Arun Gaonkar

Linkedin: https://www.linkedin.com/in/arun-gaonkar/
Website: https://arungaonkar.github.io/

Email : agaonka@ncsu.edu
Mobile : +1-919-349-6169

EDUCATION

North Carolina State University — Raleigh, NC
Masters, Computer Science; GPA: 3.96 — *Aug 2021 - May 2023*

PES University — Bangalore, India
Bachelor of Technology, ECE; GPA: 3.64 (8.57/10.0) — *Aug 2016 - Jun 2020*

SKILLS SUMMARY

- **Languages**: Python, C/C++, Java, MATLAB
- **Frameworks**: HTML, JavaScript, CSS, REST, Jenkins, Apache
- **Databases**: MySQL, PostgreSQL, AWS DynamoDB, PowerBI
- **Tools & Technologies**: TensorFlow, PyTorch, Git, GitHub, Jupyter Notebook
- **Cloud Technologies**: Azure, Kubernetes, Docker

EXPERIENCE

North Carolina State University — Raleigh, NC
Research Assistant - Social AI Lab — *Jan 2023 - May 2023*
- Curated a new dataset to identify and extract the presence of prejudice in social media for a case study.
- Manually annotated each sentence and its target with a team of 3 for higher consistency and accuracy.
- Planning to analyze the trend and leverage transformer-based models for extracting prejudiced sentences.

North Carolina State University — Raleigh, NC
Research Assistant - Innovative Educational Computing Lab — *Sept 2022 - Jan 2023*
- Led data collection procedures and conducted interacitve sessions to understand learning patterns.
- Assisted in data wrangling and data labeling and improved agent's performance in solving algebra equations.

LexisNexis Risk Solutions — Alpharetta, GA
Intern - Data Scientist - HPCC Systems — *May 2022 - Aug 2022*
- Analyzed behavior of the HPCC_Causality toolkit on synthetic and real-world datasets for causal discovery.
- Leveraged discretization technique to increase the speed of probabilistic dependence tests by 10-fold.
- Discovered causal relations using Bayesian networks, conditionalization and drew causal models with 9+ variables.

ACADEMIC PROJECTS

- **Context Based Sarcasm Detection (Natural Language Processing)**: Created a dataset of 28,000+ news entries, utilized NLP techniques including embedding, tokenization, and Bi-LSTM & RoBERTa models to achieve 96% classification accuracy for context-based sarcasm detection, resulting in improved understanding of sarcasm in language.
- **Relative Pose with Azure Functions (Cloud Integration)**: Developed an API endpoint in Azure functions that is activated through HTTP requests. Calculated the relative position and orientation of the camera by analyzing images. Integrated with Blender to estimate the relative pose and track camera motion, and thereby enhancing camera calibration.
- **Example-Bot: Personalized Code Assistant (ChatBot & DevOps)**: Built a server-based chatbot using Ansible and Git bash for CRUD operations on MongoDB, designed to help developers create, store, and retrieve customized code snippets and API examples while following Scrum and Agile methodologies. Achieved 96% test coverage by implementing unit testing with Chai and Mocha, and deployed a seamless CI/CD pipeline.
- **GitHub Graphics Bot (Web Application)**: Designed a web application incorporated with a chatbot to visualize GitHub data. Extracted, transformed, and generated visualizations of GitHub issues, collaborators, and contributions by integrating the CI-CD pipeline with Ansible, applicable in tracking teams' progress and managing resources.
- **Wildfire Data Analysis & Cause Prediction (Data Analysis & Visualization)**: Maximized wildfire cause prediction accuracy to 93% by leading a team of 3 in building an end-to-end machine learning solution and analyzing 1.88 million records using ETL & data visualization techniques, potentially enabling effective and efficient fire prevention & response measures.
- **Brain Tumor Image Classification (Image Processing)**: Achieved 92% accuracy on MRI image classification for brain tumors by developing deep learning models with Bi-LSTM and CNN using TensorFlow and optimizing hyperparameters, leading to potential improvements in early detection and treatment.

CERTIFICATES

- IBM Data Analyst Professional Certificate
- Microsoft Azure Data Scientist Associate (DP-100) Professional Certificate