RAHUL **RAWAT**

Assistant Manager- AI & Analytics

+91-7404422001

Rahulrawat178@gmail.com

Noida, Uttar pradesh, India

Dynamic Assistant Manager with 7+ experience in Big Data Engineering, Analytics, and Knowledge Graphs. Demonstrated expertise in Machine Learning, Deep Learning, and Generative AI, with a focus on applications in Computer Vision and Natural Language Processing. Adept at leading teams to develop AI-powered applications, leveraging large language models (LLMs) for innovative solutions and real-world problem-solving. Passionate about transforming data into actionable insights and driving technological advancements in AI, contributing to business growth and efficiency.

EXPERIENCE

Assistant Manager / Jan 2022 - Present
BusinessNext / Noida

- Design and implement end-to-end AI solutions, integrating machine learning, deep learning, and generative AI technologies to solve complex business problems.
- Collaborate with stakeholders to understand business needs and translate them into technical requirements and AI-driven solutions.
- Architect and manage cloud-based infrastructures (AWS, GCP, Azure) to support data processing and AI model deployment.
- Design and implement machine learning and deep learning models, ensuring they are scalable and optimized for performance.
- Develop automated workflows for continuous integration and deployment (CI/CD) of AI models.

Data Scientist / April 2021 - Jan 2022
NCR Corporation / Gurugram

- Identify valuable data sources and automate collection processes
- Undertake preprocessing of structured and unstructured data
- Analyze large amounts of information to discover trends and patterns.
- Present information using data visualization techniques.
- Propose solutions and strategies to business challenges
- Design data modeling processes to develop predictive models and perform custom analysis.

Machine Learning Engineer / Oct 2018 - April 2021
Acidaes Solutions pvt ltd / Noida

- To study and transform data science prototypes. Research and develop Machine Learning applications according to requirements.
- Creating and maintaining Big Data pipelines to assemble large and complex datasets.
- Handling structured and unstructured data to analyze large amount of information to discover trends and patterns.
- Extract data from a variety of relational and non-relational databases, manipulate, explore data using quantitative, statstical and visualization techniques.

EDUCATION

M.TECH Software Engineering
2014 - 2016 | UIET, KUK
8.86 CGPA

B.TECH Computer Engineering
2009 - 2013 | KITM, Kurukshetra
67.47%

12TH
2008 - 2009 | CBSE
60%

10TH
2006 - 2007 | CBSE
72.8%

SKILLS

Langauges: Python, Java
Databases: SQL, MinIO , Mongo db, Cassandra, Clickhouse, Neo4j, Apache Hadoop

Frameworks and Toolkits: Pyspark, Apache Beam, Apache Airflow, Scikit-learn, Tensorflow, Dask, LLAMA, SpaCy , PyTorch , NLTK , Transformers, OpenCV, Kafka, Docker, AWS, GCP, Kubernetes, Elastic stack(ELK) , Tesseract, Keras, Dialogflow, Rasa NLU, gRPC, Nginx, Docker, CrewAi, Autogen, LangChain, PEFT, LoRA

PUBLICATIONS

Published and presented a paper in 5 th International Conference on Reliability , Infocom Technologies and Optimization (ICRITO'2016) (Trends and Future Directions), IEEE Conference Amity University, Noida.

Published a paper in STM journals titled as "Assessment of various chunking techniques for de-duplication in Big Data".The Business Intelligence Analyst Course 2020 via Udemy

Published a paper in IJESRT Journal titled as "Understanding and Configuring Hadoop: To Handle the large amount of Data"

Data Analytics Consultant / Feb 2017 - Oct 2018
Constalytics / Mohali

- Research and development in the field of Machine Learning, Big Data, Natural Language Processing and Knowledge Graphs.
- Creating Big Data pipelines for preprocessing and to generate insights and turn them into actionable recommendations.
- Development of data science prototypes and transform them into different business usecases.
- Deployment of microservices using Docker in GCP and maintaining databases like Mongodb, Hadoop, SQL and Neo4j
- Developing microservices based architecture and deploying using Docker, Kuberenetes and AWS.

PROJECTS

Prompt Builder

The Gen AI-based platform enhances AI model performance and user experience by designing, developing, and optimizing prompts.
Role:

- Develop and optimize prompts to enhance Large Language model performances
- Implement prompt engineering pipeline to create high-quality, accurate, and engaging AI-generated content.
- Deploy large language models into the platform, ensuring seamless compatibility and efficiency.
- Implement techniques to reduce hallucination in AI-generated content, ensuring accuracy and reliability.

Account Receivable Prediction

Accounts Receivable Understand the factors of successful collection efforts. You can Predict which customers will pay fastest and recover more money and improve collections efficiency.
Role:

- Perform exploratory data analysis to find trends and insights from the datasets.
- Create big data preprocessing pipelines using pyspark.
- Design machine learning processes to develop predictive models and perform survival analysis to find when invoices will be paid.
- Create microservices to deploy them using docker.

DataNext Platform

DataNext is the collaborative data science platform for data scientists, data analysts and engineers to explore, prototype, build and deliver different data flows more efficiently. This platform involves Data Integration, Data preparation, Data Exploration, Machine Learning, Model Deployment and Automation of various campaigns
Role:

- Create machine learning pipelines for regression and classification tasks.
- Create big data preprocessing pipelines.
- Create microservices to deploy them using docker and kubernetes.
- Create database connectors and maintain databases like cassandra, Mongodb, MinIO, trino s3, click house etc.

DISSERTATION

Title: Bucket based data de-duplication technique for big data storage systems

Description: This technique aims to optimize the storage space for the big data storage systems. The main objective was to find the better de-duplication technique to make storage efficient and to remove duplicate data from storage systems.

HOBBIES

Zeal to Learn new things.
Playing Cricket
Travelling

LANGUAGES

Hindi
English

BotNext Platform

This a natural language understanding platform that makes it easy to design and integrate a conversational user interface into Mobile Apps and Web Apps etc. Using botnext, you can provide new and engaging ways for users to interact with your product.

Role:

- Develop NLU engine to extract custom entities and intents.
- Create microservices to deploy them using docker and kubernetes.
- Integrate Retrieval-Augmented Generation (RAG) capabilities to enhance the conversational interface with real-time, relevant information retrieval.
- Optimize the NLU engine to leverage RAG for improved response accuracy and user engagement.

Information Extraction and Retrieval Platform

Unstructured text processing platform which helps to analyze the text using Named Entity extraction, sentiment analysis and Topic extraction. This platform helps to create Knowledge Graph from text by using Dbpedia to extract possible relationships between Entities.

Role:

- Create pipeline for data preprocessing and from unstructured data.
- Create classification model for sentiment analysis and topic extraction model.
- Create knowledge graph from entities and relationships extracted from dbpedia.
- Deployment using docker and google compute engine.user engagement.

Data Unification and Integration Platform

This is data unification and integration platform which connects data from silos and helps to discover relationships between data using knowledge graph. Platform helps to store data in Graph database and graph analytics capabilities allow to explore insights using Machine learning and Graph algorithms like community detection, centrality, path analytics and collaborative filtering etc.

Role:

- Create big data pipeline for connecting various data sources. .
- Create machine learning and graph analytics pipelines to find trends and patterns from the data. .
- Create microservices and manage them using docker and google compute

Declaration

I hereby declare the above details are correct to the best of my knowledge and belief.

Place: Noida Rahul Rawat