208 Williams St
Ithaca, NY 14850
(+91) 9820208300

Kabir Walia
kabirwalia8300.github.io

github.com/kabirwalia8300
linkedin.com/in/kabirwalia
ksw63@cornell.edu

Education

Cornell University – GPA: 3.68/4.00 (Dean's List)   Aug 2018 – May 2022, Ithaca, NY
College of Engineering – B.S Computer Science with a minor in Operations Research and Management Science

Relevant Coursework – Natural Language Processing, Machine Learning for Intelligent Systems, Computer Vision, Algorithms, Database Systems, Computer System Organization, Optimization, Probability and Statistics, Data Structures, Discrete Structures

Professional Experience

Amazon – Research Engineer Intern   Jun 2020 – Present, Bangalore, India
Amazon India Machine Learning team
- Owned development and integration of an automated data labeling workflow (Active Learning) into an internal AutoML tool used by 50+ non-technical business teams globally
- Decreased costs of labeling data for text-classification tasks by over 65% while maintaining 0.95 precision and recall metrics
- Improved upon classic active learning algorithms like uncertainty-based sampling and query by committee by combining them with diversity strategies like k-Means clustering and representative sampling
- Built an interactive UI for facilitating the labeling task using Streamlit, AWS Sagemaker, Ground Truth, Step Functions, S3

Cornell NLP Group – Research Lead   Sept 2019 – Present, Ithaca, NY
Managed a team of 4 undergraduate researchers in Prof. Claire Cardie's natural language processing lab
- Explored transformer and attention-based generative neural network methods for the counterargument generation task
- Developing a novel architecture leveraging the recurrence mechanism of Transformer-XL and conditional language modelling system of the CTRL model to emulate argumentation strategies like card-stacking for producing higher quality counterarguments

Mu Sigma – Software Engineering Intern   Jun 2019 – Aug 2019, Austin, TX
India's largest data-driven management consulting firm
- Led a team of 3 interns to setup the 1st US-based AI interactive lab to showcase technological capabilities during client demos
- Implemented anomaly detection algorithms like recursive residuals, achieving 71% accuracy on a client's simulated IoT data
- Developed an automated office assistant, *muNerva*, using Flask, Alexa Voice Services, Redis, and OpenCV

Cally – Data Analyst (Remote)   May 2019 – Sept 2019, Ithaca, NY
New York -based startup combining team science and data science to revolutionize team dynamics
- Analyzed 7+ factors identified by Cornell researchers as stimuli of negative mood dissemination through teams
- Increased employee participation (15%) by restructuring surveys based on regression models used to predict response rates

Leadership

Cornell Cup Robotics – CS Team Lead   Feb 2019 – Present, Ithaca, NY
Oversaw development of a distributed system to control Minibot, an educational robot to teach children to code using Google Blockly
- Collaborated with Systems, Electrical and Mechanical sub-team leads to define workflows and implement objectives for Minibot
- Managed an 18-member software team to complete computer vision, database management, and computer networking projects
- Optimized the single-camera vision system by decreasing sensor to GUI lag from ~35s to ~5s by restructuring code base
- Regulated public GitHub, led scrums and code reviews, and pitched marquee features to C–suite of partnering ed-tech startup

Cornell University – Teaching Assistant and AEW Instructor   Aug 2019 – Present, Ithaca, NY
CS 4786/5786 (Machine Learning for Data Science), CS 1110 (Python) & MATH 2940 (Linear Algebra)
- Selected as 1 of 2 sophomore teaching assistants for a graduate level machine learning course with 180+ students
- Guided office hours, created and graded assignments, and handled questions on Piazza to assist students' learning experience
- Lectured on linear algebra concepts in a weekly workshop to a class of 18 students (on campus and on Zoom)

Projects

Indian Institute of Management (IIM) – COVID-19 Modeling   Apr 2020 – June 2020, Mumbai, India
- Chosen as part of a 4-student team from a pool of 250+ applicants to collaborate with an IIM professor for a research project
- Applied Bayesian statistics to expand on the SIR epidemiological model to improve estimation of model parameters
- Prepared a quantitative report to be submitted to the Indian government with quarantine/testing recommendations

TextDirect – SMS-based routing (Send "TextDirect" to (+1) 256-488-4302)   Dec 2019 – Jan 2020, Ithaca, NY
- Launched a navigation service providing step by step Google Maps directions as text messages (Twilio, Google Cloud Platform)
- Aimed at benefitting feature phone users and users in areas where internet/mobile data penetration is limited

Skills & Activities

Programming: Python | Java | Scala | PyTorch | TensorFlow | AWS | Google Cloud Platform | Huggingface | Git | Hadoop | Docker
General: Leadership | Design Thinking | Project Management | Agile | Scrum | Resource Planning | Teaching
Activities: Theta Tau Professional Engineering Fraternity | Cornell European Business Society | Intramural Soccer
Interests: Linguistics | Tennis | Acting | War Documentaries | Game Theory | Quentin Tarantino | Cooking | Manchester United