Week 10

Wentao Gao

Causality for Time series prediction

How to improve the performance of prediction in time series using causality.

We need to know the pipeline of time series prediction.

And then thinking which part in pipeline can we improve by the causal relation between variables

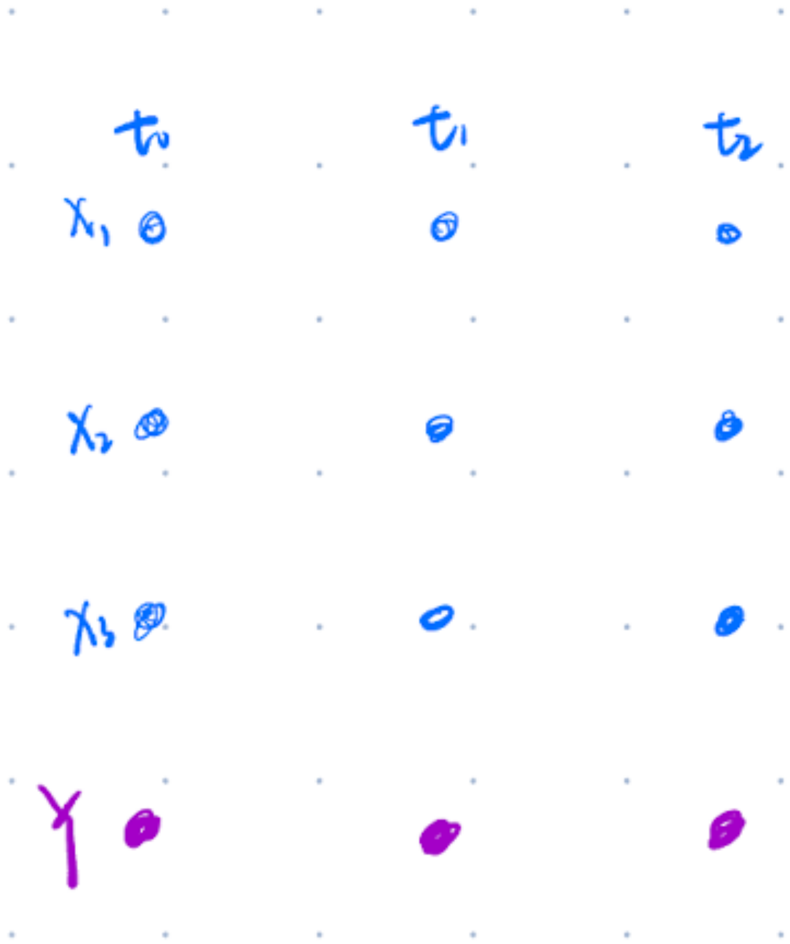
Problem

The time series forecasting problem can be formulated as:

Predicting future M steps(day or month) precipitation $\boldsymbol{Y_{t+1:t+M}}$. Given the previous L steps of observations $\boldsymbol{Y_{t-L:t}}$ and the covariates(air temperature …) $\boldsymbol{X_{t-L+1:t}}$.

Input: $\boldsymbol{Y_{t-L:t}}$, $\boldsymbol{X_{t-L+1:t}}$

Output: $\boldsymbol{Y_{t+1:t+M}}$

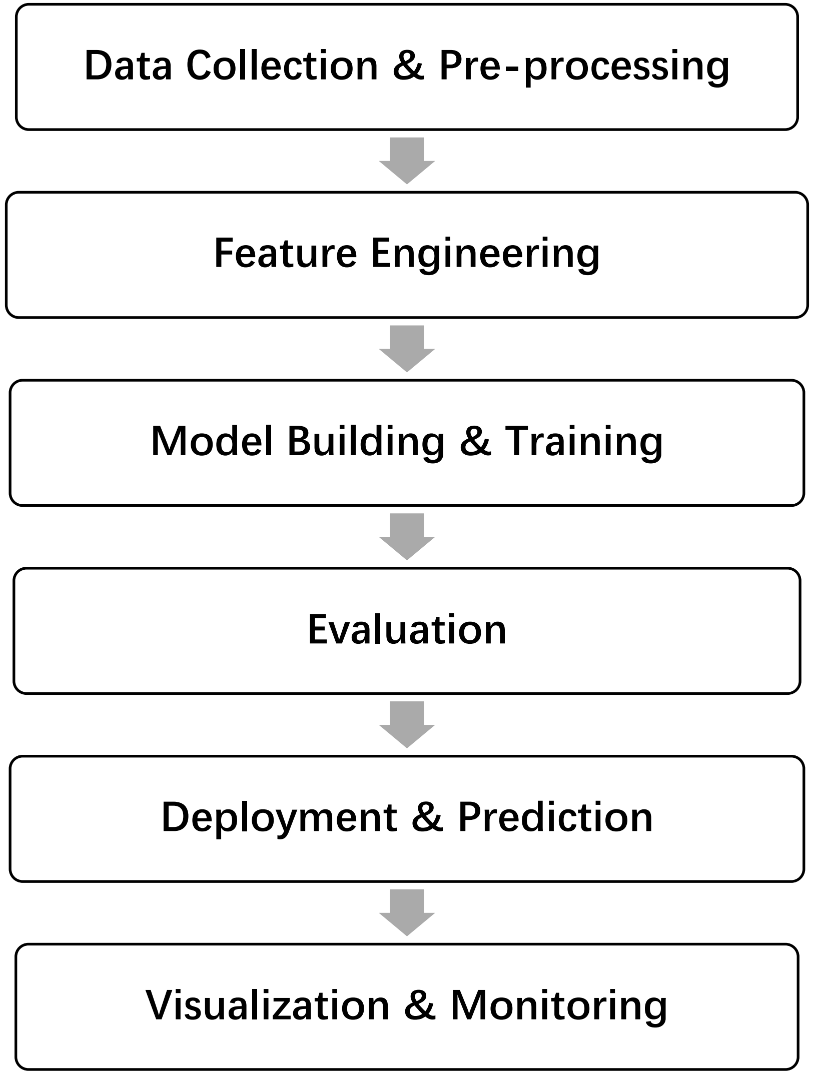


To what extent do X and Y at past times play a role in the prediction of Y at future times

How do these variables X affect Y and what are the relatively real dynamic causal mechanisms .

Climate change is a very complex physical and chemical process that follows certain relationships, such as the laws of thermodynamics, and how to incorporate these established relationships into causal diagrams

Pipeline



Temporal alignment

Normalization/Standardization

Time series features   Rolling statistics

Seasonality

Model Selection

Model Training

Metrics

Monitor and retrain if necessary.

Feature Engineering

Using causality to help feature selection

Model Building

Add causality module into model structure(LSTM, Transformer)

Evaluation

Loss function with causal relation

Deployment

Causality-based Feature Selection

Predictive Performance:
Identifying genuine causal features can enhance the performance of predictive models.

Interpretability:
Understanding the true causal relationships between features and the target variable offers deeper insights, helping us to better interpret the predictions of models.

Robustness:
Feature selection might produce more robust models, especially when faced with unseen data or data from different distributions.

Method

Constrained based method

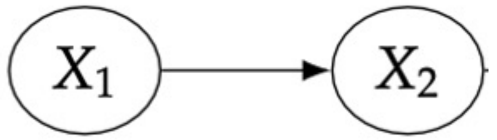
This method is based on conditional independence tests

Scored based method

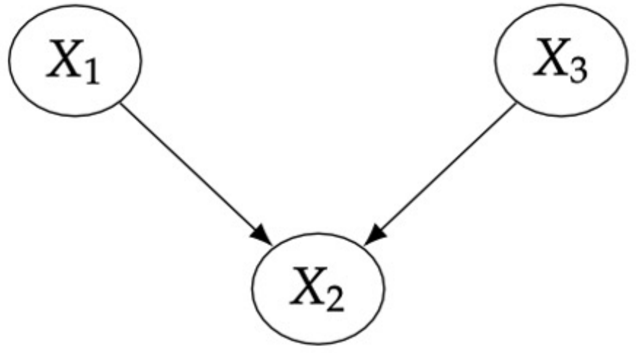
This type of method employs score-based BN structure learning algorithms to learn the MB or PC of the class variable instead of using independence tests

Basis of Constraint-based Methods

Proposition 3.1:
In a Bayesian Network (BN), if node *Vi* is a parent or child of *Vj*, then *Vi* and *Vj* are not independent given any subset of *V* excluding *Vi* and *Vj*.



Proposition 3.2:
This describes the relationship between a node and its spouses in a BN. If *Vi* is a spouse of *Vk* and *Vj* is their common child, there exists a subset making *Vi* and *Vk* independent unless *Vj* is included.



Basis of Bayesian Network Learning

Objective:

Score-based Bayesian Network (BN) learning algorithms aim to identify a Directed Acyclic Graph (DAG) structure that best matches a dataset, *D*. This match is measured by a scoring function, which evaluates the fitness between the DAG and *D*.

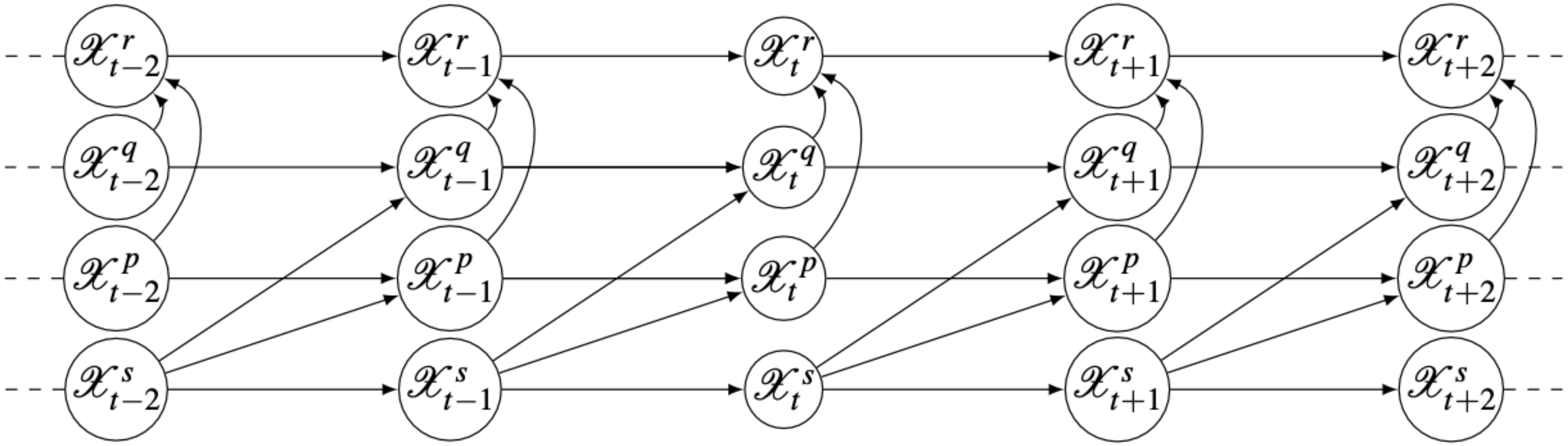
Scoring Functions:

Notable scoring functions include K2, BDeu, BDe, MDL/BIC, AIC, and MIT. These functions determine the "goodness" of a structure. An essential property of these functions is "decomposability," meaning they can aggregate local scores to compute the global score of a DAG.

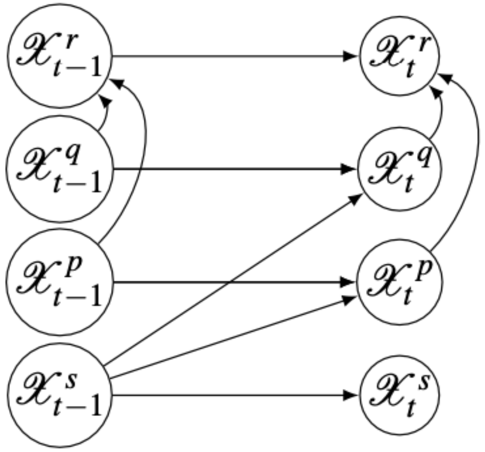
Main Idea:

Score-based algorithms leverage decomposable scoring functions to learn the DAG of selected features. They can differentiate between parents and children of a variable, which constraint-based methods cannot.

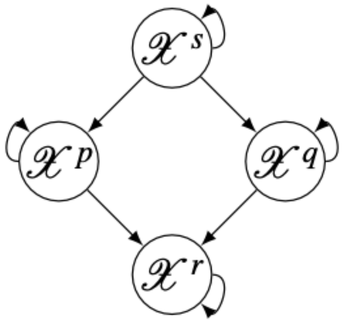
Causal graph for time series



Full Time Causal Graph (a)



Window Causal Graph (b)



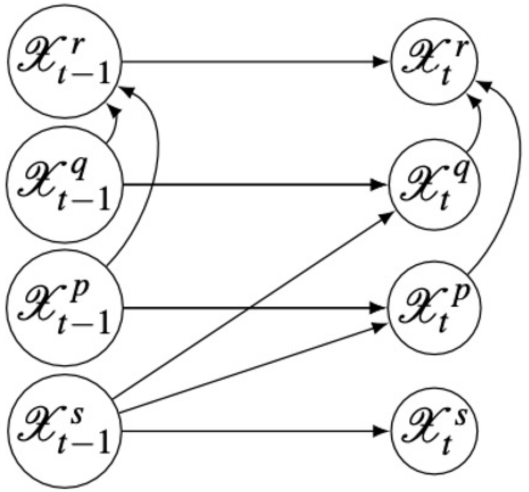
Summary Causal Graph (c)

Assumption: Consistency Throughout Time

- We suppose that our causal graph follow mechanism over time

Definition 7 (Consistency Throughout Time) *A causal graph $\mathscr{G} = (V, E)$ for a multivariate time series $\mathscr{X}$ is said to be* consistent throughout time *if all the causal relationships remain constant in direction throughout time.*

When consistency throughout time satisfied. Full time causal graph can be simplified as **Window Causal Graph.**



Window Causal Graph (b)

However, Summary causal graph cannot have a good description of the causal relation between variables because it cannot express the info of time lags.

Research proposal plan

- **Papers:**

  Causal discovery in time series
  Causality for feature selection in time series prediction
  Model structure modifying(review LSTM and Transformer)

- **Extra:**

  Preparing dataset(Area and Whole world saperately)
  Finish the code library structure