

# Anirban Laha

Mila - Quebec AI Institute, 6666 St Urbain St, Montreal, Quebec H2S 3H1, Canada.

☎ (+1)514-691-5065 | ✉ anirbanlaha@gmail.com | 🌐 anirbanl.github.io | 📄 anirbanlaha | 🐦 anirbanlaha

## Interests

Multi-Agent Reinforcement Learning, Deep Learning, Game Theory, Natural Language Generation.

## Education

### Montreal Institute for Learning Algorithms (Mila), University of Montreal

Montreal, Canada

#### PHD CANDIDATE IN COMPUTER SCIENCE

Sep. 2019 - Present

- *Courses:* Probabilistic Graphical Models, Advanced Structured Prediction and Optimization, Excursions in RL (audit).
- *Advisor:* Prof. Aaron Courville. | *Research Topic:* Learning to achieve cooperation and communication in Multiplayer Games.

### Indian Institute of Science (IISc), Bangalore

Bangalore, India

#### MS IN COMPUTER SCIENCE AND AUTOMATION

Jul. 2011 - Aug. 2013

- *Courses:* Machine Learning, Data Mining, Probability, Linear Algebra, Optimization, Game Theory. | *Advisor:* Dr. Shivani Agarwal.
- *Thesis Topic:* "Machine Learning and Rank Aggregation Methods for Gene Prioritization from Heterogeneous Data Sources". [Link](#).

### Jadavpur University, Kolkata

Kolkata, India

#### BACHELOR OF COMPUTER SCIENCE AND ENGINEERING - 9.05/10

Jul. 2006 - Apr. 2010

- *Final Year Project:* "Solving the Single-Source Shortest Path Problem for Dynamic Graphs". | *Advisor:* Dr. Chandan Mazumdar.

## Work Experience

### IBM India Pvt. Ltd.

Bangalore, India

#### RESEARCH SOFTWARE ENGINEER

Nov. 2015 - Aug. 2019

- *Projects:*
  1. **Natural Language Generation (NLG)** - Applications of seq2seq methods for structured data summarization and query-based abstractive summarization; Sparse attention models for enhancing interpretability through explicit controls.
  2. **IBM Project Debater** - Led the deep learning based efforts. Sequence classification used for claim and evidence detection in argument mining; Punctuation and Case prediction for unformatted ASR output produced from human debate recordings.
  3. **Machine Learning for Creativity** - Story generation from incoherent descriptions; ML Framework for evaluating creativity.

### Microsoft India R&D Pvt. Ltd.

Bangalore, India

#### APPLIED SCIENTIST

Aug. 2013 - Nov. 2015

- *Project:* Experimentation of new Click Prediction models for Ads Optimization in Microsoft Bing Ads for US and INTL Markets.
- *Domain:* Machine Learning, Information Retrieval, Ranking, Map-Reduce.

### Amazon Development Center India Pvt. Ltd.

Chennai, India

#### SOFTWARE DEVELOPMENT ENGINEER

May. 2010 - Jul. 2011

- *Project:* Development of automated real-time scalable Pricing System for Amazon Kindle eBooks for markets like US, UK and DE.

## Selected Publications

1. Laha A.<sup>†</sup>, Jain P.<sup>†</sup>, Mishra A.<sup>†</sup>, Sankaranarayanan K. (2019). *Scalable Micro-planned Generation of Discourse from Structured Data*. In **Computational Linguistics** journal 2019.
2. Surya, S., Mishra, A., Laha, A., Jain, P., Sankaranarayanan, K. (2019). *Unsupervised Neural Text Simplification*. In **Association for Computational Linguistics (ACL) 2019**.
3. Laha A.<sup>†</sup>, Chemmengath S.A.<sup>†</sup>, Agrawal P., Khapra M., Sankaranarayanan K., Ramaswamy H.G. (2018). *On Controllable Sparse Alternatives to Softmax*. In **Neural Information Processing Systems (NeurIPS), 2018**.
4. Nema P., Shetty S., Jain P., Laha A., Sankaranarayanan K. and Khapra M. (2018). *Generating Descriptions from Structured Data Using a Bifocal Attention Mechanism and Gated Orthogonalization*. In **NAACL-HLT, 2018**.
5. Jain P., Laha A., Sankaranarayanan K., Nema P., Khapra M. and Shetty S. (2018). *A Mixed Hierarchical Attention based Encoder-Decoder Approach for Standard Table Summarization*. In **NAACL-HLT (Short), 2018**.
6. Nema P., Khapra M., Laha A., Ravindran B. (2017). *Diversity driven Attention Model for Query-based Abstractive Summarization*. In **Association for Computational Linguistics (ACL) 2017**.
7. Pahuja V.<sup>†</sup>, Laha A.<sup>†</sup>, Mirkin S., Raykar V., Kotlerman L., Lev G. (2017). *Joint Learning of Correlated Sequence Labelling Tasks Using Bidirectional Recurrent Neural Networks*. In **INTERSPEECH 2017**. <sup>†</sup> Equal Contribution.
8. Laha A. and Raykar V. (2016). *An Empirical Evaluation of various Deep Learning Architectures for Bi-Sequence Classification Tasks*. In **International Conference on Computational Linguistics (COLING) 2016**.

## Talks and Tutorials

---

- [28 Jul, 2019] Delivered half-day tutorial at ACL 2019. **Storytelling From Structured Data and Knowledge Graphs: An NLG Perspective**. Venue: Association for Computational Linguistics (ACL) 2019 in Florence, Italy. Details: [Link](#).
- [07 Mar, 2018] **Generating Natural Language Descriptions from Structured Data**. Venue: IBM Research-IISc Workshop on Knowledge and Learning on March 07, 2018 organised at IISc. Details: [Link](#).
- [20 Jan, 2018] **Tutorial on Natural Language Processing and Generation**. Venue: Technical Talk on the broader theme of 'Cognitive Analytics and NLP' as part of the annual science and cultural fest at IISc, Bangalore.

## Awards and Achievements

---

- **Top 50% reviewer** in NeurIPS 2019 conference.
- IBM India **Distinguished Paper Award** for consecutive years **2017** (for paper at ACL 2017) and **2018** (paper at NeurIPS 2018).
- Part of **IBM Project Debater** team which received news coverage for a **live debate** ([NYT](#) | [Forbes](#) | [The Guardian](#)).
- **RANK 1** in the selection interviews of IISc, Bangalore (2011) - [Link](#). | **All-India Rank 82** in GATE CS 2010 (out of 107086 - 99.92 percentile).
- **RANKED in Top-3** during BE in Computer Science department of Jadavpur University, Kolkata (2010).
- Awarded the prestigious **Jagadish Bose National Science Talent Search Senior Scholarship** in 2006.

## Skills

---

<b>Programming Languages</b>	Python, Java, C++, C#, C
<b>ML/DL/NLP Frameworks</b>	PyTorch, Tensorflow, Jax, Numpy, Scikit-Learn, MATLAB, Theano, Keras, NLTK, Spacy, Allen-NLP
<b>Spoken Languages</b>	English (fluent), Bengali (native tongue), Hindi (fluent), French (beginner)

## Other Publications/Patents

---

- Patel K., [Laha A.](#), Chemmengath S. A., Mishra A., Bhattacharyya P., Sankaranarayanan K. (2019). *Match My Explanations: Improving Neural NLP Using Gold Explanations*. [Under Review].
- Agrawal S., Sankaran A., [Laha A.](#), Shrivastava D., Chemmengath S. A., Sankaranarayanan K. (2019). *What is Deemed Computationally Creative?*. In **IBM Journal of Research and Development**, 2019.
- Shrivastava D., Chemmengath S.A., [Laha A.](#), Sankaranarayanan K. (2017). *A Machine Learning Approach for Evaluating Creative Artifacts*. In **SIGKDD Workshop on Machine Learning for Creativity (ML4Creativity)**, 2017.
- Jain P., Agrawal P., Mishra A., Sukhwani M., [Laha A.](#), Sankaranarayanan K. (2017). *Story Generation from Sequence of Independent Short Descriptions*. In **SIGKDD Workshop on Machine Learning for Creativity (ML4Creativity)**, 2017.
- [Laha A.](#), Dey P., Bandyopadhyay S., Mukherjee S., Vallam R.D., Garg D., Ray R., Narahari Y. (2012). *Prediction Markets: Connections with Proper Scoring Rules*. **Technical Report**.
- Mukherjee S., Roy B., [Laha A.](#) (2010). *An Efficient Cryptographic Hash Algorithm (BSA)*. In **National Workshop on Cryptology**, 2010.
- **Inventor of SIX patents** filed by IBM at the US Patents and Trademarks Office (USPTO).

## Selected Projects

---

- **Interpretability of Natural Language Classification Models**: We study different approaches for interpretability like LIME, Integrated Gradients, Right for Right Reasons in the tasks of Sentiment Classification, Textual Entailment (classification) as well as POS Tagging and Neural Machine Translation. Our work proposes using ground truth explanations (say from SentiWordNet for sentiment classification) as a feedback for correction of the classification model. For the Sentiment Classification, we also create a challenge dataset based on adversarial perturbations, which will be released soon. This work is currently under review.
- **Controllable Sparse Alternatives to Softmax**: We study techniques for converting a real vector to a probability distribution. We propose a unified framework which leads to understanding of existing approaches incl. softmax. Our framework also enables providing explicit controls to vary sparsity of the output distribution. Our proposed convex loss functions enable achievement of greater sparsity for multi-label classification task. We also show encouraging results for sparse attention in NLG tasks. [NeurIPS '18 Paper](#).
- **Structured Data to Text Generation**: This can be categorized into two paradigms - (a) Supervised End-to-end attention based seq2seq approaches for summarization of tabular data ([NAACL '18 Paper 1](#), [Paper 2](#)), and (b) Unsupervised coherent description generation from tabular data, in a way that is scalable and adaptable to newer domains and does not require parallel data [CL '19](#).
- **Query-based Abstractive Summarization**: In this work, we propose a model for the query-based summarization task based on the encode-attend-decode paradigm with two key additions (i) a query attention model (in addition to document attention model) which learns to focus on different portions of the query at different time steps (instead of using a static representation for the query) and (ii) a new diversity based attention model which aims to alleviate the problem of repeating phrases in the summary. [ACL '17 Paper](#).
- **Sequence classification used for claim and evidence detection in argument mining**: We empirically evaluated multiple RNN and CNN architectures for bi-sequence classification and applied them for context-based claim and evidence detection. [COLING '16 Paper](#).

- **Punctuation and Case prediction for ASR output:** Automatic Speech Recognition (ASR) engines are good at recognizing words from human speech; however, the output from ASR needs to be formatted with punctuation and case to be useful for downstream NLP tasks. In this work, we employed a multi-task approach to simultaneously produce punctuation and casing labels for a stream of words. This solution enabled mining of claims and evidences from human debate recordings. [INTERSPEECH '17 Paper](#).
- **Story generation from incoherent descriptions:** We proposed the new task of generating a coherent story from disconnected text snippets. For this, we implement an RNN based encoder-decoder solution to produce comprehensive coherent story. [Paper](#).
- **ML Framework for evaluating creativity:** We identified important metrics for measuring creativity and proposed a regression-based learning framework for evaluating these metrics and combining them to produce a creativity score for an artifact. [Paper](#).
- **Gene Prioritization from Heterogeneous Data Sources:** In this work, we use graph-based learning-to-rank methods to learn a ranking of genes from each individual data source represented as a graph, and then apply rank aggregation methods to aggregate these rankings into a single ranking over the genes. | *Advisor:* Dr. Shivani Agarwal. [MS Thesis](#).
- **Investigation of the incentive compatible scoring rules for prediction markets:** We applied Prediction Markets in IISc Prediction League (to predict winners of DLF Indian Premier League 2012), an IISc-wide web application based on open-source Zocalo framework | *Advisor:* Prof. Y. Narahari. [Project Report](#).

## Organizational Activities

---

- **PC member** of ICML 2020, AAAI 2020, NeurIPS 2019, ACL 2018-20, EMNLP 2019, and COLING 2018.
- **Reviewer** of AAAI 2019, NeurIPS 2017-18, VLDB 2017.
- **PC member** of SIGKDD Workshop on Machine Learning for Creativity ([ML4Creativity](#)), 2017.
- **Co-organizer** of [TagMe 2014](#), a nation-wide Machine Learning Contest (organized by IISc Bangalore and Microsoft).
- **Lead Organizer** of [TwitMiner 2013](#), a nation-wide Machine Learning Contest (organized by IISc Bangalore and Amazon.com).

## Extra Curricular Activities

---

- *Hobbies:* Travelling, Photography, Reading Fiction (Crime, Suspense Thrillers, Science-Fiction).
- *Sports:* Running, Badminton, Table Tennis, Lawn Tennis, Pool, Cricket.

## References

---

1. Prof. **Aaron Courville** (Associate Professor, University of Montreal), [aaron.courville@gmail.com](mailto:aaron.courville@gmail.com).
2. Dr. **Karthik Sankaranarayanan** (Senior Technical Staff Member, IBM India Pvt. Ltd.), [kartsank@in.ibm.com](mailto:kartsank@in.ibm.com).
3. Prof. **Mitesh M. Khapra** (Assistant Professor, CSE, Indian Institute of Technology (IIT), Madras), [miteshk@cse.iitm.ac.in](mailto:miteshk@cse.iitm.ac.in).