AARAV KALKAR
+91 88264 93349
aaravkalkar@gmail.com
LinkedIn
GitHub

EDUCATION

SRM Institute of Science and Technology, Kattankulathur, India, Sept 2020 – June 2024
Bachelor of Technology in Computer Science & Engineering (Specialization in Big Data Analytics)
Undergraduate CGPA: 9.01
WES GPA: 3.92/4
Total Semester Credits: 110
Relevant Coursework: Calculus and Linear Algebra, Advanced Calculus and Complex Analysis, Data Structures and Algorithms, Object Oriented Design and Programming, Computer Organization and Architecture, Design and Analysis of Algorithms, Operating Systems, Discrete Mathematics for Engineers, Machine Learning I, Data Science, Artificial Intelligence, Deep Learning, Database Management Systems, Probability and Queueing Theory, Formal Language and Automata, Computer Networks, Applications of Nanotechnology, Big Data Tools and Techniques, Business Intelligence and Analytics, Database Security and Privacy

INTERNSHIP EXPERIENCES

Cloud Foundations Virtual Internship at Amazon Web Services Academy, Feb 2023

- Achieved AWS Certified Cloud Foundations validation through participation in the Cloud Foundations Virtual Internship at Amazon Web Services Academy in February 2023, demonstrating an exceptional grasp of AWS Cloud essentials, services, and terminologies
- Completed 48 intensive hours of coursework on SQL and Javascript, supplemented by regular module exams to solidify expertise and validate knowledge
- Completed two capstone projects, showcasing proficiency in developing customer management websites and crafting query statements for database tables

PROJECT EXPERIENCES

Image Denoising using Neural Networks Project, Aug 2023 – Nov 2023

- Conducted a comprehensive review of deep learning- based image denoising methods, including CNNs, autoencoders, Noise2Noise, and Noise2Void
- Explored the innovative Noise2Void approach, revolutionizing image processing by training directly from noisy images without corresponding clean images, gaining valuable insights into deep learning methodologies
- Analyzed the strengths, weaknesses, and real-world applications of different denoising techniques, contributing to a nuanced understanding of image processing methodologies
- Accepted for publishing in ICAIIR 5.0

Video Denoising, Jan 2024

- Conducted an in-depth exploration of innovative video denoising techniques, leveraging adaptive methods without relying on deep learning.
- ·Presented at the AEM 2024 conference and accepted for publication in the IJARSE journal.
- ·Developed adaptive filtering techniques for spatial, temporal and DCT-Color denoising, ensuring superior performance across diverse spatial contexts.
- Implemented aggressive optimization techniques in temporal denoising, maximizing performance on resource-limited machines.
- · Experimental evaluations demonstrated the efficacy of the approach in diverse, challenging scenarios, including dynamic scenes.
- Completed the entire project on Google Colab without the need for a local machine, enhancing reproducibility and accessibility for researchers and future improvements.

Market Basket Analysis Project, Sept 2023

- Acquired and processed the Instacart dataset by performing data cleaning, removing duplicates, and formatting it for the chosen algorithm
- Implemented the selected association rule mining algorithm (FP-Growth, Apriori, or Eclat) in Python, executing the algorithm on the preprocessed dataset to derive frequent item sets and association rules
- Evaluated and analyzed the outcomes, emphasizing the generated rules and their support, confidence, and lift scores

Restaurant Management System Project – Apr 2023

- Developed an interactive web application facilitating direct customer ordering from the website, streamlining the ordering process and reducing reliance on restaurant staff
- Utilized PHPMyAdmin as an administration tool and SQL server host, alongside XAMPP as a comprehensive web server stack package
- Implemented advanced functionalities including VIP table classification, customizable control levels for administrators and customers, integration of coupon codes during checkout, and a comprehensive billing system tracking all past transactions

Handwritten Digit Classification using NLP Project – Mar 2023

- Created a Convolutional Neural Network (CNN) for the classification of handwritten digits extracted from images or documents, leveraging training data from the MNIST dataset
- Developed an automated methodology for recognizing handwritten numbers, facilitating the identification of handwritten digit strings
- Conducted a comparison of classifiers including KNN, PSVM, neural networks, and CNNs, evaluating performance metrics such as accuracy, processing time, sensitivity, and specificity across various parameter configurations

Resume Parser, Mar 2023

- Completed a python notebook that is capable of categorizing students via a resume parser, automatically segmenting them into distinct fields including name, projects, and skills
- Implemented natural language processing (NLP) techniques to extract data from various file formats such as doc, pdf, and docx across a range of fields

Sentiment Analysis Project, Jun 2022

- Completed an incremental NLP model that utilizes the iMDB dataset, comprising over 50,000 movie reviews, enabling users to search for movie titles and submit reviews
- Developed an online system to automate posting and storage of reviews, facilitating movie rating based on user sentiment
- Demonstrated proficiency in machine learning and natural language processing, along with proficiency in Google Firebase and Dialogflow tools

Age and Gender Classification Model, Feb 2021

- Utilized the Python OpenCV module to detect faces in images or webcam feeds, employing a trained model
- Evaluated the model's performance metrics using accuracy for gender classification and mean absolute error (MAE) for age prediction, achieving 81% accuracy for gender classification and an MAE of 4.5 years for age prediction on the test dataset
- Gained extensive expertise in computer vision, deep learning methodologies, and model deployment, with a specialized focus on age and gender classification tasks

Bike-Sharing Demand Prediction App Project – Jul 2024

- Built a predictive app using R Shiny to estimate bike-sharing demand based on weather data
- Integrated the OpenWeather API to fetch real-time weather data, performing data wrangling and analysis in R
- Applied regression models to predict bike demand and visualized the results with interactive plots and maps
- Deployed the app on shinyapps.io, showcasing skills in web-based data applications and interactive visualizations
- Documented the project and created a GitHub repository for version control and sharing

SpaceX Falcon 9 Landing Outcome Analysis – Jul 2024

- Collected and processed data on rocket launches using the SpaceX API, performing data wrangling and formatting for clean datasets.
- Implemented helper functions to extract information about boosters, launch sites, payloads, and cores using multiple API endpoints.
- Conducted exploratory data analysis to identify factors influencing the success of first stage landings, visualizing findings with matplotlib and seaborn.
- Developed and compared machine learning models, including support vector machines, decision tree classifiers, and k-nearest neighbors, to predict landing outcomes.
- Evaluated model performance and prepared a report summarizing findings for stakeholders.
- Completed the project on Google Colab, gaining experience in cloud-based Python development and API integration.

Health Inequities Dashboard, **Jul 2024**

- Developed an interactive dashboard to monitor and analyze health inequities using public health data.
- Utilized data from the CDC and WHO APIs to visualize trends, correlations, and combined analyses of mortality rates and life expectancy.
- Key features:
  - Trend Analysis: Visualize mortality rates over the years.
  - Correlation Analysis: Explore age-adjusted mortality rates and deaths.
  - Heatmap Analysis: Show age-adjusted mortality rates by state and year.
  - Life Expectancy Analysis: Visualize average life expectancy by country.
  - Combined Data Analysis: Integrate CDC and WHO data to analyze the relationship between deaths and life expectancy.
- Insights:
  - Revealed significant disparities in health outcomes across different regions.
  - Identified geographic disparities in mortality rates and life expectancy.
  - Provided comprehensive understanding and need for targeted public health strategies.

Healthcare Variation Analysis Using HCAHPS, June 2024

- **Objective**: Analyzed HCAHPS Hospital Ratings Survey data to identify healthcare Variation patterns.
- **Data Source**: Medicare.gov (2018-2020).
- **Key Tasks**:
  - Cleaned and preprocessed data.
  - Conducted exploratory data analysis (EDA).
  - Created visualizations to highlight survey response rates, patient ratings, and state-specific insights.
- **Insights**:
  - Found that most hospitals have a 30% survey response rate and are rated 3-4 stars.
  - Identified geographic disparities in patient ratings.
  - Highlighted variations in hospital ratings within Minnesota.

Work-Life Balance and Health Outcomes Analysis, May 2024

- **Objective:** Analyzed data from IPUMS NHIS to explore the relationship between work-life balance and health outcomes.
- **Data Source:** IPUMS NHIS website , samples: (2016-2019).
- **Key Tasks:**
- Cleaned and preprocessed data.
- Conducted exploratory data analysis (EDA).
- Applied regression and classification models.
- Created visualizations to illustrate health status, work hours, and depression frequency.
- **Insights:**
- Longer working hours slightly correlated with poorer health.
- Depression frequency varied significantly across employment statuses.
- Non-employed individuals generally reported better health.

PROFESSIONAL SKILLS DEVELOPMENT

Google Data Analytics Course, 2022 – 2024

- Completed an 8-course series aimed at comprehending the practices and methodologies utilized by junior or associate data analysts in their daily responsibilities, culminating in a comprehensive case study as part of the final course
- Acquired essential analytical skills encompassing data cleaning, analysis, and visualization, alongside mastering tools such as spreadsheets, SQL, R programming, and Tableau
- Developed proficiency in visualizing and presenting data findings through dashboards, presentations, and commonly used visualization platforms

Supervised Machine Learning Course, Coursera, 2022

- Attained a comprehensive understanding of supervised machine learning techniques, specializing in regression analysis and continuous outcome prediction
- Familiarized with prominent machine-learning libraries and tools utilized for regression and classification tasks, enhancing proficiency in practical application
- Developed skills encompassing supervised machine learning, regression analysis, practical machine learning application, utilization of machine learning libraries, and problem-solving in real-world scenarios

Oracle Database Foundations, Oracle, 2022

- Developed proficiency in utilizing the Oracle Database Management System (DBMS) and SQL (Structured Query Language) for effective database management and querying
- Acquired knowledge of database design principles, normalization techniques, and the implementation of data integrity and security measures within an Oracle database environment
- Gained practical experience in creating, modifying, and managing various database objects, enhancing hands-on skills in database administration

AWS Academy Cloud Foundations, 2022

- Attained comprehensive knowledge of cloud computing principles, covering key components, benefits, and deployment models
- Developed proficiency in troubleshooting common issues and optimizing AWS resources to achieve cost- effectiveness and enhance performance

Visualizing Filters of a CNN using TensorFlow, 2021

- Successfully finished a six-module course supplemented with two knowledge check tests and a capstone project
- The capstone project entailed downloading a provided model, collecting output from every convolutional layer, training the loop, and subsequently visualizing the outcomes
- Acquired skills in deep learning and the implementation of artificial neural networks

IBM Applied Data Science Capstone

Coursera

- Demonstrated proficiency in data science and machine learning techniques using a real-world dataset.
- Applied skills in data collection, data wrangling, exploratory data analysis, data visualization, model development, and model evaluation.
- Developed machine learning models including support vector machines, decision tree classifiers, and k-nearest neighbors using Python.
- Evaluated the results of machine learning models for predictive analysis, compared their strengths and weaknesses, and identified the optimal model.
- Gained skills in GitHub, Jupyter Notebook, K-Means Clustering, and Data Science Methodology.

IBM Data Science with R - Capstone Project

Coursera

- Wrote a web scraping program to extract data from an HTML file using HTTP requests and convert the data to a data frame.
- Prepared data for modeling by handling missing values, formatting and normalizing data, binning, and turning categorical values into numeric values.
- Interpreted data with exploratory data analysis techniques by calculating descriptive statistics, graphing data, and generating correlation statistics.
- Built a Shiny app containing a Leaflet map and an interactive dashboard, and created a presentation on the project to share with peers.
- Gained skills in Data Science, Linear Regression, Data Visualization, R Programming, and Exploratory Data Analysis.