# Rashik Islam

17.06.1994  |  Seefelder Strasse 34  |  13053 Berlin  |  Germany

+49 177 8638354  |  rashik.islam.cs@gmail.com  |  -LinkedIn  |  -GitHub

*Proficient in data-driven applications with a passion for solving real-world problems with predictive learning, deep learning, statistical modeling, and creating the feedback loop of learning mechanisms.*

## Work Experience

**Nov 2021 - Feb 2024**

**DATA SCIENTIST** *Fresenius Group (Curalie GmbH), Berlin, Germany*
- On top of the diagnosis model, a trust model was developed to automate laboratory tests. This improved detection accuracy by about 3-5%.
- Participated actively in the creation of a multilingual knowledge base of diseases, symptoms, risk factors, and lab testing
- Created and implemented HPO workflow for tuning machine learning models
- Built pipelines for monitoring deployed ML models
- Actively involved in the execution of statistical analysis and in-depth error analysis in order to promote a culture of learning from mistakes by utilizing incremental development cycles and feedback loops

**Dec 2020 - Aug 2021**

**MASTER'S THESIS** *Charité – Universitätsmedizin, Berlin, Germany*
- Construction of a data pipeline to clean/extract data from >100.000 clinical notes/reports
- Presented multi-modal deep learning strategy to predict Kidney graft rejection that enhances ML models by fusing neural network architectures
- Integrated both textual and tabular data which elevated graft loss prediction which resulted an improvement of 3-6% F1-score compared to baseline

**Mar 2019 - Mar 2021**

**WORKING STUDENT, DATA ANALYTICS** *FUNKE Digital GmbH, Berlin*
- Engaged in the development of eight chatbot applications in collaboration with end-users, i.e. journalists/department leads, that harvest data from different sources to generate weekly and monthly reports.
- Forecasting algorithms for 30 portals were created with an aim to enhance market understanding and maximizing return on investment.
- POC was conducted to determine whether an article should be placed behind a paywall system comparing several ML algorithms

**May 2016 - Sep 2018**

**SOFTWARE RESEARCH ENGINEER** *Codemen Solutions Inc, Dhaka, Bangladesh*
The company provides end to end IT solution for its clients and its edge being integrating smart data, AI with embedded solutions.
- Developed an application for steel rod factories to automate the process of rod counting through image processing technique
- Involved in developing Proof Of Concepts i.e. ArgosAI (in collaboration with premiseHQ), a digitized data stream activated through video cameras that can be used to count objects, manage the use of spaces.
- Actively took part in building ETL pipeline from different sources using Rest API

# Education

| | |
|---|---|
| May 2024 -TBD | **PH.D. IN ATMOSPHERIC SCIENCES** *University Of Houston, Houston, USA [Expected Start: May 2024]* |
| Oct 2018 - Oct 2021 | **MASTER OF SCIENCE (M.SC.) IN DATA SCIENCE** *Berlin University of Applied Sciences, Berlin, Germany* [CGPA: 1.4]<br><br>**Thesis**: "A multi-modal deep learning strategy for kidney graft loss prediction on imbalanced dataset."<br><br>**Projects**:<br>• Vision based absolute Pose estimation of Autonomous Vehicle using deep learning, Url<br>• Sudoku puzzle solver from images using deep learning, Url |
| Apr 2012 – Apr 2016 | **BACHELOR OF SCIENCE (B.SC.) IN COMPUTER SCIENCE** *American International University Bangladesh (AIUB), Dhaka, Bangladesh* [CGPA: 3.7 (equivalent to 1.5 German CGPA)]<br><br>**Thesis**: "Automated Portable System for Bacteria Colony Counting & Contamination Measurement of Water Using Image Processing Technique" |

# Publications/Personal Projects

| | |
|---|---|
| Jan 2024 | **EVALUATING THE EFFICACY OF DEEP LEARNING AND HYBRID MODELS IN FORECASTING PM2.5 CONCENTRATIONS IN TEXAS: A 7-DAY PREDICTIVE ANALYSIS** *2024* preprint doi |
| Jan 2023 | **HYBRID MODEL OUTPERFORMED INDIVIDUAL MODELS IN PREDICTING DROUGHTS IN A SEMI-ARID REGION OF BANGLADESH** *2023* preprint doi |
| Oct 2020 | **POTENTIAL OF ARIMA-ANN, ARIMA-SVM, DT AND CATBOOST FOR ATMOSPHERIC PM2.5 FORECASTING IN BANGLADESH** *2021* Atmosphere Journal Vol. 12, number 1, pages 100 doi |

# Skills

| | |
|---|---|
| Programming Languages | **3+ years**: Python, **1+ years**: R | SQL, **1 year:** C++ | Javascript | HTML | CSS |
| Database | **1+ years:** MySQL | Microsoft SQL Server | NoSQL |
| Machine Learning | **1+ years:** Linear Regression | GLM | Decision Tree | Random Forest | GBM | SVM | ARIMA | Neural Network | LSTM | GRU | Transformers |
| Tools & Frameworks | **1+ years:** ClearML | Matplotplib | Numpy | Pandas | Scikit-Learn | PyTorch | Keras | Hugging Face | NLTK | OpenCV | Shapely | Matplotlib | Plotly | ggplot | dplyr | Docker , **1 year:** RapidMinder | Pyspark | Kafka| Streamlit |
| Others | **1 year:** Google Cloud Platform | AWS | FastAPI| Flask **1- year:** Trifacta, **1+ years:** Anaconda | Jupyter Notebook | R Studio | Visual Studio, **2+ years:** Git | Trello |
| Languages | **Bangla:** Native, **English:** Fluent, C2, **German:** Beginner, A1 |