KABIR AHUJA

+91 (946)-334-4066 ◇ kabirahuja2431@gmail.com
LinkedIn ◇ GitHub
Google Scholar

EDUCATION

Birla Institute of Technology and Science (BITS), Pilani *August 2015 - May 2019*
B.E. (hons) in Chemical Engineering
(advised by Dr. Ajaya Kumar Pani)
Overall GPA: 9.45/10 (Department Rank 1), Graduated with Distinction

EXPERIENCE

Udaan.com, Bangalore July 2020 - Present
Data Scientist

- A Cart Based Recommendation System using a Session-Aware model that would recommend items based on the ones that have already been added to the cart. Observed a 4% adoption of the widget with 5% increase in the order value.
- Developed models for User Personalized Recommendations using algorithms like Item-KNN, Hybrid Model with WARP Loss and Lambda-MART. The recommendations were deployed in the For-You-Page in the app as well as in the form of personalized widgets. Observed 10% adoption.
- Currently working on Repeat Recommendations for a Buy-It-Again type of widget to recommend items that a user purchases repeatedly at the appropriate time.

Microsoft Research, Bangalore January 2020 - July 2020
Research Intern *Advisor: Dr. Navin Goyal*

- Worked on analyzing and contrasting the learning capabilities of **Recurrent Neural Networks and Transformers in recognizing Context Free Languages in finite precision**.
- Investigated the extent of generalization exhibited by these models in acquiring the underlying rules of a formal language and how it is impacted by architectural choices in a model like depth, number of attention heads, the type of Positional Encodings used etc.
- Developed visualizations for the intermediate representations and attention weights of these models to better understand the algorithms learned by these for solving the given tasks.

Indian Institute of Science, Bangalore June 2019 - December 2019
Research Assistant *Advisor: Prof. Partha Talukdar*

- Worked on developing a sequence to sequence model for **Syntactically Controlled Paraphrase Generation** using Bidirectional LSTM, Tree LSTM and Pointer Generator Networks in Pytorch.
- Incorporated the desired syntax using Constituency Parse Trees and implemented a gated mechanism to selectively choose leaf node representations while decoding.
- Trained on Quora Question Pairs (QQP) and ParaNMT datasets, our model outperforms existing methods by a significant margin on metrics like BLEU, ROUGE, METEOR and Tree-Edit Distance.

Massachusetts Institute of Technology August 2018 - December 2018
Visiting Student Researcher *Advisor: Prof. William H. Green*

- Designed two different Reinforcement Learning based approaches for **optimizing molecular geometries**.
- Implemented Self attention networks, Permutation equivariant networks and Gated recurrent units (GRUs) to define the policy and value networks and trained them using Proximal Policy Optimization(PPO).
- Achieved better performance (30% fewer steps on average) than the state of the art optimization schemes for geometry optimization like BFGS, L-BFGS, FIRE etc.

Ecozen Solutions, Pune May 2018 - July 2018
Data Science Intern

- Implemented Density Estimation and Recurrent Neural Networks to determine anomalies in the equipements in real time.
- Utilized Random Forests, and Logistic Regression algorithms to model other problems like Refrigerant Leak Detection and Solar Panel Dust detection.
- The algorithms implemented were pushed in the companys products Ecofrost and Ecotron.

PUBLICATIONS

1. **Kabir Ahuja**, William H Green, Yi-Pei Li. **Learning to Optimize Molecular Geometries Using Reinforcement Learning**. *Journal of Chemical Theory and Computation.*
2. Satwik Bhattamishra, **Kabir Ahuja**, Navin Goyal. **On the Practical Ability of Recurrent Neural Networks to Recognize Hierarchical Languages**. *Proceedings of the 28th International Conference on Computational Linguistics.*
3. Satwik Bhattamishra, **Kabir Ahuja**, Navin Goyal. **On the Ability and Limitations of Transformers to Recognize Formal Languages**. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP).*
4. Ashutosh Kumar, **Kabir Ahuja**, Raghuram Vadapalli, Partha Talukdar. **Syntax-guided Controlled Generation of Paraphrases**. *Transactions of the Association for Computational Linguistics (TACL).*
5. **K. Ahuja**, AK Pani **Software sensor development for product concentration monitoring in fed-batch fermentation process using dynamic principal component regression**. *2018 International Conference on Soft-computing and Network Security (ICSNS).*

NOTABLE ACHIEVEMENTS

- Best Paper Award in COLING 2020 for our paper: **On the Practical Ability of Recurrent Neural Networks to Recognize Hierarchical Languages**
- Recipient of university's Merit Cum Need Scholarship awarded to top 2% students of a batch.
- Appeared among top 10 finalists in the Flipkart Grid Challange 2019 originally participated by 6000 teams.
- Among top 1 percentile of the 1,400,000 candidates who appeared for the JEE-Mains exam (2015).

PERSONAL PROJECTS

Fine Tuning BERT for Sentiment Analysis
Code *Tutorial*

- Using Transformers library in Pytorch, implemented a BERT based model for classifying movie reviews in the Stanford Sentiment Tree Bank.
- Achieved an accuracy of 82.6% and 88.3% with no fine-tuning and with fine-tuning respectively on the dev set.

Question Answering System using Deep Learning
Code

- Implemented a Machine comprehension system using Bidirectional GRUs and different attention mechanisms like Self Attention, Scaled Dot Attention, Bidirectional Attention Flow in Tensorflow
- Stanford SQUAD dataset was used for training the model. Achieved 0.65 F1 score on Dev set.

SKILLS

Programming Languages

(i) Python (ii) C/C++ (iii) Java (iv) MATLAB

Libraries and Frameworks

(i) Pytorch (ii) Tensorflow (iii) RDKit (iv) Numpy (v) Matplotlib

TEACHING EXPERIENCE

Neural Networks and Fuzzy Logic January 2018 - May 2019
Teaching Assistant

- Helped the course be more hands-on by introducing programming assignments for the course encompassing topics like K-nearest neighbours, Logistic Regression and Deep Autoencoders.
- Hosted a Kaggle competition where students were asked to implement a Recommendation System.
- Conducted workshops on Python programming, Neural Networks in practice and Deep Learning frameworks to provide a practical flavour to the concepts taught in class.

National Service Scheme August 2015 - May 2017
Volunteer

- Gave ∼120 hours of lectures on Physics, Chemistry and Mathematics to underprivileged students.
- Prepared assignments and conducted tests to help them score well in their exams.

TUTORIALS

- Adversarial Robustness (slides)
- Policy Gradient Methods in Reinforcement Learning (slides)(code)
- Variance Reduction in Policy Gradients (slides)(code)
- Neural Networks in Practice (slides)

EXTRA-CURRICULAR

- **Blogging** - Post tutorials on different Machine Learning concepts under the Startup Publication at Medium.
- **Movie and TV reviews** - Actively review newly released films and TV shows.