

# RAVI PRATAP

## Data Scientist

### Phone

+91 9973225589

### Email

ravipratap366@gmail.com

### Personal Portfolio

<https://ravipratap366.github.io/Portfolio/>

### Github

<https://github.com/ravipratap366>

### LinkedIn

<https://www.linkedin.com/in/ravi-pratap-7a8052165/>

📍 India

## SUMMARY

As a proficient Data Scientist with a computer science background, I bring 2.6 years of experience in solving complex business problems using Data Analytics, Machine Learning, Natural Language Processing, Computer Vision, and Deep Learning technologies. My training in Big Data Analytics from CDAC has equipped me with a solid foundation in analyzing and interpreting large datasets. With a passion for delivering actionable insights, I am dedicated to leveraging data-driven solutions to drive business growth and enhance customer experiences

## PROFESSIONAL EXPERIENCE:

**Data Scientist** 📅 **August 2022 - Present**

**Revoquant AI** 📍 **Gurugram, India On-site**

### ○ INFRARED(Product)

Developed an innovative solution leveraging Hugging Face's libraries, OpenAI's Language Model (LLM) and unsupervised machine learning techniques to enhance audit, fraud detection and deployed the solution using AWS EC2-ECR instance with CI/CD pipelines using GitHub actions.

#### Interactive Chatbot Interface: InfraBot

Developed an interactive chatbot interface using models text-davinci-003, LLaMa2 to provide users with a user-friendly way to interact with the system. The chatbot can answer queries related to audit processes, fraud detection methodologies, and data analysis techniques, enhancing user experience and accessibility.

#### PDF and Excel Data Querying using LLM

Integrated OpenAI's Language Model (GPT-3) and Hugging Face's libraries like (google/flan-t5-xxl) to create a powerful PDF and Excel querying system. This feature enables the system to extract and process information from unstructured PDF documents and Excel spreadsheets, facilitating efficient data analysis and validation.

#### Video-to-Text Transcription with Whisper

Utilized OpenAI's Whisper to convert audio content from videos into accurate text transcripts. This transcription process enhances the comprehensiveness of data available for analysis and audit purposes, improving overall accuracy.

### ○ Anomaly Detection using Unsupervised Machine Learning

Implemented unsupervised machine learning algorithms to identify anomalies in financial and transactional data. Employed techniques like Isolation Forest, DBSCAN, to automatically detect unusual patterns and potentially fraudulent activities.

### ○ 4ToSS(Product)

Leveraged advanced analytics and machine learning techniques to improve operational efficiency and drive better business outcomes.

Developed a process mining tool using Exploratory Data Analysis and Machine Learning algorithms to optimize the process from purchase requisition to invoice verification and identify the happy path of the product life cycle.

Conducted date and capacity planning using data analysis techniques and machine learning algorithms to optimize the timeline and capacity of the machines involved in a machine project.

Created Power BI plots of sales to aid in production planning and inventory control, and made informed decisions about production planning and inventory control to reduce waste and optimize inventory levels.

## Education

### ● PG-Diploma in Big Data Analytics

Sunbeam Infotech Private Limited (CDAC)

(2021)

Pune

CDAC provided expertise in Big Data technologies, Analytics, Machine Learning, programming languages, and cloud computing, and enhanced collaboration skills.

### ● Btech in Computer Science Engineering

Greater Noida Institute of Technology (AKTU)

2019

Delhi

## TECHNICAL STACK

### ● Data Science | Machine Learning | Deep Learning | Natural Language Processing | Computer Vision | Time Series Analysis

Python, OpenAI, Hugging Face, Data Visualization, Supervised Learning and Unsupervised Learning, Generative AI Predictive Modeling, Bagging and boosting, Linear and Logistic Regression, CNN, RNN EDA, Feature Engineering, Feature Selection & extraction, Hyperparameter optimization. OpenCV, Object Detection, Object Classification, Object Segmentation, Probability, Statistics

### ● Frameworks and Python Packages

LangChain, NumPy, Pandas, Matplotlib, Seaborn, Scikit-Learn, Tensorflow, PyTorch, Keras, Feature engine, Category Encoders, nltk, Web Scraping, Streamlit, Flask

### ● Programming Languages

Python, C Programming, HTML, CSS, Javascript, SQL

### ● Databases

MySQL, MongoDB

### ● Big Data and Other Tools

Hadoop, Spark, Tableau, PowerBI, Git, Canva, Excel, KNIME

### ● Cloud Deployment

AWS EC2-ECR, AWS elastic beanstalk, GCP, Heruko, Github

○ Inventory Detection

Developed a smart fix using YOLOV8 to help Coca-Cola tally non-Coca-Cola products in their fridges and to check and verify the accuracy of their inventory.

PROFESSIONAL EXPERIENCE:

AI/ML Engineer April 2021 - August 2022  
XCaliber Infotech Pvt. Ltd. Pune, India On-site

○ Humming Bird (AT&T)

Developed a predictive model using Machine Learning techniques such as Random Forest and Gradient Boosting, Time Series techniques such as Arima and LSTM, and Sentiment Analysis to forecast inventory for reverse logistic and predict prices for B2B and B2C markets.

Utilized Time Series techniques such as Arima and LSTM for inventory and price forecasting, which enabled the identification of patterns, trends, and outliers in the data.

Incorporated Sentiment Analysis techniques such as BERT to analyze customer feedback and social media sentiments, and used it as a feature to adjust the price predictions, resulting in an improved pricing strategy.

Developed an Object Detection Model using Yolo and ResNet-UNet for accurate cosmetic grading of smartphones and integrated it into the refurbishing process to improve efficiency and accuracy

MACHINE LEARNING POCS

- Designed and developed a machine learning model to predict stock prices using historical data, financial news, and feature engineering techniques (ARIMA, LSTM, Random Forest) with performance evaluation (mean squared error, accuracy).
- Developed an Optical Measurement System to calculate the length of pipes captured by a drone using image processing and computer vision techniques.
- Utilized a drone camera to capture images of the pipes, and applied image processing techniques such as image enhancement, feature extraction, and object detection to analyze the captured images.

COURSE AND CERTIFICATIONS

- Machinelearning.ai - Machine Learning Specialization,( Python
- Deeplearning.ai - Deep Learning Specialization
- Deeplearning.ai - Natural Language Processing Specialization
- CS231n - Stanford University,