

# Preethi Jayakumar

+44 7425668496 • Nottingham NG72GJ, United Kingdom • preethijay2198@gmail.com • [LinkedIn](#).

Data Analyst with 3 years of experience specialising in SQL, Python, Spark, and machine learning. Proven track record with clients like AIG Insurance, advancing from Programmer Analyst to Associate Developer. Successfully led and managed major BPM applications, significantly improving operational efficiencies and client satisfaction.

## SKILLS

- SQL (MySQL, PostgreSQL)
- Python (pandas, Numpy, Matplotlib)
- Pyspark
- Tableau, Looker, Microsoft PowerBI
- Excel, Google sheets
- Microsoft azure (Databricks, Azure Data lake)
- Data Warehouse (Big Query)
- Docker (containerization)
- Git, Jenkins
- Project management Tools (Rally, Jira)

## PROJECTS

<b>Big Data Meets Oceanography: Scalable Forecasting of Sea Surface Temperature</b>	2024
• Utilised ERA5 reanalysis data and CDS API for real-time retrieval, processing with Apache pySpark.	
• Implemented Facebook's Prophet and Random Forest Regressor (Apache Spark) for SST forecasting, achieving MAE of 0.164 (Prophet) and 0.161 (RF Regressor).	
• Deployed all services in a Docker cluster, using Apache Airflow for automation and Apache Cassandra for distributed storage, with inferential data stored in PostgreSQL which was integrated to tableau for Dashboard Visualisation.	
<b>DengAI: Predicting Disease Spread</b>	2024
• Conducted EDA on environmental datasets, addressing missing values with forward filling and KNN imputation. Developed and evaluated machine learning models (SVR, RF, GLM) for dengue case prediction.	
• Used PCA for feature engineering and assessed models with MAE and R-square metrics, with negative binomial GLM achieving the best accuracy. Calculated predicted dengue cases using the optimised GLM.	
<b>PCR and RFS using ML Prediction</b>	2023
This project is for early cancer detection using Machine learning models	
• Handled missing values using KNN imputation, selected features with a correlation matrix, and normalised using Min-Max scaling.	
• Used Logistic Regression, Decision Tree, SVM, and RF models with features selected via ANOVA F-test.	
• SVM excelled in RFS prediction, while Random Forest was best for PCR prediction.	
<b>Rummikub Game Using Python</b>	2023
• Awarded a distinction for developing a Rummikub game by leveraging advanced Python programming skills.	
<b>Statistical Analysis on Kittiwake Birds Features</b>	2023
• Awarded a distinction for using in-depth statistical analysis like hypothesis test, linear and log models for predicting breeding pairs with confidence intervals.	

## WORK EXPERIENCE

### Cognizant Technology Solutions Ltd, India

<b>Developer / Associate</b>	Jan 2023 – July 2023
• Spearheaded the development and support of 2 BPM applications, leading a team of 5 and enhancing team efficiency by 30% through bug resolution and implementing project management tools like Rally.	
• Collaborated with a cross-functional team through Scrum meetings, consistently delivering high-quality software and data features bi-weekly.	
• Optimised reporting databases and schemas, improving report generation speed by 40%.	
• Automated management reports and integrated email and SQL connections, cutting manual effort by 40% and reducing reporting errors by 15%.	
• Performed code fixes and application updates using Git and Jenkins, ensuring seamless implementation of database or application changes.	

**Programmer Analyst**

Oct 2020 – Jan 2023

- Tested and migrated major applications, ensuring data integrity and achieving a seamless operational transition with zero downtime.
  - Resolved critical tickets involving user application data and software bugs, improving SLA compliance by 15% and maintaining high user satisfaction.
  - Streamlined data collection and manipulation processes with the Operations and Business Analytics teams, enhancing data quality and reducing processing time by 20%.
  - Designed comprehensive diagrams for use cases and workflows, reducing project timelines by 15% and improving project clarity.
  - Created and maintained documentation for Pega applications, ensuring high standards of code quality and usability.
- 

**EDUCATION****Master of Science in Computer Science (Artificial Intelligence)**

Sep 2023 – Present

University of Nottingham, United Kingdom

- Key Coursework: Applied Statistics and Probability, Machine Learning, Big Data Technologies, Data Science with Machine learning.
- Awarded distinction in Programming (Python), Big Data technologies and Data Science with ML modules.

**Bachelor of Technology in Information Technology**

Apr 2016 – May 2020

Rajalakshmi Engineering College, India

- Awarded Special Prize for project on "Anomaly Detection in Chest X-ray using AI".
  - Key Coursework: DBMS, computer networks, mobile computing, computer programming (C, python, Java)
- 

**Certifications:** Google Data Analytics Certified Senior System Architect (CSSA) , Certified System Architect (CSA).