Computing technologies with enhanced skills of building prototypes and designing algorithms.
- Research knowledge with a demonstrated history of working on Weather Forecasting using Neural Net Architecture Search Algorithm.
- Experience with tools Docker, Tableau, Jupyter, GitHub, Cloud platforms (Microsoft Azure).
- Proficient in Python (Pandas, NumPy, NLTK, Matplotlib, Seaborn, Scikit-learn, Regular Expression), PySpark, SQL, Linux, CUDA.
- Professional experience with Data warehouse, Data Analysis, Exploratory Data Analysis.
- Successful background of managing Databases and Cloud Platforms by leveraging infrastructure and operational knowledge for four clients in past work experience.

PROFESSIONAL EXPERIENCE

Data Science, Redapt Inc   May 2022- Present

- Build a predictive model for non-profitable organization using decision tree on Azure ML Studio
- Collaborated closely with clients to ascertain data architecture and analysis needs to design custom analytics reports
- Analyzed data from Azure Blob Storage to create reports using PowerBI tool.
- Researched and prepared a presentation on Quantum Computing for FinovateFall conference.

Research Assistant, UMBC, MD (PI: Dr. Milton Halem)   Aug 2021- Apr 2022

- Developed a Deep Neural Network architecture using python for time series microphysics parameters, to reduce the computational time by replacing components in NASA Unified-Weather Research Forecasting model (NU-WRF).
- Trained and tested Auto-Kera's Neural Net Architecture Search (NAS) algorithm on WRF-CHEM models microphysics output over US nested domains at a coarse and high-resolution data with RMSE of 0.015.
- Analyzed 3 TB data stored in Hadoop distributed file system to retrieve essential parameters of microphysics to predict the precipitation with 40% speedup.

Senior Analyst, Capgemini Technology Service LTD, INDIA   Mar 2018- Nov 2019

- Maintained and administered databases and virtual environments in VMware and Microsoft Azure platform for 10 clients.
- Delivered in-depth training, imparting knowledge of best practices to the 10+ new recruits as Account Lead.

- Organized system infrastructure documentation and operating procedures, strengthening overall team performance.

EDUCATION.

Master of Professional Science - Data Science   **May 2023**
University of Maryland, Baltimore County (UMBC), Baltimore, MD

Master of Technology - Computer Science and Engineering   **May 2019**
Shivaji University, MH, INDIA

Bachelor of Engineering – Computer Science and Engineering   **May 2017**
Shivaji University, MH, INDIA

ACADEMIC PROJECTS

Prediction of Drug Binding Affinity of Protein Using Spark ML   **May 2022**
Extracted biological data for a single protein target. Used PySpark classification models to predict binding between sample drug compounds and target protein. Compared multiple models to find optimum solution for prediction.

Understanding the X-ray Diffraction Data of Protein Structures to Predict Resolution   **Mar 2022**
Filtered X-ray diffraction data for available protein structures from RCSB Protein Data Bank. Compared Logistic regression with regularization and decision tree models with hyperparameter tuning using grid search cross validation method. Predicted the resolution of protein structure with 85% accuracy.

WhatsApp Chat Analysis   **Jan 2022**
Analyzing the WhatsApp groups chat dataset using Natural Language Processing (NLP). To get the insight of how the age difference of people in a group affects the characteristics of the chat and the activities in the group.

Optimal Number of Cluster Identification using Robust K-means Algorithm.   **May 2019**
Developed a Robust K-means Algorithm to identify optimal numbers of clusters in protein sequences by removing noise clusters. Measured goodness of clusters using Silhouette Coefficient.
Publication:
- o Patil S U Nuli U A (2018), A Review of Clustering and Clustering Quality Measurement. International Research Computer Engineering in Research Trends.
- o Nuli U A, Patil S U (2019), Optimal Number of Cluster Identification using Robust K-means Algorithm. International Research Journal of Engineering and Technology.

CERTIFICATION
- Python Specialization (Retrieving, Processing and Visualizing Data), Coursera
- Fundamentals of Visualization with Tableau, Coursera
- C2090-930 IBM SPSS Modeler Version 18(V3).