GIRISH WANGIKAR

(+1) 984-944-3688 | Raleigh, NC | gwangik@ncsu.edu | linkedin.com/in/girish-wangikar | github.com/GirishWangikar | Portfolio

EDUCATION

North Carolina State University, Raleigh **Aug 2022 - May 2024**

Master of Science, Electrical and Computer Engineering **GPA - 3.95**

Coursework: Topics in Data Science, Advance Machine Learning, Automated Learning and Data Analysis, Pattern Recognition, Neural Networks and Deep Learning, Object Oriented Design and Development

Rajarambapu Institute of Technology, Maharashtra, India **Jul 2018 - Jul 2022**

Bachelor of Technology, Electronics and Telecommunication **GPA - 3.83**

TECHNICAL SKILLS

Programming Languages: Python (Proficient), SQL (Fluent), R (Fluent), MATLAB (Fluent), C++(Intermediate)

Frameworks/Libraries: PyTorch, TensorFlow, Keras / OpenCV, Pandas, Seaborn, Scikit-image, Scikit-learn, SciPy, NumPy, Matplotlib

Tools: TensorFlow, PowerBI, Git, Docker, Kubernetes, AWS, PyTorch

Miscellaneous: Effective Data Visualization, Data Wrangling With Attention to Detail, Strong Mathematical Skills (Statistics, Probability Theory, Geometry) and Knowledge of Machine Learning Models

PROFESSIONAL AND RESEARCH EXPERIENCE

Research Intern **Raleigh, North Carolina**

Sozzani Lab **Jun 2023 - Present**

- Assisting in the ongoing development and implementation of **RNN** for analyzing genomics **time-series** data in the context of gene regulatory networks and create prediction models to further enhance accuracy over previous methods.
- Continuously conducting data gathering, sorting, and analysis of **30,000+** genomics data samples using **SQL** to ensure the ongoing quality and usability of genomics data for further investigation.
- Currently involved in the addition of an **AutoEncoder** framework tailored for **predictive modeling** of Differentially Expressed Genes (DEG's) in plant genomes, aiming to improve the team's ability to identify the function of these important genes in plant yield.
- Produced comprehensive **research paper** style documentation detailing the **algorithms** and **deep learning** methodologies employed, with the aim of future publication.

Data Science Intern **Maharashtra, India**

Exposys Data Labs **Jan 2022 - Mar 2022**

- Conducted in depth **ETL** operations on **500GB** of raw **unstructured data** to identify key factors influencing car sales and developed **predictive model** to forecast car sales by analyzing historical data and market conditions.
- Used machine learning techniques, including regression analysis and time series forecasting to uncover insights and patterns within dataset.
- Worked collaboratively with **cross-functional** teams to translate data findings into actionable recommendations while also creating data visualizations using **PowerBI** to support data-driven decision making.

ACADEMIC PROJECTS

- **Deep Learning - Deepfake Images and Video Detection Algorithm | Python (TensorFlow, PyTorch)** **Feb - May 2023**
  - Generated **120,000** fake images from **CelebA** real images dataset for different Generative Adversarial Networks (**GANs**).
  - Leveraged **Amazon S3** for centralized storage of 120,000 fake images, ensuring data integrity and **version control**.
  - Utilized **Amazon EC2** GPU instances to expedite model training, with **Accuracy of 99.37%** and **F1 score of 98.97%**.
- **Sensor Data Analysis - Terrain Identification using LSTM | Python, SQL (TensorFlow)** **Jan - Feb 2023**
  - Applied **Bidirectional LSTM** to identify terrain by making use of accelerometer and gyroscope measurements.
  - Addressed dataset imbalance by assigning class-specific weights to mitigate the impact of uneven class distribution.
  - Achieved a notable **89.3% accuracy** on the test set and collaborated closely with a team member from the Department of Statistics at NCSU to enhance the project's statistical and analytical aspects.
- **Computer Vision - 2-D Object Detection for Autonomous Vehicle | Python (Keras, TensorFlow)** **Nov - Dec 2022**
  - Utilized YOLOv3 to train a Machine Learning model on **10,000+ images** for Object Detection in Autonomous Vehicles.
  - Compared the performance with other algorithms using model Frames Per Second (FPS) and mean Average Precision (mAP).
- **Algorithm Development - Diabetes Disease Detection | Python (PyTorch)** **Aug - Dec 2021**
  - Employed **SVM** classification algorithm to ascertain diabetes presence using a locally collected dataset.
  - Utilized patient data encompassing age, BMI and glucose levels for training, achieving a classification A**ccuracy of 89%**.

CERTIFICATIONS & EXTRACURRICULARS

- Secured 3rd place in the Machine Learning track at the 3rd annual N.C. PSI **Hackathon** as a member of team of four students
- Google Cloud Training – Data Engineering, Big Data and Machine Learning on GCP Specialization Certificate
- TensorFlow Developer Certificate – Built models in TensorFlow to apply on Image Recognition, Object Detection and NLP
- NVIDIA - "Deep Learning Fundamentals", "Image Segmentation Techniques", "Time-Series Data Modeling with RNN"
- Co-authored the research paper "Diabetes Detection - An Application of Machine Learning in Healthcare Industry" published in International Research Journal of Modernization and Engineering Technology & Science, Volume 4, Issue 6.