Ananth Balashankar

ananth@nyu.edu • ananthbalashankar.github.io

Education

- **New York University**
  Ph.D candidate in Computer Science, GPA of 3.985/4 — *2017–Present*
  'Enhancing Robustness through Domain Faithful Machine Learning'
  Advised by Prof. Lakshminaryanan Subramanian and Dr. Alex Beutel (Google AI)
- **IIT Kharagpur**
  B.Tech and M.Tech in Computer Science, GPA of 9.6/10 — *2009–2014*
  Institute Silver medal for best academic performance. Rank 1 out of 98.

Work Experience

- **Google AI** — **New York**
  Student Researcher (part-time), sponsored by Google Student Research Advisorship Program — *Sep 2019–Present*
  - Improved counterfactual robustness over demographic groups for toxicity detection and gendered pronoun resolution NLP tasks in BERT-based models
  - Improved sample efficiency for domain generalization of sentiment analyses and question paraphrasing tasks using controlled text generative models
- **Google** — **Mountain View**
  Software Engineer, Google Play Apps Recommendations Team — *Dec 2015–Oct 2017*
  - Designed ML models with the Google Brain team to improve personalized recommendations of Apps in Google Play store - Authored a Google Research blog post and filed a patent on this work
  - Deployed the first multi-task wide and deep learning and reinforcement learning ranking model with the DeepMind team using Tensorflow (library was open sourced later) to optimize multiple business objectives
  - Built Play Store personalization infrastructure to allow scalable and faster A/B testing of machine learning models by a team of 15 engineers optimizing diversity, safety, engagement and top-line business metrics
  - Primary oncall maintaining the global serving of Google Play Store's personalization stack
- **Google** — **London**
  Software Engineer, Google Play Developer Console Team — *Oct 2014–Dec 2015*
  - Built the IARC Apps rating integration, open alpha-beta testing framework for Google Play developer console

Research Experience

- **New York University**
  Advised by Prof. Lakshminarayanan Subramanian and Dr. Alex Beutel — *Oct 2017–Present*
  Thesis Direction: Enhancing robustness through domain faithful machine learning applied in the domains of natural language, privacy, socio-economics and health.
- **Google Health**
  Summer Research Intern, advised by Dr. Alex Beutel, Narayan Hegde and Apaar Saadhwani — *May-Aug 2019*
  - Improving robustness of medical diagnostic models in Pathology
- **Google AI**
  Summer Research Intern, advised by Dr. Alyssa Lees and Dr. Chris Welty — *May-Aug 2018*
  - Subgroup Fairness via the lens of Causality and Confounding Variables
- **IIT Kharagpur**
  Advised by Prof. Niloy Ganguly — *2013–2014*
  - Stability Analysis of Location based Landmarks
- **Google Summer of Code**
  Advised by Prof. Bart Massey, Portland State University — *May–July-2013*
  - Built open source library for real-time hyperlocal business analytics in shopping malls

- **University of Calgary**
  Advised by Prof. Behrouz Far (MITACS) *May–July-2013*
  - Developed robust distributed systems for detecting emergent behavior
- **Microsoft Research India**
  Advised by Dr. Vishnu Navda (Mobile Systems and Networks group) *May–July-2012*
  - Developed Signal Strength Aware Application Scheduler for Mobile Phones

Publications

- Enhancing Neural Recommender Models through Domain-Specific Concordance. Ananth Balashankar, Alex Beutel, Lakshminaryanan Subramanian. **WSDM '21**.
- Can We Improve Model Robustness through Secondary Attribute Counterfactuals?. Ananth Balashankar, Xuezhi Wang, Ben Packer, Nithum Thain, Ed Chi and Alex Beutel. **EMNLP '21**.
- Fine-grained prediction of food crisis using news. Ananth Balashankar, Samuel Fraiberger, Lakshminaryanan Subramanian. **Accepted at INFORMS '21, IC2S2 '21, Under Review at Science Advances**.
- Learning Faithful Representations of Causal Graphs. Ananth Balashankar, Lakshminarayanan Subramanian. **ACL' 21**.
- VACCINE: Using Contextual Integrity for Data Leakage Detection. Yan Shvartzshnaider, Zvonimir Pavlinovic, Ananth Balashankar, Thomas Wies, Lakshminarayanan Subramanian, Helen Nissenbaum and Prateek Mittal. **WWW '19**.
- Identifying Predictive Causal Factors from News Streams. Ananth Balashankar, Sunandan Chakraborty, Samuel Fraiberger, Lakshminarayanan Subramanian. **EMNLP '19**
- Pareto Efficient Fairness for Skewed Subgroup Data. Ananth Balashankar, Alyssa Lees, Chris Welty, Lakshmi Subramanian. **ICML '19**- AI for Social Good Workshop. **[Best Paper Award]**
- Reconstructing the MERS Disease Outbreak from News. Ananth Balashankar, Aashish Dugar, Lakshmi Subramanian, Samuel Fraiberger. **ACM COMPASS '19**.
- Fairness Sample Complexity and the Case for Human Intervention. Ananth Balashankar, Alyssa Lees. CHI '19 - Bridging the Gap Between AI and HCI Workshop.
- Unsupervised Word Influencer Networks from news streams. Ananth Balashankar, Sunandan Chakraborty, Lakshmi Subramanian. ACL '18 Workshop on Economics and Natural Language Processing (ECONLP).
- Towards Applying Open Domain Question Answering to Privacy Policies. Yan Shvartzshnaider, Ananth Balashankar, Thomas Wies, Lakshminarayanan Subramanian ACL '18 Workshop on Machine Reading for Question Answering (MRQA)
- Causal Inference from News Streams. Ananth Balashankar, Sunandan Chakraborty, Samuel Fraiberger, Srikanth Jagabathula, Lakshminarayanan Subramanian. ICML '18 Workshop on Machine Learning for Causal Inference, Counterfactual Prediction, and Autonomous Action (CausalML).
- Catalyzing Inclusive Agricultural Transformation using Machine Learning. Samuel Fraiberger, Ananth Balashankar, Lakshmi Subramanian. World Bank Technical Report '18.
- (Stable) virtual landmarks: Spatial dropbox to enhance retail experience. Swadhin Pradhan, Ananth Balashankar, Niloy Ganguly and Bivas Mitra. COMSNETS '14.

Working Papers

- Spatio-temporal modeling of urban air quality using low-cost monitors. Shiva Iyer, Ananth Balashankar, William Aeberhard, Ulzee An, Sameeksha Jain, Sujoy Bhattacharya, Guiditta Rusconi, Anant Sudarshan, Rohini Pande, Lakshmi Subramanian. **Under Review at Nature Sustainability**
- Localized Pollution Hotspots: Inferences from a Three-year Fine-grained Air Quality Monitoring Study in Delhi. Shiva Iyer, Ananth Balashankar, Rohini Pande, Anant Sudarshan, Lakshminarayanan Subramanian. **Under Review at Nature Climate Change**
- Learning Overlap-Aware Temporal Prediction Models. Ananth Balashankar, Srikanth Jagabathula, Lakshminarayanan Subramanian. Under Review at AAAI '21.
- Beyond The Text: Analysis of Privacy Statements through Syntactic and Semantic Role Labeling. Yan Shvartzshnaider, Ananth Balashankar, Vikas Patidar, Thomas Wies, Lakshminarayanan Subramanian. Under Review at AAAI '21.
- Improving Robustness through Pairwise Generative Counterfactual Data Augmentation. Ananth Balashankar, Xuezhi Wang, Yao Qin, Ben Packer, Nithum Thain, Ed Chi and Alex Beutel.

- Predicting Angiographic Disease Status: Where to draw the line between demographically decoupled and jointly trained models?. Ananth Balashankar, Alyssa Lees, Srikanth Jagabathula, Lakshminaryanan Subramanian.

Teaching Experience

Course assistant and lab instructor for graduate level courses at NYU:

- **Big Data and ML Systems** (CSCI-GA. 3033-016, Spring 2019). : Taught a class of over 100+ MS in Computer Science and Computer Engineering, Entrepreneurship and Innovation (MS-CEI) students, and designed labs on Spark distributed ML computing platform, PageRank algorithm, deep learning neural network models for text processing, image recognition, graph learning, multi-arm bandits, recommender systems and healthcare inference. This class covered introductory and recent concepts in big data and machine learning systems. The class focused on a broad spectrum of big data computational problems, algorithms and platforms.
- **Foundations of Networks and Mobile Systems** (CSCI-GA. 2630-001 and 002, Fall 2021): Taught a class of 100+ students from Tech MBA, MS-CEI programs with hands-on lab sessions on internet technologies like DNS, HTML, JavaScript,SQL, PHP, React, etc. This was an introductory course that exposed students to the fundamentals of computer networks and mobile systems. The class began with introductory concepts of network protocols across different layers of the network stack including routing, transport, naming, addressing and connected them to the core building blocks of the Internet. Further, the class focused on networking concepts in the evolution of Web-based systems, providing an introduction to data-center networks, clouds and next-generation networks.

Honors and Skills

- **Awards:**
  - Google Student Research Advisor Program Fellowship (2019-2022)
  - Spot bonus for research contributions in the Google Responsible AI team - 2021
  - NYU Harold Grad Memorial Prize for promising Ph.D achievement - 2019
  - Best Paper Award at ICML AI for social good conference workshop - 2019
  - MacCracken Fellowship (2017-22)
- **Patent:** Signal-aware data transfer in cellular networks. Vishnu Navda, Ramachandran Ramjee, Sahil Suneja, Ananth Balashankar. US Patent 8843169
- **Mentorship:**
  - Mentored 4 Master's students at NYU: Aashish Dugaar (2018-19), Vikas Patidar (2019-20), Sriram Ramesh (2021-22) and Sudharsana Kannappan (2021-22) in their Master's research projects.
  - Mentored Civis India, an NGO working on providing access and gathering feedback to drafts of public policies in India, through their grant application process for the Google AI for social impact challenge 2021.
  - Mentored Lev Proleev, a software engineering summer intern in an ML project at Google in 2016.
- **Co-curricular Skills:** General secretary for organizing technical fests and captain of the hardware, software, quiz and product design teams at IIT Kharagpur.