DIFFUSION MODELS FOR CAUSAL FORECASTING

Since 2014, Wiremind has positioned itself as a tech company transforming the world of transport and events with a 360° approach combining UX, software, and AI.

Our expertise lies primarily in optimizing and distributing our clients' capacity. We work on various projects such as ticket forecasting and pricing, 3D optimization of air freight or scraping competitor prices. Our applications are the preferred tool of companies such as SNCF, United Airlines, Qatar Airways or even PSG to visualize, analyze and optimize their capacity.

Dynamic and ambitious, we strive to maintain our technical DNA which is the engine of our success. The company, profitable and self-financed since its creation 10 years ago, is mainly composed of engineers and experts and currently supports the growth of our business model based on "software-as-a-service" solutions.

Numerous Machine Learning (ML) tasks are forecasting problems, used to make downstream decisions. A prototypical example is the prediction of the impact of prices on a demand for goods, later used to optimize revenue. With such predictions, downstream optimization procedures expect the model to generalize outside of the observational distribution, since different decisions yield different observed distributions. Causal models, which aim to learn the structural causal models underlying the data generation process, are a natural for such use-cases. *This project aims to combine deep learning and causal inference to develop causal forecasting models over multiple time-steps.*

In recent work Crasson et al. (2024), we developed an approach based on orthogonal learning to fit causal forecasting models with daily effects. That is, we look at the impact of a price change at time-step t, on the demand at that same time-step. In practice however, may want to treat the entire sequence of actions as our treatment. For instance, if prices at times $1 : t-1$ were particularly high, we can expect a lot of pent-up demand, and increased sales at t. We also expect large changes in price to affect demand.

The orthogonal loss from Crasson et al. (2024) supports categorical treatments, so we can consider each sequence of treatments as one category to apply the theory. However, instantiating such a high-dimensional treatments would be intractable. Instead, we propose a sampling-based approach in which estimates of the treatment and outcome probability in the training data is replaced by two samples from this distribution. We propose to use diffusion models to model this distribution, which is used to generate samples for orthogonal learning, fitting a causal model for sequences of treatments. We refer to Song et al. (2020) for a more detailed introduction to diffusion models, and also mention the recent work Cardoso et al. (2024), which could offer relevant perspectives on learning a causal model.

Building on this framework, we will then seek to leverage our diffusion models to create confidence tunnels in the forecasts. Intuitively, knowing high-density sequences of treatments in the training data can help us assess the confidence we should have in a forecast: forecasts on sequences of treatment that are often observed in the training data are trustworthy; conversely, if a sequence to predict on was very unlikely in the training data, we should not trust the forecast. We aim to develop a theoretical understanding of our diffusion models to assess the likelihood of a given sequence under the training data distribution to assess the reliability of downstream forecasts.

REFERENCES

Gabriel V Cardoso, Yazid Janati El Idrissi, Sylvain Le Corff, and Eric Moulines. Monte carlo guided diffusion for bayesian linear inverse problems. In *ICLR International Conference on Learning Representations*, 2024.

Thomas Crasson, Yacine Nabet, and Mathias Lécuyer. Training and evaluating causal forecasting models for time-series. *arXiv preprint arXiv:2411.00126*, 2024.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2020.