Seminar Hora Informaticae   Institute of Computer Science, Prague

Aditi Kathpalia, Department of Complex Systems, Institute of Computer Science, CAS:

Causality and Machine Learning.

Despite the recent success and widespread applications of machine learning (ML) algorithms for classification and prediction in a variety of fields, they face difficulty in interpretability, trustworthiness and generalization. One of the main reasons for this is that these algorithms are building black-box models by learning statistical associations between the given 'input' and its 'output'. Decisions made solely based on 'associational learning' are insufficient to provide explanations and hence difficult to be employed in real world tasks requiring transparency and reliability. To overcome these limitations of ML algorithms, researchers are moving towards 'causal machine learning' by aiding ML decision-making based on causal reasoning and understanding. We will discuss 'the science of causality', its requirements in ML and possible means of integration with ML. We will also compare different ML algorithms based on their performance in learning temporal order/ structure in single time series as well as their ability to classify coupled pairs of time-series based on their cause-effect (or driver-driven) relationship.

References:

[1] Kathpalia, A., & Nagaraj, N. (2021). Measuring causality: The science of cause and effect. Resonance, 26, 191-210.

[2] Pearl, J., & Mackenzie, D. (2018). The book of why: the new science of cause and effect. Basic books.

[3] NB, H., Kathpalia, A., & Nagaraj, N. (2022). Causality Preserving Chaotic Transformation and Classification using Neurochaos Learning. Advances in Neural Information Processing Systems, 35, 2046-2058.

[4] Kathpalia, A., Charantimath, K. P., & Nagaraj, N. (2021). Learning Generalized Causal Structure in Time-series. arXiv preprint arXiv:2112.03085.

———————————————

Aditi Kathpalia is currently a postdoctoral researcher at the Department of Complex Systems, Institute of Computer Science of the Czech Academy of Sciences in Czech Republic. Her research interests include causal inference and causal machine learning, complex systems, information theory and computational neuroscience.

———————————————

HORA INFORMATICAE (meaning: TIME FOR INFORMATICS) is a broad-spectrum scientific seminar devoted to all core areas of computer science and its interdisciplinary interfaces with other sciences and applied domains. Original contributions addressing classical and emerging topics are welcome. Founded by Jiří Wiedermann, the seminar is running since 1994 at the Institute of Computer Science of the Czech Academy of Sciences in Prague.

https://www.cs.cas.cz/horainf