Ahmed Medhat

amedhatm@gmail.com

http://amedhat.github.io

SUMMARY

I have been blessed with a diverse network science and machine learning research career within both academia and industry. Despite not having completed a PhD degree, I have worked in highly selective network science labs at Oxford and Facebook, traditionally only admitting PhD graduates. My research career involved a combination of network science, computational social science, graph learning and the engineering of supervised ML systems and datasets.

My research in Oxford involved using community detection methods to study wikipedia editor networks. While my research within startups and Facebook involved using network science and graph learning to study and create products for large scale social and financial networks, such as researching how social network structure shapes cognitive biases and social participation decisions, and building network and geographic datasets that aided WHO and UNHCR in response efforts to the Covid Pandemic and the Ukrainian refugee crisis, respectively. I have also built extensive experience in the end-to-end process for building predictive ML models, especially in creating algorithms and systems for ground-truth generation of visual and natural language labels, and learning from weakly supervised data.

RESEARCH INTERESTS

Graph Learning, Network Science, Deep Learning, Computational Social Science, Network Medicine, Network Neuroscience, Computational Biology, Weak Supervision

EDUCATION

Oxford University, Oxford, UK
Master of Science, Computer Science (Network Science)   **Oct 2011**
Thesis: "A Network Modeling Approach to Assisting Collaboration in Large Scale Online Environments"
Thesis Grade: 68 (eq. 3.86 GPA)

American University in Cairo, Cairo, Egypt
Bachelor of Science, Electronics Engineering   **Jun 2009**
GPA: 3.81 (*summa cum laude*)

RESEARCH EXPERIENCE

Senior Research Scientist, Network Science, *Facebook*   **Dec 2021 - Present**
Senior member of the network science team, within Facebook's Core Data Science Department. A highly selective research group, with researchers traditionally coming from competitive computer science and computational social science graduate programs.

- Used GNNs and community detection methods for scalable node attribute prediction tasks in networks with billions of nodes and tens of billions of edges.

- Built models to improve conversational health by using whole graph embeddings to detect content integrity issues in comment threads.

- Designed a new multi-layer centrality metric that scales to billion-node, multi-billion edge scale graphs, and benchmarked it against state of the art methods.

- Used network generation models to simulate the impact of the friendship paradox and feedback related perception biases on content creation activity in news feed based social platforms. Resulting in a paper submitted to the Web Conference 2023. As well as 3 further potential research work-streams.

Senior Research Scientist, Data for Social Good, *Facebook* **Dec 2020 - Dec 2021**
Data Science for Social Good is a team of scientists within the Computational Social Science Group, with diverse backgrounds, including machine learning, network science, political science, epidemiology and economics, who are focused on leveraging Meta's mobility data to build datasets and tools to aid in natural and man-made disaster relief.

- **Key Project:** Implemented algorithms for de-biasing geographical datasets derived from Facebook data to make it more representative of global on-ground populations. This required combining granular population density maps, wealth estimates and user demographics to re-weigh the sample of people with location data to be globally representative. This model was used for creating representative human co-location maps and population displacement maps during crises.
- **Impact**: The model had sizable global impact as it was utilized for disaster response in relation to the two key crises of the last 2 years, with colocation data being used for responding to the Covid Pandemic by WHO, and displacement maps being used for responding to the Ukrainian refugee crisis by the UNHCR, to track refugee numbers and distribution across countries. With the model achieving displacement estimates within 1-2% of real on-ground sample estimates.

Principal Data Scientist, Sharing Ecosystems, *Facebook*   **Oct 2015 - Dec 2020**
During my tenure, I was responsible for initiating and leading a variety of projects that aimed to understand key drivers of content sharing behavior and creation of products to fulfill people's sharing needs. Specifically as related to the sharing of original, personal content.
- **2015-2018 Key Project**: Investigated the complex question of what drives people to share less or more content on Facebook. This work was interdisciplinary in nature, drawing on graph learning, network science, economics, mass communication theory, causal inference and network experimentation, to quantify the degree to which different factors contribute to a person's decision to share content. Such as how the friendship paradox shapes a person's sharing rates due to them perceiving their friends receiving more feedback than they actually do.
- **2017-2020 Key Project**: Ran a 3 year effort to create ground-truth data and classifiers for understanding content creation motivations across most of the world's languages.

  - Designed a taxonomy of content creation motivations and secured funding of $20M+ to a build a group for labeling content with that taxonomy.
  - Hired, trained and managed 600+ labelers, working across 3 continents and labeling content across 15 different languages.
  - Designed methods to adequately anonymize and obfuscate private data, to ensure people's privacy was not compromised during the labeling process.
  - Implemented an algorithm for detecting potential adversarial labeler behavior and using active learning to improve label quality
  - Implemented weak supervision methods to expand training set size.
  - Used inverse propensity matching to re-weigh labeled data to be representative of global sharing rates.
  - Worked with ML engineers to internationalize our models across 200+ languages by employing a translate-train learning model, which was shipped across multiple product verticals in the Facebook App.

Chief Data Scientist, *DueDil*   **March 2012 - October 2015**
As a member of the founding team, I built the company's data science capabilities from the ground up. Helping it become one of the top Financial Startups in Europe, and growing the company to over 100 employees and clients to over a million businesses in the process. By the time I left the company, I was managing two applied research teams of around 8 physics, bio-engineering and computer science PhD grads, and data engineers.
- Created a machine learning team that conducted applied research to create novel business datasets. Solving problems such as matching company networks to bank transaction networks using node embeddings.
- Created an analytics team that used causal inference and controlled experimentation techniques to analyze product performance and predict user preferences.

- Published highly influential research related to the UK economy, including research on migrant entrepreneurs that went viral across the UK media, up to being cited by the prime minister's office in political debates.

Researcher, Oxford Internet Institute   **Oct 2011 - Dec 2013**
Analyzed the provenance of knowledge production on Wikipedia, to answer whether the content discussing different cultural groups was produced by and representative of them, as opposed to being produced by people not representing that culture. Work involved a mixture of computational social science, network science, applied machine learning and geo-mapping. Such as utilizing personalized page rank for entity disambiguation and community detection to infer editor locations and ethnicities.

PUBLICATIONS   **Medhat, Ahmed**, and Shankar Iyer. "The Friendship Paradox and Social Network Participation." arXiv preprint arXiv:2211.05288 (2022). *Submitted to TheWebConf 2023*

Iyer, Shankar, Brian Karrer, Daniel Citron, Farshad Kooti, Paige Maas, Zeyu Wang, Eugenia Giraudy, **Ahmed Medhat**, P. Alex Dow, and Alex Pompe. "Large-Scale Measurement of Aggregate Human Colocation Patterns for Epidemiological Modeling." medRxiv (2022). *In final revisions for the Epidemics Journal*

Graham, Mark, Bernie Hogan, Ralph K. Straumann, and **Ahmed Medhat**. "Uneven geographies of user-generated information: Patterns of increasing informational poverty." Annals of the Association of American Geographers 104, no. 4 (2014): 746-764.

Medhat, A. "A Network Modelling Approach to Ranking Collaboration in Large Scale Online Environments" (2012). Workshop on Information in Networks. New York, NY.

PRESS COVERAGE   *Fortune*, "Data scientists are using the most annoying feature on your phones to save lives in Ukraine", 2022.

Facebook Research Blog, "Making our displacement maps more representative ", 2021.

Financial Times, Independent, Telegraph, Huffington Post and BBC on Contribution of Migrant Entrepreneurs to the UK Economy, March 2014

Guardian and Huffington Post on Wikipedia Language Maps, November 2011

OTHER EXPERIENCE   **Angel Investor & Venture Partner**, *Ada Ventures*   **May 2015 - Present**
Driven by my research career and technical interests, I've done over 15 start-up investments and advisory engagements. Many of the founders I've supported built successful companies with hundreds of employees. This has allowed me to form an understanding of how fundamental and applied computational research can drive the invention of new technologies that can transform healthcare, finance, the climate and beyond.

SKILLS   **Programming Languages**
Fluent in Spark, R, Python, C, C++, Java, SQL/Hive

Frameworks
Fluent in PyTorch, TensorFlow, Hadoop/MapReduce, GraphViz

COURSEWORK (OXFORD)   Probability & Computing, Intelligent Systems, Machine Learning, Computational Linguistics, Information Retrieval, Reasoning about Information Update, Computers in Society.