

+880-1309078826  
rafat.joy99@gmail.com

# Rafat Ashraf Joy

## Data Analyst

linkedin.com/in/rajoy99  
github.com/rajoy99

### EXPERIENCE

#### ML Engineer (Intern)

Pioneer Alpha

Jul 2020 — Sep 2020

Dhaka, Bangladesh

- Deployed models to web using Flask web framework.
- Gained hands-on experience of Scikit-learn, Keras, PyTorch libraries.

### EDUCATION

**Bachelor of Science, Computer Science and Engineering,**  
Shahjalal University of Science and Technology, GPA: 3.50/4.00

Jan 2018 — Dec 2021

### SKILLS

<b>Programming Languages</b>	Python, C++, Scala, SQL
<b>ML Libraries</b>	Pytorch, Tensorflow, Scikit-learn, Xgboost, Catboost
<b>Data Visualization Libraries</b>	Matplotlib, Plotly
<b>Databases</b>	MongoDB, MySQL
<b>Web Frameworks</b>	Flask
<b>Softwares</b>	Tableau
<b>Others</b>	Linux, Git, Bash

### PUBLICATIONS

1. Joy, R. A. *Fine Tuning the Prediction of the Compressive Strength of Concrete: A Bayesian Optimization Based Approach* in *2021 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)* (2021), 1–6.
2. Joy, R. A. *An Interpretable Catboost Model to Predict the Power of Combined Cycle Power Plants* in *2021 International Conference on Information Technology (ICIT)* (2021), 435–439.

### PROJECTS

#### Customer Churn Prediction Web App

Scikit-learn, Flask, Heroku

- A multi-layer perceptron classifier model runs in the backend of this web app to predict customer churn in context of the telecom industry. The model has attained 96 % accuracy on the test dataset. This project has been deployed to Heroku utilizing Flask web framework.

#### Estimating Cloud Data Center Workload

Keras, Scala, Optuna, Statsmodel

- The dataset used in this project was taken from: TU Delft's business critical cloud workload time series data. First, the data was smoothed using 'savitzsky golay' filtering; this step removed the noise. ARIMA was applied to estimate the workload using Statsmodel. Then, several deep learning methods : LSTM, RNN, TCN(Temporal Convolutional Networks) were applied on the data. Finally, all of the deep learning models were hyper parameter tuned using Optuna.

#### Detection of COVID-19 from Raman spectroscopy

Scikit-learn

- This project applies a LASSO-regularized logistic regression model to detect Covid-19 from Raman spectroscopy data. As the number of features was very large compared to number of observations, we resorted to LASSO regression to avoid overfitting. The model is initialized with a liblinear solver along with L1 (LASSO) penalty, and achieved 97 % accuracy on the test dataset.