+91 8848432038
ikya bhavan, nilamel,kollam
datas293@gmail.com

Ashik N

Data Science Engineer

Tableau Public: ashik.nazar/vizzes
github.com/ashiknazar/
linkedin.com/in/ashik-n-22b36b152/

"I am passionate about building a career in data science, with a strong interest in research, algorithm development, and Natural Language Processing (NLP). I am particularly keen on learning how to generate text in more desired and contextually relevant ways, while also valuing the importance of production-level systems and their real-world applications."

Education

B.Tech ,Computer science and Engineering 2013 — 2017
College Of Engineering and management Punnapra, Alappuzha

Skills

Programming Languages	Python,R,Sql, PHP,Javascript,css,java
Ai/ml and data science Libraires	Scikit-learn, TensorFlow, PyTorch, Keras,Open CV, Pandas, NumPy
Algorithms	Regression, Classification, Clustering, Time Series Analysis, NLP (Natural Language Processing), Computer Vision
Techniques	Supervised Learning, Unsupervised Learning, Deep Learning, Data Preprocessing, Model Evaluation, Feature Engineering
Data Visualization	Matplotlib, Seaborn, Plotly
Web Development Frameworks	Django,Flask,CodeIgniter, Laravel
Databases	MySQL, PostgreSQL, MongoDB
Tools-Technologies	Version Control: Git Data Engineering:Big Data (Basic Knowledge) Cloud Computing: Basic knowledge of Cloud (AWS/GCP)
Other Skills	Business Intelligence: Power BI, Tableau Project Management: JIRA

work Experience

Data Science Engineer **Feb 2024 — Present**
Ethqan Technologies pvt ltd *sreekaryam, Trivandrum*

- Completed academic projects for clients, delivering tailored data science solutions.
- Conducted webinars and classes to educate students on AI and data science.
- Trained and mentored data science interns on various projects and tools.
- Developed Django web applications, integrating data science models for clients.
- Held client meetings to provide business analytics and insights through visualizations.

Data Science Intern **Apr 2023 — Dec 2023**
Oracuz infotech pvt ltd *muttada, Trivandrum*

- Maintained effective communication with seniors to develop data science projects.
- Worked on deep learning projects, applying advanced algorithms to solve real-world problems.
- Contributed to business data analysis projects, providing insights through data visualization and reporting.
- Gained exposure to cloud implementation, learning how to deploy data science models in the cloud.

Web Developer **Jun 2018 — Jan 2020**
Maltasoft Pvtltd *Palayam, Trivandrum*

- Developed web applications using CodeIgniter and Laravel frameworks.
- Managed the backend of an online education platform using core PHP.
- Created multiple ERP dashboards to streamline business operations.

PHP intern **Jul 2017 — Jan 2018**
Bell Technolabs *Eranakulam, kerala*

- Developed web applications using core PHP and CodeIgniter frameworks.
- Worked with JavaScript and CSS to enhance front-end functionality.
- Gained experience with SQL for database management and optimization.

+91 8848432038
ikya bhavan, nilamel,kollam
datas293@gmail.com

Ashik N

Data Science Engineer

Tableau Public: ashik.nazar/vizzes
github.com/ashiknazar/
linkedin.com/in/ashik-n-22b36b152/

Projects

Text Summarization with Pegasus

https://github.com/ashiknazar/textSummarizer

- Developed a text summarization system leveraging the Pegasus model fine-tuned on the SAMSum dataset.
- Preprocessed dialogue data with tokenization using Hugging Face's tokenizer and transformed inputs for model training.
- Fine-tuned the Pegasus model using the Transformers library, applying advanced training arguments like gradient accumulation and warmup steps.
- Implemented a ROUGE metric evaluation pipeline to assess the quality of the generated summaries.
- Saved the fine-tuned model and tokenizer for deployment, enabling real-time summarization through a simple pipeline interface.
- Demonstrated the system by summarizing conversations into concise outputs and compared performance against human-written summaries.

Students Performance analysis

https://github.com/ashiknazar/student_performance

- Analyzed a dataset to identify key factors affecting student performance.
- Performed data preprocessing tasks, including handling missing values, feature scaling, and encoding categorical variables.
- Explored the data using visualizations (e.g., histograms, scatter plots) to identify patterns and correlations in student performance.
- Applied machine learning algorithms (e.g., linear regression, decision trees) to predict student performance based on various features.
- Tuned model parameters using techniques like Grid Search and Cross-validation to optimize performance and prevent overfitting.

Diabetes Prediction with Machine Learning

https://github.com/ashiknazar/diabetesPredictor

- Developed a diabetes prediction system using machine learning models trained on the Pima Diabetes dataset.
- Preprocessed the dataset by handling missing values, normalizing numerical features, and encoding categorical variables.
- Built and fine-tuned various machine learning models, including Logistic Regression, Random Forest, and Gradient Boosting, to identify individuals at risk for diabetes.
- Optimized model hyperparameters using grid search and cross-validation to achieve better performance.
- Evaluated model performance using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC.
- Saved the best-performing model for deployment, enabling real-time diabetes risk prediction via a streamlined interface.
- Demonstrated the system by predicting diabetes outcomes for new data inputs and validated results against clinical benchmarks.

Breast cancer clustering analysys

https://github.com/ashiknazar/breast_cancer_cluster

- Analyzed the breast cancer dataset to identify key features influencing clustering outcomes.
- Performed Exploratory Data Analysis (EDA) to gain insights into feature distributions and relationships.
- Conducted feature selection and analysis, identifying the most significant variables for clustering.
- Applied clustering algorithms (e.g., K-Means) to group similar data points based on the analyzed features.
- Validated clustering results by evaluating the quality of the clusters and their relevance to breast cancer classification.
- Visualized clustering results using various plotting techniques (e.g., pair plots, heatmaps) to present insights effectively.

Insurance Data Analysis

https://github.com/ashiknazar/insurance_data_analysis

- Performed data analysis on insurance datasets to identify key features influencing insurance renewal rates.
- Conducted exploratory data analysis (EDA) using R and Jupyter Notebook, visualizing trends and correlations in the data.
- Used various statistical plots (e.g., histograms, scatter plots) to analyze the relationship between customer features and renewal decisions.