

# Farjana Sultana Samia

*Software Engineer(Natural Language Processing)*

Website ♦ LinkedIn ♦ Github ♦ [samia.cse.miu@gmail.com](mailto:samia.cse.miu@gmail.com)

## OBJECTIVES

---

To pursue a challenging and research-oriented career in the fields of Artificial Intelligence, Machine Learning, Deep Learning, Data Science, Natural Language Processing.

## RESEARCH INTERESTS

---

Artificial Intelligence(AI) — Machine Learning(ML) — Natural Language Processing(NLP) — Data Science(DS)

## STANDARDIZED TEST

---

TOEFL iBT : Total Score: 68 (Reading-11, Listening-20, Speaking-20, Writing-17)

Revised GRE : Total Score: 296 (Verbal-139, Quantitative-157, Analytical Writing-3.0)

## PROFESSIONAL EXPERIENCE

---

**Software Engineer (NLP)** /Technometrics Limited *December, 2022 - Present*

- Developed Bengali Multi-label Topic classification model using pre-trained BERT. which can predict 17 different topics with 85 percent accuracy.
  - Multi-class, BERT, Flair, Pytorch, Transfer Learning, Fasttext
- Developed a Bengali aspect-based sentiment analysis model to predict different polarity scores based on different aspects from the same sentence.
  - PyABSA, Predicting aspect and sentiment of that aspect, Writing Annotation Guideline
- Developed a social media and news media data collecting bot to extract all types of data from Facebook and several news portals and analysis application, including Word Cloud, Trending Topic, Most Active User, Viral Post Analysis, etc. This bot can automatically complete scraping tasks as per the scheduled timetable, and store the data in MongoDB. The collected data can be made available through an API as required.
  - Selenium, MongoDB, Scraping, Multiprocessing
- Developed a scraping pipeline that can scrape 25 Bengali blogs and 15 news portals.
  - Scrapy, MongoDB

**Assistant Teacher (Math & Science Teacher)** /Uttara Model Academy(UMA) *January, 2022 - November 2022*

- Conducting 5th to 8th grade classes
- Creating Math and Science questions

**IT Expert (IT Support)** /Bonnie Store (E-commerce site) *20 hours per week(Remote job)*

- Solving Bonnie Store website technical issues
- Developing new themes
- Website Designing

## TECHNICAL SKILLS

---

- **Programming Languages:** C/C++, C#, R, Matlab, Python

- **Libraries:** TensorFlow, Keras, Numpy, OpenCV, Matplotlib, Pandas.
- **Web Technologies:** HTML, CSS, Bootstrap, Asp.net, Flask, FastAPI
- **Web Scraping:** Scrapy, Selenium, Request, BeautifulSoup
- **Database:** MySQL, MongoDB

## PROJECTS

---

### Project 1 CIFAR-10 - Object Recognition in Images

[GitHub](#)

- Object Recognition in CIFAR-10 Images(CIFAR-10 dataset consists of 60000 32x32 colour images in 10 classes).
- Using Python Keras library.

### Project 2 Data Visualization using Seaborn (Breast Cancer Diagnostic data.)

[GitHub](#)

- Features are computed from a digitized image of a fine needle aspirate (FNA) of a breast mass. They describe characteristics of the cell nuclei present in the image.
- Observing all Pair-wise Correlations.
- Observing the Distribution of Values and their Variance with Swarm Plots.
- Using Joint Plots for Feature Comparison Using Python Seaborn library.

### Project 3 Flask based University CGPA system website

[GitHub](#)

- This semester grading system is using on Python Flask for software developing.
- I've used here Bootstrap and java scripts for designing the website.

### Project 4 Basic Face Detection using CNN

[GitHub](#)

- I have applied a simple Transfer Learning model using python on my nephew's happy and sad photos
- It can detect my nephews happy and sad face properly.

### Project 5 Developed a Bengali news scraper

[GitHub](#)

- It can scrape news articles from prothomalo.com, banglatribune.com, jagonews24.com and samakal.com

### Project 6 Developed a location-based product scraper

[GitHub](#)

- It can scrape product information from Amazon.com

### Project 7 Developed a twitter data scraper

[GitHub](#)

- It can scrape twitter profile information from twitter.com

### Project 8 Flower category prediction using transfer learning

[GitHub](#)

- Resnet50, VGG19, Early Stopping

### Project 9 Red blood cell counting from images of blood cells using transfer learning

[GitHub](#)

- Faster RCNN, SSD, YOLO

## RELEVANT COURSEWORKS

---

### Undergraduate Courses:

- Structured Programming (C), OPP (Java), Data Structures, Algorithms, Statistics and Probability, Database Systems, Operating Systems, Software Development, Computer Graphics, Internet and Web Programming (HTML, CSS, Java Script, PHP), Computer Networking, Artificial Intelligence, Neural Networks and Fuzzy Systems, Pattern Recognition, Computer Vision and Robotics.

### Courses:

- **Ishraak Solutions Limited** -Machine Learning

- **Coursera** - Data Science Math Skills

## EXTRA-CURRICULAR ACTIVITIES

---

- **Microsoft Bangladesh**, *Brand Ambassador* (A Certificate of Recognition Brand Ambassador. Microsoft targets the youth and provides a platform for the people)(Aug 2018 - Dec 2018)
- **Young Bangla**, *Campus Ambassador* (With the motto “Connecting the Dots ”, becomes a young talent to promote, share and spread out youth towards my community.)(Aug 2018 - Dec 2018)
- **Women Techmakers**, *Volunteer* (Spreading and Developing women are technologically strong through this team )(Jan 2017 - Mar 2019)

## RESEARCH EXPERIENCE

---

### Bangla Text Document Categorization

- Working on a Natural Language Processing (NLP) based research work. Automatic document categorizing is an emerging research area in the field of NLP. In this categorization, Using various Deep Learning and Machine Learning Models.

### Undergraduate Research Work (Using Python)

- **Thesis Title:** Prediction of COVID-19 from Chest X-ray Images Using CNN And Pre-Trained Models
- **Brief :** Comparing Convolutional Neural Network And seven Pre-Trained Models performance using a single data-set which is ”COVID-19 Radiography Database”. Among all pre-trained models we got highest accuracy result from VGG-16 and EfficientNetB0 got the lower accuracy result.

## PUBLICATION

---

- ”A Resource-efficient NLP-based Online Bengali News Categorization Model with Outlier Removal”, (**Under Review**)
  - Worked on a Natural Language Processing (NLP) based research work. Automatic document categorizing is an emerging research area in the field of NLP. In this categorization, used various Deep Learning and Machine Learning Models.

## EDUCATION

---

**B.Sc. in Computer Science and Engineering** ( Feb 2017 - Jan 2022 )

Manarat International University(MIU)

**CGPA:** 3.90 out of 4.00

**Thesis:** Prediction of COVID-19 from Chest X-ray Images Using CNN And Pre-Trained Models

**Higher Secondary Certificate** ( 2014 - 2016 )

Shaheed Bir Uttam Lt. Anwar Girl’s College(SAGC)

**GPA:** 5.00 out of 5.00

**Secondary School Certificate** ( 2012 - 2013 )

Uttara High School and College(UHSC)

**GPA:** 5.00 out of 5.00

## TOOLS/SOFTWARE

---

- Github, VS Code, SPSS, Jupyter Notebook, Pycharm, Matlab, Photoshop, SQL Server, OS (Ubuntu, windows)
- **Document Processing:** LaTeX, MS Office, Adobe Illustrator, OriginPro.

## REFERENCE

---

### **Naimul Haque**

Lecturer

Department of Computer Science & Engineering

Ashulia model town, Dhaka-1341, Bangladesh

E-mail: [naimul@manarat.ac.bd](mailto:naimul@manarat.ac.bd)

Phone: +880 1939023802

### **Tareq Al Muntasir**

Head of NLP,

Hishab

Technologies Ltd

Email : [muntasir2000@gmail.com](mailto:muntasir2000@gmail.com)

Phone: +880 1673110720