# Avijit Ghosh

## Northeastern University

linkedin.com/in/evijit github.com/evijit
evijitghosh@gmail.com evijit.github.io

Complex Networks | Computational Sociology | Algorithmic Fairness | Machine Learning | Natural Language Processing

## Education

Present **Northeastern University, Boston**
Ph.D. in Computer Science

2019 **Indian Institute of Technology, Kharagpur**
B.Tech. in Chemical Engineering
M.Tech. in Financial Engineering
Minor in Computer Science
CGPA: 8.77/10

2014 **Modern English Academy, Barrackpore**
ISC Class XII
Percentage: 94.5%

2012 **St. Claret School, Barrackpore**
ICSE Class X
Percentage: 97.8%

## Publications

**1. "Analyzing Political Advertisers' Use of Facebook's Targeting Features"** 2019
**Avijit Ghosh**, *Giridhari Venkatadri, Alan Mislove*
at the Workshop on Technology and Consumer Protection (ConPro) 2019, San Francisco, California, USA.

**2. "Public Sphere 2.0: Targeted Commenting in Online News Media"** 2019
*Ankan Mullick,* **Avijit Ghosh***, *Ritam Dutt*, Sayan Ghosh*, Abhijnan Chakrabarty*
at the European Conference on Information Retrieval (ECIR) 2019, Cologne, Germany.

**3. "SAVITR: A System for Real-time Location Extraction from Microblogs during Emergencies"** 2018
*Ritam Dutt, Kaustubh Hiware,* **Avijit Ghosh**, *Rameshwar Bhaskaran*
at the WWW 2018 workshop on Exploitation of Social Media for Emergency Relief and Preparedness (SMERP) - 2018 Lyon, France.

**4. "Molecule2Vec: Vector Space Representation of Organic Molecules for prediction of properties using Deep Neural networks"** 2018
**Avijit Ghosh**, *Debasis Sarkar*
at the European Congress of Chemical Sciences (EUCHEMS) 2018, Liverpool, UK.

**5. "WebSelect: A Research Prototype for Optimizing Ad Exposures based on Network Structure"** 2016
**Avijit Ghosh**, *Agam Gupta, Divya Sharma, Uttam Sarkar*
at the Workshop on Information Technology and Systems (WITS) 2016, Dublin, Ireland.

* Equal contribution

## Internship Experience

**May – July 2019** — **Visiting Researcher**
### LIG, University of Grenoble Alps
Study of how news companies promote different items on social media, investigating possible patterns of differential use.
Supervisor: Dr. Oana Goga
Facebook | Twitter | Advertisements | News Media | Opinions | Transparency

**May – July 2018** — **Visiting Researcher**
### Northeastern University, Boston
Study of advertiser behavior and targeting patterns on Facebook by using the ad reach information obtained from Facebook's ad transparency feature and the personal targeting dataset from Propublica's Facebook ad dataset, aided with controlled ad placement experiments.
Paper accepted at **ConPro 2019.**
Supervisor: Prof. Alan Mislove
Facebook | Propublica | Advertisements | Behavior Patterns | Transparency | Controlled experiments

**Dec 2017 – Jan 2018** — **Data Science Intern**
### Ernst and Young, Gurgaon
Automatic PDF report generation system by reading data from company database. Fraud likelihood prediction analyzing the credit history of consumers provided by client companies.
Supervisor: Gaurav Jain
Web server | Automation | Database | Classification | Fraud Detection

**May – July 2017**

**Summer Intern**

### Xerox Research, Bangalore

- Implemented XTrack, a Smart Vehicle Tracking and Battery usage minimizing Algorithm.
- Uber Surge Price Prediction using Spatio-Temporal techniques like the Neural Hawkes and Recurrent Marked Temporal Point Process. Was given the **Best Internship Project** award.

Supervisors: Narendra Annamaneni, Poorvi Agrawal

Android | Bluetooth | Handshake | Algorithm | Uber | Transportation | Point Process | Tensorflow

**Apr – Aug 2016**

**GSoC Student**

### Google Summer of Code - OpenMRS

Replaced the HTML XForms system used in the Android app with native generated forms using the Forms REST Api and added offline form saving. Configured Travis CI to automatically build and push the apk to play store.

Supervisors: Rafal Korytkowski and Robert O'Connor

Android | Travis.CI | Play Store | Rest API | Database | Open Source

## Research Experience

### Thesis Work

**Masters Thesis Ongoing**

**Modeling Connectedness of Firms in Financial Markets with Heterogeneous Agents**

- Goal: To discover connectedness and study heterogeneous agents in an experimental framework of ASM (Artificial Stock Market)
- Decomposition of volatility spillover or variance through networks.
- To create models of interconnectedness using network theory.

Supervisor: Professor Abhijeet Chandra , VGSOM, IIT Kharagpur. Co supervised by FNA.fi team, UK

Complex Networks | Finance | Failure Prediction | Artificial Stock Market

**Bachelors Thesis 2017-18**

**Vector Space Representation of Organic Molecules to predict aqueous solubility**

- Converted 3D Molecules to a Vector Space Model using Doc2Vec.
- Using information from IUPAC and other literature, created an exhaustive database of aqueous solubility data.
- Using the trained molecule vectors and the solubility data, train ML and Deep Learning Algorithms to predict solubility within an error of 0.3 g/litre.

Supervisor: Professor Debasis Sarkar, Chemical Engineering, IIT Kharagpur

Doc2Vec | Deep Learning | Regression | Chemistry

### Faculty supervised projects

**July – Dec 2018**

**Automated Extraction of Catchwords from Legal Documents**

- Automated catchword identification using both unsupervised and supervised techniques.
- The proposed unsupervised methodology uses graph centrality measures to rank the phrases.
- We also propose a supervised technique of extracting catchwords by formulating the catchword extraction as a sequence labelling task using CRF and Bi-LSTM models.
- Manuscript submitted to **SIGIR 2019.**

Supervisor: Professor Saptarshi Ghosh, CSE IIT Kharagpur

NLP | Legal Document | Sequence labelling | Catchphrase extraction

**Nov 2015 – Jan 2018**

**Data Driven Disaster Response Systems using Social Media**

***Project: Savitr - Realtime location extraction during emergencies***

- Developed a system called Savitr (presented at **WWW-SMERP 2018**) that leverages the information posted on the Twitter microblogging site to monitor and analyse emergency situations.
- Employed NLP techniques to infer the locations mentioned in the microblog text, in an unsupervised fashion and display it on a map-based interface.
- The system achieves a F-score of 0.79, significantly faster than other comparable methods.

***Project: Classification and Summarization of tweets during a disaster event***

- Developed an improved SVM Classifier to separate disaster related tweets into Situational and Non Situational Classes, using sentiment detection, dependency graphs and linear patterns.
- Built a software Demo called DISSUM using Flask to showcase the working of the above. This was selected and showcased at **IBM Day Conference 2016**, IIT Kharagpur.

Supervisor: Professor Saptarshi Ghosh and Professor Niloy Ganguly, CSE IIT Kharagpur

Twitter | Geolocation | Dash | Python | NLP | ML | Feature extraction | Classification | Summarization | Disaster

**Mar – May 2016**

**Maximizing the reach of advertisements based on Network Structure**

- Built a Graph of websites by scraping traffic information from Alexa.
- Designed a tool (named Webselect) to select the best subset of websites to maximise the reach of advertisements, within budget and demographic limits.
- Used Genetic Algorithm to optimize the selection problem as the original problem is NP-Hard.

Supervisors: Professor Uttam Sarkar, MIS, IIM Calcutta and Professor Agam Gupta, IIM Rohtak

Complex Networks | Web Crawling | Advertisements | Genetic Algorithms | Flask | Python

## Miscellaneous Projects

- **"Selective Commenting for Online News Media"** – Automatically position user comments against relevant news article paragraphs. **Accepted at ECIR 2019.**
  NLP | Deep Learning | Web Design
- **"Using Global Vectors in Social Interaction Network for Song Recommendation."** – Independent work, extended abstract submitted to CompNet 2018.
  Complex Network | Facebook | Social recommender | Music
- **"Bias detection in Google Search Autocomplete."** – under Prof. Alan Mislove, NEU, 2018.
  Bias | Discrimination | Algorithmic Fairness
- **"How News and Word of Mouth Affects Stock Price."** – System to find relevant news articles causing price fluctuations. Silver at Inter IIT Tech Meet 2017.
  Finance | Market prediction | ML
- **"Regression and Time Series Modelling: US Job Index"** – Analyzed factors that affected the job index of USA over the past 30 years. Prediction accuracy of 84%.
  Regression | Time series | Forecasting
- **Android Apps** – Free and Paid apps on the Google Play Store and freelancing for a startup (Truckerrs).
  Android Development | UX Design
  Play Store Profile: https://goo.gl/Gbgt9C

## Awards and Grants

### Academic

2018 **SGSIS Challenge Grant**
*Awarded the SGSIS Challenge Grant worth INR 1 Million for Masters Thesis. Only 9 projects from the Institute qualified.*

2012 **Mamraj Agarwal Rashtriya Puraskar**
*Conferred by the Governor of West Bengal for ranking 5th in India in the ICSE Exams.*

2010 **NTSE Scholar**
*Qualified for the National Talent Search Examination Scholarship conducted by NCERT.*

### Technical

2019 **Institute Order of Merit – Technology**
*Awarded by the Technology Students Gymkhana.*

2017 **Inter IIT Tech Meet – Kanpur**
*Silver Medal in the Stock Market Analysis Event.*

2016 **Inter IIT Tech Meet – Mandi**
*Gold Medal in the Software Development Event.*

2015 **YU App Challenge**
*(3rd Prize) – at National Level App Making Competition by Micromax.*

2014 **Flipkart Hackathon**
*Podium finish in the Flipkart Hackathon organised at IIT Kharagpur.*

### Extracurricular

2017 **5th Position – Inter Hall General Quiz**

2015 **Bronze Medal – Open IIT Bengali Elocution**

2011 **Runners Up Team – The Frank Anthony Memorial All India Inter School Debate – Regional Level**

2010 **8th Rank – Albert Barrow Memorial All India Interschool Creative Writing Competition**

## Positions of Responsibility

2014 – 19 **Advisor, Kharagpur Open Source Society**
*Advises a team of student coders who organize events to network and spread awareness about Free and Open Source culture.*

2016 – 17 **General Secretary Technology, Vidyasagar Hall of Residence**
*Handled a tech budget of INR 130K over the academic year.*

2014–15 **Associate Manager, Entrepreneurship Cell**
*Organized Entrepreneurship Drive in Bhubaneswar. 2000 students attended.*

2017 **Mentor, Google Summer of Code**
*Advised a student to finish set goals and monitored pull requests.*

## Skills

**Languages:** C, Java , BASIC, Python and R
**Data Science:** Hadoop, Spark, ML
**Android Development:** Play Store,freelance
**Web:** HTML, CSS, Flask, Mysql, Javascript
**Version Control Systems and CI :** Git, Travis, AppVeyor, Codacy
**Design:** Photoshop, Illustrator, Justinmind, LaTeX

**Relevant Subjects:** Algorithms, Artificial Intelligence, Machine Learning, Social Computing, Information Retrieval, Computer Architecture and Operating Systems, Regression and Time Series, Financial Analytics, Natural Language Processing, Scalable Data Mining.