SOUMYAJIT SAHA

https://www.linkedin.com/in/soumyajit-saha-8b47a2182/ | Salt Lake City, Utah | +1 (385)-358-5011 | soumyajitsaha1997@gmail.com
https://github.com/Soumyajit-Saha

EDUCATION

MS, Computer Science | University of Utah, USA Aug 2021-Present

- **CGPA: 4.00/4.00.**
- **Relevant Coursework**: Advanced Algorithms, Computer Architecture and Natural Language Processing.

B.Tech, Computer Science and Engineering | Maulana Abul Kalam Azad University of Technology, India Jul 2020

- **DGPA: 9.48/10.00.**
- **Relevant Coursework**: Data Structures & Algorithms, Object Oriented Programming, Formal Language and Automata Theory, Design and Analysis of Algorithms, Principles of Computer Programming, Operating Systems, Database Management Systems, and Computer Networks.

SKILLS

- **Languages:** Java, Python, C, MATLAB, JavaScript, HTML/CSS.
- **Technical Domains:** SQL, SSRS, Optimization, Pattern Recognition, Feature Selection, Machine Learning, and Development.
- **IDE:** Jupyter, Spyder, Python IDE, Visual Studio Code, IntelliJ.

WORK EXPERIENCE

Graduate Teaching Assistant | University of Utah Aug 2021-Present

- Conducting doubt clearing sessions, grading and reviewing projects and assignments for Master's in Software Development program. [**Android, Java, Data Structures and Algorithms,** and **IntelliJ**]

Programmer Analyst | Cognizant Technology Solutions Sep 2020-Jul 2021

- Designed and developed Healthcare System interfaces and reports for "Blue Cross and Blue Shield of Minnesota". [**Teradata, SSRS, PowerBI,** and **WhereScape**]
- Assisted the development team to secure a position among **Top Business Intelligence** accounts by increasing the revenue by **10%**.

Business Intelligence Intern | Cognizant Technology Solutions Feb 2020-Jun 2020

- Accomplished assignments and training in MSBI domain. [**MySQL, MSSQLServer, SSIS, SSRS, PowerBI, Data Warehousing,** and **Microsoft Azure**]
- Secured a performance score of **95%**.

Research Intern | Jadavpur University CMATER lab Dec 2018-Jan 2019

- Implemented various wrapper-based algorithms and developed a **Harmony Search** based technique for the feature selection for generic classification models.
- Selected relevant features, augmented the **Classification Accuracy** using meta-heuristic algorithms and reduced the model building time.

RELEVANT PROJECTS

Information Extraction from Business Acquisition Documents Oct 2021-Nov 2021

- Executed data cleaning on real life raw text data by incorporating word and sentence level tokenization.
- Extracted relevant syntactic, lexical, semantic and statistical features from the cleaned data and performed vectorization.
- Performed text level classification by training the model using Logistic Regression with hyper parameters tuning to extract different role fillers, thereby obtaining an **F-score** of **0.55** on test dataset. [**NLTK, scikit-learn,** and **spaCy**]

A Bi-stage Feature Selection Approach for COVID-19 Prediction using Chest CT images Sep 2020-Oct 2020

- Proposed a bi-modular hybrid model to detect COVID-19 from chest CT images.
- Trained a convolutional neural network model using CT images and extracted features from the model.
- Incorporated ensemble of two filter methods namely Mutual Information and ReliefF for initial screening of features and applied wrapper-based **Dragonfly Algorithm** for further selection of relevant features. Tested the Model on two open access databases namely SARS-CoV-2 CT images and COVID-CT datasets. [**scikit-learn, TensorFlow,** and **Keras**]
- Attained **Prediction Rates** of **98.39 %** and **90.0%** on the said datasets respectively.
  Link: https://link.springer.com/article/10.1007/s10489-021-02292-8

Recognition of Online Handwritten Bangla and Devanagari Basic Characters: A Transfer Learning Approach. Apr 2020-May 2020

- Incorporated transfer learning models, namely, **VGG-16**, **ResNet50** and I**nception V3** for recognition of online handwritten Bangla and Devanagari characters using online handwritten character image datasets. [**TensorFlow, Keras,** and **SciPy**]
- Trained the models from scratch and using pre-trained weights; and imposed external challenges by augmenting the training datasets.
- Generated **Classification Accuracy** of **99.68%** and **99.97%** for Devanagari and Bangla datasets respectively.
  Link: https://link.springer.com/chapter/10.1007%2F978-981-16-1092-9_45

Feature selection for Facial Emotion Recognition using Cosine Similarity based Harmony Search Algorithm Jan 2020-Mar 2020

- Implemented a feature selection technique called **Supervised Filter Harmony Search Algorithm** based on Cosine Similarity and minimal redundancy maximal relevance concept using Pearson's Correlation Coefficient.
- Tested the model on two benchmark facial emotion recognition datasets, namely the Radboud faces databaseand the Japanese female facial expression. Considered 5 feature descriptors including uniform local binary pattern, horizontal-vertical neighborhood local binary pattern, Gabor filters, histogram of oriented gradients and pyramidal histogram of oriented gradients for Feature extraction. [**MATLAB**]
- Achieved highest **Classification Accuracy** of **97.79%**, **Precision** of **98.6%**, **Recall** of **97.8%,** and **F-measure** of **98.19%.**
  Link: https://www.mdpi.com/2076-3417/10/8/2816

Online Handwritten Bangla Character Recognition Nov 2019-Dec 2019

- Extracted geometrical, topographical and spatial features for online handwritten Bangla character recognition using Fréchet distance and incorporated feature selection using **Filter based Harmony Search Algorithm**. [**OpenCV, MATLAB,** and **Weka**]
- Attained **Classification Accuracy** of **98.33%** and **52%** reduction in the dimension.