

SATAKSHI DESHMUKH

(959) 248-9601 | satakshi.deshmukh@uconn.edu | <https://www.linkedin.com/in/satakshi-deshmukh/>

Experienced Data Professional skilled in SQL, Python and big data technologies. Expert in developing predictive models and SQL reports improving reporting efficiency by 40%. Led cross-functional teams and implemented ETL processes to enhance data-driven decision-making. Proven track record in boosting client satisfaction and business growth by innovative data solutions.

EDUCATION

UNIVERSITY OF CONNECTICUT, SCHOOL OF BUSINESS, HARTFORD, CT

December 2024

Master of Science in Business Analytics and Project Management, Concentration Data Science

CGPA 3.9/4

RAJIV GANDHI PROUDYOGIKI VISHWAVIDYALAYA, INDIA

June 2021

Bachelor of Technology, Computer Science & Engineering

CGPA 8.1/10

TECHNICAL SKILLS

- **Programming Languages** SQL, Python, R, C, C++, Java, SAS, DAX
- **Data Analytics** Pandas, NumPy, EDA, Data Cleaning, Data Visualization, Feature Engineering, PCA, Data Modeling
- **Data Science** Machine Learning, NLP, Deep Learning, Clustering, LLM, Time Series Forecasting, TensorFlow, PyTorch
- **Data Engineering** ETL, Hadoop, Pig, Hive, Spark, MapReduce, Data Warehousing, Data Pipelining, Redshift, Big Query
- **Statistics** Hypothesis Testing, Probability Distribution, Regression Analysis, ANOVA, Chi-Square Test
- **Software & Tools** Excel, Power BI, Tableau, Git, Visual Studio, SAS, JMP, Docker, Streamlit, Kafka, Kubernetes
- **Cloud Technologies & Databases** Snowflake, Microsoft Azure, AWS, GCP, SQL Server, Oracle, MySQL, Maria DB
- **Project Management Skills** Agile Methodologies, JIRA, MS Project, Scrum Practices

PROFESSIONAL EXPERIENCE

Software Engineer, Yardi Software India Pvt Ltd, India

July 2021 – July 2023

- Implemented **ETL** with Hadoop and AWS for data acquisition and preparation, driving insights for client projects.
- Developed **predictive model** achieving 90% accuracy in forecasting tenant rent default risk and lease renewal decision, leveraging historical tenant data and machine learning algorithms in Python.
- Developed and optimized 25 **custom reports** using SQL server, enhancing data retrieval and reporting efficiency by 40%.
- Created **stored procedures and triggers** using SQL demonstrating strong analytical and problem-solving skills.
- **Analyze data issues** reported by clients and construct 50 data fix packages using SQL ensuring accuracy and reliability.
- Led global cross-functional team collaboration to resolve custom report-related issues, resulting in substantial increase in client satisfaction, demonstrating effective communication skills and teamwork.

Data Science Intern, Adinath Poly Plast Pvt Ltd, India

April 2021 – June 2021

- Developed 88% accurate **predictive model** for sales of marketplace to optimize inventory management using Python.
- Identified critical sales **KPIs** for product lines, leveraging Excel and Python to transform data into actionable insights.
- Collaborated on developing **Power BI dashboard** to visually represent KPIs, enhancing Amazon sales strategy.

Software Developer Intern, Technosoft Informatics, India

May 2019 – July 2019

- Developed backend for the 'Arihant Transformers' website using **Java** enterprise edition framework (Spring, Hibernate)

ACADEMIC PROJECTS

- Engineered **Health Insurance Premium Predictor** web app, achieving 0.85 R^2 with Gradient Boosting, performed data analysis, feature engineering and implemented ML regression algorithms using **NumPy, Pandas, Seaborn, Scikit-Learn**.
- Developed **Power BI dashboards** for HR analytics and Spotify Songs KPIs, and **Tableau dashboard** for Amazon Prime analysis, delivering critical insights and driving data-driven decisions in diverse domains.
- Led **sentiment analysis** and customer satisfaction level prediction on 34,000 Amazon reviews using **NLP, Deep Learning** and **ML**, achieving 93.7% F1 score with **LSTM**. Uncovering key insights through topic modelling by LDA and NMF **clustering**.
- Engineered data analysis **pipelines** on Amazon using **Hive and Pig** for 49,590 movie ratings dataset, performing advanced queries and managing **AWS Hadoop** cluster operations to identify top movies and produce rating histograms.
- Developed **LLM-powered Gen AI app** transforming natural language into SQL queries, facilitating data analysis and enhancing tailored business needs via connection to personal **MySQL** database using **SteamLit, Gemini Pro, LangChain**.
- Conducted **A/B testing** on Facebook and AdWords campaigns using **Pandas, hypothesis testing**, and **regression analysis**, enhancing client ROI by demonstrating superior platform's cost-effectiveness, leading to a 30% increase in conversions.
- Developed CNN classifier to **diagnose pneumonia** with **OpenCV, Keras, TensorFlow** from preprocessed chest X-ray dataset comprising 5,840 images, enhancing pediatric healthcare through improved disease detection with 92% test accuracy.
- Developed property management database, conducted **data modeling** and SQL-driven EDA to produce actionable insights.