

Sidhant Jain

Delhi NCR, India · 9654393773
sidhantjainoff@gmail.com | sj020.github.io

EDUCATION

Post Graduate Program in Data Science	Praxis Business School	80.98%	2022 — 2023
Bachelors in technology in Mechanical Engineering	Shiv Nadar University	7.78	2016 — 2020

EXPERIENCE

PwC India | Data Science Associate

July 2023 – Present

- Collaborated on the development of a data migration pipeline, ensuring smooth transition from **MS-SQL to the Azure Data Lake**, resulting in improved data retrieval efficiency and minimized downtime.
- Worked on developing a bespoke **alert system** utilizing the **SMTP protocol** to promptly notify stakeholders of any data migration failures, including details of the affected table, ensuring timely response and resolution.
- Engaged in a development initiative by crafting a **finance chatbot**, serving a user base of 20k employees, harnessing the capabilities of a pretrained **Llama2** model to achieve state of the art performance.

Cognizant Technology Solutions | Analyst

July 2021 - Sep. 2022

- Designed a destination **recommendation system** tailored multinational luxury hotel, harnessing collaborative based techniques to guide guests in selecting their ideal destination.
- Implemented **MLflow for monitoring** and versioning, complemented by **Airflow for orchestration**, to streamline the management of machine learning workflows.
- Developed a **REST endpoint** that delivers results in JSON format, facilitating seamless integration with their in-house **web application**.

Cognizant Technology Solutions | Analyst

July 2020 - July 2021

- Developed a specialized trading system focused on agricultural products like fertilizers, pesticides.
- Obtained industry-level training in Java, Spring Boot, HTML, and CSS, gaining technical expertise in these areas.

PROJECTS

News Genie

- Led a team of 5 members in the development of **headline generation and text summarization** web application, employing cutting-edge techniques to enhance user experience and content comprehension.
- Utilized **transfer learning** with **transformer** based models, fine-tuned on Economic Times dataset.
- Assessed the model's performance using the **ROUGE score** and achieved the highest score of 74%.
- Engineered a **batch prediction pipeline** designed to generate predictions during periods of user inactivity, ensuring continuous data processing.

Home Credit Default Risk

- Engaged in a project centred around classification, specifically aimed at **predicting the risk of customer default**.
- Utilized a diverse set of classification algorithms, which encompassed **decision trees, logistic regression**, as well as ensemble techniques like Random Forest and AdaBoost.
- Conducted model performance assessment using AUC ROC score, ensuring the reliability and effectiveness of predictive capabilities.

TECHNICAL SKILLS

- Languages:** Python, Java
- Libraries:** Pandas, Numpy, Sklearn, PyTorch, Seaborn, Flask
- Databases:** MySQL, MongoDB
- Visualizations:** Power BI
- Miscellaneous:** Azure Data Factory, Azure Databricks, Docker, Streamlit, FastAPI, Kedro, HTML/CSS