SHNEHI KARKI
San Mateo, CA • 571-239-9133• karkishnehi@gmail.com • Github • Linkedin

PROFESSIONAL SUMMARY

Proven track record in collecting, cleaning, and preprocessing large datasets, performing exploratory data analysis, and developing sophisticated data visualizations. Demonstrated ability to deploy machine learning models in production environments, ensuring scalability and operational efficiency. Strong collaborative skills, with experience working in cross-functional teams and communicating findings to stakeholders through comprehensive reports, presentations, and interactive dashboards. Proficient in a wide array of programming languages, databases, and analytical tools, including Python, R, SQL, Tableau, and PowerBI. Committed to continuous learning and staying updated on industry trends and best practices to deliver high-impact solutions in data analysis and data science.

SKILLS

Programming Languages and Databases: Python, R, SQL, HTML, CSS, MySQL, PostgreSQL
Software Tools: Git, Jupyter Notebook, Google Colab, RStudio
Visualization Tools: Tableau, PowerBI, Matplotlib, Seaborn, Plotly, Bokeh, ggplot2, Geopandas, Folium, R Shiny, Streamlit
Analytical Techniques: Statistical Analysis, Machine Learning Techniques, Data Mining, Predictive Analytics, Deep Learning, Neural Network, Model Deployment & Validation, A/B Testing, Time Series Analysis, Natural Language, Statistical Framework, Statistical Tools, Algorithm Development, Statistical FrameWork

RELEVANT EXPERIENCE

Knowledge Scale, Manassas, VA April 2023 – Feb 2024
Data Analyst Intern

- Spearheaded the redesign of data structures utilizing advanced methodologies with MySQL and PostgreSQL, enhancing system accessibility and reliability, thereby achieving a notable 15% boost in data processing efficiency.
- Led targeted data analysis projects using Python and SQL in Jupyter Notebook environments, applying machine learning and statistical techniques that drove strategic business decisions and resulted in significant product enhancements.
- Crafted advanced data visualizations and interactive dashboards using Tableau and PowerBI, which improved stakeholder decision-making efficiency and satisfaction, evidenced by a 25% increase in positive feedback.

Omdena, Remote Feb 2023 to Present
Junior Data Scientist **(Contract)**

- Deployed and scaled machine learning models, including predictive analytics tools, using Python and Streamlit in production environments, enhancing business operations and meeting stringent scalability demands.
- Generated in-depth analytical reports from machine learning experiments using Python (Pandas, Matplotlib, Seaborn), delivering insights that bolstered strategic planning and led to a 30% improvement in business responses.
- Collaborated with cross-functional teams to design and implement AI-driven tools using Scikit-learn and Streamlit, aligning with strategic goals and boosting operational efficiency by 20%.

PROJECTS

The Blue Zone Project April 2024 – May 2024

- Analyzed diverse datasets including health, lifestyle, social, and environmental factors using Python, R, and SQL, to pinpoint patterns indicative of "blue zones", areas where individuals experience longer and happier lives.
- Leveraged Python for data analysis and machine learning, particularly utilizing deep learning algorithms, to achieve accurate predictions and insights.
- Collaborated with a multidisciplinary team to integrate findings into actionable strategies for promoting well-being in identified regions.

Predicting Alzheimer's Disease Stages Using Deep Learning with CNN Mar 2024

- Evaluated the model's performance using Python and TensorFlow, achieving a testing accuracy of 91.38%, demonstrating its efficacy in categorizing Alzheimer's disease stages.
- Utilized neural network techniques with Keras and TensorFlow to build the model, focusing on deep learning, and measured its performance using key metrics like accuracy, area under the curve (AUC), and F1-score.
- Collaborated with healthcare professionals to validate model predictions and explore potential clinical applications, contributing to advancements in Alzheimer's disease research and diagnosis.

Classifying Space Debris Feb 2024

- Conducted thorough statistical analyses using Python to identify patterns in space debris data, enhancing the predictive model's robustness and improving accuracy by 15%.
- Deployed the predictive model on Streamlit, enabling interactive user engagement and providing real-time performance visualization, which increased user interaction by over 30%.

- Presented findings and methodologies at industry conferences, contributing to knowledge sharing and professional development in the field of space debris analysis.

EDUCATION

Eastern University, St. Davids, PA
Master of Science in Data Science
Anticipated Graduation: October 2024
GPA: 4.0
Relevant Coursework: Introduction to Foundation of Data Science, Introduction to Statistical Modeling, Analytics with R, Data Manipulation, Data Visualization, Applied Machine Learning

Southern New Hampshire University, Manchester, NH
Master of Science in Data Analytics
GPA: 4.0
Relevant Coursework: Enterprise Data Management, Decision Methods and Modeling, Presentation and Visualization of Data, Optimization and Risk Assessment, Predictive Analytics, Advanced Data Analytics, Project Management

University of Massachusetts at Lowell, Lowell, MA
Bachelor of Science in Information Technology
GPA: 3.90
Courses: Introduction to Python Programming, C++, Java Development, HTML CSS, Database Modeling, Data and Database Management with SQL

University of Southern Maine, Portland, ME
Bachelor of Science in Biology
GPA: 3.5
Courses: Calculus I, Calculus II, Biostatistics, Statistics

LEADERSHIP AND ACHIEVEMENTS

- Distinguished Scholar, Southern New Hampshire University: Awarded for outstanding academic performance in data science.
- Project Lead, Omdena: Led a team of 5-6 in developing a predictive machine learning model to forecast urban growth in Africa, enhancing decision-making and project collaboration.