SEJAL KANKRIYA

(716) 305-0091 | sejaldee@buffalo.edu | LinkedIn | GitHub | Portfolio

EDUCATION

M.S. in Data Science, University at Buffalo, The State University of New York **Aug 2022 - Feb 2024**
Coursework: Computer Vision and Image Processing, Statistical Learning, Probability, Data Modelling & Query Languages

B.E. in Computer Engineering, Vishwakarma Institute of Information Technology **Aug 2017 - Jun 2020**
Coursework: Design & Analysis of Algorithm, Artificial Intelligence & Robotics, Data Analytics and Machine Learning

PROFESSIONAL EXPERIENCE

Research Assistant, University at Buffalo, Buffalo, NY **May 2023 – Dec 2023**

- Coordinated with 5-member team in applying GNNs for hospital monitoring, resulting in a 27.4% enhancement in patient movement prediction accuracy and optimizing resource allocation
- Adopted Natural Language Processing (NLP) techniques to process text-based comments, achieving patient outcome prediction accuracy of 89%

Programmer Analyst, Cognizant, India **Jan 2020 - Jul 2022**

- Performed exploratory data analysis using Tableau, Jupyter Notebook, and Python (Pandas, NumPy, and Bokeh) to craft 10+ insightful, interactive visualizations, facilitating better decision-making and client understanding
- Deployed a standalone pipeline through CI/CD to build and integrate test microservices with Docker, AWS, and CDK
- Collaborated with stakeholders to translate business requirements into efficient data models, reducing storage space by 25% by minimizing data redundancy
- Managed and optimized database operations for over 8 million records in PostgreSQL and Oracle
- Applied Agile methodologies to streamline inquiry processes, reducing response time by 15% and providing support for 50+ weekly inquiries across Confluence, Jira, and Bugzilla
- Managed and mentored four new joiners with onboarding, knowledge transfer, and assistance with code review

Data Analyst, ExpertsHub, India **May 2019 – Jul 2019**

- Streamlined supply chain operations by incorporating a deep learning model that predicted demand fluctuations with 95% accuracy, using TensorFlow and Python
- Leveraged MongoDB to administer large datasets, efficiently handling over 2TB of data
- Accomplished 20% reduction in stockouts and overstock situations by incorporating predictive analytics with Tableau

Data Science Intern, PixeLink Technologies, India **Jan 2019 – Feb 2019**

- Spearheaded development of an AI Chatbot, leveraging NLP and ML algorithms, improving patient interaction efficiency by 25% and decreasing manual queries by 34%
- Engineered and launched a Kafka-based streaming data pipeline that processed over 1 million messages per day, enabling real-time data analysis and cutting data latency by 40%

PROJECTS

AcademiQ AI | *Python, LLM, NLP, Flutter, Django*

- Developed an AI-powered teaching assistant that handles real-time academic inquiries, boosting engagement through personalized interactions, utilizing a custom LLM and integrating with Flutter and Django for cross-device compatibility
- Achieved significant reduction in response times for student inquiries, with 50% decrease in wait times for answers, by optimizing the LLM model's processing capabilities and improving backend efficiency

Satellite Image Dehazing | *Python, PyTorch, OpenCV, TensorFlow, NumPy, Matplotlib, Seaborn*

- Built and implemented AOD-Net image dehazing model on a 1000-image unstructured dataset. Refined existing model, boosting PSNR 20% and SSIM 10%
- Directed a team to develop a real-time image dehazing application utilizing Python and OpenCV, expediting processing time from ~10 seconds to ~4 seconds per image

360-degree object detection and assistance for visually impaired people | *Python, SQL, Java, TensorFlow, NumPy*

- Led implementation "Hope," a robotic assistant attaining 92.8% accuracy in AI-based object detection for visually impaired users, employing ResNet18 and NLP for intuitive navigation
- Improved real-time data processing speeds through efficient cloud-based image handling on AWS EC2, resulting in a 7-second decrease in object recognition time
- Utilized AWS EC2, S3 and AWS IoT to deploy cloud-based services, ensuring smooth initiation and ongoing functionality

Time-To-Event Analysis in Clinical Trials | *R, Data Analysis (dplyr, survival analysis), ggplot2*

- Conducted time-to-event analysis for a clinical trial employing Kaplan-Meier estimator, revealing 13.3% increase in estimated post-diagnosis survival probability for new drug

RESEARCH PUBLICATION

International Journal for Research in Applied Science & Engineering Technology, Volume 8, Issue VI, June 2020
Technical Paper Number (IJRASET29247): Impact of COVID-19 on Indian Economy

SKILLS

Programming Languages: Python, R, Java, C++, MATLAB, SQL, Bash, HTML5, Angular, JavaScript
Libraries and Frameworks: PyTorch, TensorFlow, Keras, scikit-learn, Flask, Matplotlib, Seaborn, dplyr, NLTK, HuggingFace
Big Data Technologies and Cloud Computing: Hadoop, Spark, Hive, PySpark, Microsoft Azure, Amazon AWS, GCP
Advanced AI and Analytics: Predictive Modelling, Statistical Modelling, Time Series Analysis, Natural Language Processing, Deep Learning, LLMs (GPT-3), Prompt Engineering, A/B Testing, Computer Vision
Visualization Tools: Tableau, Power BI, Microsoft Excel, Looker
Database Technologies: PostgreSQL, MySQL, MongoDB, SQLite, IBM DB2