

Shubham Wankar

 shubhamwankar23@gmail.com	 Bengaluru, IN	 www.linkedin.com/in/shubham-wankar/	 8978996233
---	---	---	--

Data Scientist

SUMMARY

Dynamic and results-driven Data Scientist with over 7+ years of professional experience and a Master's Degree in Data Science from Liverpool's John Moores University. Experienced in leveraging data-driven insights to optimize business performance, with a strong focus on data science specialization. Currently building Generative AI solutions at Accenture for a Global Banking client and developing RAG models. Seeking opportunity to grow in the field of Gen AI.

KEY SKILLS

- Data Analysis • Sentiment Analysis • NLP • Mathematics • Exploratory Data Analysis (EDA) • Generative AI
- Statistical Analysis • Project Management • Database Management • Problem Solving • Critical Thinking • Version Control

TECHNICAL SKILLS

Programming Languages: Python, SQL

Tools: Tableau, Power BI, Looker Studio, Microsoft Office, Microsoft Excel, Git, Github, Bash

Database: PostgreSQL, Microsoft SQL, MongoDB

Machine Learning: Regression, Classification, Clustering, Decision Trees, Random Forest, SVM, Neural Networks

Deep Learning: ANN, CNN, RNN, LSTM and GRU

Frameworks: Tensorflow, PyTorch, Langchain, Flask, Django, FastAPI, Streamlit, OpenAI API, HuggingFace

Libraries: NumPy, Pandas, Sci-kit Learn, Matplotlib, Seaborn, OpenCV, Requests, Beautiful Soup, NLTK, Spacy, Transformers, AutoEncoders

Cloud Platforms: Microsoft Azure (Azure AI Services, Azure AI Search, Azure OpenAI Service), Databricks

PROFESSIONAL EXPERIENCE

Senior AI/ML Engineer	Sep '24 - Present
Accenture	Bengaluru, IN
Implemented Semantic Re-Ranking approach to rank the retrieved chunks for internal RAG model improving relevancy by 20% Global Banking Client:	
<ul style="list-style-type: none">• Integrated complex Azure Promptflows with AI Load Balancer to reduce the latency to 5 secs for an internal HR tool	
Data Scientist	Jun '23 - May '24
Invictus Performance Lab	Bengaluru, IN
<ul style="list-style-type: none">• Led data pipeline design, boosting data availability and reducing processing time by 20%• Automated reporting via Python, cutting manual effort by 40% and improving client satisfaction by 15%• Developed interactive dashboards with Looker Studio, enabling stakeholders to make data-driven decisions, resulting in a 15% improvement in individual performance• Streamlined workflows, increasing efficiency by 30% through automation of daily tasks• Built predictive model with XGBoost, achieving 89% accuracy in forecasting client readiness• Mentored junior team members and interns, fostering their growth and contributing to a 20% increase in team productivity	
Senior Risk Analyst	Feb '17 - Oct '22
Amazon	Hyderabad, IN
<ul style="list-style-type: none">• Managed stakeholder relationships to minimize the risk of intellectual property rights infringement, leveraging data-driven insights to drive decision-making.• Optimized the Standard Operating Procedure, enhancing clarity and reducing working time by 20% through data analysis and process improvement.• Conducted knowledge-sharing meetings to improve team productivity, resulting in a 15% increase in efficiency.• Collaborated cross-functionally to identify and mitigate risks posed by recurring bad actors, utilizing data analysis techniques to inform preventive measures.• Trained and mentored 25 junior team members, achieving an overall team productivity of 95% and a quality rating of 98%.	

PROJECTS

Text 2 SQL LLM App	May '24 - Aug '24
<ul style="list-style-type: none">• Created a robust Text to SQL application utilizing Gemini 1.5 Flash & Langchain to transform natural language questions into SQL queries and return results as natural language responses	

<ul style="list-style-type: none"> Developed a Streamlit-based chat interface to facilitate real-time interaction with a SQL agent, enabling seamless dynamic querying and response generation 	
Next Word Prediction LLM Finetuning	Feb '23 - Mar '23
Zummit Infolabs	Bengaluru, IN
<ul style="list-style-type: none"> Finetuned Distil-BERT model for next word prediction on a custom dataset, resulting in a Perplexity Score of 8.25 	
Customer Churn Prediction	Jun '22 - Jul '22
<ul style="list-style-type: none"> Engineered an XGBoost PCA model achieving 82.7% accuracy for churn prediction at a Telecom Company Utilized a Logistic Regression Model with comprehensive feature selection to identify 6 critical factors influencing customer churn 	

Internships	
Data Scientist	Jan '23 - May '23
Zummit Infolabs	Bengaluru, IN
<ul style="list-style-type: none"> Engineered the optimization of Distil-BERT model on a dataset of over 100k empathetic dialogues, culminating in the development of an empathy-driven next word suggestion model Led and supervised a team of 30 interns, successfully delivering 10 deployable ML models tailored for specific use cases 	
Data Science and Business Analytics Intern	Oct '22 - Oct '22
The Sparks Foundation	
<ul style="list-style-type: none"> Constructed a hybrid CNN-GRU Stock Price Prediction model using custom neural networks by analysing 2 different datasets Created a Dashboard and a Storyboard for effectively visualizing Global Terrorism Data using Tableau 	
Data Science Intern	Jun '22 - Aug '22
Corizo.in	
<ul style="list-style-type: none"> Performed detailed exploratory data analysis and applied feature engineering to 1000+ rows of data to gain actionable insights Made an End-to-End Wine Quality Prediction model and deployed it using Flask API with 88% accuracy 	

Education	
Masters, Data Science	Oct '21 - Jun '23
Liverpool John Moores University, UK	
Post Graduate Diploma in Data Science	Jun '20 - Jul '21
IIIT Bangalore	
Bachelor of Technology (BTech)	Aug '12 - Jul '16
J.B. Institute of Engineering and Technology	Hyderabad, IN

Certifications	
<ul style="list-style-type: none"> HackerRank SQL Intermediate Certification Generative AI Engineering with Databricks 	