Akarsh Raj

M.Sc. AI @ FAU, Germany

- Nürnberg, Germany, 90431
- +4915754741079
- akarshraj.ds@gmail.com

Profiles

- in akarshrajsingh7
- ⬤ akarshrajsingh7
- ⬤ akarshrajsingh

Technical Skills

Generative AI, Large Language Models(LLM), RAG (Retrieval Augmented Generation), Python, ML, DL, NLP, Computer Vision, Pytorch, Tensorflow, OCR, Optimization, DataBricks, Flask, XgBoost, C++, R, SQL, beautifulsoup, HTML, CSS, pyspark, Airflow, H2O, pandas, scikit-learn, MS Excel, MS Powerpoint

Languages

English

Hindi

German

Summary

6.5+ years of experience in the field of Data Sciences and AI. Has extensive hands-on experience with *NLP, tabular and Computer Vision problems*. Currently ranked within the top 4 percent in the Competitive Tier on Kaggle. Technical Speaker at multiple data science meetups.

Education

Friedrich-Alexander-Universität Erlangen-Nürnberg   October 2023 - Present
Nürnberg, Germany   M.Sc. Artificial Intelligence
🔗 https://www.fau.de/

University of Petroleum & Energy Studies   July 2013 to May 2017
Dehradun, India   B.Tech. Computer Science
83.2%
🔗 https://www.upes.ac.in/

Experience

Siemens Healthineers   July 2024 - Now
Working Student (LLMs in Finance)   Forchheim, Germany
🔗 https://www.siemens-healthineers.com/

- Developing a pipeline for automatic **Financial Report Generation** leveraging **Large Language Models** (LLMs) for the same. Also, making sure that the LLMs don't hallucinate. Achieved via **Prompt Engineering** approaches like COT and Few Shot examples and some internal **fine-tuning** of the backbone.

Evalueserve   April 2022 - July 2023
Lead Engineer (Data Science)   Gurugram, India
🔗 https://www.evalueserve.com/

- Infinity AI : Developed a **RAG (Retrieval Augmented Generation)** tool on top of GPT based models for abstract question answering, information extraction, and more from personal or organisational documents using natural language. Improved firm's reasearch capabilities by over 30% (w.r.t TAT and Quality). Also implemented **masking of PII data**, processing of tabular data in the input datasets

- Canvass AI - Designed & developed a NLP system for researchers which **web scraped google results** (SERP-API) based on a query and generates various **text analytics** (NER, n-gram analysis, sentence level clustering, machine comprehension, etc.) on the results, increasing researcher's efficiency by over 40%

- FinTex: Developed an application to **extract tables** from Financial Reports. Worked on digital as well as scanned PDFs. Solution architecture involved **extensive OCR techniques, table localisation and detection algorithms**, to extract refined tables

- Worked with stakeholders to develop quarterly roadmaps based on impact, effort and efficiency

ZS Associates   October 2018 to March 2022
Senior Data Scientist   Gurugram, India
🔗 https://www.zs.com/

- Copay Anomaly detection: Developed **anomaly detection engine** to detect fraudulent pharmacies / prescribers using **autoencoders** & **isolation forest** identifying fraudulent pharmacies who were responsible for $5M+ valued **fraudulent transactions** annually. Supplemented the model through SHAP explanation model

- Receipt Detection using Deep Learning: Developed an **object detection model using Faster-RCNN** to detect receipts in the image; Implemented. Transfer-Learning and achieved a **MAP of 85 % in production**

- Topic modelling - Developed a **topic model** which tagged topics in notes taken down by **medical field agents**. Used unsupervised model approach which uses transformer architecture

- Clinical Trial Optimisation: Designed and developed a solution for **clinical trial optimisation** using **linear programming** satisfying thousands of constraints on clinical trial sites

- **Time Series Forecasting** of new to brand (NBRx) numbers for a medical drug using Prophet, also tackled & incorporated COVID impact into the forecast

Xceedance   May 2017 to September 2018
Data Scientist   Gurugram, India
🔗 https://xceedance.com/

- First year **loss prediction model** for one of the largest US based insurance firm which reaped them 3M+ profit, annually. Also developed an API using Flask for the same

- Introduced Text Field Classification System assisting Underwriting process, minimizing process time from 20 Minutes/Entry to an instant using Python, Machine Learning, and **OCR techniques** for one of the **Fortune 10** client.

Honors & Awards

- Kaggle Days Meetup : **Opening Speaker** (June, 2019) - *Opening speaker of the first ever Kaggle meetup in India conducted in Delhi-NCR. Talked at few other similar events*

- *Dream Team Award* (Evalueserve) - *Awarded for the developing NLP system for Evalueserve employees which improved their research analyst's efficiency by over 40%*

- **Project Champion** 2019 & 2020 (ZS Associates) - *Awarded for the best project across all offices at ZS*

- **Rising Star** (Xceedance) - *Awarded on annual day celebration of the firm, for commendable contribution to the growth and success of Data Science team at Xceedance*