

---

**TERRENCE TRICCO**, Memorial University

*Synthetic generation of multi-modal discrete time series using transformers*

Synthetic data is artificially generated data that contains the same properties as the real data it mimics. There are many benefits to synthetic data. For instance, it allows for the generation of significant quantities of data, more than may be possible through real data collection. Similarly, rare classes or labels can be generated which may be useful for data augmentation or transfer learning. Synthetic data may also be valuable where there exists privacy or security concerns. In this work, we have used a transformer-based model to generate multi-modal, discrete time series data with application to personal financial data. Our data contains multiple classes of events in a highly irregular temporal sequence, where each event may operate on its own timescale with simultaneous dependence upon other events. We have built our architecture and encoded our features specifically to handle multiple patterns found within date and time features. Our transformer-based results have been compared with results from the generative adversarial network (GAN) model, DoppleGANger.