UMESH. U | +44 7918572183|umeshuddar1997@gmail.com

|https://github.com/umesh660 | https://medium.com/@umeshuddar1997

Experience

Repyute Networks Pvt. Ltd. | Data Scientist | Feb'22 – Sep'23

- Developed Python code to automate web scraping utilizing APIs and Selenium, enabling an online loan service provider to extract and deliver valuable data. Reduced manual effort by 80% and accelerated data delivery by 50%.
- Wrote and implemented efficient functions utilizing Python to parse URLs, tokens, and pay slips, resulting in a 40% increase in data processing efficiency.
- Designed and meticulously trained a machine learning model with an impressive 93.8% accuracy, enabling precise prediction of loan repayment likelihood. This model significantly enhanced risk assessment and informed data-driven decision-making processes.
- Leveraged advanced computer vision techniques to automate the seamless extraction of data from uploaded documents, achieving an exceptional accuracy rate of 98.2%. This innovation remarkably streamlined document processing, enhancing efficiency and precision.
- Managed project tasks and issues effectively using JIRA, ensuring timely and organized project execution.
- Maintained version control and facilitated over 16 code deployment using Git, enhancing codebase stability.
- Revamped and standardized the existing codebases, resulting in a 20% improvement in code efficiency and readability, making code maintenance more seamless.
- Collaborated closely with team leads and managers to optimize project workflow, ensuring a 15% increase in project execution efficiency and timely delivery.

Etutor World Corporation| Business Analyst | Sep'21 – Jan'22

- Led the weekly operational task of data validation, screening, and reporting, consistently handling large datasets in alignment with stringent business rules, ensuring data accuracy and compliance.
- Allocated daily target and objective to sales representatives over ZOHO CRM platform.
- Designed and executed the SQL Queries to validate the Database.
- Generated territory alignment and assigned the representative accordingly that gained 18% more reach and 15% cost-cutting.
- Planned product backlog of LMS Software, wrote feature specifications and user stories.
- Created management reports, MIS report, marketing reports and other Ad-hoc internal/ external report.
- Conducted A/B Testing of the product, leading to a significant increase in the click rate by 25%.

Advith Consulting LLP| (Data Management) | Dec'20 – Aug'21

- Managed and maintained a database, created reports, and prepared a data backup strategy, resulting in a 98% data availability rate
- Analyzed business requirements from user input and existing systems, leading to a 20% improvement in system efficiency and accuracy.
- Prepared testing scenarios, test cases, and test environments, contributing to a 15% reduction in software bugs and enhancing overall system reliability.
- Test case planning , creation and execution for User Acceptance, Integration and performance
- Expertly interpreted complex data, transforming it into insightful analytical reports that provided actionable strategies, leading to a 25% increase in business performance and growth.
- Acted as an interface between business units, technology teams, and support teams, facilitating a 30% improvement in communication and collaboration efficiency.

Skills

- **Data Science skills : -** Machine learning **,** Data Analysis, Visualization of Data Insights, Pattern &Trend identification, Time series data, A/B Testing
- **Product Development:** - Feature prioritization, Quality Assurance and testing, Product Launch.

1. R Studio(ggplot2, dplyr, tidyr, randomForest, glm, Shiny)
2. Visualization tools: - PowerBI, MS Visio, Excel, VBA
3. Google Analytics, Statistics, Probability, MySQL, Metabase
4. Python (Numpy, Pandas, Matplotlib, Seaborn, Keras, Tensorflow, ComputerVision).
5. Langchain, llm, ChromaDB

Projects

- **Image Classification Model Development**

  Developed an image classification model using Python in a Google Colab environment leveraging TensorFlow and Keras libraries. The project aimed to classify images into two categories: "Happy" and "Sad".

  Key Contributions:

  Data Preprocessing: Implemented data preprocessing steps including image extension validation and removal of unsupported image files. Ensured data integrity and compatibility for model training.
  Model Development: Constructed a convolutional neural network (CNN) architecture using Keras Sequential API. Designed the model with convolutional layers, max-pooling layers, and dense layers to extract features and perform classification tasks effectively.
  Model Training: Utilized TensorFlow's image dataset API to split the data into training, validation, and test sets. Trained the model using the Adam optimizer and binary cross-entropy loss function. Monitored model performance on validation data to prevent overfitting.
  Model Evaluation: Evaluated model performance metrics, including precision, recall, and binary accuracy, achieving a commendable accuracy of 93.5% on the test set.
  Prediction Demonstration: Provided an example of model usage for making predictions on new images. Demonstrated the application of the trained model to classify images as "Happy" or "Sad" based on learned features.
  Outcome: Successfully developed an image classification model achieving high accuracy in distinguishing between "Happy" and "Sad" images. Contributed to advancing the understanding and application of deep learning techniques for image analysis tasks.

- **Web Content Analysis: Extracting Insights from Website Textual Data**

  Utilized advanced NLP techniques with LangChain, particularly leveraging the LangChain OpenAI module, to address diverse text processing challenges. Key highlights of the project include:

  Module Integration: Integrated essential modules from LangChain and other libraries to streamline text processing workflows.
  API Key Configuration: Configured environment variables for secure access to OpenAI's powerful API, ensuring seamless integration into the project environment.
  Library Management: Managed library dependencies effectively by installing and updating the LangChain OpenAI library to ensure compatibility and access to the latest features.
  Module Utilization: Leveraged various modules from LangChain OpenAI to perform tasks such as document loading, text chunking, embeddings, and vectorization, demonstrating proficiency in utilizing advanced NLP tools.
  Model Initialization: Initialized OpenAI's language model instance with tailored parameters, fine-tuning settings like temperature and maximum tokens for optimal performance in text generation tasks.
  Data Acquisition: Employed an asynchronous HTML loader to efficiently fetch and load data from specified URLs, enabling seamless integration of external text sources into the project pipeline.
  Text Processing: Employed techniques such as text chunking to break down large HTML documents into manageable segments, facilitating efficient processing and analysis of textual data.
  Vectorization and Similarity Search: Utilized OpenAI embeddings and FAISS for vectorization of document chunks and performed similarity search operations, enabling efficient retrieval of documents related to specified query strings.
  Result Visualization: Implemented functionality to print documents similar to the provided query, providing actionable insights into the textual data corpus.
  This project exemplifies a comprehensive approach to NLP, showcasing proficiency in utilizing cutting-edge tools and methodologies to address complex text processing challenges effectively.

Education

- Master of Science in Data and Decision Analytics | University of Southampton | Sep’23 - Present

- Bachelor of Engineering in Mechanical Engineering | BMS College of Engineering, Karnataka | 8.2 CGPA | 2015-2019