



# Stebin George

DATA SCIENTIST/ANALYST

## Profile

Accomplished Data Scientist with a passion for delivering valuable data through analytical functions and data retrieval methods. Proficient in predictive modeling, data processing, data mining algorithms, and scripting languages, including Python and R. Committed to helping companies advance by helping them to develop strategic plans based on predictive modeling and findings.

## Education

### M.SC DATA SCIENCE, Christ University, Lavasa, Pune

AUGUST 2021 – JUNE 2023

- Percentage : 80.33
- Class Representative for the year 2021-22
- Published a Research paper in IEEE Titled "Predicting Employee Attrition Using Machine Learning Algorithms."

### B.SC PHYSICS, Mahatma Gandhi College, Iritty, Kannur

JUNE 2018 – MAY 2021

- Graduated with 95.76% with a 7th rank
- Cleared JAM Exam
- Member of Mock parliament and got third prize in State Competition

### Higher Secondary, St Mary's HSS, Edoor, Kannur

JUNE 2016 – MARCH 2018

- Graduated with 98.75 %
- Awarded Inspire Scholarship

## Internships

### Data Science Intern, EISystems Technologies, Ghaziabad, Uttar Pradesh

FEBRUARY 2022 – MAY 2022

- Got training on advanced Python and Data Visualization Tools.
- Completed the internship and submitted a project titled "A Study On Air Quality of Chennai."
- Tools : Excel , Pandas, Sklearn, Matplotlib, Seaborn

## Certifications

- Google Data Analytics Professional Certificate
- Inferential statistics, Duke University
- Python(basics), HackerRanks
- Introduction to machine learning, Duke university
- Natural language processing, Udemy
- AWS cloud certificate

## Projects

- Employee Attrition Prediction Using Machine Learning Algorithms

## Details

Kannur

8289971305

[stebingeo17@gmail.com](mailto:stebingeo17@gmail.com)

NATIONALITY

India

## Links

[www.linkedin.com/in/stebin-george](https://www.linkedin.com/in/stebin-george)

## Skills

Python

R

Machine Learning

SQL

Data Analysis

Database Management

Amazon AWS

Microsoft Excel

Hadoop

Flask

## Languages

English

Hindi

Malayalam

## Hobbies

Cycling

Cricket

Reading

- Developed and deployed a predictive model to Predict the Attrition Rate in an organization
- Used different ML Algorithms and Hyper Tuned the models to select the best predictive model
- Tools: **Flask, Pandas , Sklearn**
- **Extracting collocations from a text corpus and finding the similarity between them**
  - Through this NLP project, collocations were able to extract from a text corpus, and the similarity index was ranked based on the statistical measures.
  - Tools: **Spacy, NLTK , Pandas, Text-Preprocessing(Stemmer, Lemmatization, Vectorization)**
- **Predicting the Position of Football Player using Machine Learning Algorithms**
  - Created and deployed a Flask model that predicts the football player's position given the attribute.
  - Dataset: [https://www.kaggle.com/datasets/stefanoleone992/fifa-20-complete- player- dataset](https://www.kaggle.com/datasets/stefanoleone992/fifa-20-complete-player-dataset).
- **A Study On Air Quality of Chennai**
  - This project uses machine learning models to predict air quality in chennai.
  - The output obtained showed that Random Forest regressors were performing better in accuracy.