# KUNAL  KOLHE

✉ kunalkolhe333@gmail.com

---

## Professional Summary

- A Data Consultant having 5.5 years of total industry exposure and 4+ years in data science, proficient in python, large volume data analytics, automations, machine learning and deep learning techniques with data processing, modeling, data insights, production support and end to end data management using ETL.
- Strong experience in data analytics, developed and configured ML models in Regression, Classification, Clustering with model trainings, validations, matrices and L3 production support using notebooks, airflow.
- Delivered data-driven AI solutions with process automations, data management from various data sources to data warehouse as GBQ, API, SFTP, flat files, web scrappers, documents to resolve business challenges.
- Expertise knowledge on core ML methodologies including Regression, Classification, Clustering, Pandas, NumPy, Scikit-learn and familiar with TensorFlow, Keras, Gensim, NLP, OpenAI tools, LLM, GPT, data scraper, AutoML.
- Developed, re-configured and monitored end to end production data pipelines on airflow, analyzed bugs and fixed data issues on priority using python, GBQ, dashboards with ADOC releases.
- Involved in scrum calls, client's weekly meetings, overseen junior's assignments and recommended changes in places.
- Highly motivated individual, team player, quick learner, hunting new emerging stacks and targeting leadership assignments.

## Technical Stacks

- **Proficient –** Python, Pandas, NumPy, Scikit learn, PyCaret, Regression, Classification, Clustering, Seaborn, SQL.
- **Familiar –** TensorFlow, Keras, NLP, ChatGPT, LLM, LangChain, OpenAI, , Gensim, scraping, C/C++.
- **Cloud / ETL Tools –** Airflow, Big Query, S3, Vertex AI, GCS, SNS, DAG, Composer, Streamlit, Gradio, AWS, GCP.
- **IDE / Platforms –** Notebooks, PyCharm, VS Code, Jupyter lab, Spider, Linux, Window, Jira, Git.

## Work Experience

- Software Engineer at GSPANN Technologies | *Oct 2022 – Current*
- Data Science Developer at Softqube Technologies | *Feb 2022 – Oct 2022*
- Software Engineer at 63 Moons Technologies | *May 2021 – Oct 2021*
- Software Engineer at KPIT Technologies | *Oct 2018 – Jan 2021*

## Education

- PG-DAC in Computer Science | CDAC Kharghar, Mumbai *(2018)*
- B-Tech in Electrical Engineering | BATU University, Raigad *(2017)*

## Participations

- Technical DevFest event | *(GDG, India)*
- Microsoft AI classroom series program | *(Microsoft, India)*
- Machine learning with explainable AI | *(TMLC Academy)*
- AWS SA cloud foundation | *(Sunbeam, Pune)*

## Achievements

- First rank in Carrom at GSPANN Premier League
- Gold Medalist at University Carrom Championship

## Projects

**Title –** DRL data analytics and data management with ETL data pipelines and ML on cloud.

**Client –** Dr. Reddy's Lab| 20+ Team size

**Technologies –** Python, Machine Learning, GCP, Airflow, Big Query, GCS, Vertex AI, Notebooks, DAG, web API.

---

**Description –** DRL data analytics using ML and custom data pipelines including data insights on demand, data extraction, ingestion, preprocessing, transformation, live monitoring, supporting, managing data with reporting to business heads and stakeholders on demand.

**Responsibilities –**

- Developed, enhanced, monitored end-to-end 60+ data pipelines on GCP cloud composer with validations and mail alert mechanism.
- Involved in ML modeling as Clustering, RFC models, training and configuring, validations as per business needs.
- Analyzed and resolved production data issues with L3 support, ADOC releases and RCA on priors.
- Designed custom scripts using python to automate existing workflows and reduced 1.5 hours on validations.
- Involvement in business meetings, scrum calls, architect reviews, handling juniors, reviewed and finalized deliverables items with data flows and documentation checks.

---

**Title –** Hospital Insurance Fraud Claims Detection
**Client –** Confidential | 3+ Team size
**Technologies –** Python, Machine Learning, Notebook, Plotly, Streamlit.

---

**Description –** Automate the insurance claiming process where it identifies the hospital insurance claim is valid or invalid using machine learning solutions. Model classifies the insurance claims in verification process at insurance company.

**Responsibilities –**

- Performed in-depth EDA, features engineering, data transformation, data insights using python, statistical techniques and provided matrices to stakeholder.
- Used Random Forest Classifier, data balancing techniques for imbalanced nature, stratified K-Folds as cross validation and achieved model accuracy of 92% with F1-score 88%.
- Managed juniors and reviewed solutions, provided end-to-end demo with data dictionaries, cloud deployment POC.

---

**Title –** Hotel Revenue Portfolio Enhancement
**Client –** Kriya RevGEN | 7+ Team size
**Technologies –** Python, Machine Learning, FB Prophet, Notebook, Automations, SQL, Web scraping.

---

**Description –** It provides hotel revenue management solutions which helps in managing hotel forecasts and budgets, pricing strategies, brand initiatives and maintain revenue growth in portfolios. We have delivered automated AI solutions on dynamic hotel pricing using ML strategies, forecasting and yield methods.

**Responsibilities –**

- Designed and implemented dynamic price prediction action plans for hotels with yield management, global rate competitors using Random Forest regression model 90% avg. accuracy.
- Analyzed 30+ hotel properties including revenue data, hotel data, customer data and shared detailed EDA reports with statistical findings, patterns and recommendations to management for revenue BI decisions.
- Reviewed existing manual data scrapers and optimized it by automation on airflow environment.
- Attained stakeholders and external data vendors team, provided end-to-end demos with data dictionaries and cloud deployments.

---

**Title –** Forecasting the asset prices for stockbrokers and vendors.
**Client –** Confidential | 10+ Team size
**Technologies –** Python, Machine Learning, Deep learning, NN, Notebooks, Web API, AWS, C++, Jira, Bitbucket.

---

**Description –** ODIN is a capital market product which offers trading in multi assets like equity, commodity, settlements, stock analysis and assets valuation for end-users and stockbrokers. It focuses on AI based POC which shows stock analytics and forecasting values using neural network modeling.

**Responsibilities –**

- Researched top gainer and looser capital market competitors with NSE indices NITFY 50, NIFTY MIDCAP 50 in banking, pharma, IT sectors on daily, weekly, monthly basis and shared data insights via Amazon S3.
- Developed AI solutions for stock data analysis, scripts price predictions POC using LSTM model with avg. RMSE less than 3% under organizational policies for product enhancement.
- Automated Jira process, work-logs, validations using python, selenium and minimized team review efforts by 0.5 hours.