Shefali Singh

LinkedIn • Github

PROFESSIONAL SUMMARY

Data-driven professional with 2+ years of experience in QA and data analytics. Skilled in Python, SQL, Excel, and machine learning. Proven track record of enhancing testing efficiency by over 30% and delivering predictive models using ML and deep learning to solve real-world challenges. Adept at automating workflows and building data pipelines to improve operational performance.

EDUCATION

Master of Engineering (Computer Science) 2022 - 2024
Thapar University Grade: 9.07/10.0

Bachelor of Engineering (Computer Science) 2016 - 2020
Chitkara University Grade: 9.04/10.0

Intermediate (Class XII) 2015 - 2016
Nav Yug Higher Secondary School Grade: 84.6%

Matriculation (Class X) 2013 - 2014
Nav Yug Higher Secondary School Grade: 94.0%

EXPERIENCE

QA analyst September 2019 - July 2022
ORANGE BUSINESS SERVICES *Gurugram, Haryana*

- Automated testing pipelines across 10+ GUI and API applications, resulting in a 40% faster test cycle.
- Reduced manual effort by 30% through Python automation scripts for repetitive regression task execution.
- Utilized JIRA, Selenium, and API tools to streamline testing, improving issue resolution rates by 25%.

INTERNSHIPS

Data analyst / Machine Learning Engineer September 2023 - June 2024
LG SOFT INDIA *Greater Noida , UP*

- Leveraged ML models to forecast AC energy consumption, boosting efficiency by 18% through strategic optimization recommendations; evaluated model performance using precision, recall, and F1-score for accuracy
- Used Tesseract OCR for automated Korean text extraction and translation from design diagrams, aiming to reduce design review time by 25%.
- Executed NLP-driven market analysis by scraping 15K+ records, performing feature extraction and sentiment classification, influencing 3 product roadmap decisions

PROJECTS

Math & Reasoning Assistant: May 2025

- Engineered a Streamlit application leveraging Gemma-2 and LangChain to serve as an interactive Math & Reasoning Assistant.
- Designed and implemented tools within LangChain for real-time problem-solving capabilities in mathematics and reasoning.
- Achieved a high accuracy rate of 95% in solving diverse math problems, demonstrating strong computational and logical reasoning prowess.

CV Parsing & Matching May 2025

- Engineered a CV parsing and matching pipeline leveraging BERT and GPT models to accurately extract skills from resumes.
- Implemented a similarity scoring mechanism to effectively match extracted skills with job descriptions.
- Achieved high performance metrics: 85% precision and 88% recall in matching candidates to relevant job roles.

• **Retrieval-Augmented Generation (RAG):** March 2025

- Developed a Streamlit-based Retrieval-Augmented Generation (RAG) application for efficient Q&A over 50+ PDF documents.
- Integrated LangChain, FAISS, and Llama3 to build a robust RAG pipeline, enabling intelligent document querying.
- Achieved over 90% accuracy in Q&A performance, demonstrating highly effective information retrieval and generation.

Chinook Music Sales Analysis February 2025

- Analyzed over 9,000 music sales records to extract key business insights.
- Utilized SQL to identify top-performing genres and inform strategic marketing decisions.
- Drove a 15% reduction in customer churn and enhanced customer targeting through data-driven recommendations.

Zomato Restaurants Analysis (Github) (Demo) October 2024

- Conducted a comprehensive analysis of Zomato restaurant data using Excel, identifying key trends and patterns in customer ratings, cuisine preferences, and geographical distribution.
- Utilized advanced Excel functions and pivot tables to evaluate restaurant performance, highlighting top-performing eateries and areas with potential for growth based on customer reviews and sales data.
- Leveraged data visualization tools in Excel to create interactive dashboards, enabling stakeholders to easily track and compare restaurant metrics, customer satisfaction levels, and market trends in real-time.
- Tools used: Excel, PowerPoint, google spreadsheets.

Optimizing Energy Consumption of AC through Machine Learning February 2024

- The project utilized machine learning to optimize energy efficiency in LG's AC product line by analyzing user behavior data and environmental variables.
- It aimed to develop predictive models for dynamic AC settings, enhancing energy efficiency and user comfort.
- Tools used: Python, TensorFlow,scikit-learn, Matplotlib,Jupyter notebook

Tesseract OCR-based computer vision system to translate Korean text on PCB and microprocessor designs into English. January 2024

- Developed a Tesseract OCR-powered computer vision system to extract and translate Korean text from PCB and microprocessor designs into English.
- Streamlined design interpretation and accelerated development workflows through automated text recognition and translation.
- Tools used: Python, OpenCV, Tesseract OCR, Google Translate API.

Web scraping, Sentiment analysis, and Feature engineering September 2023

- Strengthened market positioning by conducting in-depth competitive analysis using web scraping to gather actionable data.
- Applied sentiment analysis and feature engineering to derive insights, driving strategic decision-making.
- Tools used: Python, Beautiful Soup, Selenium, NLTK, TextBlob, spaCy.

CERTIFICATIONS

Professional Certificate Course in Data Science & AI Newton School September 2024

- Currently pursuing a Professional Certificate in Data Science & AI, focusing on data analysis, machine learning, artificial intelligence, and statistical modeling.
- Developing expertise in advanced data techniques to extract insights and address complex business challenges.

Natural Language Processing (NLP) in Python with 8 Projects UDEMY (Link) June 2024

- Certified in Natural Language Processing, specializing in text classification, summarization, sentiment analysis, and tokenization.
- Gained practical experience in word embeddings, text generation, and data visualization.

The Complete Python Bootcamp From Zero to Hero in Python Udemy (Link) April 2024

- Certified in Python, mastering core concepts like data structures, OOP, and essential skills like file handling, debugging, and unit testing.
- Developed strong command-line skills and a robust understanding of fundamental Python programming principles.

Machine Learning A-Z: AI, Python & R + ChatGPT Prize [2024] UDEMY (Link) March 2024

- Certified in "Master Machine Learning on Python & R."
- Expertise in ML model development, predictive analytics, RL, NLP, and dimensionality reduction.

SKILLS

Computer Languages: Machine Learning, Python, SQL, C++
Data Tools: TensorFlow, NumPy, Power BI
Software Packages: Virtualization, Pandas, Matplotlib, Linux, MySQL, Excel, OpenCV
Additional Courses: Data Structure
Soft Skills: Verbal/nonverbal communication, Written communication, Leadership, Time management, Decision-making, Research, Teamwork
Others: AWS, Jira, Microsoft Office, Spreadsheet, LLM, Generative AI, Tableau