Varun Vuppala
Data Scientist
varun.v@mymailwork.com | +1 (347)-806-7967 | San Marcos, CA

SUMMARY

- Data Scientist with 3+ years of experience in Data Extraction, Modeling, Wrangling, Statistical Analysis, Machine Learning, and Data Visualization.
- Expert in Agile, Waterfall, and SDLC methodologies, with a deep understanding of statistical techniques such as Hypothesis Testing, ANOVA, and Time Series analysis.
- Well-versed in Machine Learning algorithms and Predictive Modeling, including Regression Models, Decision Trees, XGBoost, Random Forests, Sentiment Analysis, Naïve Bayes Classifier, and SVM.
- Understanding Python with SciPy Stack packages, including NumPy, Pandas, SciPy, and Matplotlib.
- Skilled in crafting impactful data visualizations using Tableau, Power BI, and Microsoft Excel.
- Experienced with cloud platforms like AWS and GCP, and adept at managing databases including MySQL, SQL Server, and MongoDB.
- Expert in data cleaning and modeling, delivering actionable insights through strong communication and problem-solving skills.

EDUCATION

Master in Computer Science | California State University, San Marcos, CA

Bachelor in Information Technology | CVR College of Engineering, Hyderabad, Telangana

SKILLS

Methodologies: SDLC, Agile, Waterfall

Language: Python, R, SQL, SAS, SPSS

IDEs: Visual Studio Code, PyCharm, Jupiter Notebook

Statistical Methods: Hypothetical Testing, ANOVA, Time Series

Machine Learning: Regression analysis, Bayesian Method, Decision Tree, Random Forests, Support Vector Machine, Neural Network, Sentiment Analysis, K-Means Clustering, KNN, Classification, SVM, Naive Bayes, Natural Language Processing (NLP), LLM, CNN, XGBoost, Deep Learning

Packages: NumPy, Pandas, Matplotlib, SciPy, ggplot2, Scikit-Learn, PyTorch, TensorFlow, Keras, Spark

Visualization/Other Tools: Tableau, Power BI, Jira, Microsoft Excel

Cloud Technologies: AWS, GCP, Azure

Database: MySQL, SQL Server, Oracle, MongoDB

Software/Other Skills: Data Cleaning, Data Wrangling, Critical Thinking, Communication Skills, Presentation Skills, Problem-solving, Decision Making, EDA, Communication Skills, Databricks, Data Visualization, Predictive Analytics, Pattern Recognition, JMP, Data Integrity, Quantitative Data, Data Science, Statistics, Statistical Analysis, Data Analytics, Data Modeling, Big Query, Snowflake, Data Analysis, Data Mining, SAP, Mathematics, Computer Science, Programming

Operating System: Windows, Linux

WORK EXPERIENCE

Synchrony Financial, USA | Data Scientist   Jan 2024 - Current

- Led Agile ceremonies, including sprint planning, daily stand-ups, sprint reviews, and retrospectives, ensuring iterative development and fostering a culture of continuous improvement. Proficient in SQL with expertise in complex queries, database management, and optimization.
- Enhanced sales forecasting accuracy by 30% through hypothesis testing and time series analysis, driving data-informed decision-making.
- Spearheaded developing and deploying a text classification model, working with cross-functional teams to extract actionable insights from healthcare datasets, enhancing clinical decision-making and patient care.
- Engineered predictive models using advanced machine learning algorithms, such as linear and multivariate regression, Naive Bayes, Random Forests, K-means clustering, KNN, PCA, and regularization, for robust data-driven insights.
- Leveraged AWS EC2 (Elastic Compute Cloud) instances for running data processing tasks, model training, and inference, utilizing scalable computing resources to handle variable workloads.
- Developed NLP models to analyze customer feedback, uncovering key themes and sentiment trends for product enhancement.
- Managed end-to-end ETL processes, orchestrating the extraction, transformation, and loading of data from multiple sources into the data warehouse, ensuring seamless data availability and consistency.
- Crafted compelling data visualizations with Tableau and Matplotlib, translating complex datasets into intuitive, decision-friendly formats.
- Collaborated with database administrators to optimize MySQL performance, refining query execution plans for enhanced processing speed.

Borland Software, India | Data Scientist   Jan 2020 - July 2022

- Executed rigorous testing and validation of data science models and solutions before deployment, adhering to structured Waterfall methodologies to ensure reliability and accuracy.
- Executed A/B testing to evaluate the effectiveness of marketing campaigns and product features, providing actionable insights that drove significant optimizations. Strong understanding of Data Science principles, including machine learning models and statistical analysis.
- Optimized large-scale data processing with Apache Spark, cutting processing times by 40% and boosting data throughput.
- Applied advanced machine learning algorithms such as Naïve Bayes, KNN, Logistic Regression, Random Forest, and XGBoost to predict loan defaults, improving risk management strategies accurately.
- Improved model accuracy by 20% by deploying deep learning models using TensorFlow and PyTorch.
- Automated end-to-end data pipelines using Azure Data Factory and Azure Databricks, enhancing operational efficiency and significantly reducing manual errors.
- Automated real-time reporting by integrating Power BI with Excel and other data sources for accurate data-driven insights.
- Ensured data integrity and availability with robust backup and recovery strategies for SQL Server databases.