PRADEEP RAJ PRABHU RAJ

Chicago, Illinois, USA
+1-312-998-3116 pradeepraj18062002@gmail.com https://pradeeprajpraburaj.carrd.co/ https://github.com/Pradeep-18062002

Summary

Aspiring Data Science Intern with a strong foundation in machine learning, and AI. Proven ability to develop innovative projects, such as a CNN-based pet facial expression detection system and a bias-free music recommendation system. Internship experience in building secure data-sharing platforms, enhancing data integrity and automating workflows. Eager to apply technical skills to solve real-world problems in a dynamic team environment.

Skills

- **Programming Languages**: Python, Java, C, C++, SQL
- **Web Development**: HTML, CSS, JavaScript, React, Angular
- **Database**: MySQL, MongoDB, PostgreSQL
- **Machine/Deep Learning**: Classification, Clustering, PyTorch, CNN, RNN, Transformers, Generative Models, TensorFlow
- **Tools**: Git, Google Colab, Jupyter Notebook, PySpark, Hadoop, Visual Studio Code, Selenium, Scikit-learn, Day.js
- **Data Visualization Tools**: MS Excel, Pandas, Matplotlib, Seaborn, PowerBI
- **Software Testing Frameworks**: Jasmine Framework
- **API Development**: FastAPI, Postman, Pydantic

Work Experience

KaaShiv InfoTech | *Software Engineer Intern* **Feb 2022 - Mar 2022**

- Designed and developed a secure data-sharing platform using Java, incorporating advanced encryption techniques like Attribute-Based Encryption (ABE) and Proxy Re-Encryption, enhancing data integrity by 30%.
- Optimized system performance by implementing efficient algorithms for real-time file encryption and key management, using tools such as IntelliJ IDEA and JUnit for development and testing, resulting in a 30% improvement.
- Automated file-sharing processes using SMTP for real-time notifications, reducing manual intervention by 40%, resulting in streamlined workflows and enhanced user experience.

Projects

ProPublica COMPAS Bias Analysis & Mitigation **February 2025**

- Analyzed the ProPublica COMPAS dataset (10,000+ records) to detect racial bias in risk assessment predictions.
- Implemented bias mitigation techniques including Batch Processing and Counterfactual Data Augmentation, reducing disparate impact by 35%.
- Performed data analysis using Pandas, Matplotlib, and Seaborn, identifying a 42% disparity in false positive rates between racial groups.
- Visualized fairness metrics such as Demographic Parity, Equal Opportunity, and Equalized Odds, improving overall model fairness by 40%.

Bias-Free Music Recommendation System **Aug 2024 - Nov 2024**

- Developed a music recommendation system using CNN, Word2Vec, and Node2Vec, reducing popularity bias by 50% and ensuring fair exposure for lesser-known artists.
- Fetched 28,743 artist, album, and track records using optimized API calls to Last.fm, improving data extraction efficiency by 40% and supporting robust metadata for the recommendation system.
- Reduced cold-start problem impact by utilizing hybrid models that integrated content-based and metadata-based embeddings, achieving accurate recommendations for new users and items, and reducing its impact by 40%.
- Validated results using comparative analysis with state-of-the-art algorithms and visualized outcomes using Matplotlib and Seaborn, increasing recommendation precision by 30%.

Pet Facial Expression Detection **Aug 2023 - Nov 2023**

- Developed a deep learning model to classify pet facial expressions using convolutional neural networks (CNNs) with Python and PyTorch, achieving accuracy of 92%.
- Achieved increased model robustness and scalability for diverse real-world scenarios by using dropout layers and L2 regularization, reducing overfitting by 40%.
- Validated the model's generalizability by testing it across unseen datasets, achieving 90% accuracy using Scikit-learn, showcasing its versatility.

Automated Real-Time Police Radio Data Extraction **Jan 2023 - Feb 2023**

- Automated the real-time extraction and download of Chicago Police Department's radio communications using Python and Selenium, reducing manual effort by 90%.
- Configured Selenium WebDriver with ChromeDriver to track and capture newly posted audio files, enhancing real-time monitoring and achieving 100% data capture accuracy.

Education

University of Illinois at Chicago **Present**
Master of Science, Computer Science (GPA: 3.66/4) *Chicago, Illinois*

Sri Venkateswara College of Engineering **Aug 2020 - Jul 2024**
Bachelor of Engineering, Computer Science (GPA: 3.36/4) *Chennai, India*