# Somnath Rakshit

somnath@utexas.edu · (+1) 737-333-1713 · somnathrakshit.github.io

## Education

**The University of Texas at Austin**, Texas, USA 2019 – Present

*M.S.* in Information Studies

**Jalpaiguri Govt. Engineering College**, Jalpaiguri, India 2014 – 2018

*B. Tech* in Computer Science and Engineering (CSE), GPA: 8.68
**Thesis:** Detection and Localisation of Diabetic Retinopathy
**Advisor:** Dr. Dipak Kumar Kole

## Experience

**Centre of New Technologies, University of Warsaw,** Warsaw, Poland

*Researcher* Bioinformatics and Machine Learning Feb 2019 – Present

Applying various machine learning techniques to extract complex multidimensional dependencies to make sense of vast quantities of biological data generated by high-throughput experiments.

**Istituto di Informatica e Telematica, Consiglio Nazionale delle Ricerche** Pisa, Italy

*Visiting Researcher* Bioinformatics and Natural Language Processing May 2019

Dataset preparation and development of a novel way to preprocess text in medical domain for improved classification accuracy.

**Cyware Labs** Bangalore, India

*Software Engineer* Machine Learning July 2018 – Nov 2018

Development of automated tools in the domain of natural language processing.

- Developed a method to assign probability to news articles based on their importance using deep learning.
  **Technologies used** - Tensorflow, Django, Celery, Elasticsearch.
- Developed a method to eliminate news which are already similar to published news.
  **Technologies used** - Gensim, Scikit-learn, Django, Elasticsearch.
- Developed methods to cluster similar news.
  **Technologies used** - Tensorflow, Django, Elasticsearch.

*Software Intern* Machine Learning May 2017 – July 2017

Developed various ways to augment the task flow using machine learning.

- Developed a text summarization tool using the TensorFlow library as backend to summarize any given news article using an unsupervised approach by using the TextRank algorithm.
  **Technologies used** - Tensorflow, Django, Celery.
- Developed an algorithm to determine the trending keywords based on a particular time frame. Also, made changes to prune or add to the search list accordingly.
  **Technologies used** - Gensim, NLTK, Django, SQL.
- Built an Android app, using logistic regression, to detect and notify the user about the spam/phishing/unsafe links that are sent through text messages in various popular messaging applications like WhatsApp.
  **Technologies used** - Tensorflow, Django, Celery, MySQL. **URL** - `Google Play Store`

## Projects

Geograpy3 Sep 2018

- 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.

**Technologies used** - NLTK.

Identifying Land Patterns from Satellite Imagery in Amazon Rainforest Jan 2018 - Mar 2018

- Multi label classification of land patterns in Amazon Rainforests using Keras. The results obtained demonstrate state of the art performance.
  **Technologies used** - Keras, OpenCV, matplotlib.

Detection and Localisation of Diabetic Retinopathy Jun 2017 - May 2018

- Automated classification and localization of diabetic retinopathy was performed using Keras in fundus images.
  **Technologies used** - Keras, OpenCV, matplotlib.

## SKILLS

- Programming Languages: Python, Java, R, C, C++, Matlab
- Framework: PyTorch, Tensorflow, Keras, NLTK, Numpy, scikit-learn, Pandas, matplotlib, Django, Git
- Databases: Elasticsearch, MySQL

## MISCELLANEOUS

- LinkedIn: https://www.linkedin.com/in/somnathrakshit/
- Kaggle: https://www.kaggle.com/somnathr
- GitHub: https://github.com/somnathrakshit
- Codechef: https://www.codechef.com/users/somnathrakshit
- Languages: English - Fluent, Bengali - Native speaker

## PUBLICATIONS

### Identification of Four miRNAs by Analysing Multi-View miRNA-seq Data of Stomach Cancer

Somnath Rakshit, Namrata Pant, Sushmita Paul, Indrajit Saha "Identification of Four miRNAs by Analysing Multi-View miRNA-seq Data of Stomach Cancer", IEEE TENSYMP, June, 2019, Kolkata, India (Accepted)

### Deep Learning for Detection and Localization of Thoracic Diseases using Chest X Ray Imagery

Somnath Rakshit, Indrajit Saha, Dariusz Plewczynski, "Deep Learning for Detection and Localization of Thoracic Diseases using Chest X Ray Imagery", 18th International Conference on Artificial Intelligence and Soft Computing, June, 2019, Zakopane, Poland

### Integration of miRNA-seq and Clinical data to Identify Breast Cancer Subtype specific miRNAs using Survival Analysis and their Interactions with Genes and Transcription Factors

Michał Denkiewicz, Somnath Rakshit, Indrajit Saha, Dariusz Plewczynski, "Integration of miRNA-seq and Clinical data to Identify Breast Cancer Subtype specific miRNAs using Survival Analysis and their Interactions with Genes and Transcription Factors", Frontiers in Genetics, 2018 (Communicated)

### Detection of Diseases in Potato Leaves using Transfer Learning

Soumik Ranjan Dasgupta, Somnath Rakshit, Dipak K. Kole, Dhiman Mondal, "Detection of Diseases in Potato Leaves using Transfer Learning", CIPR, January, 2019, Kolkata, India (Accepted and Presented)

### Deep Learning for Integrated Analysis of Breast Cancer Subtype Specific Multi-omics Data

Somnath Rakshit, Subha Shankar Chakraborty, Indrajit Saha, Dariusz Plewczynski, "Deep Learning for Integrated Analysis of Breast Cancer Subtype Specific Multi-omics Data", IEEE TENCON, July, 2018, Jeju Island, South Korea

### Machine Learning for Object Labelling

Indrajit Saha, Somnath Rakshit, Tanay Ghosh, "Machine Learning for Object Labelling", IEEE TENCON, July, 2018 Jeju Island, South Korea