Somnath Rakshit

+1 (512) 798-3552| somnath@utexas.edu
somnathrakshit.github.io | *linkedin.com/in/somnathrakshit/* | github.com/somnathrakshit/

EDUCATION

Master of Science, Information Studies (Data Science Track), May/2021
The University of Texas at Austin | GPA: 3.85/4 **| Courses taken:** Mathematical Statistics for Applications, Linear Models, Data Mining, Introduction to Machine Learning, Database Management | **Teaching Assistant:** User Generated Content Analytics (Fall 2019), Data Science Principles (Spring 2020 & 2021), Advanced Predictive Modeling (Fall 2020).

Bachelor of Technology, Computer Science and Engineering, May/2018
Jalpaiguri Government Engineering College, India | **GPA**: 8.68/10
Courses taken: Artificial intelligence, Data Structures, Calculus, Discrete Mathematics, Probability and Statistics, Design and Analysis of Algorithms, Object Oriented Programming

EXPERIENCE

Research Assistant, Intelligent Data Exploration and Analysis Laboratory, UT Austin, May/2020 – Aug/2020
- Privacy-preserving federated learning using healthcare data (Mentor: Prof. Joydeep Ghosh). Worked on computable COVID-19 subphenotype generation with tensor factorization.

Research Assistant, Centre of New Technologies, University of Warsaw, Jan/2019 – Aug/2019
- Project - Ranking genetic biomarkers for breast cancer subtypes (Mentors: Dr. Indrajit Saha and Dr. Dariusz Plewczyński). Developed a novel ranking algorithm using multi-objective genetic algorithms. Additionally, showed the relation between up/down regulation of genes and survival probability of a population

Software Engineer – Machine Learning, Cyware Labs, July/2018 – Nov/2018
- Clustered similar articles and ranked by articles' importance using CNNs resulting in 2x no. of articles selected.
- Determined trending keywords using Named Entity Recognition from news articles.

SKILLS

Programming Languages: Python, Java **| Frameworks**: Scikit-learn, Numpy, Pandas, Matplotlib, Spacy, PyTorch, Tensorflow, Keras, Git, Django, Spark, Hadoop **| Databases**: SQL, Elasticsearch

PROJECTS AND PUBLICATIONS

Sequence parameter selection for MRI parameter mapping, **Mar/2020 (Ongoing)** - Wrote an open source GPU accelerated simulator in PyTorch for the Extended Phase Graph Algorithm for certain sets of sequences. After successfully parallelizing the simulator, we increased its speed by almost 100,000 times. This work was submitted as an abstract in ISMRM virtual conference. Guide: Dr. Jon Tamir

VizWiz Challenge, Mar/2020 - Developing a method to caption images taken by people who are blind using a multi-modal dataset containing text and image data. Guide: Dr. Danna Gurari

Deep Stock Predictions, Mar - May/2020 – Proposed a data driven approach for optimal selection of window length and multi-step prediction length, and added analyst calls as technical indicators to a multi-stack Bidirectional LSTM strengthened by the addition of Attention units. Guide: Dr. Edison Thomaz

geograpy3, Sep/2018 (Ongoing) – geograpy3 is a Python library that is used to extract place names from a URL or text, and add context to those names, e.g., distinguishing between a country, region or city. It can be installed from PyPi using pip and has over 38,000 downloads.

Detection and Localisation of Diabetic Retinopathy, Jan - May/2018 – Classification and localization of diabetic retinopathy was performed using Resnet model in fundus images. The trained model was interpreted using GradCAM method to generate heatmaps showing regions of importance as determined by the trained model.