Hanna Paul

Starting a career after maternity

Personal Information

hannapaul8585@gmail.com

linkedin.com/in/hanna-paul

github.com/hanna-paul

+91 9496761310

Bangalore, India

Professional Skills

- Python
- ML Algorithm
- MS Excel
- MySQL
- MongoDB
- Pandas
- Numpy
- Sklearn
- Pyspark

Education

Master of Computer Applications (Correspondence)
IGNOU
2017-present

Bachelor of Science in Mathematics
Mahatma Gandhi University
2009-2012

Certificates

Certified Programming Professional & Master Data Science Program in GUVI (2021)

Certificate of Completion Data Science Foundations: Fundamentals
Linkedin learning (2021)

Certificate of Participation in The Guinness World Record Event- Most users to take an online computer Programming lesson in 24 hrs
GUVI (2021)

Languages

English
Hindi
Malayalam

Objective

Seeking opportunities to join a company that can help me in enhancing my skills, strengthening my knowledge, and realising my potential. I am willing to explore a wide variety of opportunities that can help me gain perspective.

Projects

Electricity load prediction using ML Algorithm

- Created a prediction model to predict the electricity load of a northern Indian state. Load forecasts are extremely important for energy suppliers, ISOs, financial institutions, and other participants in electric energy generation, transmission, distribution, and markets.
- Model pipeline consists of data cleaning, processing, visualizing, training the data using different regression models. Selected a model based on the R-squared value and then predicted the load value on the test data.
- Identified the data as a supervised regression problem. Trained data using the regression models. Cross validation on dataset is done. For each model R-squared is calculated and compared. Based on the comparison chose the best model and predicted the load.
- *Tools: Python, pandas, numpy, sklearn, matplotlib, seaborn*

IPL Data Analysis

- Analysed the last 12 season's IPL dataset effectively and found some interesting insights. Used sqlite relational DB and created a database and loaded it with the dataset. Created a class to expose generated result-set to data consumer.
- Loaded the data using pyspark then interpreted the data using methods like select, filter, sort etc.
- *Tools: Python, Pyspark, Google Colaboratory, Sqlite3, pandas*

COVID 19 Data Visualization

- Created a visualization for viewing the current confirmed, recovered and death status of the whole world's Covid data by placing the cursor on the map. Map is updated automatically with the latest data keeping the visualization up-to-date.
- *Tools: Python, pandas, geoviews*