**MADAN NEUPANE**

[**mxneu17@gmail.com**](mailto:mxneu17@gmail.com) (**936)-701-1048 Mountain View, CA**

<https://www.linkedin.com/in/m-neupane-ds/>

**PROFESSIONAL SUMMARY**

Seasoned Data Scientist with extensive experience in Data Analysis, Visualization Techniques and Predictive Analytics. Skilled in Python, SQL, Tableau and Big Data Technologies and a good Problem Solver with a proven track record at Bank of America and Cigna Health. Aspiring to leverage my data science expertise in a dynamic, innovative environment. Aiming to contribute significantly to data-driven decision-making, utilizing Advanced Analytics and Machine Learning.

**EDUCATION**

**Master of Science in Statistics, 2023 |** Sam Houston State University, Huntsville, TX

**Bachelor of Science in Electrical Engineering, 2018** | Tribhuvan University, Lalitpur, NP

**PROFESSIONAL EXPERIENCE**

**Bank of America, Dallas TX  Oct 2022 - Present**

**Data Scientist**

* Cleaned and Preprocessed complex datasets using SQL, R, and Python, laying a solid foundation for insightful data analysis.
* Implemented various statistical algorithms, including Multivariate, Linear, and Logistic Regression, for accurate risk estimation in welfare dependency contexts.
* Created and fine-tuned predictive models for risk management using Linear Regression, Random Forests, and SVM leading to a 90% efficiency.
* Utilized NLP techniques like topic modeling, text classification, and chatbots to improve customer service efficiency.
* Executed model monitoring, recalibration, and continuous improvement processes maintaining an accuracy of 96%.
* Led fraud detection initiatives using anomaly detection algorithms and analyzed transaction patterns improving the process and business gains by 35%.
* Collaborated with Data Engineers on ETL processes and optimized SQL queries for effective data extraction.
* Conducted data analysis using Hive and RedShift for efficient data retrieval and exploration.
* Utilized Spark Core, Spark SQL, and Spark machine learning libraries for handling big data projects.
* Applied data imputation and preprocessing techniques using Scikit-learn for optimal model training.
* Engineered features, normalized data, and encoded labels using Scikit-learn preprocessing tools.
* Automated manual processes using Spark, Python, and AWS for increased operational efficiency.
* Actively collaborated with cross-functional teams for successful project execution and delivery.
* Utilized Tableau and PowerBI for comprehensive and interactive data visualization.
* Integrated emerging technologies like Kubernetes, Docker, and TensorFlow to stay current in the data science field.
* Developed and implemented predictive models using machine learning algorithms such as linear regression, classification, multivariate regression, Naive Bayes, Random Forests, K- means clustering, KNN, PCA and regularization for data analysis.
* Adopted Agile and Scrum methodologies for efficient project management and team collaboration.
* Experienced with a focus on Big data, Deep Learning, Machine Learning, Image processing or AI.

**Environment:** *SQL, R, Python, Multivariate Regression, Linear Regression, Logistic Regression, Linear Regression, Random Forests, SVM, NLP, Hive, RedShift, Spark Core, Spark SQL, Spark MLlib, Scikit-learn, AWS, Tableau, PowerBI, Kubernetes, Docker, TensorFlow, Naive Bayes, K-means clustering, KNN, PCA, Agile, Scrum.*

**Cigna Health, Houston TX       Sep 2021 – Oct 2022**

**Data Scientist**

* Developed and optimized predictive models using Python Scikit-learn for fraud detection.
* Project management in financial data analytics, from initial planning and problem definition to the development, training, evaluation, and deployment of predictive models by 60%.
* Cleaned and manipulated complex datasets to create the data foundation for further analysis and the development of key insights (MS SQL server, R, Tableau, Excel).
* Utilized advanced machine learning techniques including Random Forest, Gradient Boosting, Multivariate Regression, and Logistic Regression.
* Skilled in big data processing with Spark SQL and Python on AWS; proficient with Hive and Spark for data manipulation.
* Conducted comprehensive data cleaning and preprocessing using Pandas and Numpy to ensure data integrity.
* Expert in data visualization and reporting using Tableau, SQL, and Excel; proficient in data blending and preparation.
* Performed data manipulation, data preparation, normalization, and predictive modeling. Improve efficiency and accuracy by evaluating models in Python and R.
* Collaborated with data engineers, business analysts, and internal stakeholders in an Agile environment.
* Streamlined data analysis and model optimization workflows with Python.
* Improved fraud prediction performance by using random forest and gradient boosting for feature selection with Python Scikit-learn.
* Analyzed and solved business problems and found patterns and insights within structured and unstructured data.
* Developed MapReduce / Spark Python modules for predictive analytics & machine learning in Hadoop on AWS.

**Environment:** *Python, Scikit-learn, MS SQL Server, R, Tableau, Excel, Random Forest, Gradient Boosting, Multivariate Regression, Logistic Regression, Spark SQL, AWS, Hive, Spark, Pandas, Numpy, MapReduce, Hadoop*.

**NCELL, Kathmandu, NP Sep 2017 – Jul 2021**

**Software Developer**

* Skillfully designed and developed robust telecom software applications, leveraging Python and Java for optimal functionality and efficiency.
* Expertly applied critical telecom protocols such as SIP, RTP, and VOIP to develop and enhance key features in telecom software.
* Managed complex networking and communication tasks with proficiency in TCP/IP, UDP, and HTTPS, ensuring seamless data transmission.
* Efficiently managed telecom data using various database systems like MySQL, PostgreSQL, and NoSQL, ensuring data integrity and accessibility.
* Developed and maintained intricate telecom billing systems, focusing on accuracy, reliability, and user-friendly interfaces.
* Implemented various features using popular frameworks such as Spring, Django, and Flask, enhancing the software's functionality and user experience.
* Utilized Git for effective source code management, maintaining a high standard of code quality and collaborative development
* Employed Jenkins for CI/CD, ensuring rapid and reliable deployment of software updates and new features
* Expertly debugged and troubleshot software applications, identifying and resolving issues promptly to maintain optimal performance
* Diligently complied with various telecom standards including 3G, 4G, LTE, and VoLTE, ensuring high-quality voice and data services
* Contributed significantly to the provision of high-quality voice and data services, maintaining standards in line with the telecom industry's best practices.

**Environment:** Python, Java, SIP, RTP, VOIP, TCP/IP, UDP, HTTPS, MySQL, PostgreSQL, NoSQL, Spring, Django, Flask, Git, Jenkins, 3G, 4G, LTE, VoLTE.

**TECHNICAL SKILLS**

|  |  |
| --- | --- |
| Languages | *Python, SQL, Scala (for Spark), PySpark, R* |
| Skills | *Machine Learning (Regression, Classification, Clustering), Deep Learning (Neural Networks, NLP), Bayesian Learning, Optimization, Pattern Identification, Behavioral Modeling, Statistical Modeling* |
| Tools | *Tableau, Power BI* |
| Data Framework | *Hadoop, HDFS, Map Reduce, Kafka, Hive, Apache Spark* |
| Cloud Technologies: | *AWS (lambda, S3, Glue, EC2, EMR, Redshift), Azure (Data Lake Gen2, Data factory, Azure SQL), Snowflake, Databricks, GCP* |

**REFERENCE AVAILABLE UPON REQUEST`**