Causality Identify via Bayesian Inference and Explainable DNN Prediction

研究生:郭家宏 指導教授:陳添福 博士 國立陽明交通大學資電亥客與安全碩士學位學程

In the wide-range use of deep neural networks (DNN), the explainable and robustness of DNN become increasingly important. We not only deploy a feasible model to the real world but also clear how the model makes decisions. Our work is motivated by real-world cases with high-dimensional numerical data in the industries. We encode the original data into Bayesian networks, which consist of directed acyclic graphs (DAG) and conditional probability distributions to perform Bayesian inference by these explainable Bayesian networks. Due to the characteristic of the

Bayesian theorem, our method is more explainable and robust than conventional DNN models.

Overview



- Improve the performance on numerical data with high dimensions features
- Find out the causality between models input and output without extra models

• Make deep neural networks more explainable

Evaluation

Ground truth relation between input data (upper row) and target (bottom row)

