Emaad Manzoor — *Application to the NBER Economics of AI PhD Workshop, 2020*
emaad@cmu.edu / emaadmanzoor.com

Past Experience

I have developed machine-learning methodology for streaming anomaly detection in networks [1], purchase forecasting [2], high-dimensional anomaly detection [3] and automatic expansion of taxonomies [4]. These are published in top-tier machine learning venues, are well-cited[1] and have been taught in advanced computer science coursework[2]. I also have work on interpretable concept discovery from text under review at a top-tier quantitative marketing journal [5].

Current Research Interest: Causal Inference from Text

I am interested in econometric applications enabled by machine-learning, specifically in applications where text plays an important role as the outcome, the treatment, a control, a moderator or an instrument. Textual data is best dealt with non-parametrically due to its high dimensionality and unstructured nature. Hence, econometric applications that rely on text often use machine-learning. However, with machine-learning comes several challenges in (i) consistent estimation and valid inference, (ii) interpretability, and (iii) sensitivity analyses. My research agenda focuses on the causal inference process using machine-learning, with the goal of resolving the aforementioned problems in the specific context of textual data.

I have the following research currently underway:

1. *Influence via Ethos: On the Persuasive Power of Reputation in Deliberation Online:* I use text as a control and double machine-learning to identify the effect of reputation on successful persuasion. This is under review at Management Science.
2. *Identifying Biases in the Language of Peer-Review:* I use a difference-in-differences setup with two reputed machine-learning conferences, one of which switched to double-blind reviewing, to identify the effect of visible author identities on disparities in the language of peer-review. Text is an outcome in this context.
3. *Interpreting Linguistic Heterogeneity:* In this work, I develop a method to interpret heterogenous treatment effects where text is a moderator.

---

[1] https://scholar.google.com/citations?user=TcMyxM0AAAAJ

[2] Advanced Computer Security (CS-563): https://courses.engr.illinois.edu/cs563/sp2020/assignments.html
Advanced Data Mining (EECS-598-008): https://web.eecs.umich.edu/~dkoutra/courses/W19_598/

References

[1] Emaad Manzoor, Sadegh M. Milajerdi, Leman Akoglu. Fast Memory-Efficient Anomaly Detection in Streaming Heterogeneous Graphs. *ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2016*.

[2] Emaad Manzoor, Hemank Lamba, Leman Akoglu. xStream: Outlier Detection in Feature-Evolving Data Streams. *ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2018*.

[3] Emaad Manzoor, Leman Akoglu. RUSH! Targeted Time-limited Coupons via Purchase Forecasts. *ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2017.*

[4] Emaad Manzoor, Dhananjay Shrouty, Rui Li, Jure Leskovec. Expanding Taxonomies with Implicit Edge Semantics. *The Web Conference, 2020*.

[5] Dokyun Lee, Emaad Manzoor, Zhaoqi Cheng. Focused Concept Miner (FCM): Interpretable Deep Learning for Text Exploration. *Under review at Marketing Science (third round)*.

Emaad Manzoor

emaad@cmu.edu / emaadmanzoor.com

EDUCATION

2016 — 2021 **Carnegie Mellon University**
Ph.D. in Information Systems (Heinz College)
Advisors: Dokyun Lee, George H. Chen

2015 — 2016 **Stony Brook University**
Ph.D. in Computer Science (transferred)
Advisor: Leman Akoglu

2013 — 2015 **King Abdullah University of Science and Technology**
M.S. in Computer Science
Advisor: Panos Kalnis

2008 — 2012 **Birla Institute of Technology and Science (BITS) Pilani**
B.E. in Computer Science

SELECTED AWARDS

2020 McKinsey & Company PhD Fellowship
2019 INFORMS e-Business Cluster Best Paper, finalist
Snap Inc. Research Fellowship, finalist
Jane Street Depth First Learning Fellowship, finalist
2018 Marketing Science Institute Research Grant, co-PI ($10,000)
2016 — 2018 NSF Student Travel Grant for ACM SIGKDD Conferences ($3,050)

SELECTED COURSEWORK

Spring 2020 **Modern Causal Inference**, Edward Kennedy, CMU (Statistics)
Spring 2020 **Foundations of Causal Inference**, Edward Kennedy, CMU (Statistics)
Spring 2019 **Advanced Machine Learning Theory**, Pradeep Ravikumar, CMU (Machine Learning)
Fall 2018 **Advanced Introduction to Machine Learning**, Nina Balcan, CMU (Machine Learning)
Spring 2018 **Statistical Machine Learning**, Larry Wasserman, CMU (Statistics)
Spring 2018 **Econometrics II**, Akshaya Jha, CMU (Heinz)
Fall 2017 **Convex Optimization**, Aarti Singh & Pradeep Ravikumar, CMU (Machine Learning)
Spring 2017 **Econometrics I**, Edson Severnini, CMU (Heinz)
Spring 2017 **Advanced Probability Overview**, Jing Chen, CMU (Statistics)
Fall 2016 **Microeconomics**, Brian Kovak, CMU (Heinz)
Fall 2016 **Intermediate Statistics**, Larry Wasserman, CMU (Statistics)

PUBLICATIONS

WWW 2020 **Expanding Taxonomies with Implicit Edge Semantics**
Emaad Manzoor, Dhananjay Shrouty, Rui Li, Jure Leskovec
The Web Conference, 2020

KDD 2018 **xStream: Outlier Detection in Feature-Evolving Data Streams**
Emaad Manzoor, Hemank Lamba, Leman Akoglu
ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2018

KDD 2017 **RUSH! Targeted Time-limited Coupons via Purchase Forecasts**
Emaad Manzoor, Leman Akoglu
ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2017

KDD 2016 **Fast Memory-Efficient Anomaly Detection in Streaming Heterogeneous Graphs**
Emaad Manzoor, Sadegh M. Milajerdi, Leman Akoglu
ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2016

PAPERS IN PROGRESS

2020+ **Identifying Linguistic Biases in Peer-Review**
Emaad Manzoor, Nihar B. Shah

2019+ **Influence via Ethos: On the Persuasive Power of Reputation in Deliberation Online**
Emaad Manzoor, George H. Chen, Dokyun Lee, Michael D. Smith
Under review at Management Science (first round)

Preliminary versions presented at:
- Statistical Challenges in e-Commerce Research (SCECR), 2020
- International Conference on Computational Social Science (IC2S2), 2020
- ISMS Marketing Science Conference, 2020
- INFORMS Annual Meeting, 2020
- Conference on Information Systems and Technology (CIST), 2019
- ISMS Marketing Science Conference, 2019
- INFORMS Workshop on Data Science, 2019

2018+ **Focused Concept Miner (FCM): Interpretable Deep Learning for Text Exploration**
Dokyun Lee, Emaad Manzoor, Zhaoqi Cheng
Under review at Marketing Science (third round)