Chintan Vekariya

Data Scientist | Machine Learning Enthusiasts | Cloud Enthusiast

Mannheim, Germany, +49 175 9649316, chiintan.vekariya@gmail.com, LinkedIn, Portfolio

---

PROFILE

Passionate Data Science master's candidate and AI Engineer with a drive for innovation in fast-paced environments. Expertise in developing and implementing cutting-edge AI solutions, optimizing product functionality, and driving data-driven decision-making processes for impactful solutions.

EMPLOYMENT HISTORY

Sep 2022 — May 2024

Associate Software Engineer, Cloudoffis Technologies LLP
Ahmedabad, India

- **Project 1:** "Netra"
- **Responsibility:** In my role as the architect behind the NETRA Document Classification Engine for SMSF Audit Software, I ensured adherence to classification standards while developing and consistently updating the NETRA pipeline. My responsibilities also included implementing strategic logic changes to accommodate a wide array of document types, conducting thorough data exploration and preprocessing, training and optimizing machine learning models, and creating machine learning deployment pipelines for sub-components to adhere to CI/CD using AWS Code Pipeline and Jenkins.
- **Accomplishment:** Designed a pipeline incorporating 60+ transformers and scikit classifiers, expanding support to 23 new document categories through enforcing SMOTE for imbalanced data handling. Achieved a confidence increase of up to 95% for 14 categories, resulting in a notable 25% increase in overall confidence.
- **Tools and Technology:** Python, NLP, ML, Sklearn, Pandas, Linux, AWS, Git, Jenkins

- **Project 2:** "Tathya"
- **Responsibility:** Spearheaded the development of an automated data extraction pipeline for more than ten financial documents, utilizing LLM as the foundational framework.
- **Accomplishments:** Fastened the data extraction pipeline development by eliminating static extraction using Tabula and Camelot to dynamic extraction using Large Language Models.
- **Tools and Technology:** Python, Generative AI, LLM, Langchain, Linux, AWS, Git, Nginx

---

INTERNSHIPS

May 2022 — Jun 2022

Machine Learning Intern, Cloudoffis Technologies LLP
Ahmedabad, India

- **Project:** "Document Classification using CNN variants"
- **Responsibility:** Developed a predictive model for PDF classification utilizing Transformers, for textual contextual data-based PDF Documents.
- **Accomplishment:** Through rigorous testing on dummy data, I achieved a 15% increase in confidence levels compared to their previous classification models.
- **Tools and Technology:** Python, Transformers, Git, Datasets, PyTorch

Jun 2021 — Jul 2021

Research and Development Intern, Cloudoffis Technologies LLP
Ahmedabad, India

- **Project:** "Intelligent Document Management using AI"
- **Responsibility:** Researched and developed a signature detection pipeline for legal Trust Deed documents verification.
- **Accomplishment:** Provided them with the easy to integrate module for signature detection from the document.
- **Tools and Technology:** Python, Computer Vision, Git, Roboflow, LabelImg

---

EDUCATION

Feb 2024 — Present

Masters in Data Science, University of Mannheim
Mannheim, Germany

Jul 2019 — May 2023

Bachelor of Technology in Computer Science and Engineering, Charotar University of Science and Technology
Anand, India

Achievement: Received GOLD Medal for highest grades during Graduation.

Projects: Netra - i

Description: A Pi-operated device with an embedded camera for real-time navigation. Predictions are instantly converted into speech, enabling seamless navigation for users, especially those with visual impairments.

Tools and Technology: Python, Computer Vision, Deep Learning, Raspberry Pi, Linux, IoT devices

PROJECTS

Jul 2023 — Dec 2023

CloudSync

- **Description:** CloudSync revolutionizes information access by merging innovation and technology. Using RAG architecture and a chatbot interface, it offers users natural, conversational interactions with tailored documents, eliminating manual retrieval for efficiency in a centralized manner.
- **Tools and Technology:** Python, Langchain, Git, Streamlit, OpenAI, LLM (Llama 2 and GPT-3.5), sentence-transformers, Visual Studio, Qdrant.

Mar 2023 — Jun 2023

Falcon

- **Description:** Falcon is a project that offers holistic model interpretation and monitoring in machine learning. It combines SHAP and LIME techniques to provide a versatile framework for understanding models across tabular data and textual data domains. With a focus on both model explainability and interpretability, Falcon empowers users to gain deep insights into their models' decision-making processes.
- **Tools and Technology:** Python, JS, HTML & CSS, Nginx, MySQL, Transformers, Scikit-Learn, Flask, Evidently, Linux (Server).

SKILLS

- **Programming Languages:** Python, SQL, R.
- **Software Tools:** PyCharm, Jupyter Notebook, Spyder, Google Collab.
- **Version Control:** GitHub, Git, Bitbucket.
- **Machine Learning Skills & Packages:** Regression, Classification, Clustering, Generative Modelling, Fine-Tuning, Scikit-Learn, Model Quality Assessment and Testing, Pandas, Numpy, Matplotlib, Transformers, Langchain.
- **OS & Others:** Linux, Bash, AWS, MS Azure AI.
- **Organizational:** Time Management, Multi-Tasking, Agile.
- **Interpersonal:** Collaboration and Team Work.
- **Problem Solving:** Active Listening, Analysis and Decision Making.

LANGUAGES

English   Highly proficient

ACHIEVEMENTS

- Achieved GATE'CS All India Rank 1308 in 2023.
- Achieved GATE'CS All India Rank 1466 in 2022.
- Cloud Core Badge holder from IBM.

CERTIFICATES

- Introduction to Machine Learning in Production (Coursera).
- Configuration Management and the Cloud (Coursera).
- R Programming (Coursera).
- Introduction to Containers, Kubernetes, and OpenShiftV2 (Cognitive Class)

REFERENCES

Mudra Vora from **Cloudoffis Technologies LLP**

mudra.v@cloudoffis.com.au

Jaimin Mungalpara from **Simform**

Jaimin.ml2001@gmail.com

Chintan Vekariya, 2024