

Abhinav Yadav

☎ +91 9140435720 ✉ abhi9v2204@gmail.com in <https://www.linkedin.com/in/abhii9v>

PROFESSIONAL SUMMARY

Data Scientist with experience developing and deploying machine learning models for insurance and advertising analytics, forecasting, and business optimization. Skilled in Python, SQL, statistical modeling, and time series forecasting with hands-on experience building end-to-end ML pipelines, monitoring model performance, and implementing automated retraining workflows. Experienced in NLP and LLM-based systems, with a strong focus on converting business problems into scalable data science solutions.

EXPERIENCE

Data Scientist, *Cognitio Analytics Inc*

Sep 2025 – Present

- Developed **time series forecasting models** to predict **12-month PLPM (Per Life Per Month) cost** for an insurance client using regression and boosting approaches, achieving **less than 10% MAPE**.
- Implemented an automated **model monitoring framework** to track prediction accuracy, detect **data and concept drift**, and evaluate model stability over time.
- Built scheduled **retraining pipelines** that automatically trigger model retraining and deployment when monitoring thresholds or performance metrics degrade.
- Performed **exploratory data analysis (EDA)** and feature engineering on large healthcare/insurance claims datasets using Python and SQL to improve forecasting reliability.
- Collaborated with actuarial, product, and engineering teams to integrate cost forecasting outputs into analytics systems for premium planning and risk management.

Data Scientist, *ShyftLabs, Noida*

Jan 2024 – Aug 2025

- Built and deployed **session-level recommendation systems** for cross-sell and up-sell opportunities, enabling personalized ad delivery for registered and anonymous users.
- Developed multiple **time series forecasting models** for impressions, requests, and clicks using stacked models (**Linear Regression, XGBoost, LightGBM**) achieving **MAPE under 10%**.
- Designed a scalable **semantic search engine** using transformer-based embeddings and **Elasticsearch vector search** for improved query relevance.
- Built **CTR prediction models** to estimate click probability and dynamically rank advertisements.
- Developed a **hybrid search architecture** combining vector embeddings (paraphrase-MiniLM) with LLM-based query intent detection.
- Conducted data analysis and feature engineering using Python, Pandas, and SQL to improve model performance and campaign optimization.

Data Analyst & Research Intern, *Intellipaat, Bangalore*

Aug 2023 – Dec 2023

- Conducted statistical analysis on learner activity data to identify behavioral trends and provide recommendations for platform improvements.
- Applied clustering algorithms (**K-Means, DBSCAN**) for student segmentation and cheating detection, achieving **80% automated detection accuracy**.
- Mentored students in machine learning fundamentals and practical implementation of algorithms.

EDUCATION

Lovely Professional University

Aug 2020 – Jul 2024

B.Tech in Computer Science

Relevant Coursework: Machine Learning, Deep Learning, Natural Language Processing, Statistics, Database Systems

TECHNICAL SKILLS

- **Programming:** Python, SQL, PySpark, Git
- **Machine Learning:** Linear Regression, Logistic Regression, Random Forest, Gradient Boosting, XGBoost, LightGBM
- **ML Frameworks:** Scikit-learn, TensorFlow, PyTorch
- **Time Series:** Forecasting, Feature Engineering, Model Evaluation
- **NLP / GenAI:** Transformers, LLMs, Embeddings, Prompt Engineering
- **Data Processing:** Pandas, NumPy, ETL Pipelines
- **Model Deployment:** Docker, Flask, Gunicorn, MLflow
- **Databases:** PostgreSQL, Elasticsearch, MongoDB
- **Visualization:** Tableau, Matplotlib, Seaborn
- **Statistics:** Hypothesis Testing, A/B Testing

CERTIFICATIONS

- **Natural Language Processing Specialization** — DeepLearning.AI
- **Python for Data Science and AI** — IBM