

Sanjay Sathish Kumar

📍 Dublin, Ireland 📞 0899751053 ✉ sanjay.sathish0604@gmail.com [in LinkedIn](#) [GitHub](#)

Summary

Results-driven computer science graduate with a deep passion for interesting futuristic technologies such as data science and machine learning and industrial expertise in automation testing. Proficient in Python, SQL, REST API, and Robot Framework. Proven ability to streamline development, testing processes, and deliver reliable, scalable software solutions in Agile environments.

Technical Skills

Programming Languages: Python, C, SQL, JavaScript, HTML, CSS

Data Analytics Tools: Tableau, Pandas, PySpark, Zoho Analytics

Cloud & DevOps: AWS, VMware, Terraform, Docker, Jenkins, CI/CD

Software Development & Testing: Robot Framework, Git, Gerrit, REST API, Postman, Jira

Operating Systems: Linux

Core Concepts: Data Structures & Algorithms, Machine Learning, Data Analysis, Web Development, Cloud Computing

Experience

DISYS India Private Limited

Aug 2021 – Nov 2021

Developer

Chennai, India

- Collected and managed COVID-19 vaccination data for 1,000 employees across multiple Indian states using SQL queries.
- Created Zoho Analytics dashboards visualising vaccination trends across 6 states where company branches are located, enabling HR to make return-to-office decisions.
- Analysed existing SQL queries to identify opportunities to reduce query response time by 10 seconds.
- Analysed company's employee data to create reports with employee performance matrix and derived insights on the company's recruitment strategy.
- Used SQL to extract and analyse candidate application and onboarding data, helping HR identify delays and optimise the recruitment process.

Gigamon

Apr 2022 – Mar 2023

Software Engineer

Chennai, India

- Resolved a legacy defect by removing the 256-character limit on name aliases across all features, and implemented automated tests to replace manual configuration, improving testing efficiency and consistency.
- Led training sessions for 20+ interns on automation testing with Robot Framework.
- Performed API testing and automated API test cases using Postman and Test Automation Framework, ensuring feature reliability and reducing manual efforts by 30%.
- Optimised test execution by implementing cache memory, reducing test run times by 5 minutes per execution.
- Utilised Gerrit for code review and version control, ensuring code quality, tracking changes, and streamlining collaboration in automation script development.

Projects

Diabetes and Heart Disease Prediction - Python, Random Forest, Gradient Boost, Adaptive Boost, SMOTE

- Developed machine learning models to predict diabetes and heart disease, achieving AUC-ROC scores of 81.4% and 81.3%, enhancing early risk detection for 200,000+ patients.
- Used various data augmentation techniques, with SMOTE proving the most effective, enhancing model performance by correcting class imbalance and boosting recall and F1 scores by 73% for diabetes and 66% for heart disease.
- Conducted age-specific and demographic-based analysis, enhancing model reliability across diverse populations.
- Applied feature selection techniques (ANOVA F-test, SHAP values, and correlation analysis), leading to a 7% improvement in AUC-ROC using the F-test.
- Optimised ensemble models such as AdaBoost, Random Forest, and Gradient Boosting improving model accuracy by 15%.

Budget-Based Restaurant Finder Web Application - AWS, HTML, Flask, PySpark, Pandas

- Designed and deployed a restaurant finder web application on AWS, utilising S3 for data storage and EC2 for hosting, handling large-scale datasets for restaurant recommendations.
- Built a responsive and interactive webpage using HTML, CSS, and JavaScript, integrating it with Flask APIs for seamless data retrieval and dynamic restaurant recommendations.
- Processed a Swiggy dataset containing over 80,000 entries, applying PySpark to clean data by standardising currency values, removing duplicates, and handling missing entries.
- Developed an interactive filtering system for 1,000+ restaurants, enabling users to search by city, budget, and food preferences, improving query efficiency by 40%.
- Optimised application performance by deploying backend services on AWS EC2 instances, ensuring fast response times and smooth data retrieval.

Education

Dublin City University <i>M.Sc. in Computer Science (Data Analytics)</i>	Aug 2023 – Sep 2024 Dublin, Ireland
Anna University, R.M.D. Engineering College <i>B.E. in Computer Science (GPA: 7.6/10)</i>	Sep 2017 - Jun 2021 Chennai, India

Achievements

- "The Data Analyst Course Complete Data Analyst Bootcamp 2021" from Udemy.
- "Introduction to HTML5, CSS3, and JavaScript" from Coursera.
- "SQL Applied in Data Science" from edX.
- "Python Basics for Data Science" from edX.