

Lahari Poddar

Berlin, Germany

homepage : <https://laharipoddar.github.io/>

ACADEMIC QUALIFICATION

Qualification	University	Year	CPI/Max CPI
PhD (pursuing)	National University of Singapore	2019 (expected)	4.5/5
M.Tech in CSE	IIT Bombay	2013	9.0/10
B.Tech in CSE	WBUT	2011	9.14/10

RESEARCH INTERESTS

Data Mining, Machine Learning, Information Retrieval, Natural Language Processing, Probabilistic Graphical Models, Deep Learning, Bayesian Networks, Text Mining, Opinion Mining, Credibility Analysis

RESEARCH PUBLICATIONS

- **Lahari Poddar**, Wynne Hsu, Mong Li Lee "Predicting User Reported Symptoms Using a Gated Neural Network", ICTAI (2019)
- **Lahari Poddar**, L. Neves, W. Brendel, L. Marujo, S. Tulyakov, P. Karuturi "Train One Get One Free: Partially Supervised Neural Network for Bug Report Duplicate Detection and Clustering", NAACL (2019)
- **Lahari Poddar**, Wynne Hsu, Mong Li Lee, Shruti Subramaniyam "Predicting Stances in Twitter Conversations for Detecting Veracity of Rumors: a Neural Approach", ICTAI (2018) , **Best Student Paper Award**
- Kishaloy Halder, **Lahari Poddar**, Min-Yen Kan "Cold Start Thread Recommendation as Extreme Multi-label Classification", XMLC for Social Media, WWW (2018)
- **Lahari Poddar**, Wynne Hsu, Mong Li Lee "Author Aware Aspect Topic Sentiment Model to Retrieve Supporting Opinions from Reviews", EMNLP (2017)
- Kishaloy Halder, **Lahari Poddar**, Min-Yen Kan "Modeling Temporal Progression of Emotional Status in Mental Health Forum: A Recurrent Neural Net Approach" , WASSA, EMNLP (2017)
- **Lahari Poddar**, Wynne Hsu, Mong Li Lee "Quantifying Aspect Bias in Ordinal Ratings Using a Bayesian Approach", IJCAI (2017)
- Brijesh Bhatt, **Lahari Poddar**, and Pushpak Bhattacharyya " IndoNet: A Multilingual Lexical Knowledge Network for Indian Languages", ACL (2013)

INDUSTRY EXPERIENCE

Snapchat Research [California, USA] **Research Intern** [Aug '18 - Dec '18]

- Developed a partially supervised deep learning model for duplicate bug reports tracking and automatic clustering as part of the **NLP Research** team to save huge amount of engineering cost.
- Wrote production level code, managed a team of annotators and submitted a research publication

Amazon [Berlin, Germany] **Applied Science Intern** [May '18 - July '18]

- Worked in **NLP Research** team to improve a user's shopping experience on the website
- Developed a hierarchical word embedding model that can capture word semantics and aid in domain specific aspect terms mining using terabytes of user reviews

Microsoft [Bangalore, India] **Software Development Engineer** [July '13 - July '14]

- Worked in **Relevance and Ranking** team in **Bing Ads** as part of the core-ad selection group
- Designed and implemented algorithms for search query rewrite to map user intention into the advertisers' bidding space and serve relevant advertisements using NLP techniques and Machine Learning models to increase revenue

Information Reliability of User Generated Content on the Web **PhD Thesis**
Guided by: Prof. Wynne Hsu and Prof. Mong Li Lee [March '15 - till date]

- User generated content about products and services in the form of reviews are often diverse and even contradictory.
- Designed and developed machine learning models to quantify latent user biases, estimate intrinsic quality of items, modeling consensus of opinions and detecting veracity of rumors.
- Worked on Probabilistic Graphical Models, Bayesian Inference and Deep Learning Models for solving the above objectives on real world datasets.

SCHOLASTIC ACHIEVEMENTS

- Received **Best Student Paper Award** at ICTAI 2018 [2018]
- Received **Research Achievement Award** from School of Computing, NUS [2018]
- Won **Judges' Choice Award** in Yahoo! Hackathon [2012]
- Ranked among **Top 10%** students in IIT Bombay-CSE Department [2011-2013]
- Awarded the **Best Student Award** by Institute of Engineering & Management [2011]
- Secured **All India Rank 72** out of 136,027 candidates in GATE [2011]

Professional Academic Services

- Invited Program Committee member for: ACL 2019, AAAI 2019, SIGIR 2018, JCDL 2018, MR2AMC 2018.
- Invited reviewer for NAACL, JCDL, SIGIR, AAAI, VLDB, IEEE Transactions on Systems, Man and Cybernetics.
- Selected for and served as a student volunteer in EMNLP 2017, CIKM 2017

Teaching/Mentoring Experience

- Teaching Assistant for CS1020 Data Structures and Algorithms, Sem II 2014-2015
- Teaching Assistant for CS3201 Software Engineering Project, Sem I 2015-2016, Sem II 2015-2016, Sem I 2016-2017
- Hosted and mentored students for research internships at NUS

RESEARCH PROJECTS

Finding Popular Travel Destination From Tweets

Guided by: Prof. Chua Tat-Seng

[January '15 - April '15]

- Developed an API that uses millions of tweets from Twitter users all across the globe to provide a dynamic list of trending destinations emerging from live reactions of people on social media.
- Used a range of NLP tools and machine learning models along with efficient data processing framework to build a scalable, and real time system.

Multilingual Multiword Expressions

Guided By: Prof. Pushpak Bhattacharyya

M.Tech Project

[July '12 - June '13]

- Developed a multiword expression(MWE) extraction engine by employing a combination of rule based (to leverage linguistic features of Indian languages) and statistical (for collocations) filters to detect MWE for all major Indian languages.

Information Retrieval from Social Media

Guided by: Prof. Ooi Beng Chin

[August '14 - November '14]

- Developed a system for streaming live tweets in batches from Twitter, extracting named entities from them on-the-fly and finally storing them in a distributed inverted index.
- Given a query, performed selection and ranking of tweets based on its temporal and contextual information.

Building a lexical resource for Indian Languages using Universal Words

Guided By: Prof. Pushpak Bhattacharyya

[July '12 - Nov '12]

- Converting the isolated Hindi-UW dictionary to LMF model, the ISO standard for Natural language lexicons in order to ensure reusability using NLP techniques to merge two heterogeneous volumes of interlingual lexemes

ACADEMIC INTERNSHIP

GETALP, UJF, France

Guided By: Prof. Christian Boitet

[May '12 - July '12]

- Developed algorithms to unify heterogeneous volumes of resources and connected 18 Indian Languages to English, French and Russian through a volume of interlingual lexemes
- Interlingual Machine Translation and Information Extraction can greatly benefit from these linkages. It also provides a platform to build Hindi-UNL converter and deconverter

SOFTWARE SKILLS

- Programming Languages: C, C++,C#, JAVA, Python (Tensorflow, Keras, Scikit), Shell
- Big-Data Analytics : Hadoop, Scope, ElasticSearch
- Databases : MySQL, PostgreSQL
- Tools : SVN, Git, Maven
- Cloud Technologies : Amazon EC2, Google Cloud Platform