YOGENDRA SINGH

DATA SCIENTIST

9302276794 | yogendrasingh.forwork@gmail.com | www.linkedin.com/in/yogendrasingh14

- IIT Delhi Alumnus offering around 3 years of in-depth experience in executing data-driven solutions to increase efficiency & accuracy and assisted the business to undertake strong decisions
- Currently serving as a Data Scientist at Halliburton Development Center India, partnering with a wide range of Product Service Lines (PSLs) to develop and implement Machine Learning solutions that drive significant impact across the organization
- Skilled in conducting full lifecycle of data science (end-to-end) including data gathering via, data transformation, data analysis, data modelling, and production ready deployment
- Identified, developed and implemented appropriate statistical techniques, algorithms, and Deep Learning/Machine Learning Models to create new, scalable solutions that resolved business challenges
- Excellence in supporting the development of innovative business solutions to improve product performance and reduce cost by leveraging combination Data Science & Data Engineering skills
- Proficiency in leveraging the latest data and visualization tools and software to provide the business with insights on customers and make recommendations for improvement
- Talented Analyst with exceptional background in utilizing data from diverse information systems to build tools and forecasting models that remarkably improve organizational decision-making capabilities
- Team-based management style with skills in determining company's mission & strategic direction, capable of leading & motivating individuals to maximize levels of productivity

SKILLS

Technical Skills: Data Science & analytics, Deep Learning, Machine Learning (ML) & Artificial Intelligence (AI), Statistics, Statistical Analysis, Predictive Modelling, Insights, Process Optimization, Data Visualization, Time Series Analysis, Anomaly Detection, Natural Language Processing (NLP), Embeddings, Transformers, LLM, A/B Testing

Tools and Technologies: Programming: Python, SQL, Big Data: Apache Spark, Cloud: Azure, ETL: Databricks, AutoML, MLOps, Keras, Tensorflow, Scikit-learn, Plotly, Matplotlib, Pandas, Numpy, Scipy, GIT, Langchain, llama-index

Soft Skills: Communication, Team player, Collaboration, Problem-solving, Leadership

PROFESSIONAL EXPERIENCE

Data Scientist Oct, 2023 – Present
Halliburton Development Center India Bangalore, KA

- Building Forecasting Models using various Machine Learning and creating informative and visually appealing data visualizations
- Developing and optimizing complex SQL pipelines, streamlining data extraction, transformation, and loading processes, resulting in reduction in data processing time and improved data accuracy for critical business decision-making
- Maintaining clear and organized documentation of analysis methodologies, findings, and insights
- Worked on large data sets effectively in order to find patterns and provide data – driven solutions

Gen AI Powered Q&A platform:

- Innovated a Gen AI powered Questions & Answers platform catering to natural language queries.
- Incorporated Retrieval Augmented Technique (RAG)
- **Tech Stack:** NLP, Vector Embedding, Vector Databases, Large Language Model (LLM)

RTS Lithology Prediction Project:

- Successfully deployed a lithology prediction model on Halliburton's real-time data streaming platform and significantly enhanced accuracy of lithology analysis for more informed decision-making
- **Tech Stack:** Python, Data wrangling, PySpark, Databricks, Azure blob storage, REST API

Associate Data Scientist July, 2021 – Sept, 2023
Halliburton Development Center India Bangalore, KA

Borehole Holdup Prediction Project:

- Engineered a holdup prediction tool utilizing machine learning techniques, overcoming limited data availability by amalgamating simulation and lab data, and reduced error by 57%
- **Tech Stack:** Python, Regression ML Models, Data Extraction, Data Wrangling, Feature Engineering, Visualization

Fluid Analysis Project:

- Led a project employing machine learning for classification of fluids and then prediction of fluid properties and resulting in a 32% optimization over the previously deployed solution
- **Tech Stack:** Python, ML Models, Neural Network, Pickle

Transmission Failure Prediction:

- Developed a transmission failure prediction system by harnessing time series data, crafting a tailored cost metric for robust model evaluation, ensuring the reliability and precision of the predictive model
- Performed extraction, transform and load (ETL) using pyspark
- **Tech Stack:** Python, Data wrangling, Feature engineering, PySpark, Databricks, Azure blob storage, AutoML, Recurrent Neural Network (LSTM)

EDUCATION

- M. Tech. in Engineering Analysis and Design
  Indian Institute of Technology, Delhi (IIT Delhi)
  2019 – 2021
  Office of Career Services – Nucleus Coordinator

PATENTS

- Borehole Holdup Prediction Using Machine Learning and Pulsed Neutron Logging Tool Data, U.S. Patent serial number 18/194,887, filed on April 3, 2023. Status: Patent Pending
- Transmission Failure Prediction by Leveraging a Production Impact Evaluation Metric, U.S. Patent serial number 18/511,180, filed on November 16, 2023. Status: Patent Pending

CERTIFICATIONS

- Microsoft Certified: Azure Data Scientist Associate
- Generative AI with Large Language Models course certification (by DeepLearning.ai on Coursera)
- Neural Networks and Deep Learning certification course (through Coursera by DeepLearning.ai)