# Kabir Walia

208 Williams St
Ithaca, NY 14850
(+1) 571-831-1162

kabirwalia8300.github.io

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

## Education

**Cornell University** – GPA: 3.76/4.00 (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¶, Machine Learning for Data Science¶, Algorithms, Learning with Big Messy Data, Computer System Organization, Linear Algebra¶, Probability and Statistics, OOP and Data Structures in Java, Discrete Structures, Introduction to Computing using Python¶
(¶ = teaching assistant)

## Professional Experience

**Microsoft** – Incoming Data & Applied Science Intern                                          Nov 2020 – Present,  Bellevue, WA
*Business Applications Group (BAG)*
- Joining the BAG team at Microsoft as a Data & Applied Science intern for Summer 2021

**Amazon** – Research Engineer Intern                                          Jun 2020 – Sept 2020,  Bangalore, India
*Amazon India Machine Learning team*
- Owned development and integration of an automated data labeling workflow (Active Learning) into an internal AutoML tool to streamline business workflows for 50+ non-technical teams globally
- Decreased costs of labeling data for text-classification tasks by over 65% while maintaining 0.95 precision and recall metrics
- Benchmarked BALD (Bayesian Active Learning by Disagreement) against classical uncertainty and diversity–based sampling methods on 8 Amazon Business datasets – laying the foundation for further internal research on neural active learning
- Built an interactive UI for facilitating the labeling task using Streamlit, AWS Sagemaker, Ground Truth, Step Functions, S3

**Cornell NLP Group** – Undergraduate Researcher                                          Sept 2019 – Present,  Ithaca, NY
*Member of 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 modeling 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

## 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
- Expanded on the SIR epidemiological model to account for quarantine and improved parameter estimation for Indian data
- 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

- **Software:** Python | Java | C | Scala | PyTorch | TensorFlow | AWS | SQL | Databricks | Hadoop | Docker | Git | Bash | Linux
- **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 | Interior Design | Harmonica | Cooking | Manchester United