# SUMITH JOHN

Parasuvaikal, Thiruvananthapuram, Kerala
sumithjohn1888@gmail.com | +91 9947788586 | www.linkedin.com/in/sumith-john-708122282 | https://github.com/Sumithjohn

## SUMMARY

Motivated and detail-oriented Junior Data Scientist with a strong foundation in machine learning, predictive modeling, and data analytics. Proficient in Python, SQL, Scikit-learn, and TensorFlow, with hands-on experience in data preprocessing, model optimization, and AI deployment. Skilled in applying statistical analysis and deep learning techniques to extract insights and develop data-driven solutions. Passionate about building scalable AI applications that enhance decision-making. Seeking an entry-level role to apply my technical skills, analytical thinking, and problem-solving abilities in a dynamic data-driven environment.

## TECHNICAL SKILLS

**Programming Languages:** Python, SQL
**Machine Learning & AI:** Scikit-learn, TensorFlow, Keras, XGBoost, NLP
**Techniques:** Deep Learning, Supervised Learning, Neural networks
**Data Analysis & Visualization:** Pandas, NumPy, Matplotlib, Seaborn
**Databases:** MySQL
**Tools & Frameworks:** Streamlit, Jupyter Notebook
**Soft Skills:** Problem Solving , Analytical Thinking, Adaptability, Collaboration

## INTERNSHIPS

**Data Science Intern | Srishti Innovation Computer System PVT.LTD. | July 2024 - Present**

- Analyzed and processed large datasets using Python, Pandas, and NumPy.
- Applied statistical methods and machine learning techniques to derive actionable insights.
- Developed data-driven models and visualizations using Scikit-Learn, Matplotlib, and Seaborn

**Machine Learning Intern | ACUTRO Technologies PVT.LTD. | June 2023 - July 2023**

- Developed and optimized machine learning models using Scikit-Learn, TensorFlow, and Pandas.
- Applied feature engineering, hyperparameter tuning, and model evaluation to enhance accuracy.
- Conducted exploratory data analysis (EDA) and preprocessing

## PROJECTS

**Heart Disease Prediction using Machine Learning.**

https://github.com/Sumithjohn/Heart-Disesase-Prediction.git

- Developed a high-accuracy (85.2%) predictive model for heart disease detection using KNN, Decision Tree, and Random Forest, applying hyperparameter tuning for optimization.
- Implemented advanced data preprocessing techniques, including feature scaling, missing value imputation, and outlier detection, to improve model accuracy and performance.
- Built and deployed an interactive Streamlit application for real-time heart disease prediction, enabling users to input data and receive immediate AI-driven feedback based on the trained model.

**Human Stress Detection using Machine Learning**

https://github.com/Sumithjohn/Human-Stress-Detection-using-Machine-Learning.git

- Developed a machine learning model for human stress level prediction using physiological data (heart rate, body temperature, respiration rate, etc.) with Logistic Regression, KNN, Random Forest, and XGBoost.
- Implemented advanced data preprocessing, including missing value imputation, feature selection, and correlation analysis, to enhance model accuracy and performance.
- Built and deployed an interactive Streamlit application for real-time stress prediction, allowing users to input data and receive instant AI-driven insights based on the trained model.

## EDUCATION

**B.E. Computer science and Engineering** | Noorul Islam Centre for Higher Education | 2020 - 2024

## LANGUAGES

- Malayalam
- English
- Tamil