Shubham Yedekar

"Data Scientist | Machine Learning | Patient Engagement Analytics"

Hartford, CT — 860-380-6594 — yedekarshubham7188@gmail.com — https://www.linkedin.com/in/shubham-yedekar/

Summary

Data Scientist with expertise in machine learning, predictive modeling, and advanced analytics, skilled in Python, SQL, and Tableau. Passionate about driving healthcare innovation by leveraging scalable data solutions to optimize patient engagement and improve health outcomes.

Work Experience

Hexaware Technologies – Data Centric Team, Mumbai, India

Software Engineer   *Aug 2022 – Aug 2023*

- Designed and deployed predictive models using XGBoost and Random Forest, automating healthcare adjudication workflows and reducing operational delays by 40%.

- Implemented automated ETL pipelines in Python and SQL, standardizing the processing of large healthcare datasets and enhancing data integrity by 25%.

- Developed anomaly detection models to identify abnormal patterns in claims data, reducing rework by 20% and improving system accuracy.

- Conducted detailed analyses of structured and unstructured claims data, delivering actionable insights to optimize cost positioning and provider performance metrics.

- Collaborated with cross-functional teams to design monitoring tools for tracking operational workflows, improving policy adherence and claims management processes.

Associate Software Engineer   *Sep 2021 – Aug 2022*

- Automated healthcare claims validation workflows using Python-based frameworks, reducing processing time by 50% and increasing scalability.

- Built logistic regression models to improve system performance predictions, enhancing anomaly detection and increasing predictive accuracy by 15%.

- Designed comprehensive data dictionaries and standardized reporting definitions, streamlining claims workflows and reducing delivery delays by 20%.

- Developed a Tableau-based tracking tool to visualize claims and appeals KPIs, enabling stakeholders to monitor metrics in real time and make data-driven decisions.

Projects   `https://github.com/ShubhamRSY`

- **Provider Selection Analytics:** Developed Python-based models using clustering algorithms (K-Means, DBSCAN) to analyze provider performance data, optimizing cost efficiency and identifying top-performing providers based on quality and cost metrics, resulting in more informed and effective provider selection.

- **Claims Adjudication Optimization:** Built machine learning algorithms with decision trees and ensemble methods (Random Forest, XGBoost) to automate claims adjudication, reducing adjudication time and improving decision transparency by predicting which claims should be processed automatically and which require manual review.

- **Auto Insurance Claim Cost Prediction:** Designed a Poisson regression model (a type of GLM) to forecast claim costs based on factors like severity and frequency, improving the accuracy of claim cost predictions by 15%, leading to better budgeting and resource allocation for insurance claims.

- **Operational Efficiency Dashboard:** Created interactive Tableau dashboards to visualize claims and appeals data, enhancing real-time operational transparency and enabling stakeholders to quickly identify workflow inefficiencies, improve resource allocation, and optimize decision-making.

- **NLP-Based Sentiment Analysis:** Built a sentiment analysis model using Bidirectional LSTM to classify healthcare sentiment from textual data in claims and appeals, achieving 90% accuracy and providing actionable insights into appeals trends for more informed prioritization of cases.

Education

Master's in Data Science — University of Connecticut, Storrs, USA — GPA: 3.8/4.0	*Aug 2023 – Dec 2024*
Bachelor of Engineering in Electronics — University of Mumbai, India — GPA: 3.5/4.0	*July 2017 – May 2021*

Certifications

- **Certified Data Scientist (CDS)**

- **Google Cloud Professional Data Engineer Certification**

- **Advanced Machine Learning with TensorFlow on Google Cloud**

- **Tableau Desktop Certified Associate**