

Harshith Sakala Santhosh | Data Analyst

E: sakala.harshith@gmail.com | M: +919100833723 | [LinkedIn](#) | A: Hindupur, India

SUMMARY

Awarded distinction in M.Sc. Applied Data Science and B. Tech Computer Science, specializing in innovative data and web solutions. Demonstrated success in boosting drug treatment outcomes and driving user engagement through effective website development. Skilled in Python, predictive modeling, and advanced data analysis.

SKILLS

- **Programming Languages:** Python, R
- **Predictive Modeling:** Linear Regression, Logistic Regression, SVM, Random Forests and XG Boost.
- **Data Visualization:** Plotly, Matplotlib and Power-BI.
- **SQL:** MySQL and SQL Server(T-SQL).
- **Web Technologies:** HTML, CSS, Bootstrap, Flask, Django, AWS Elastic Beanstalk, Code pipelines and S3.

EDUCATION

M.Sc. Applied Data Science | University of Central Lancashire | 2021-2022 | Distinction

- **Modules:** Artificial Intelligence and Machine Learning (75%), Visual Information Processing (70%) and M.Sc. Project (81%).
- **Dissertation:** Leaf Disease Detection and Segmentation Through Transfer Learning.
- **Achievement:** Best Data Science Student of the Year 2022

B. Tech Computer Science and Engineering | Lovely Professional University | 2017-2021 | Distinction

- **Modules:** Python Programming (85%), DBMS (88%), Java Programming (73%) and Capstone Project (95%).
- **Dissertation:** Hospital Cypher – [A Smart Surveillance System.](#)

WORK EXPERIENCE

Freelance Web Developer | Krish Chem | July 2023 – Oct 2023 | Remote, India

- Conducted comprehensive requirement analysis by visiting the company's office and conducting in-depth interviews with the company's manager, resulting in a thorough understanding of the client's needs and preferences.
- Developed and launched a responsive fertilizer's company website using HTML, CSS, Bootstrap, and Flask, resulting in a visually appealing and user-friendly interface resulting in 500+ users accessing in the first month of deployment.
- **Skills and Technologies Used:** HTML, CSS, Bootstrap, Flask (Python), Contentful (CMS) and AWS (EC2 and route 53).

Data Science Intern | Data Glacier | Feb 2023 – Jun 2023 | Remote, UK

- Collaborated and led a team of 3 data science interns to analyze patient data and identify key factors impacting drug persistency, leading to the creation of an accurate prediction model resulting in a 15% increase in treatment efficiency.
- Conducted data manipulation, exploratory data analysis, and developing statistical modeling methods to optimize drug treatment efficiency for diabetes and tuberculosis. Reduced dimensionality of dataset by employing feature reduction techniques, resulting in selection of top 20% features, and providing better data insights.
- Implemented predictive machine learning models like SVM, Deep Learning etc. resulted in 10% increase in model prediction efficiency compared to previous models, leading to 15% better efficiency in drug treatment.

- **Skills and Technologies Used:** Python, Chi-Squared Tests, T-tests, SVM, Random Forests, ANN's and Plotly.

PROJECT EXPERIENCE

[Portfolio Website](#) | Personal Project | Dec 2024- Jan 2025

- Crafted a responsive portfolio website using HTML, CSS, Bootstrap, JavaScript, Django, and SQLite, designed for modern scalability and seamless user interaction.
- Enhanced user interactions with an AI-powered chatbot, utilizing the RAG AI technology with Lang flow and Astra Vector Database for nuanced, intelligent responses that adapt to user queries.
- Elevated content delivery through advanced AI, offering a richer, more engaging user experience by seamlessly integrating retrieval and generative capabilities, reflecting a cutting-edge approach to user engagement.
- **Optimized performance with sophisticated data handling**, employing SQLite to ensure rapid response times and a smooth interface, supporting the complex demands of innovative AI features.
- **Skills and Technologies Used:** HTML, CSS, Bootstrap, Django (Python), SQLite, Lang flow, Astra Database (Vector Database).

[Leaf Disease Detection and Segmentation Through Transfer Learning](#) | University of Central Lancashire | Jul 2022- Sept 2022

- Developed and led a team of two in implementing state-of-the-art deep learning models, including Mask RCNN for disease instance segmentation and Efficient Net B1 and Inception Net V3 for leaf disease detection. This resulted in accurate disease spot detection and achieved a 71% Mean Average Precision (MAP) by Mask-RCNN on a diverse dataset.
- Achieved impressive results with an overall increase in performance by 40% compared to classification models used in previous similar research projects. Conducted data manipulation, exploratory data analysis, and developing statistical modeling methods to optimize drug treatment efficiency for diabetes and tuberculosis. Reduced dimensionality of dataset by employing feature reduction techniques, resulting in selection of top 20% features, and providing better data insights.
- **Skills and Technologies Used:** NumPy, Pandas, Mask RCNN, Efficient Net B1, Inception Net V3, TensorFlow, Matplotlib, Scikit-Learn, Python and Google Colab.

[Hotel Cancellation Analysis](#) | University of Central Lancashire | Feb 2022- Mar 2022

- Led a team in Providing data-driven insights to enhance Average Daily Rate (ADR) and reduce cancellation rates for hotels and resorts.
- Collected and preprocessed a comprehensive hotel dataset to ensure data quality and accuracy for analysis. Analyzed various features such as booking lead time, customer demographics, to understand their impact on cancellation rates and identified significant correlations and patterns influencing cancellation behavior.
- Formulated data-driven suggestions to reduce cancellation rates and enhance hotel revenue.
- **Skills and Technologies Used:** Python, Pandas, Plotly.

CERTIFICATIONS AND ARTICLES

- | | |
|--|----------------------|
| ● IBM Data Science Professional Certification | Sept 2023 - Feb 2024 |
| ● Heart of Statistics and Powerhouse of Data Science: Data Distributions | Jun 2024 - Jun 2024 |
| ● 30 Days of Learning AI/ML through LinkedIn articles | Mar 2024 – Present |

References Available on Request