Sankarshan Mridha

email id: sankarshan7@gmail.com, mo: +91 9836245170/8902655437

Summary

Masters in **Machine Learning** from IIT Kharagpur and 1 year prior work experience. Prior work includes **spatio-temporal data-mining** and **sentiment analysis**. He also has extensive knowledge of **Data Structure** and **Algorithms**. Looking for positions involving algorithmic data science and/or related product development.

Technical Skills

Languages:	**C, C++, Python** (Scikit-learn, **Numpy, Pandas**), **Matlab**, R
Web Development:	HTML, CSS, JavaScript, Bootstrap
Databases:	MySQL
Other tools:	WEKA, CVX, MOSEK

Education

M.S. in Machine Learning and Transportation Network — January 2015 – Present
Indian Institute of Technology, Kharagpur
Thesis: Learning Problems in Transportation Network
Advisor: Prof. Sourangshu Bhattacharya and Prof. Niloy Ganguly
CGPA: 9.15/10

B.E. in Computer Science and Engineering — July 2009 – May 2013
Jadavpur University
CGPA: 8.62/10

Higher Secondary (+2) Examination — May 2009
Barasat P.C.S Govt. High School
Percentage(%): 87.2

Secondary Examination — May 2007
R.K.Mission Boys' Home High School, Rahara
Percentage(%): 93.80

MS Thesis

Estimation of Transportation Network Parameters

Objective: Identification and management of heavy traffic routes during major events in an urban area (e.g. Pride March in New York City)

Dataset: Crowd-sourced transportation data (public) of NYC (165 million trips); OpenStreetMap (OSM)

Contributions: We modeled city roads as a spatio-temporal transportation network. Estimation of a link's (road segment) mean travel time and its variance were posed as linear regression problems and solved using constrained optimization techniques in CVX. Results showed $\sim 14\%$ improvement over *Independent Link* (no correlation between links) model. [Languages Used: Python, Matlab]

Projects

Understanding Product Quality using Sentiment Analysis of Product Review (Term Project in course: Information Retrieval)

Contributions: In this project, we proposed methods for sentiment analysis of user reviews from e-commerce product pages. The proposed method used Standford's NLTK for text pre-processing followed by the LingPipe Classifier for polarity estimation of user reviews. The resulting polarities can be used to augment average ratings of products to provide a more informative shopping experience. [Languages Used: Java; Tools Used: Apache Solr, Nutch web crawler]

Understanding Personal Content in Quora (Term Project in course: Complex Network)

Objective: Analysis of how the popular question-answer (QA) site Quora (~ 80K QAs, ~ 47K user profiles) is transforming from a generic QA site to a more personal topic related QA site.

Contributions: We used psycholinguistic analysis tools like LIWC (emotional, cognitive and structural components analyser) and other similar statistical measures to analyse the nature of Quora's content over time. We discovered that Quora's users have shifted towards querying more personal content. For e.g., our results showed that personal topics had ($\sim 22.4\%$) more edits than general topics and that $\sim 32.25\%$ of these personal queries are done anonymously. [Languages Used: Python]

Skin Lesion Segmentation (B.E final year project)

Contributions: Development of a computer-based skin disease (like leprosy) diagnostic system which is capable of detecting the lesion contours automatically for image-based clinical assessment. The solution

involved the application of Gaussian smoothing on input gray scale images for noise reduction followed by the Canny Edge Detection Algorithm to detect lesions. [Languages Used: C]
(Advisor: Prof. Mita Nasipuri)

Work Experience

Junior Research Fellow October 2014 – Present
Indian Institute of Technology, Kharagpur
Research Group: Complex Network Research Group (CNeRG)

Programmer Analyst August 2013 – July 2014
Cognizant Technology Solutions India Pvt. Ltd.
Objective: Architecture design and development of a sophisticated data-warehousing solution for Mercury Insurance Group (leading US based insurance company). Developed platform was used for acquisition of raw unstructured policy data and transformation into meaningful structured information for future analysis.

- Contributions:
  -Responsible for end-to-end delivery of the project from gathering of requirements, specifications, to development and delivery.
  -Designed and implemented end-to-end architecture using ETL tool Informatica, SQL Server and shell scripts.

Teaching Assistant: Programming and Data Structure Theory (2016-2017 Spring), Smart-phone Computing (2016-2017 Fall), Operating System (2015-2016 Spring), Social Computing (2015-2016 Fall), Programming and Data Structure Lab (2014-2015 Spring)

Publications

Mining Twitter and Taxi Data for Predicting Taxi Pickup Hotspots
Sankarshan Mridha, Sayan Ghosh, Robin Singh, Sourangshu Bhattacharya and Niloy Ganguly, the 9^{th} IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Sydney, Australia 2017.

Awards

Best Academic Demo
CommBox: Utilizing Sensors for Real-Time Cricket Shot Identification and Commentary Generation – Ashish Sharma, Jatin Arora, Pritam Khan, Sidhartha Satapathy, Sumit Agarwal, **Sankarshan Mridha**, Satadal Sengupta, Niloy Ganguly, in Demos and Exhibits Track of the 9^{th} IEEE International Conference on Communication Systems and Networks (COMSNETS), Bangalore, India, January 2017.

Contact

Address	1 Bengal Chemical Hridaypur, Kolkata-700127
Mobile	+91 983-624-5170/890-265-5437
Email	sankarshan7@gmail.com
LinkedIN	https://in.linkedin.com/in/sankarshanmridha
Web	http://cse.iitkgp.ac.in/~smridha/