Ameya Godbole

ameyag416@gmail.com | +1 (617) 676-8873
ameyagodbole.github.io | LinkedIn://ameyag416 | Github://ameyagodbole

EDUCATION

UNIVERSITY OF SOUTHERN CALIFORNIA — Los Angeles, CA | Aug 2021 - Present
PhD in Computer Science

UNIVERSITY OF MASSACHUSETTS AMHERST — Amherst, MA | Aug 2018 - May 2020
MS in Computer Science (GPA: 4.0/4.0)

IIT (INDIAN INSTITUTE OF TECHNOLOGY) GUWAHATI — Guwahati, India | Aug 2014 - May 2018
B.Tech in Electronics & Communication Engineering (Major GPA: 9.15/10)
Minor in Computer Science & Engineering (Minor GPA: 8.8/10)

PUBLICATIONS

[1] R Das*, **A Godbole** *, A Naik, E Tower, M Zaheer, H Hajishirzi, R Jia, A McCallum. **"Knowledge Base Question Answering by Case-based Reasoning over Subgraphs"** . ICML 2022
[2] R Das, M Zaheer, D Thai, **A Godbole** , E Perez, JY Lee, L Tan, L Polymenakos, A McCallum. **"Case-based Reasoning for Natural Language Queries over Knowledge Bases"** . EMNLP 2021
[3] R Das, **A Godbole** , N Monath, M Zaheer and A McCallum. **"Probabilistic Case-based Reasoning for Open-World Knowledge Graph Completion "** . Findings of EMNLP 2020
[4] R Das, **A Godbole** , S Dhuliawala, M Zaheer and A McCallum. **"A Simple Approach to Case-Based Reasoning in Knowledge Bases"** . AKBC 2020 [Best Paper Runner-up]
[5] **A Godbole** *, D Kavarthapu*, R Das*, Z Gong, A Singhal, H Zamani, M Yu, T Gao, X Guo, M Zaheer and A McCallum. **"Multi-step Entity-centric Information Retrieval for Multi-Hop Question Answering".** MRQA-EMNLP 2019 [Best Paper]
[6] **A Godbole** *, R Das*, M Zaheer, S Dhuliawala and A McCallum. **"Reasoning over Chains of Facts for Explainable Multi-hop Inference"** . TextGraphs-EMNLP 2019 [Shared task 1st place entry]
[7] **A Godbole** *, S Bhat* and P Guha. **"Progressively Balanced Multi-class Neural Trees"** . NCC 2018

PREPRINTS

[1] **A Godbole** , R Jia. **"Benchmarking Long-tail Generalization with Likelihood Splits"** . arXiv Oct 2022
[2] **A Godbole** , A Dalmia and S Sahu. **"Siamese Neural Networks with Random Forest for detecting duplicate question pairs"** . arXiv Jan 2018

EXPERIENCE

INFORMATION EXTRACTION AND SYNTHESIS LABORATORY — Jun 2020 - Jun 2021 | Amherst, MA
Research Fellow

- Contributor to the **OpenReview** conference platform most recently used to host **ICLR 2021** .
- Incorporated a language model based system trained on citation and authorship graphs into **OpenReview Expertise** . This system generates affinity scores between submitted papers and available reviewers.
- Added features to the fairness-constrained matching algorithms of **OpenReview Matcher** , which solve an optimization problem to assign papers for review given pre-computed affinity scores.

SRI INTERNATIONAL — May 2019 – Aug 2019 | Menlo Park, CA
Machine Learning Intern

- Member of a team of researchers from the **Artificial Intelligence Center (AIC)** participating in the **DARPA** program: Radio Frequency Machine Learning Systems **(RFMLS)** .
- Developed a simulator of the physical system for faster experiment turnaround time.
- Applied and benchmarked **reinforcement learning** & **imitation learning** based approaches to control an antenna array for RF monitoring showing improvements over existing baselines.

CENTRE FOR DEVELOPMENT OF ADVANCED COMPUTING May 2016 – July 2016 | Pune, India
Software Development Intern

- Designed and contributed to a molecular dynamics simulator at CDAC.
- Implemented the principles of **parallel computing** with MPI to make a simulator capable of utilizing the processing capabilities of a CPU cluster for particle dynamics simulations.

PROJECTS

CASE-BASED REASONING (CBR) [Code]
IESL, UMass Amherst Jan 2020 – Apr 2022

- Applied case-based reasoning to **knowledge base completion** by retrieving reasoning rules from similar entities and applying them to the query entity, resulting in a non-parametric, interpretable, and performant system.
- Implemented a probabilistic framework to weight rules using prior co-occurrence frequency and precision as a reasoning rule. The resulting system matches/outperforms baselines.
- Developed an online setting with **dynamically growing knowledge graphs** and demonstrated the advantage of our framework over previous methods.
- Applied a neuro-symbolic CBR approach to **semantic parsing** achieving competitive performance on several knowledge base question answering tasks. This system is capable of generating programs with **new KB relations without any further training** by incorporating a few human-labeled examples in the case memory.
- Extended this system to perform analogical reasoning using graph neural networks. The system learns knowledge-base question answering **(KBQA) just from denotations** requiring less human annotation effort.

DATA CENTER ENERGY PORTFOLIO OPTIMIZATION
Dr. Mohammad Hajiesmaili & Dr. Philip Thomas, UMass Amherst Aug 2019 – Apr 2020

- Framed the task of energy procurement for data centers as a reinforcement learning problem. The agent manages an onsite battery to meet instantaneous energy demand while minimizing supply costs.
- Demonstrated that **imitation learning** approaches perform better than **reinforcement learning** for this task.

ENTITY-CENTRIC INFORMATION RETRIEVAL
IESL, UMass Amherst Jan 2019 – May 2019

- Developed a document retrieval technique that uses information of entities present in the initially retrieved evidence to learn to 'hop' to other relevant evidence.
- In a setting, with more than **5 million** Wikipedia paragraphs, our approach leads to significant boost in retrieval.
- The retrieved evidence also increased the performance of an existing QA model (without any training) on the HotpotQA benchmark by **10.59 F1** .
- Won **1st place at TextGraphs 2019** by applying the same principles to Explanation Regeneration.

PROGRESSIVELY BALANCED MULTI-CLASS NEURAL TREES [Code]
Dr. Prithwijit Guha, Dept. of EEE, IIT Guwahati Aug 2017 – May 2018

- Proposed and tested an entropy impurity based objective function for incorporating a learnable perceptron into the decision tree framework.
- The learned classifier achieves comparable accuracy with fewer test time computations than an MLP.

QUORA QUESTION PAIRS (KAGGLE) [Website][Report]
Collaborator: Aman Dalmia Apr 2017 - Jun 2017

- Trained a Siamese Gated Recurrent Unit (GRU) RNN over sentence pairs to detect duplicate questions.
- Our team secured a position in the top 25% among 3000+ teams on Kaggle.

SERVICE

REVIEWER

- ACL Rolling Review: Jan 2022, Sep 2022
- Empirical Methods in Natural Language Processing (EMNLP): 2022
- Neural Information Processing Systems (NeurIPS): 2021 (emergency), 2022
- International Conference on Machine Learning (ICML): 2022

COURSEWORK

GRADUATE

- AI: Artificial Intelligence, Reinforcement Learning, Probabilistic Graphical Models, Machine Learning, Automated Knowledge Base Construction
- Systems: Distributed & Operating Systems
- Theory: Algorithms for Data Science, Advanced Algorithms

UNDERGRADUATE

- Machine Learning: Spoken Language Systems, Computer Vision, Pattern Recognition & Machine Learning
- Electronics & Communication: Advanced Topics in Random Processes, Information Theory & Coding, Image Processing, Communication Networks
- Mathematics: Mathematical Techniques for Control and Signal Processing, Linear Algebra, Mutivariable Calculus, Differential Equations

SCHOLASTIC ACHIEVEMENTS

- Secured merit-based change of discipline from Electronics and Electrical Engineering to Electronics and Communication Engineering in July 2015
- Secured All India Rank 1893 in JEE Advanced 2014 (out of 126k)
- Secured All India Rank 547 in JEE Mains 2014 (Percentile score: 99.87)
- Qualified for the state level of the Regional Mathematics Olympiad by securing top position in the Mumbai Regional stages of 2013 and 2012

TECHNICAL SKILLS

Programming Languages	Python, C++
Frameworks/Libraries	PyTorch, TensorFlow*, MATLAB*
Miscellaneous	Numpy, Pandas, scikit-learn, OpenMP*, MPI*

* Elementary proficiency