

SHRIPRASAD KHARPE

Postgraduate in Big Data Analytics

91-7057069222 shrikharpe22@gmail.com LinkedIn GitHub Hinjewadi, Pune-411057

Profile

Postgraduate in Big Data Analytics with hands-on experience in **Python, SQL** and **Spark**, specializing in data preprocessing, **machine learning**, **generative ai** model development, and dashboard creation using **Power BI**. Proficient in **AWS** and other analytics tools, with a solid grasp of handling structured and semi-structured data formats such as CSV and JSON.

Education

Centre of development for advance computing	Mumbai-410210
Postgraduate in Big Data Analytics	08/2024 - 02/2025
SPPU University, Sinhgad College of Engineering	Pune-411041
Bachelor of engineering	07/2020 - 07/2024

Professional Training Experience

Centre for Development of Advanced Computing (CDAC)	Mumbai
Data Science Trainee	Aug 2024 – Feb 2025

CDAC is the premier R&D institution under MeitY, Govt. of India, specializing in advanced computing and emerging technologies.

- Completed hands-on training in Big Data Analytics with focus on machine learning, data engineering, and cloud deployment.
- Built predictive models using Python, TensorFlow, and PyTorch for healthcare and financial use cases.
- Processed large datasets using Apache Spark, Hive, and Hadoop; performed EDA and feature engineering.
- Developed dashboards with Power BI and Tableau to visualize insights.
- Deployed end-to-end solutions on AWS and Azure platforms.

Projects

Brain Tumor Detection and Classification System	01/2025 - 02/2025
---	-------------------

- Developed a CNN-based tumor detection model using **6,000+ MRI images**, achieving **67% accuracy** to support faster diagnosis and assist doctors in decision-making.
- Built a responsive web interface using **Flask**, **HTML**, and **CSS** for MRI scan input and result display
- Integrated **OpenAI** for generating insights like symptoms, possible causes, and recommended actions based on predictions
- Deployed on **AWS** to enable scalable, real-time access.

Ollama-Powered Chatbot Platform (FastAPI & Streamlit)	05/2025 - 05/2025
---	-------------------

A full-stack chatbot platform for generative AI using modern web technologies

- Developed a full-stack **Generative AI** chatbot using **Streamlit** (frontend) and **FastAPI** (backend) for conversational AI with local **LLMs** (Mistral via Ollama).
- Engineered a RESTful API with FastAPI and LangServe to expose **LangChain** runnables, facilitating modular and scalable LLM inference
- Implemented **robust client-server communication** between Streamlit and FastAPI via **requests**, including comprehensive error handling

SBIN Stock Data Analysis	12/2024 - 01/2025
--------------------------	-------------------

- Analyzed SBIN stock trends to identify short-term trading opportunities using **Spark RDD**, **Python (Pandas, NumPy, Matplotlib)**, and Streamlit
- Assessed market sentiment and highlighted trade signals, providing actionable insights via GitHub
- The solution addresses the need for **real-time** stock analysis, helping traders capitalize on market movements and improve trading strategies

Programming Languages: Python, SQL

Machine Learning & Deep Learning: Scikit-learn, TensorFlow, PyTorch, Keras, Convolutional Neural Networks (CNNs)

Natural Language Processing (NLP) & LLMs: Text Preprocessing, Tokenization, Transformer Models

Big Data Technologies: Hadoop, Apache Spark, Apache Hive, Big Data Analytics

Cloud Platforms & Services: AWS (EC2, S3, IAM, Lambda, Redshift, SageMaker, VPC), Microsoft Azure

Data Visualization & BI Tools: Tableau, Microsoft Power BI, Matplotlib, Seaborn

Databases: MySQL, MongoDB

Data Science & Statistical Analysis: Hypothesis Testing, Regression Analysis, Statistical Modeling

Data Manipulation & Scientific Computing: NumPy, Pandas

Operating Systems: Linux