+91-9110771246

Vallabhapurapu Karuna

vallabhapurapukaruna@gmail.com

Linkedin Hyderabad, Telangana

EXPERIENCE: 2 .9Years

ZABDA Technologies Pvt Ltd, HYD — D*ata Scientist*

December 2021 – February 2024

- Applied clustering analysis to unstructured data (extraction, transportation, integration) by writing a machine learning algorithm
- Implemented an NLP model to classify text data into different topics
- Developed and fine-tuned machine learning and deep learning models to achieve optimal accuracy in structured and unstructured data.
- Demonstrated proficiency in supervised and unsupervised algorithms, applying them to diverse problem domains.
- Conduct exploratory data analysis (EDA) to identify trends, patterns, and anomalies.
- Predictive Modelling & Machine Learning:

  Build, train, and deploy machine learning models to solve business problems (e.g., classification, regression, clustering).

  Fine-tune models, evaluate their performance, and improve prediction accuracy.

  - Data Visualization:

  Create clear and compelling visualizations to communicate data insights to stakeholders.

  Design and implement dashboards to monitor business performance in realtime.
- Use statistical techniques to analyse data, interpret results, and draw actionable conclusions.
- Optimize machine learning algorithms for scalability and efficiency in production environments.
- RAG & Graph RAG for creating chatbots and Assistant Creations
- Managed the complete machine learning life cycle, from data collection and preprocessing to model deployment and monitoring.
- Successfully addressed regression and classification challenges to meet specific business needs.
- Effectively handled unbalanced datasets, selecting and implementing appropriate models.
- Utilized a deep understanding of Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN) to tackle complex tasks.
- Managed and processed data in various formats, including Excel, CSV, and TSV files.
- Implemented RNN-LSTM models for tasks such as nextword prediction and text summarization in Generative AI.
- Measured document similarity using Cosine similarity, Jaccard similarity for various applications.
- Conducted time series analysis for temperature

and milk production forecasting using LSTM models.

- Basic understanding of Docker for ML models into deployment.
- Creating ML model deployment on FLASK.
- Extract entities from documents (CV&JD) using RAG-LLM application ● Create chatbots using pretrained LLMs of PHI-3/GPT Models/LLAMA3 models.
- Created trends of dashboards using Tableau.

SKILLS

Languages code: Microsoft

SQL Server, Python (Numpy , Pandas, etc.)

Technologies: Data Visualization tools (PowerBI , Tableau(Basic), Amazon Quick sight), MSSQL, Docker(Basic).

Concepts: Business Intelligence (BI), Statistical Analysis Machine Learning Algorithms,

AWS Administration, Deep Learning , Supervised and Unsupervised Algorithms ,ML Life Cycle Management,

Regression and Classification ,Open-cv image processing, LSTM Time series forecasting,

Hobbies

YouTube, Travel, Games.

AWARDS

- **HackRack -> SQL (Basic)**
- **AWS Cloud Practitioner** frameworks using python.
- Demonstrated proficiency in Python libraries such as NumPy, pandas, Kera's, Scikit-Learn, SciPy, OpenCV, Matplotlib, and Seaborn.

Insignia Consultancy Solutions, Remote — *Data Scientist*

March 2024 – September 2024

- CV parsing using NLP
- JD parsing using NLP
- Creating custom chatbots using LLAMA-3,PHI-3 LLM models using Google collab
- Basics of OLAMA, LM-studio for running LLM models into local systems ● Learn about BERT and Transformer models.
- Basics of LLM models
- Career assistant Chatbots, AI Interviewer using RAG.

PROJECTS

1.NACH Mandidate key-value extraction using Textract

- Automating the process of extracting Handwritten fields from Nach Mandates.

  Below mentioned fields are extracted using AWS Textract:

- Key Roles & Responsibilities:
- Collected NACH mandate files and stored in AWS S3.
- Triggered S3 into AWS Lambda for writing the Python code.
- Called the AWS Textract in Lambda function for key-value extraction.
- Lambda Triggered each NACH file in the S3 looped.
- Textract extracted all KV pairs with confidence score
- Filter outed the required KV fields and stored it into AWS Dynamo DB

2.Customer Personality analysis using clustering algorithm Python, pandas

- Project uses clustering algorithms to segment and comprehend several client groups by analysing the personalities of the customers. Businesses can better serve customers and allocate resources by customising their marketing and services based on behavioural trends. Insights from the project enable more tailored and efficient consumer engagement. Server – Flask.

3.Quivvr Project (AI-Based Recruitment Platform)

- This is like a recruitment platform; it deals with communication between employees and company portals for hiring employees.

Responsibility's:

>> I have done extraction, substitution, parsing CVs and JD (job portals)

>>Creating pipelines for CV parsing & JD parsing using standard libraries (NLP) >>Finding accuracies between actual dataset (what contains in cv) and predicted dataset (cv parsed output) using different similarity methods)

LANGUAGES English, Telugu

EDUCATION

NVR College Of Engineering and Technology Affiliated to JNTUK : Tenali, AP — *Bachelors*
2018-2021 - Electronics and Communication Engineering GPA:8.3

Govt. Polytechnic College For Women Affiliated to SBTET AP :Guntur, AP — *Diploma*
2015-2018 - Electronics and Communication Engineering -

GPA:9.1

SSC-10TH ZPH SCHOOL-Kollipara, AndraPradesh-Guntur(dist.)
2015-GPA:9.2