Arunaggiri Pandian Karunanidhi

231 Quinby Road, Apartment D, Rochester, New York 14623

585-303-4622 arunaggiri.pandian@gmail.com linkedin.com/in/arunaggiri-pandian github.com/Arunaggiri-Pandian

Education

University of Rochester **Aug 2022 – Dec 2023**
Master of Science in Data Science; **GPA**: 3.7 / 4.0 Rochester, New York

Anna University **Aug 2018 – Jun 2022**
Bachelor of Engineering in Electronics and Communication Engineering; **GPA**: 3.9 / 4.0 Tamil Nadu, India

Experience

Micron Technology **May 2023 – Present**
Data Scientist Intern Boise, Idaho

- Engineered a **data preprocessing pipeline** for approximately **62,000** die failure images to optimize feature extraction for the pre-trained **Xception** model.
- Implemented a **Sequential K-Means Clustering** technique on extracted image features, achieving a **20%** improvement in accurately identifying diverse failure clusters.
- Innovated the **Dispersed Density Score (DDS)**, a novel metric for cluster quality assessment, which enhanced the efficiency of failure analysis procedures by an impressive **35%**.

University of Rochester Medical Center **Jan 2023 – Aug 2023**
Graduate Research Assistant - Computer Vision Rochester, New York

- Employed **Convolutional Neural Network** and **transfer learning**, achieving a **15%** error reduction in predicting weight of newborns.
- Optimized prenatal care by utilizing specialized pre-trained models such as **U-Net** for **fetal ultrasound images**.
- Leading a team of **four**, engaged in constructing a novel architecture aimed at **outperforming** the existing models, thereby enhancing **precision** in fetal weight prediction.

Delta Electronics **Feb 2022 – Jun 2022**
Data Analyst Intern Bangalore, India

- Improved the **First Pass Yield Rate** in the production line by analyzing unit failure data utilizing **SQL** and **Python**.
- Presented insights through **dashboards** to testing managers, resulting in a **13%** reduction in **failure analysis costs**.
- Implemented a **comparative analysis** to identify the most efficient component examination method, reducing total testing cycles and saving approximately **60 labor hours** monthly.

University of Galway **Mar 2021 – Jan 2022**
NLP Research Intern Galway, Ireland

- Designed and implemented a transformer-based sentiment analysis model with a **soft voting classifier**, resulting in a **12%** improvement in classification accuracy for **code-mixed data**.
- Proposed a novel **semi-supervised** approach to categorize misogynous memes into four overlapping categories.
- Elevated elderly speech recognition accuracy by **17%** through the development of a **speech corpus** for ASR training.

Publications

[**1**] Identification of Indian Medicinal Plants from Leaves using Transfer Learning Approach. [**View**]
[**2**] Development and Deployment of a Machine Learning Model for Automatic Heart Failure Prediction. [**View**]
[**3**] A Comparative Study on Plant Classification Performance Using Deep Learning Optimizers. [**View**]
[**4**] A Novel Convolutional Neural Network Architecture to Diagnose COVID-19. [**View**]
[**5**] Transformer Based Sentiment Analysis in Dravidian Languages. [**View**]
[**6**] A Reliable Technique for Sentiment Analysis on Tweets via Machine Learning and BERT. [**View**]

Projects

Public Opinion on Student Loan Forgiveness | Python, NLP, Twitter

- Analyzed **public opinion** on the Student Loan Forgiveness program using data **scraped** from Twitter.
- Performed **sentiment analysis** using **VADER**, determining that 51.5% of responses were positive, 31.1% were negative, and 17.4% were neutral towards the program.
- Utilized **LDA** for **topic modeling** to uncover reasons behind sentiment categorization and understand public perception.

Graduate Housing Marketplace | MySQL, HTML, CSS

- Developed and implemented a **database system** for the graduate student housing market at the University of Rochester.
- Improved housing transactions with a **user-friendly platform** for students to buy and sell.
- Streamlined **inventory management**, leading to a **75%** reduction in manual paperwork, and improved overall **tracking** and **monitoring** of available items.

Technical Skills

Languages: Python, R, MySQL, Oracle, SAS, MATLAB
Data Expertise: Statistical Modeling, Data Analysis, Machine Learning, Data Mining, Deep Learning, Data Visualization, ETL Development, Natural Language Processing, Prompt Engineering
Tools and Frameworks: PyTorch, Scikit-learn, Keras, NLTK, TensorFlow, Git, NumPy, Pandas, Matplotlib, Seaborn, Tableau, PySpark, Databricks, HuggingFace, ChatGPT