Akhil Bukkapuram

Raleigh, NC | +1 (919)-931-8510 | bakhil@ncsu.edu | linkedin.com/in/akhil-bukkapuram

EDUCATION

Master of Science in Electrical Engineering **GPA:** 4.0/4.0
North Carolina State University, Raleigh August 2022 - May 2024
Courses: Topics in Data Science, Advanced Topics in Machine Learning, Resource Dependant Neural Networks, Accelerating Deep Learning, Random Processes, Digital Imaging Systems, Cloud Computing Technology, Pattern Recognition, Object Oriented Design and Development

Bachelor of Technology in Electronics and Communications Engineering **GPA:** 8.17 /10.0
PES University, Bangalore, India August 2018 - July 2022
Relevant Courses: Machine Learning, Deep Learning, Random Processes, Digital Image Processing, Signal Processing, Linear Algebra, Control Systems, Python, C
Minors: Computer Science (Database Management Systems, Data Structures, Design and Analysis of Algorithms, Operating Systems)

SKILLS

Programming Skills: Python (Scikit-Learn, Keras, NumPy, TensorFlow, PyTorch, Matplotlib), C, C++, R, MATLAB
Database skills: PostgreSQL, MySQL, Snowflake
Cloud & Toolkits: AWS, Spark, Docker, Git, Tableau, Spreadsheets
Machine Learning: ANN, CNN, RNN, GAN, Regression, Random Forests, KNN, Transformers

EXPERIENCE

Research Assistant, NC State University, Raleigh May 2023 - Present
- Investigated the effect of sample influence on pruning and generalization in neural networks.
- Analyzed existing literature on model re-use, domain shift, and reliability in machine learning, identifying three potential research areas for further exploration.
- Collaborated with fellow student researchers by reviewing their proposed research directions and offering constructive feedback, leading to the development of stronger research proposals.

Data Quality Intern, ABB, Bangalore January 2022 - March 2022
- Conducted profiling and cleaning activities to ensure data accuracy, completeness, and consistency.
- Supported the IT team to create an automated dashboard using Excel and Tableau that provided real-time visibility into inventory levels, reducing manual errors by 80%.

Summer Intern, Centre for Data Science and Applied Machine Learning, PES University August 2020 - January 2021
- Designed a deep learning network for medical diagnosis, achieving a sensitivity of 95% and a specificity of 92% on a dataset of 10,000 medical images by utilizing over-sampling and cost sensitive learning techniques.
- Conducted an ablation study on the network to evaluate the contribution of individual model components to performance, resulting in a 3.5% improvement in accuracy by removing redundant layers.
- Suggested and implemented pruning and quantization that reduced the size of the model by 70% while maintaining 95% of the original accuracy.

Data Analysis Intern, Sunrise Technologies, Bangalore May 2020 - August 2020
- Improved data quality and integrity by developing cleaning and validation procedures, resulting in a 20% reduction in errors.
- Developed and automated weekly reports and dashboards using Tableau to track KPIs and provide actionable insights to stakeholders.

ACADEMIC PROJECTS

Terrain Identification from Time Series Data
- Addressed sampling mismatch in raw time-series sequential IMU measurements and applied data augmentation techniques (ADASYN and SMOTE) to handle imbalanced data for terrain identification.
- Conducted a comparative evaluation of Random Forests and Bidirectional LSTM models using different window sizes for terrain identification. Significantly improved the weighted average F1 score from 0.75 (baseline model) to 0.82.

Mapping Mangrove Vegetation in Multispectral Images
- Developed a comprehensive project plan to create a custom dataset of 20 very large size images using Landsat-8 from USGS Earth Explorer, in collaboration with Tech for Wildlife, Goa.
- Utilized advanced image-processing techniques to analyze and annotate the dataset, resulting in an accuracy rate of over 90% on the ground truth masks.
- Developed a Multi-Scale U-Net model for segmenting mangrove vegetation in satellite images with a 72% IoU metric.

Stock Market Analysis
- Processed historical data from more than 600 companies listed in the NASDAQ and S&P 500, applying data cleaning techniques to extract meaningful insights from JSON endpoints using Pandas.
- Demonstrated proficiency in Python's Data ETL, Visualization, API frameworks, and statistical methods to create informative visualizations and perform relevant analyses.

Context Based Sarcasm Detection with NLP
- Curated a news dataset of 28,000+ entries using BeautifulSoup. Pre-processed using pandas and NumPy (ETL).
- Investigated context dependency for sarcasm detection by employing embedding & tokenization by NLTK libraries.
- Developed and trained Bi-LSTM & RoBERTa models for sarcasm detection, achieved 96% classification accuracy