

Farjana Sultana Samia

Ph.D. Student, Iowa State University
[Website](#) ♦ [LinkedIn](#) ♦ [Github](#) ♦ fssamia@iastate.edu

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)

EDUCATION

Ph.D. in Computer Science (Jan 2024 – Present)

Iowa State University, Ames, IA, USA

CGPA: 4.00 out of 4.00

Research: Doing research in Natural Language Processing (NLP)
under the supervision of **Professor Dr. Qi Li**

B.Sc. in Computer Science and Engineering (Feb 2017 – Dec 2021)

Manarat International University (MIU), Dhaka, Bangladesh

CGPA: 3.90 out of 4.00

Latter Grade: Graduated with distinction; 1st position among 48 students

PROFESSIONAL EXPERIENCE

Graduate Teaching Assistant /Iowa State University January 2024 - Present

- COM S 327: Advanced Programming Techniques
 - Grading, Office hours, addressing problem-solving Q/A on Piazza platform

Software Engineer (NLP) /Technometrics Limited December, 2022 - December, 2023

- 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 /Uttara Model Academy January, 2022 - November 2022

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

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

RESEARCH EXPERIENCE

Research Experience /Iowa State University

July 2024 - Present

Data Science Lab, Iowa State University

Advisor: Professor Dr. Qi Li

- Collaborating research in Natural Language Processing (NLP)

Bangla Text Document Categorization

- *"A Resource-efficient NLP-based Online Bengali News Categorization Model with Outlier Removal"*, (On review Elsevier IPM publication)
- 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-traind models we got highest accuracy result from VGG-16 and EfficientNetB0 got the lower accuracy result.

PROJECTS

Project 1 Retrieval-Augmented Generation (RAG)

[GitHub](#)

- Using GPT3 openAI model for developing a RAG based system to improve generative AI model accuracy by integrating external factual data retrieval. Simultaneously using Funix for turning Python functions/classes into accessible web applications.

Project 2 CIFAR-10 - Object Recognition in Images

[GitHub](#)

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

Project 3 Data Visualization using Seaborn (Breast Cancer Diagnostic data.)

[GitHub](#)

- Using Joint Plots, Swarm Plots for Feature Comparison Using Python Seaborn library and Pair-wise Correlations.

Project 4 Flask based University CGPA system website

[GitHub](#)

- This semester grading system is using on Python Flask for software developing.

Project 5 Basic Face Detection using CNN

[GitHub](#)

- I have applied a simple Transfer Learning model using python on my nephew's happy and sad photos

Project 6 Developed a Bengali news scraper

[GitHub](#)

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

Project 7 Developed a location-based product scraper

[GitHub](#)

- It can scrape product information from Amazon.com

Project 8 Developed a twitter data scraper

[GitHub](#)

- It can scrape twitter profile information from twitter.com

Project 9 Flower category prediction using transfer learning

[GitHub](#)

- Resnet50, VGG19, Early Stopping

Project 10 Red blood cell counting from images of blood cells using transfer learning

[GitHub](#)

- Faster RCNN, SSD, YOLO

RELEVANT COURSEWORKS

Graduate Courses:

- COMS 5790: Natural Language Processing, COMS 6730: Advanced Topics in Machine Learning, COMS 5730: Machine Learning

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)

REFERENCE

Naimul Haque
Assistant Professor,
Department of CSE
Manarat International University
naimul@manarat.ac.bd
Phone: (+880) 1939023802

Md. Imran Hosen
Assistant Professor,
Department of CSE
Manarat International University
imranjucse21@gmail.com
Phone: (+880) 1936940835

Dr. Qi Li
Assistant Professor,
Department of CS
Iowa State University
qli@iastate.edu