

# VIVEK ANIL CHOUDHARI

Data Scientist | Machine Learning Specialist | NLP & GenAI Enthusiast | Predictive Analytics Expert

[choudhari.vivek98@gmail.com](mailto:choudhari.vivek98@gmail.com) | [9890293565](tel:9890293565) | [LinkedIn](#) | [Portfolio](#) | [GitHub](#)

---

Innovative Data Scientist with 4+ years of experience in developing and deploying Machine Learning solutions across diverse domains. Skilled in leading critical projects, optimizing data-driven workflows, and implementing scalable AI models. Proven track record of improving operational efficiency by 80% and reducing data processing time by 90% through advanced analytics and automation. Passionate about leveraging AI for predictive modeling, object detection, & decision-making to drive business impact.

## PROFESSIONAL EXPERIENCE

---

Medimaze Solutions Private Limited, Pune, MH

**Data Scientist** | May 2023 – Present

- Led AI-driven solutions to improve diagnostic accuracy and operational efficiency which will in turn reduce the diagnosis time by around 80%.
- Worked with cross-functional teams to integrate AI into existing medical workflows. Implemented cost-effective, scalable solutions, saving time and resources.
- Managed AI project lifecycles, ensuring alignment with strategic goals and compliance.
- Identified and resolved workflow challenges, driving process automation improvements.
- Maintained high standards of quality and compliance in AI models.

### Projects:

#### MRI Spine Vertebrae Segmentation (Ongoing)

- Leading a project focused on segmenting all vertebrae from sagittal MRI spine scans.
- Aiming to build an efficient AI pipeline for automatic spinal segmentation to enhance diagnostic accuracy and support early detection of spine-related pathologies.
- Involved in model development, validation, and planned integration into clinical workflows for seamless adoption.

#### Tender Summarisation Using LLM (Ongoing)

- Developing a Large Language Model (LLM) solution using Retrieval-Augmented Generation (RAG), vector databases, and advanced embedding techniques.
- Built an initial Streamlit-based UI for testing and validation purposes.
- The objective is to summarise large tender documents (400+ pages) into precise, actionable 1-2 page summaries to assist stakeholders in quick decision-making and reduce manual effort.

#### AI-ML Deep Learning Models in Chest X-Ray

- Developed deep learning models for chest X-ray pathology detection, including Cardiomegaly, Pleural Effusion, Pneumothorax, Tuberculosis, and more, with validation accuracies above 93%.
- Designed custom algorithms for clinical measurements like Cardiac Size and Costophrenic Angle to support diagnosis.
- Curated large datasets and implemented data pre-processing techniques for optimizing model performance and accuracy.
- Successfully deployed the solution on a Linux server and integrated it into the client's environment, aligning with stakeholder requirements and ensuring seamless operation.

#### YOLO Pathology Detection AI in Chest X-Ray

- Created and deployed YOLO-based models for detecting chest abnormalities such as nodules and rib fractures with validation accuracies of more than 97%, significantly improving detection rates and workflow efficiency.
- Led the integration of YOLO models into the clinical diagnostic pipeline, ensuring real-time detection and precise abnormality identification.

#### Simplified Report Generative AI (Mobile Healthcare App)

- Developed an AI chat bot to simplify medical reports for patients, explaining medical terms and findings. Initially integrated with ChatGPT API, now using OLLAMA Gemma for enhanced functionality.

#### Generative AI for Machine Protocol (CT/MRI)

- Built an AI system that guides machine technicians on correct protocols for CT & MRI scans, ensuring accurate scan and reducing human error.

#### Gender-Specific Reporting AI

- Introduced an Ollama AI to flag and prevent gender-specific reporting errors, ensuring no male-specific terms appear in female reports.

#### Generative AI for Report Labels/Tags

- Automated the tagging of patient studies (Normal/Abnormal) using AI to streamline workflow and data organization.

## Cognizant Technology Solutions, Pune, MH

### **Analyst** | July 2021 – Apr 2023

- Led development of ML models to predict customer churn for British Gas (UK), delivering actionable insights that improved retention strategy decisions.
- Built end-to-end pipelines for customer feedback analysis, applying NLP techniques like sentiment analysis, topic modelling (LDA), and keyword extraction to identify service gaps across regions.
- Designed and deployed a classification model to categorize service complaints, achieving 87% accuracy and enabling faster issue resolution.
- Conducted A/B testing simulations for customer retention campaigns, analysing experimental outcomes and recommending winning strategies using statistical inference.
- Optimized feature selection and model interpretability using SHAP and LIME frameworks, enhancing trust and explainability for business stakeholders.
- Created clustering models (K-Means, DBSCAN) for customer segmentation based on behavioural and demographic data, helping marketing teams design targeted campaigns. Collaborated on a time series forecasting project, predicting service demand trends and aiding in better workforce and logistics planning.
- Documented and presented findings through interactive dashboards (Power BI, Tableau) and technical reports, ensuring stakeholder alignment and visibility.

### **Internship Trainee** | Feb 2021 – July 2021

- Developed predictive models for mobile price classification and sentiment analysis, applying feature engineering, hyperparameter tuning, and model evaluation techniques.
- Applied clustering algorithms and custom similarity metrics to group customer datasets, improving segmentation speed and insight discovery.
- Built exploratory data analysis (EDA) dashboards using Python libraries (Pandas, Matplotlib, Seaborn), providing data-driven storytelling to support project outcomes.
- Proactively pursued certifications in Python programming, SQL, and machine learning fundamentals, demonstrating continuous learning and technical growth.

## SKILLS

---

**Programming & Tools:** Python, SQL, Power BI, Linux/Windows (Command-line scripting, basic deployments)

**Machine Learning & Deep Learning:** Regression, Classification, CNNs, YOLO Object Detection, Natural Language Processing (NLP), Generative AI (OpenAI API, Ollama)

**Libraries & Frameworks:** Pandas, NumPy, Scikit-learn, TensorFlow, Keras, Matplotlib, OpenCV

**Core Strengths:** Data Analysis, Statistical Thinking, Project Ownership, Team Collaboration, Technical Communication, Rapid Learning

## EDUCATION & TRAININGS

---

Vishwakarma Institute of Information Technology, Pune, MH

**Bachelors of Technology, Mechanical Engineering (2021)**

ExcelR Academy, Pune, MH

**Data Science & Data Analytics (2023)**

## PUBLICATIONS & CERTIFICATIONS

---

**Mathematical Modelling of an Automatic Bag Mask Valve Emergency Ventilator** (Published on IRJET)

Published research on "Mathematical Modelling of an Automatic Bag Mask Valve Emergency Ventilator," addressing emergency medical equipment needs during critical healthcare shortages.

---

GL Academy	AI for Healthcare
ExcelR Solutions	Data Science Certification
ExcelR Solutions	Data Analytics Certification
IBM	Machine Learning with Python
Udemy	Artificial Intelligence A-Z 2023
Udemy	The Data Science Course 2023: Complete Data Science Boot Camp
Microsoft Certifications	AI-900: Azure AI Fundamentals by Microsoft Certifications
Udemy	Deep Learning A-Z: Hands-On ANN
Udemy	Machine Learning A-Z: Hands-On Python & R
Udemy	Feature Engineering for Machine Learning

## PROJECT CASE STUDIES

---

**Lung & Colon Cancer Image Classification (Classification using CNN)**

- Developed a CNN architectures model & Preprocessed images using a dataset of 25,000 images across five classes.

**Disease Prediction (Regression)**

- Created a predictive model for assessing apoplexy risk using 5,000 medical study records.