Xing Han — Research Statement

2501 Speedway, EER 6.836, Austin, TX – Homepage
(512)-987-5026 • aaronhan223@utexas.edu

Research Summary

Decision-making in high-stake applications is increasingly data-driven and supported by machine learning (ML) models. The integration of such models into our every-day's life has uncovered the necessity of building trustworthy systems. Motivated by this, the overarching theme of my research focuses on improving the reliability and safety of machine learning models for different applications.

Along this direction, I mainly approach this problem via three major areas. 1) *Uncertainty quantification*: my goal is to better capture the epistemic uncertainty in common predictive modeling and classification problems. This includes both distribution-based and distribution-free methods for different use cases. 2) *Robustness*: strengthen the model's defense against adversarial attacks, data poisoning, and distribution drift, and construct certified robustness to provide safety guarantees via statistical modeling and optimization methods. 3) *Interpretability*: build human-understandable explanations for model decisions, with particular focus on non *i.i.d.* data such as time series or sequential data. Exploring their applications in healthcare and biomedicine is also an interest.

Current Work

Learning with Hierarchically Aggregated Time Series

Forecasting large-scale time series with hierarchical or grouped constraints are commonly seen in many practically important applications. In this setting, data at different aggregation levels possess distinct properties w.r.t. sparsity, noise distribution, sampling frequency, etc. A forecasting model should address both accuracies of individual time series and coherency across the hierarchy respecting any constraints that it imposes. Existing works predominately employ a two-stage approach, where base forecasts are first obtained for each time series followed by reconciliation among forecasts. These approaches assume unbiased forecasting models and are not numerically stable. To address this problem, I first designed a hierarchical time series (HTS) forecasting framework called `SHARQ` that connects reconciliation with learned parameters of forecasting models [1]. This method employs a novel objective function for each forecasting model to ensure the forecasts of adjacent aggregated levels are coherent. Meanwhile, `SHARQ` simultaneously produces multiple quantile forecasts and they are also calibrated by the predefined hierarchy. Using `SHARQ`, each time series is assigned a local model whose objective function is modified according to the hierarchy. I further improve `SHARQ` in two aspects: 1. dynamically combine point forecasts from a set of heterogeneous models using a gating network, which improves the representation power over a single local model; the objective function of `SHARQ` can also be applied on gating networks. 2. design a novel method to simultaneously produce coherent and model agnostic quantile estimations. The resulting method is called `DYCHEM`, which provides better point and quantile forecasts [2]. Building upon these works, I aim to further ameliorate the costs when a large number of HTS forecasts need to be done. Specifically, I proposed a novel approach to clustering HTS for efficient forecasting and exploratory data analysis. Since an HTS contains an inherent multilevel structure, clustering via leveraging local information from adjacent levels will lead to better performance. I developed a clustering procedure that can cope with massive HTS with arbitrary lengths and structures. This method, besides providing better insights from the data, can also be used to accelerate the forecasts of `DYCHEM` on a large number of HTS. Each time series is first assigned the forecast from its cluster representative, which can be considered as a "shrinkage prior" for the set of time series it represents. Then this base forecast can be quickly fine-tuned to adjust to the specifics of that HTS [3].

Impact and Ongoing Research These works have been summarized into three papers. `SHARQ` (AISTATS'21) has served as one of the state-of-the-art methods of HTS forecasting; `DYCHEM` (ICDMW'22) also demonstrated its superior performance and will be deployed into an industry forecasting pipeline. Along this direction, I am currently working on estimating heterogeneous treatment effects in nonstationary time series to provide counterfactual explanations to end-users under the intervention of rare events or changes in the environment.

Interpretable & Robust ML

I have investigated multiple problems around the topic of building interpretable and robust ML models. As a first step, I studied a sample-based explanation method for arbitrary models called `MFS`, which finds a subset of training samples that are most responsible for a specific prediction, where the model's decision would have changed upon removal of this subset from the training data [4]. `MFS` connects both the original decision and its counterfactuals with a small set of training samples, making this approach more interpretable to end-users. I have applied `MFS` to a variety of tasks including data poisoning detection, the training set debugging, and understanding loan decisions. Second, I studied the problem of learning monotonic models with respect to a subset of inputs, which is the desired property in many applications with fairness or security concerns [5]. I proposed a method that leverages arbitrary neural architectures and provably ensures the monotonicity of learned models over a subset of features. This method is able to learn non-trivial neural networks that can approximate arbitrary monotonic functions. Applying this to various regression and classification tasks using datasets with monotonic features achieves better performance. In addition, I also found that enforcing monotonicity provides a natural tool for enhancing the interpretability and robustness of neural networks. Lastly, motivated by constructing predictive intervals for HTS, I studied conformal prediction methods that quantify uncertainty without any distributional assumptions [6]. Most existing methods can only provide an average coverage guarantee, which is not ideal compared to the stronger conditional coverage guarantee. To approximate conditional coverage, I proposed `SLCP` that employs a modified non-conformity score (measure how bad samples "conform" to the model). The score leverages the local approximation of its conditional distribution over the input variables by kernel density estimation. `SLCP` can be extended to a general framework that unifies many existing baselines. Empirically, it also shows superior conditional coverage than the current state-of-the-art.

Impact and Ongoing Research These works have been summarized into three papers. `MFS` (IJCNN'21) has been selected as a spotlight presentation of the ICML WHI workshop. Certified monotonic neural networks have received a spotlight award at NeurIPS'20. I also publicized these ideas in multiple talks and discussions during my time in industry. Along this direction, I am currently working on making transformer models more robust to outliers and adversarial attacks via robust kernel density estimation. Extensive experiments on language models and image classification have demonstrated better results than various baselines [7].

ML Applications & Collaborations

During my graduate study, I am also fortunate to learn and interact with different collaborators on multiple ML applications. I was once involved with a multidisciplinary project and collaborated with many institutions. The project called `Tesserae`, aims to model job performance via wearable sensors paired with a smartphone app to gauge biomarkers like heart rate, sleep, physical activity, and stress [8]. This experience facilitated me to get a better understanding of how to leverage ML into interdisciplinary research. I also got a chance to work on a text style transfer problem for the automated design of slogans in e-commerce applications [9, 10]. The biggest challenge encountered was the lack of quality data, where in my specific application, a major amount of data is unlabeled short sentences with ambiguously defined text styles, which hampers good performance. This problem is being actively studied and playing an important role in deploying trustworthy ML into real applications. In addition, I collaborated with UT-Austin peers to study federated learning problems that allow each client to build a personalized model without enforcing a common architecture across clients [11]. The core idea is to use the instance-level representations obtained from peer clients to guide the simultaneous training of each client. Along this direction, I am actively looking into ML applications in other disciplines and collaborations with people from corresponding backgrounds.

Future Research Plans

With the development of information technology, a lot of life-critical data is being collected and stored in electronic forms, such as client data with extreme weather events, clinical notes, or magnetic resonance imaging. Data can take the form of images, tabular datasets, signals, or text and they often have interconnected relationships or privacy constraints (e.g., aggregation). Predictive modeling of such data is cost-sensitive. Research into novel ML methods for transformative applications in environmental science and healthcare is impactful and beneficial. In general, I am interested in developing robust models for characterizing critical events, and building interpretable and privacy-preserving ML systems to augment medical capabilities.

Recently, transformers have shown remarkable performance in vision, language, and sequential modeling. However, they are trained on a colossal amount of multimodal data collected from various sources, making model optimization a very challenging task. Furthermore, the performance of state-of-the-art models can be unstable in practical applications, and their computational efficiency can be improved. Along this line of inquiry, I am interested in designing and training transformer models that are both efficient and generalize well, especially when intersecting with domain-specific applications. Moreover, I hope these works can also be extended to diffusion models whose backbones are transformers. I am passionate about interdisciplinary research. In the long term, I look forward to addressing the challenging problems that arise when applying principled ML methods in practical scenarios across different domains.

References

Xing Han, Sambarta Dasgupta, and Joydeep Ghosh, "Simultaneously reconciled quantile forecasting of hierarchically related time series," in *International Conference on Artificial Intelligence and Statistics*. PMLR, 2021, pp. 190–198.

Xing Han, Jing Hu, and Joydeep Ghosh, "Dynamic combination of heterogeneous models for hierarchical time series," *Submitted to IEEE Transaction on Neural Networks and Learning Systems*,

2023, under review.

Xing Han, Tongzheng Ren, Jing Hu, Joydeep Ghosh, and Nhat Ho, "Efficient forecasting of large scale hierarchical time series via multilevel clustering," *Submitted to IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023, under review.

Xing Han and Joydeep Ghosh, "Model-agnostic explanations using minimal forcing subsets," in *2021 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2021, pp. 1–8.

Xingchao Liu, **Xing Han**, Na Zhang, and Qiang Liu, "Certified monotonic neural networks," *Advances in Neural Information Processing Systems*, vol. 33, pp. 15427–15438, 2020.

Xing Han, Ziyang Tang, Joydeep Ghosh, and Qiang Liu, "Split localized conformal prediction," *Submitted to Conference on Uncertainty in Artificial Intelligence (UAI)*, 2023, under review.

Xing Han, Tongzheng Ren, Tan Minh Nguyen, Khai Nguyen, Joydeep Ghosh, and Nhat Ho, "Robustify transformers with robust kernel density estimation," *Submitted to International Conference on Machine Learning (ICML)*, 2023, under review.

Suwen Lin, Stephen M Mattingly, **Xing Han**, and et. al, "Sensing personality to predict job performance," in *Proceedings of the Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, CHI EA*, 2019.

Xing Han and Jessica Lundin, "Multi-pair text style transfer for unbalanced data via task-adaptive meta-learning," in *Proceedings of the 1st Workshop on Meta Learning and Its Applications to Natural Language Processing*, 2021, pp. 28–35.

Jessica Lundin, Owen Winne Schoppe, **Xing Han**, Michael Reynolds Sollami, Brian J Lonsdorf, Alan Martin Ross, David J Woodward, and Sonke Rohde, "Machine-learning based generation of text style variations for digital content items," Aug. 4 2022, US Patent App. 17/163,162.

Disha Makhija, **Xing Han**, Nhat Ho, and Joydeep Ghosh, "Architecture agnostic federated learning for neural networks," in *Proceedings of the 39th International Conference on Machine Learning*. 2022, pp. 14860–14870, PMLR.