# Sachin Rajput

https://www.linkedin.com/in/sachin-rajput-51b402162/ | S.Rajput@2023.ljmu.ac.uk | 9717481050
https://varun56278.github.io/Portfolio/ | Delhi, India | https://github.com/Varun56278

## EDUCATION

**Liverpool John Moores University, Liverpool L3 5UX, UK** **Jan 2022-Dec 2023**
Master of Science, Machine Learning & AI
***Courses:*** *Foundations of AI, Analysis of Algorithms, Machine Learning, NLP, AI, Deep Learning*

**Guru Gobind Singh Indraprastha University, Delhi, INDIA** **Aug 2015-Jun 2019**
Bachelor of technology, Mechanical Engineering GPA: 7.4
***Courses:*** *Algorithms and Design, mathematics, Robotic, Artificial Intelligence*

## TECHNICAL SKILLS

- **Interests and Domain:** Full stack development, Knowledge Graph, Data Analysis, Machine Learning, NLP, Time Series, CNN, RNN, GBM, LightGBM.
- **Languages and Frameworks:** Python, C++, HTML, SQL, CSS, JSON, Flask, Git, Django, Jenkin, Docker, Mlib, Scoop, Zookeeper, Generative AI.
- **Databases and QL:** MySQL, MongoDB, Cassandra, TCP, big data, relational databases.
- **Libraries:** Pandas, Plotly Matplotlib, Dash, Numpy, Seaborn, Scikit-Learn, Spacy, Gensim, Keras, Pytorch, Tensorflow, Re, Beautiful Soup, NLTK, LLM
- **Tools and Infrastructure:** SageMaker, GCP, AWS, Tableau, Azure, Docker, Snowflake, Pyspark, Glue, EC2, ECS, Hadoop, apache hive, lambda, Map Reduce, Microsoft Excel, Power BI, Databricks, Azure, Azure Machine Learning Studio, OpenAI, Hugging face, BERT, GPT, PALM, LaMDA, GAN
- **Modelling Techniques:** Exploratory Analysis, Statistics, Data Mining, Data Visualization, Linear Regression, Logistic Regression, Random Forest, Support Vector Machines, K-Means, Clustering, A/B Testing, Forecasting.
- **Problem Cases:** Sales Forecasting, Customer Churn, Price Optimization, Market Mix Modelling, Computer Vision & NLP, Recommendation Engine

## WORK EXPERIENCE

- **Data Scientist | Inventico Infotech Pvt Ltd, Noida, India** **Jan 2022– July 2022**
  - Developed and implemented an automated ticket classification system leveraging Natural Language leveraging Processing (NLP) techniques quantitative analysis and Machine Learning algorithms to streamline ticket management process. I have used Topic Modelling, statistic, analytics, NLTK, Logistic Regression, Decision Tree, Random Forest, Naïve Bayes, metrics, and I have got 95% Accuracy with Logistic Regression model.
  - Work on Build PySpark and Spark Mlib, End to End AWS Sage maker blazing text Classification Model. Blazing Text can train the model on more than a billion words in couple of minutes using a multicore CPU or GPU, while achieving performance or par with SOTA deep Learning text classification. Sage maker needs unique training jobs to run. The S3 bucket and prefix to use for training and model data. Bucket, Sage Maker SDK will create a default bucket following a pre-defined naming convention in the same region. The IAM role ARN used to give Sage Maker access to your data. It can be fetched using the method from sage maker python SDK.
  - Business Context of Edtech company had s phenomenal seed A funding round. It used the money to increase its brand awareness. As the marketing spend increased, it got several leads from different sources. It had to be profitable in the long run to sustain the business. After some time, quantitative analysis of Business reduces their customer acquisition cost in the long run. The main objectives of lead scoring. Removing junk by categorizing leads based on propensity to purchase. Gain insights to streamline lead conversion and address improper targeting. I have created three pipelines for use DATA Pipeline, TRAINING Pipeline, Inference Pipeline. Our Metrics for ML model AUC score > 75%, Precision > 65% Recall > 75%.
- **Data Scientist | EzappSolution, New York, USA** **Aug 2020– Jan 2022**
  - Worked in the **Research and Development team focusing on Digital Technologies.**
  - A music streaming company in south Asia with Business Insights. They offer their services to millions of people, supported by advertising and paid subscription. The arrival of some new competitors, the company's churn rate is rising. The task is to predict the propensity of customer churn. Customers are likely to leave us within the next month. Provide any insights into why they are not continuing their subscription. The Problem could be approached with an ML solution, objective is to preventive actions to be taken before customers churns. Using Churn Data. Machine Learning Model predicted customer will likely leave the service within the next month. Before building a model we created architecture for clarity regarding the ML system. I have used ETL for Data. I have created three pipelines for use DATA Pipeline, TRAINING Pipeline, Inference Pipeline. Our Metrics for ML model AUC score, Precision, Recall and I have used Airflow and ML flow for scalability and Monitoring.
  - Collaborated in a team of 4 to build a sales forecasting model using ARIMA and Facebook prophet to investigate Sales. Collect Data from SQL Dashboard in python by plotly dash is fully dynamic and responsive on any device. I here 3 charts and one data table in this web app. We used our leadership; Skills We were influence by our previous solve problems. We were work with team sales incentives and after that we did product launches.
  - Worked on Video Frame Prediction using GAN. I used action recognition dataset, predicted future frames to analyze the evolution of prediction.
  - We worked on Cycle GAN take noise as input and convert it to an image. We converted T1 MRI images into T2 MRI images. The ideal converted images of one domain into another domain.
  - We worked on Cycle GAN Convert images from horse into zebra and convert images from the digit 4 domain to digit 9 domain.
  - Worked on object detector to detect the presence of bicycles, people and other objects that are present in an image using CV2. Similar YOLO detector SSD use a single pass to detect the presence of objects in the input image.
  - Oil Tanker Segmentation used to build a Mask R-CNN model to segment oil tankers present in satellite images, asset mapping is key of image segmentation.
  - Worked on Lung Segmentation using U-Net used on the framework to perform semantic segmentation over a collection of chest X-ray images.
  - Worked on Summarization large documents using Large Language Model using Hugging Face and open AI.
  - Worked on Document Retrieval and Question Answering with Large Language Model using Hugging Face and open AI.
- **Machine Learning Engineer | Tech Vision, New Delhi, India** **Nov 2019-July 2020**
  - **Machine Learning:** Automation of Shipment Forecast using ARIMA, facebook prophet, XGBOOST, ENSEMBLE.
  - **Recommendation:** Sentiment Based Product Recommendation Model.
  - Utilized NLP Algorithms to automate the process of identifying sentiment from customer reviews, increasing of customer reviews, increasing accuracy rate up to 95%.

## ACADEMIC PROJECTS

- **Credit Card Fraud Detection** | ***Python scripting language, Flask, PyCaret (AutoML), AWS***
  A Model Building and Prediction using Pycaret, Product Operations, production systems, Data Infrastructure, distributed systems, information systems.
- **Mobile app Rating Prediction** | ***Python, ML, Data Science, Deep Learning, Natural Language Processing***
  WebApp to generate a Mobile app rating Prediction based on a single input title given by user through ML and Deep Learning.

## PERSONAL INFORMATION

- **Gender:-** Male, **Nationality :-** Indian