Debjyoti Paul

CONTACT INFORMATION

130 S 800 E APT 101
Salt Lake City
Pin: 84102
University of Utah
USA

http://www.cs.utah.edu/~deb
deb@cs.utah.edu
https://www.linkedin.com/in/debjyotipaul385/
http://github.com/debjyoti385
+1 385 313-7219

RESEARCH INTERESTS

Large Scale Social Media Data Analytics, Machine Learning and Data Mining techniques, Deep Learning, Bayesian Learning, Text Analytics, Indexing techniques, Data Visualization.

EDUCATION

University of Utah, School of Computing, Salt Lake City, UT, USA

PhD. Student, Computer Science & Engineering, Fall 2015- *Current*

Research Advisor: Professor Feifei Li
GPA: 3.93/4.0

Indian Institute of Technology Kanpur, Kanpur, UP, INDIA

M.Tech., Computer Science & Engineering, 2011-2013

Thesis Title: *Multi-constraint Job scheduling problem in Grid*
Research Advisor: Professor Late Sanjeev K Aggarwal
Research: Efficient scheduling strategy of Jobs on Grid with energy efficient solution for defined SLAs.
CGPA: 8.67/10 (*Rank: 3*)

West Bengal University of Technology, Salt Lake, Kolkata, WB, INDIA

B.Tech., Computer Science & Engineering, 2007-2011

College: *Institute of Enginering and Management*
CGPA: 8.93/10 (*Rank:* < 10)

WORK EXPERIENCE

Amazon AI Research, May 2017 - Aug 2017, New York, United States

Research Scientist Summer 2017 Intern

- Advisors: Baris Coskun, Edo Liberty
- Mentors: Ramesh Nallapati, Zohar Karnin
- Hyperparameter optimization with MxNet.

Flipkart, July 2013 - May 2015, Bangalore, INDIA

Software Developer in Data Platform (Exceeds Expectations for performance in Jan-July 2014)

- Data Platform is data bank of *flipkart.com*, India's biggest e-commerce company.
- Created environment for Big data analytics in scalable environment, processing pipeline for batch and stream.

THESES

[2013] M.Tech Thesis, Advisor: Late Dr. Sanjeev Kumar Aggarwal

Multi-constraint Job scheduling problem in Grid

- The objective is to efficiently schedule Jobs on Grid to achieve maximum utilization of resources with energy efficient approach. Modeling real world computing & storage grid on Multi-objective Evolutionary Algorithm satisfying hard constraints on jobs constraints, resources constraints, and soft constraints on cost, and energy consumption on NSGA-II. Local optimization using Pareto optimal front, global optimization by applying mutation on population on search space. Also introduced job grouping technique with constraints to accumulate fine grained jobs which keeps processing time as low as possible.

PUBLICATIONS

Geotagged US Tweets as Predictors of County-Level Health Outcomes, 2015–2016, Quynh C. Nguyen, Matt McCullough, Hsien-wen Meng, **Debjyoti Paul**, Dapeng Li, *American Journal of Public Health, September, 2017*, *DOI:* 10.2105/AJPH.2017.303993.

Compass: Spatio Temporal Sentiment Analysis of US Election, **Debjyoti Paul**, Feifei Li, Murali Krishna Teja, Yu Xin, Richie Frost, *What twitter says!*, *23rd SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD 2017)*, Aug 13-17, 2017, Halifax, Canada. *DOI:* 10.1145/3097983.3098053.

Social media indicators of the food environment and state health outcomes, Quynh C. Nguyen, Hsien-wen Meng, Dapeng. Li, Matt McCullough, **Debjyoti Paul**, Kanokvimankul. P, Nguyen. T, Li. Feifei, *American Public Health Association, 148, 120-128.*, 2017, *DOI:* 10.1016/j.puhe.2017.03.013.

Debjyoti Paul, Sanjeev K. Aggarwal, Multi-objective Evolution based Dynamic Job Scheduler in Grid, *The 8th International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS 2014)*, IEEE, July 2nd - 4th, 2014, Birmingham, UK. *DOI:* 10.1109/CISIS.2014.50.

Manash Pal, Arnab Bhattacharya, **Debjyoti Paul**, RCached-tree: An Index Structure for Efficiently Answering Popular Queries, *ACM International Conference on Information and Knowledge Management (CIKM 2013)*, Oct. 27 - Nov. 1, 2013, San Francisco, CA, USA. *DOI:* 10.1145/2505515.2507817.

Debjyoti Paul, Sumana Basu, Punit Beriwal, IEEE, *Lightweight Security Enhancement Protocol for Radio Frequency Identification (RFID)" in SPSITM International Conference*, 2011, Kolkata, WB, INDIA

Projects (More..)

Compass -*http://estorm.org*	This project explores how Twitter can be used to analyze the popularity of Political Parties for the Presidential Election 2016. It is purely based on the data collected from Twitter. Read our Compass paper for details.
Event Aggregator	This project finds events from news articles, categorize them and gathers all articles talking about same event to a set. Each set of articles is referred as Event Entity. Information extraction process is applied to extract more information related to it. The code can process large scale data. The code will be made public later.
QuakeAnalysis -*https://goo.gl/ovcWVT*	This project is based on the seismic activity across world which widely varies in characteristic and patterns. We have found some distinguish patterns among the seismic activities and present them in an insightful manner. Check demo and wiki of Exploration & Analysis
MusicAtlas -*https://goo.gl/UH54UP*	This website is designed for music lovers to learn about music based on countries. This is a unique tool to explore and analyze the trend of music based on time frame, genre and artists. Almost all data from 19th century to till date. Please make sure you allow SSL connection from api server via this link before accessing @musicatlas.
QuestionAnswering -*https://goo.gl/QGfW2y*	Closed domain Question Answer System. The system has *Recall* of 63% and *F-score* of 43%.
Dartnews -*https://goo.gl/wMOl4C*	A street news browsing application, with an interactive GIS interface. News is organized by locality and topic. The user can explore based on topics (eg crime, politics etc) and geolocation. We used Context Dependent Geoparsing, where we attempt to find out which location is relevant to the News. Across all the topics, we saw at least 87% accuracy of topic prediction, and at least 80% accuracy of location prediction.
Intelliad -*https://goo.gl/UAnEdz*	A Social Media driven Intelligent Ad-Targeting framework using Geo-profiling. The idea is to tag all geo-location enabled tweets(available publicly) with semantic categories (say sports, politics etc.) and their sentiment (positive, neutral, negative) using text mining. To enable serving of Ads, they also need to be tagged using same categories based on their content.

Find more projects at https://www.cs.utah.edu/~deb/#project

Hackathon Trophies

[2016]	EMC2 Code	Mars Challenge Hackathon Winner (around 45 teams)
[2015]	Goldman Sachs	Air Quality Hackathon 1st Runners-up (around 30 teams)
[2014]	InMobi	Freedom Hack Worldwide Hackathon 1st Runners-up (around 160 teams)
[2013]	Yahoo	Yahoo HackU 2013 Hackathon Winner (around 40 teams)

Achievements

[2013] Ranked 3rd out of 39 M.Tech students of CSE department in Indian of Institute Technology, Kanpur

[2012] Secured All India Rank 228 in GATE 2012 among 0.16 million participants and 223 in 2011 among 0.15 million participants of Computer Science & Information Technology department.

[2011] Achieved All India Rank of 7 in Indian Space Research Organization (ISRO) recruitment exam for Junior Research Fellow among 0.2 million candidates.

References

For skillsets please visit homepage https://www.cs.utah.edu/~deb/#skills.
All references are provided upon requests.