Gaurav Arora

Curriculam Vitae

⊠ gauravarora.daiict@gmail.com 🖻 gauravaror.github.io github:gauravaror

Education

July 2018 – Master of Science (Computer Science) 87.55/100. Present The University of Melbourne Key Courses: Statistical Machine Learning, AI Planning for Autonomy, Stream Computing, Stochastic Optimisation, Web Search and Text Analysis GPA: 8.07/10. August 2008 – **B.Tech (ICT)**

July 2012 Dhirubhai Ambani Institute of Information and Communication Technology

Publications

- Gaurav Arora, Afshin Rahimi, Timothy Baldwin: A few simple techniques to overcoming catastrophic forgetting in continual learning for NLP (To be submitted to EMNLP 2020)
- · Gaurav Arora, Afshin Rahimi, Timothy Baldwin: Does an LSTM forget more than a CNN? An empirical study of catastrophic forgetting in NLP, ALTA 2019
- Parth Mehta, Gaurav Arora, Prasenjit Majumder: Attention based Sentence Extraction from Scientific Articles using Pseudo-Labeled data, arXiv preprint arXiv:1802.04675.
- Khushboo Singhal, Gaurav Arora, Smita Kumari: SMS Normalization for FAQ Retrieval, **FIRE 2011**

Selected Projects

December A few simple techniques to overcome catastrophic forgetting in continual learning setup 2019 - for NLP.

Present Master's project

Advisor: Prof. Timothy Baldwin

Advisor: Prof. Timothy Baldwin

We investigated and found following simple techniques can reduce catastrophic forgetting:

• Reducing temperature with each added task reduces forgetting.

o Adding task information to all hidden layers input helps reduce forgetting.

Feburary Using functional composition to improve dynamic mollifiers loss.

2019 – July Course project

Advisor: Felisa Vázquez-Abad As part of the course project in stochastic optimisation class, we built on the previous work proposing 2019 a dynamic mollifier based loss to avoid divergence of parameter norm observed with cross-entropy loss. Composing dynamic mollifier with quadratic function achieved accuracy comparable to cross-entropy loss retaining the benefits of dynamic mollifier loss.

December Does an LSTM forget more than a CNN?.

2018 - Master's project

- December We conducted an investigative study on catastrophic forgetting while training NLP tasks. We established 2019 the following key insights:
 - CNN forgets less than LSTM because of max-pooling operation.
 - In continual learning setup, fine-tuning pre-trained embeddings results in a sub-par performance.
 - Following curriculum learning schedule reduce forgetting.
 - Deeper networks forget more than shallow networks.

November Abstract generation using Summarization.

- 2017 June We proposed a weakly supervised abstract generation engine for scientific documents with following novel 2018 contributions:
 - Used topic modelling based context embedding technique to generate the summary.
 - Proposed attention mechanism to identify cue phrases indicative of summary worthy sentences.
 - Our method preserved the overall structure of a document in the abstract.

Experience

August 2014 – Software Developer.

Present Frisk Search Pty. Ltd.

I am part of a team responsible for building scalable search solution for enterprises. In this role, I have been responsible for machine learning component of a wide range of projects including snippet generation, image indexing, extracting tables from images, query boosting, expansion and suggestions, decoupling indexing and text extraction.

Remote

Bangalore, India

Remote

November Junior Research Fellow, IMPRINT Project.

2017 – JuneDhirubhai Ambani Institute of Information and Communication TechnologyGujarat, India2018IMPRINT project aims to build a platform for event monitoring in Indian languages. I was responsible for
building the domain classification engine and evaluation framework for the project. I worked primarily on
improving the accuracy of domain classification engine using bootstrapping methods in limited labelled
data setup.

August 2014 – CTO, Co-founder.

January 2017 Shopick, Acquire Technologies Pvt. Ltd. Delhi, India I was part of the three-member founding team building machine learning based technology solutions to solve pain points of offline retailers. We built various solution including a chatbot end-point for brands to showcase their offers and products to the customers.

August 2012 – Software Engineer.

May 2014 Hewlett-Packard ISO

Worked as a reliability engineer for printer communication team in HP ePrint Platform. I worked primarily on reducing the memory footprint of the XMPP connections and fixing memory leak issues. On multiple occasions, the leadership team recognised me for fixing critical production issues.

May 2012 – **Summer Internship**.

August 2012 Xapian, Google Summer of Code I contributed a language model based weighting scheme and evaluation framework in Xapian.

August 2011 – Research Assistant.

July 2012 Dhirubhai Ambani Institute of Information and Communication Technology Gujarat, India I worked as a research assistant in 12 universities consortium research project to build an Indian language search engine. I was responsible for building the relevance judgment tool and Gujarati stemmer.

Talks

Does an LSTM forget more than a CNN?	Sydney NLP Meetup, ALTA 2019
Hands-on Session, Machine Learning	University of Calcutta, 2018

Positions of Responsibility

Reviewer, Forum for Information Retrieval Evaluation	2019
Mentor, Google Summer of Code	2014, 2016, 2017

Skills

Expertise Natural Language Processing, Machine Learning, Information Retrieval Area

Tools and C, C++, Python, J2SE, Perl, Ruby, PyTorch, Xapian

Languages