

Pranava Kailash Subramaniam Prema

Data Scientist

Results-driven Data Scientist with expertise in data analysis, predictive modeling, and machine learning, with strong experience in Python, SQL, and cloud platforms like AWS. I leverage data to uncover actionable insights, with a history of automating processes and improving operational workflows. Passionate about applying statistical models and AI-powered solutions to solve complex business challenges.

✉ pranavakailashsp@gmail.com

📍 Guildford, United Kingdom

🌐 linkedin.com/in/pranava-kailash

📞 07909807017

📄 pranava-kailash.github.io/

🐙 github.com/Pranava-Kailash

SKILLS

Data Analysis and Management

Data Analysis , Data Modelling and Data Governance

Collaboration Tools

Git , Github and Git action

Machine Learning

Regression, Clustering, Classification and Natural Language Processing (NLP)

Programming Languages

Python (including libraries like Pandas and NumPy) and SQL

Data Science Tools

Large Language Model (LLM), TensorFlow, Matplotlib and Power BI

Development and Methodologies

API Development, Agile Methodologies, Cloud Computing (AWS), Selenium, BeautifulSoup and RAG (Retrieval-Augmented Generation)

EDUCATION

M.Sc Data Science

Univeristy Of Surrey

02/2024 - 02/2025

Surrey, United Kingdom

- **CyNER2.0 - Cybersecurity Named Entity Recognizer (NER)**
Developed a domain-specific NER model using transformer-based models (BERT, DeBERTa, DarkBERT) to identify critical cybersecurity entities, including malware, vulnerabilities, and threat actors

B.E Computer Science and Engineering

SNS College of Technology

05/2019 - 05/2023

India

- **Fingerprint Voting System** - Developed a secure and efficient voting system using an Arduino Uno board and fingerprint recognition technology, ensuring voter authentication and preventing fraudulent activities

WORK EXPERIENCE

Data Researcher / Engineer

Fleet Street Research UK - Freelance

10/2022 - 01/2024

India

- Automated data extraction pipelines with Python-based scripting, increasing data acquisition efficiency by 50%.
- Enhanced database workflows by optimizing Python libraries, leading to 10% operational improvement.
- Smoothened the process of image importation into database, hence enhancing data quality
- Migrated data pipelines between systems like MySQL and SQLite , ensuring zero downtime or data loss

INTERNSHIPS

Data Analytics Intern - Yoshops.com - India

(02/2022 - 04/2022)

- Automated data collection through web scraping tools, increasing real-time pricing accuracy by 40% and improving product performance insights.

Data Science (NLP) Intern - Forsk Coding School - India

(03/2021 - 05/2021)

- Conducted in-depth Sentiment analysis on customer reviews using Natural Language Processing (NLP), finding key areas for product improvement. Applied sentiment analysis to understand customer behavior and improve decision-making

PERSONAL PROJECTS

PDF ChatBot (03/2025 - 03/2025)

- Utilized OpenAI and LangChain to extract text, generate embeddings, and create QA chains, demonstrating expertise in natural language processing and AI-driven data analysis
- Integrated FAISS (Facebook AI Similarity Search) for efficient vector storage and similarity search, improving the accuracy and speed of question answering based on document content
- Created a Streamlit application that enables users to interactively query PDF documents, showcasing skills in front-end development and user interface design

Job Search Assistant using LangChain (03/2025 - 03/2025)

- Developed an automated job scraping tool that efficiently retrieves listings from LinkedIn, enhancing job market insights and improving the job search experience
- Utilized LangChain and OpenAI to analyze job descriptions, providing accurate insights on required skills and keywords, resulting in better candidate preparedness and CV optimization
- Created a user-friendly interface with Streamlit, enabling users to easily navigate and access detailed job profiles and insights