

RAHUL VEERAPUR

+91-6362887342 • rxv5218@psu.edu • linkedin.com/in/rahul-v-rvce999 • github.com/vrahulrvce

EDUCATION

Masters Data Analytics

Aug 2024 - Feb 2026

Pennsylvania State University, Philadelphia, PA

coursework: Applied Statistics, Data Mining, Large-Scale Database and Warehouse, Database Design, Deep Learning, Reinforcement Learning, Large Scale Databases for Real-Time Analytics, Data-Driven Decision Making, Natural Learning Processing

Bachelor of Engineering in Computer Science

Sept 2018 - Dec 2023

R.V College of Engineering, Bangalore, KA

coursework: Data Structures, AI & ML, Parallel Architecture and Distributed Programming, Software Engineering, Database and Design, Big Data Analytics using Distributed Platforms, analytics programming in python.

TECHNICAL SKILLS

Languages and Frameworks: C, Java, Springboot, SQL, SQLite, MongoDB, Python, Flask, Pandas, Plotly, Flutter, R, Maven, CUDA, ECL, Tensorflow, C++, Pytorch, PySpark.

Technologies: Git, Apache, Shell Scripting, Object Detection, Machine Learning, Artificial Intelligence, Image Recognition, HPCC, Natural Learning Processing, Deep learning, CNN, ANN, RNN.

EXPERIENCE

Chicago Mercantile Exchange

Jan 2022 – Aug 2022

Software Developer Intern

- Engineered a robust Java-based search engine across Oracle DB, MSSQL, and MySQL, efficiently managing over 300 servers, each hosting 1000 databases with 300 tables.
- Enhanced data analysis efficiency for employees by ensuring high performance and streamlined data querying processes.
- Implemented advanced scripts and queries, creating a single connection tokenization for multiple servers, boosting server performance by 14%.
- Enabled user requirement forecasting and developed a centralized console for executing queries across multiple servers simultaneously, significantly boosting productivity by 24.9%.

ACADEMIC PROJECTS

Adaptive Learning and Self-Evaluation Platform

Jan 2023 – Oct 2023

Mentored a team of three members supervising and teaching them to approach the solution while developing an NLP

- Built a NLP model using 1900 of course materials trained with hugging face and SV2 for human-like interaction
- Developed and launched a dynamic self-evaluation platform using Firebase and Flask, tracking user engagement and performance trends, which increased completion rates by 80% and provided valuable insights for statistical evaluation.
- In-depth evaluation and interface created an student retention of 78%.
- Designed an adaptive learning system to analyze student performance data, enabling personalized question delivery targeted feedback to enhance learning outcomes by 59%.

Covid19 Prediction model using basic symptoms

Aug 2021 - Dec 2021

Real-time COVID-19 prediction model using big data.

- A decision tree-based random forest classification algorithm within HPCC System's architecture; focused on Covid test prediction, achieving a 33% increase in model precision and a 19% reduction in computational overhead.
- ECL modules such as MLCORE for dataset preprocessing and accuracy metrics calculation, and Learning trees for implementing the random forest classification algorithm.
- Conducted model training with 80 data split by using several data wrangling methods on a single cluster CPU utilizing the ClassificationForest from the Learning- Trees module with a decision tree.
- Deployed the trained model onto a THOR logical file classifier which was subsequently sprayed onto the HPCC System for comprehensive analysis and real-world application.
- Showed a hit ratio of 81.7% in detecting the cases of COVID based on symptoms.

Centralized Automated Facial Recognition and Identification

Aug 2021 - Dec 2021

Developed a CNN based YOLO face detection system with CUDA achieving 92% identification accuracy and 90% reliability.

- An autonomous person detection and recognition system for enhancing campus security, accurately identifying unique faces in CCTV footage with resolutions of 360 to 480p.
- Implemented contour detection using CUDA achieving a 92% identification probability match for each person significantly improving detection accuracy.
- Integrated IOT-based cameras to add a cross-referencing layer, ensuring real-time detection and enhancing system reliability.
- Continuously optimized the CNN-based algorithm to maintain maximum detection accuracy updating it to ensure a 90% reliability rate.