

LALIT

Data Scientist | AI/ML Engineer

3.6 Years Experience

SUMMARY:

Results-driven **AI/ML Engineer** with hands-on expertise in Python, Machine Learning, Deep Learning, and NLP. Proven experience in building AI solutions, developing intelligent chatbots, and delivering data-driven insights. Skilled in using TensorFlow, Keras, OpenCV, and end-to-end project development. Strong background in data extraction, model development, and visualization.

SKILLS:

Languages/Frameworks: Python, TensorFlow, Keras, OpenCV, Beautiful Soup, Selenium, Requests, Langchain, CrewAI

Machine Learning: Regression, Classification, Deep Learning (CNN, RNN, LSTM), NLP (NLTK, Text Blob), Generative AI (OpenAI, Llama, Prompt Engineering), Large Language Models (LLM)

Databases: MySQL, MongoDB

Tools: Power BI, ETL Pipelines, Tesseract OCR

Others: Data Preprocessing, Web Scraping, Data Wrangling

PROFESSIONAL EXPERIENCE

Confidential

Software Engineer - Data Scientist | AI/ML Engineer

Aug 2022 – Present

- Engineered machine learning and deep learning models leveraging Keras, TensorFlow, and sophisticated Natural Language Processing techniques, resulting in a 25% increase in accuracy and a 30% reduction in model training time. These advancements streamlined data-driven decision-making processes and enhanced overall operational efficiency.
- Applied text-OCR techniques with Tesseract to extract valuable data from PDF files, employing advanced Natural Language Processing (NLP) methodologies. Successfully structured and stored the extracted data into Excel files, optimizing accessibility and facilitating streamlined data analysis processes.
- Utilized deep learning methodologies for intent classification in chatbot development, incorporating advanced Natural Language Processing (NLP) techniques. Integrated Large Language Models (LLM) to enhance conversational capabilities, resulting in a 20% improvement in intent recognition accuracy and a 15% reduction in response time, ultimately delivering a more efficient and intuitive user experience.
- Orchestrated the identification and extraction of valuable data sources through Selenium and Beautiful Soup resulting in a 30% increase in data collection efficiency; converted unstructured data into a structured format and achieved a 25% reduction in data processing time, facilitating the creation of insightful visualizations that unveiled critical industry patterns and trends.
- Built autonomous agent workflows using Crew AI to coordinate multiple role-specific LLM agents for complex task automation and decision-making.

Confidential

Data Scientist | AI/ML Engineer

May 2022 - Aug 2022

- Extracted data from diverse sources and synthesized datasets to analyze relationships between countries and predict forthcoming drought conditions. By leveraging advanced data querying techniques, our team identified key indicators and trends, enhancing preparedness efforts and enabling proactive mitigation strategies.
- Conducted extensive analysis of statistical data, importing and integrating datasets from over 15 sources. Utilized Python for meticulous data cleaning, preprocessing, and wrangling, resulting in a 30% reduction in data processing time. Developed insightful visualizations using Power BI, facilitating the assessment of past and present drought conditions.
- Reduced the development time by 50% by creating a scalable modular template for writing ETL pipelines in Python.
- Developed 10+ data crawlers to source macro-economic, financial, geopolitical, and demo graphical time-series data from public domain websites using APIs, Beautiful Soup, and web automation tool, Selenium, and expanding from 100K to 20M+ records.

Confidential

Data Scientist | AI/ML Engineer

Sept '21 - Mar '22

- Utilized Beautiful Soup and Selenium to scrape data from multiple airline websites, enabling the prediction of airline fares through supervised machine learning regression algorithms. Developed a robust model achieving an impressive accuracy of 89%, empowering stakeholders with precise fare predictions for informed decision-making and strategic planning.
- Extracted data from numerous second-hand car websites employing Beautiful Soup, Requests, and Selenium. Utilized supervised machine learning regression algorithms to predict the prices of used cars, achieving an accuracy of 85%. This predictive model provides valuable insights for buyers and sellers, aiding in informed decision-making within the used car market.
- Engineered web crawlers leveraging Selenium and Beautiful Soup to extract unstructured data from diverse sources including Amazon, Flipkart, Wikipedia, ICC, and more. Transformed the scraped data into structured formats and created multiple data frames for comprehensive analysis, enabling deeper insights and informed decision-making.

Education

- Bachelor of Technology from JSS Academy of Technical Education, UPTU, Noida in 2012.

Projects

Organization Chatbot Development Project:

Successfully developed a sophisticated chatbot system comprising rule-based and AI components. The rule-based bot initiates interactions, gathering personal information to generate leads, resulting in a 25% increase in lead generation efficiency. Upon lead creation, the chat seamlessly transitions to the AI bot, powered by deep learning sequential models implemented in Python. Integrated Natural Language Processing, Large Language Models (LLM), and OpenAI API to enhance the bot's performance and capabilities. The AI bot effectively identifies user intents based on dataset analysis, providing tailored responses and enhancing user engagement. This innovative approach improves customer interaction efficiency and drives organizational growth.

Neu Cup – A Urine Testing Application:

Successfully developed a sophisticated urine testing application by integrating advanced deep learning and image processing techniques. Utilized YOLO (You Only Look Once) for image segmentation and Open CV for object detection, significantly enhancing model accuracy from 30% to 89%. Leveraged Django to create a robust backend infrastructure, facilitating seamless integration and scalability. Applied data preprocessing and augmentation techniques to improve model training and performance. This comprehensive approach significantly improved diagnostic precision and reliability, resulting in a highly effective and accurate testing solution. Key technologies and methodologies used include Python, TensorFlow, Keras, machine learning, data analysis, and computer vision, ensuring the application meets industry standards for data-driven healthcare solutions.

Malignant Comment Classifier:

Developed an artificial intelligence model to classify malignant comments using Natural Language Processing (NLP) with NLTK library. Implemented decision tree classifier algorithm, resulting in a 20% improvement in classification accuracy. This enhanced model effectively identifies harmful comments, contributing to a safer online environment.

Sentiment Analysis using LSTM.

Employed Long Short-Term Memory (LSTM) in a Deep Learning model to analyze user sentiments regarding Amazon's Alexa. Integrated NLTK, Text Blob, and deep learning techniques to enhance sentiment analysis accuracy by 25%. This innovative approach provides valuable insights into user perceptions, aiding in product improvement and customer satisfaction strategies.

Certifications

- PG Program in Data Science, Machine Learning, and Neural Networks from Data Trained Education.
- Business Analytics with Tableau, Natural Language Processing with ML and Deep Natural Language Processing, Deep Learning from Data Trained Education.
- Microsoft Power BI Desktop for Business Intelligence from Udemy.
- The Complete MySQL Boot Camp 2022: Go from zero to hero from Udemy.