

# Suresh Bhojwani

## Machine Learning Lead || Principle Architect

Experienced Lead Machine Learning Engineer with over 6+ years of expertise in Data Science and Microsoft Azure Consulting. Proficient in Data Science Applications (DSA), Scalable Systems, OpenAI, Prompt Engineering, NLP, System Design, Distributed Computing, and ML/Data Pipeline Development, Collaborated with Fortune 100 companies across the United States, United Kingdom, Japan, and the Asia-Pacific region.



✉ sureshbhojwanics@gmail.com

in linkedin.com/in/suresh-bhojwani-33b335128

medium.com/@sureshbhojwani001

## EDUCATIONS

### MS IN COMPUTER SCIENCE & ENGINEERING MDSU, RAJASTHAN

2016 - 2018

INDIA

### BACHELOR IN COMPUTER SCIENCE AND ENG. MDSU, RAJASTHAN

2013 - 2016

## WORK EXPERIENCE

### Lead Machine Learning Engineer/ Lead Architect Celebal Technologies

09/2021 - Present

Remote

Celebal Technologies is a premier software services company in the field of Data Science, Big Data, and Enterprise Cloud.

#### Achievements/Tasks

- Working as Lead Machine learning engineer on various use cases collaborated with global Team developers to build and develop Optimized Machine Learning Products and Services
- Leading a team of 200 on Various Classical and NLP use cases along with Generative AI and RAG bases use cases leveraging Open AI GPT 3/3.5/4 LLM using Prompt engineering
- Various data science projects and worked Automation of tasks using Python and created end to end pipelines from scratch from algorithm design to deployment into production, Worked in E-commerce (Car Sales), Consultancies, Energy, and Supply Chain Management and healthcare domain

### Principle Consultant Celebal Technologies

04/2019 - 08/2021

Hybrid

#### Achievements/Tasks

- Worked as Solution architect for Machine learning use cases and collaborated with global Team developers and UI team to build and develop Optimized Machine learning Systems.
- Leading a small team of 50 on various data science projects and worked Automation of tasks using Python and created end to end pipelines from scratch from algorithm design to deployment into production.

### Software Consultant ML/AI Independent consultant

2016 - 2018

Remote

#### Achievements/Tasks

- Worked as machine learning engineer on classical machine learning, NLP, Deep learning projects, Deployment and integration with UI, Python Back-ends, Web scrapping, Software Support for National and International Clients as independent consultant

## SKILLS

Languages: Python, pyspark, C/C++, Java.

GenAI: LLM, Prompt Engineering, GPT 3/3.5/4, Embeddings, Langchain, LLAMA-v2, Azure cognitive services, Google cloud search, Elastic search, Azure bot framework SDK, Azure Qna maker.

Databases: SQL server, Cosmos DB, Mongo DB, Redis, Azure Blob storage, Feature Store, Delta Lake

Deployment/Cloud Tools: Azure, AWS, Databricks, Mlflow, Azure ML studio, Azure Data lake, Azure Data Factory, Nginx, Fast API, Flask API, Django, VM Linux/Windows, Azure App service/ functions, AKS, ACl, Docker, MLFlow, Kedro, Locust

## PROJECTS

### CT Miner - A smart information retrieval tool

- Created QA/Document Search system to read PDFs of tender documents and return answers to each of user's questions. Earlier stakeholders used to extract answers from 500 page PDFs, which were around/month.
- Extracting information from unstructured data files, such as PDF, Docx, emails, pictures of type Resumes, legal tenders' docs, invoices etc, designed and implemented algorithm, Search system with end-to-end Deployment
- Able to extract information from more than 100 Gb of document and Handle 5k+ concurrent users
- Tech Used :** Python, Azure Machine learning Studio, Prompt Engineering, LLM, Openai GPT 4, Azure bot Framework, Facebook FAISS Indexer, Elastic Search Indexer, ANN Search, Azure Kubernetes, ADF, Azure Linux VM, Azure SQL server, Azure cosmos DB, Azure Blob, Azure Cognitive Services, Fast API, Nginx server

### Call Center Data Analysis

- Worked on call center voice and email data classification for important business KPI's which benefits CRM top management to analyze and take informed decisions based on conversations and textual data.
- Tech used:** Azure ML studio, Azure speech to text, Non relational databases, Django, Kubernetes

### Constraints Based Optimization

- Created a Product where all the supply chain and working cost need to be managed by an intelligent automated system and reduced the overall cost by 20 percent
- Designed and deployed the algorithm for procurement, chartering from seaports, distribution in multiple plants of raw material, Automated Dispatch scheduler for scheduling companies raw material trains as automated system

### UGC moderation :

- Worked on user generated content moderation model to filter data like : comments, reviews etc. on the web portal for A car ecommerce company

## CERTIFICATES

Databricks Machine Learning professional, Lake House Fundamentals, Advance Software Eng.(DSA, System Design)

# Additional Projects

## Call Center Data Classification - CRM Decision Making

The project focused on call center data classification to support CRM decision-making. By analyzing conversations and textual data, the goal was to provide insights into customer interactions and key business KPIs. The system automated the classification process for **voice recordings and email conversations**, facilitating efficient data analysis.

### Impact and Benefits:

The project had a significant impact on CRM decision making by providing automated data classification and analysis. CRM decision makers gained valuable insights into customer interactions, sentiment analysis, and essential KPIs. This empowered them to make data-driven decisions, enhance customer satisfaction, and optimize operations

## Lead Scoring System for large E-commerce Giant

The project focused on implementing a lead scoring system to identify potential buyers from web-generated leads. Leveraging user's online viewing patterns and LMS data, the objective was to enhance the conversion rate of leads sent to dealers from 3% to 9.8%.

### Impact and Benefits:

The implementation of the lead scoring system had a **substantial impact on lead conversion rates**. By utilizing advanced data analysis and machine learning techniques, the project enabled more efficient identification of promising leads.

## OpenAI-Based KPI Bot for Excel Data Extraction

The project centered around the development of an OpenAI-powered KPI bot that utilizes GPT-3.5 for extracting information from Excel files. In the backend, prompt engineering techniques were employed to convert **user's natural queries into SQL-based queries** for efficient data retrieval.

### Impact and Benefits:

The implementation of the OpenAI-based KPI bot had a transformative impact on data extraction and analysis from Excel files. By power of GPT-3.5 and NLP, users could interact with their data using plain language queries.

## SAS to PySpark - Databricks Migration

The SAS to PySpark Code Converter project aimed to simplify the migration from SAS to PySpark on the Databricks platform. It automated the conversion of SAS code into PySpark, ensuring a seamless transition for data analytics tasks.

**Approach:** The project involved analyzing existing SAS code, developing custom code conversion algorithms, rigorous testing, and seamless Databricks integration, the converter delivered significant benefits, including time and cost savings by **eliminating manual code rewriting**.

## Constraints-Based Optimization for Supply Chain of Largest Alumina Importer

The project's primary objective was to optimize the supply chain operations for the largest alumina importer while developing a **smart scheduler for efficient import management** from various international ports into Indian ports. The implementation of constraints-based optimization techniques played a pivotal role in achieving these objectives.

### Impact and Benefits:

Optimizing cost-efficiency, operational performance, and overall competitiveness, it marked a substantial step forward in the alumina import market

## MLOps Platform Implementation on Databricks with MLflow Integration

This project focused on establishing an end-to-end MLOps platform on the Databricks platform, incorporating MLflow for model management and encompassing all stages of the MLOps lifecycle.

The goal was to automate model development, orchestrate pipelines, manage workflows, and seamlessly transition artifacts **from on-premises environments to Databricks** using distributed computing capabilities.

### Impact and Benefits:

The MLOps platform on Databricks with MLflow integration delivered key benefits, including streamlined model development via distributed computing, improved collaboration with MLflow, extensive automation throughout the MLOps lifecycle, scalability for growing data and models, and enhanced model performance through hyperparameter tuning. The seamless transition of artifacts ensure data consistency and operational efficiency.