SANDEEP KOKKULA

San Francisco, CA

+1 (346) 586 9807 | sandeepkok5443@gmail.com | [Linkedin](#)

OVERVIEW

Experienced Data Science Leader with over **10+ years** of industry experience in high-performance, worldwide scale end-to-end solutions for challenging business issues utilizing AI, ML, NLP, Gen AI, LLM and Computer Vision, which made significant revenue generation in the respective industries. Worked with diverse tools including Python, R, Azure, AWS, and IBM Cloud in multiple verticals like retail, healthcare, marketing, supply chain sectors, airline, banking, and consumer goods.

EXPERIENCE

Research, Gen AI, University of Houston   January 2024 - Present
- Pioneered the advancement of a RAG-based Recommendation system, resulting in a **15% increase in sales** within targeted segments, utilizing LangChain, Milvus, and Llama2 (fine-tuned on Lit-GPT).
- Developed a GenAI-powered NLP summarization tool capable of distilling the essence of articles and conducting comparative analyses across multiple pieces. This tool currently processes approximately **1000** articles weekly, saving users an average of **50 hours per week**. It is developed using LangChain and the GPT-4 (32k) model.

Data Science Team Lead / Manager, Accenture   June 2021 – August 2023
- Spearheaded the entire implementation of a state-of-the-art store tool for a leading FMCG client, completely transforming in-store shopping experiences for multiple outlet chains. This initiative led to a remarkable **30% increase in sales** and an impressive **50% decrease in manual workload**.
- Executed retail price optimization project leveraging regression techniques to analyze competitor prices, customer segmentation, and temporal sales data, enhancing pricing strategies and maximizing profits.
- Established and managed a high-performing team of over **10** experts, including data scientists, machine learning engineers, software engineers, and visualization specialists. This collaborative effort led to the successful delivery of exceptional results across various industries by leveraging cutting-edge AI solutions.
- Build and maintain strong customer relationships, serving as a trusted advisor to customers, resulting in significant savings of millions through AI-enabled solutions.

Lead Data Scientist, IBM   January 2018 - June 2021
- Developed a cutting-edge predictive analytics model to forecast the impact of external disruptions, such as COVID-19, on supply chain operations, resulting in a **20% reduction in potential losses** and an optimization of the end-to-end manufacturing process.
- Utilized advanced machine learning algorithms to accurately predict the time to fault for a top car manufacturing company, **improving maintenance scheduling efficiency** by **30**% and **reducing downtime** by **15**%.
- Implemented AI-driven automation solutions for handling **30** million customer calls, resulting in a **50**% reduction in response time and a **35**% **increase in overall customer support efficiency**.
- Led business analytics initiatives for senior executives, providing data-driven insights and actionable intelligence to guide strategic decision-making, consulted with clients, providing expert guidance and tailoring AI solutions to meet unique business objectives.

Senior Data Scientist, Sandvine Technologies   June 2013 - January 2018
- Devised and implemented an advanced fraud information system utilizing predictive analytics to identify and recommend strategies for reducing fraudulent activities, resulting in a **15**% **decrease in fraud-related revenue losses**.
- Analyzed database investigations alongside customers to uncover a substantial revenue loss attributed to fraudulent users, prompting immediate action plans that profited the company.
- Implemented innovative clustering methods using Python Scikit-learn to **enhance the classification accuracy of authentication protocols** by **20**%, leading to improved fraud detection capabilities and customer security.

EDUCATION

Master's in Data Science   May 2024
University of Houston, Houston, Texas
- Coursework: Machine Learning, Data Science, Deep Learning, Gen AI, LLM, Big Data, Probability and discrete Mathematics
- CGPA: 4.0/4.0, Dean's Scholarship

Bachelor's in Computer Science   May 2013
Keshav Memorial Institute of Technology, India
Coursework: Artificial Intelligence and soft computing, Data Structures and Algorithms, and Computer Networks.

SKILLS

Cloud: Amazon AWS, Microsoft Azure, IBM Cloud and GCP
Programming Language: Python, Java, Spark, R, C, C++, MATLAB
Machine Learning: Linear and Logistic Regression, SVM, KNN, Decision Trees, Random Forest, Forecasting, PCA, Predictive models and Clustering
Deep Learning & NLP : Gen AI, LLM, Lang Chain, Chain of Thought (CoT), FLAN T5, CNN, Faster R-CNN, YOLO, SSD, TensorFlow, Keras, Image processing, Pytorch, Neural network, Syntactic Parsing, Entity Parsing, Statistical features, Sentiment Analysis, Recommendation systems, Search ranking, Information retrieval.
Data Science & Miscellaneous Technologies: ETL, Data Science pipeline, CI/CD Pipeline, MLOPs, Statistics, Time series, Data analytics, A/B testing, Tableau, Power BI, Fine-tuning, vector embedding, NLP, neural network optimization, Dockers, Kubenetes
Deployment Platforms: AWS Bedrock, AWS (EC2, Lambda), Azure Functions, Hugging Face Spaces
Big Data & Databases: MySQL, NoSQL, Oracle SQL, Postgres, Hadoop Ecosystem, IBM DB2, MapReduce, Snowflake, Advanced Data Mining, Apache Spark, Excel.
Web Technologies: HTML, CSS, RESTful, AWS S3 storage, and Sage maker instances.
Tools & IDE: Jupyter Notebook, PyCharm, Eclipse, Git, Weka, IBM Bluemix.

PUBLICATIONS

Pre-Processing Glaucomatous Eye and its Classification using MATLAB and Python https://t.ly/ubaAX
Dec'15
In International Conference on Computational Methods in Engineering and Health Sciences (ICCMEH- 2015), Malaysia.
- Our segmentation algorithm was developed using a fast, hybrid-level set model, which will provide a more accurate delineation of the disk boundary than the Hough transform.
- More Accurate extraction of the optic disk and optic cup by using a modified least ellipse square model.
- More than **97**% of Glaucoma classification accuracy using SVM classifier.

AWARDS

- Awarded "Super Hero of the Year" across IBM Worldwide
- Won Compass Intelligence IoT Innovator Award in "IoT Innovator: Driverless Vehicles"
- Manager choice award for every quarter for 3 and a half years

CERTIFICATIONS

- Generative AI Professional certification, Oracle Cloud Infrastructure - 2024
- Generative AI with Large Language Models from DeepLearning.AI, Amazon Web Services - 2024
- Microsoft Azure AI Data Scientist Certified, Amazon AWS Data Scientist Certified, and IBM Data Science Professional Certified

EXTRA-CURRICULAR

- MIT (Master's): Played a crucial role in organizing and participating in TechTatva'15 at college, a four-day national-level technical fest.
- JNTU (Bachelor's): Won the trophy of the SR Champions cricket tournament in 2012 by representing KMIT and won the inter-college chess championship in 2013.