Performance Uncertainty and Ranking Significance of Early-Warning Models

Discussion

Soohyon Kim (Bank of Korea)

Nov. 6, 2020

Summary

- The small data nature of macroeconomic panel data in the early-warning models
- Three sources of data variation:
  - set of countries - events are few and heterogeneous
  - set of years - global regime shifts are rare but significant
  - set of countries' histories - experienced specific idiosyncratic shocks
- If we have to pick one approach for the data with such uncertainty, which one should we select on what criteria?
  - choosing a model based on performance uncertainty
  - by iteration of jackknife resampling and estimation to construct confidence intervals to compare performance uncertainty
- Three dimensions of dropping samples:
  - countries
  - years
  - country-year blocks

Model Stability Check and Alternative Approaches

- A learning algorithm is said to be stable if it produces consistent predictions even with some perturbation
- A common way of testing algorithm stability is to check prediction consistency through sample perturbation such as random resampling
- In this study, assessing performance uncertainty is done in a quite similar manner
  - Only difference is that it is a selection between a model and an algorithm
- However, unlike any other data set used in engineering, we are dealing with small and sparse time series
  - and a machine learning algorithm has hyperparameters
  - We can also think of an alternative approach to test model stability when we apply the performance uncertainty to choose between algorithms

Stability Check without Random Resampling

- Resampling is a way to reuse data to generate new, hypothetical samples (namely resamples) that are representative of an underlying population
  - It's used when we don't know the underlying distribution for the population
- There are terms addressing those innate uncertainties of data in certain field of science such as seismic hazard analysis: epistemic and aleatoric uncertainty
  - When data is small or sparse we say there is epistemic uncertainty
  - When data is noisy we say there is aleatoric uncertainty
- We can use alternative ML algorithms to address the degree of uncertainty using the same data set with different hyperparameters
  - Macroeconomic data sets are often too small and sparse

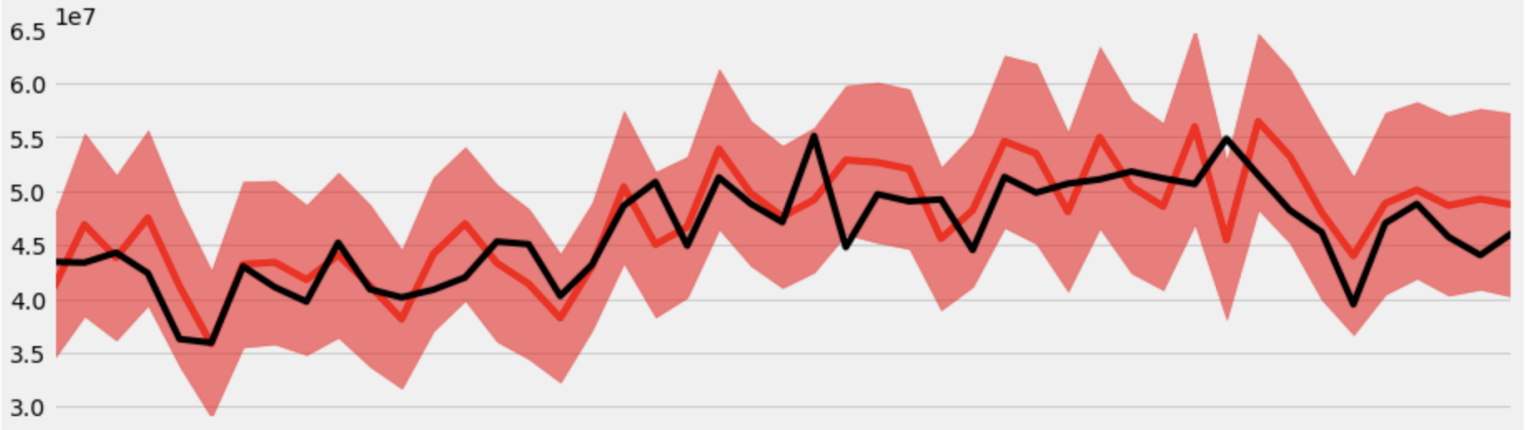
Stability Check without Random Resampling

- Performance uncertainty of early-warning system comes from small and sparse data
  - As we have epistemic uncertainty in the data, it is hard to know the true feature or distribution of that
  - The amount of information is already quite insufficient to train an algorithm
- Alternative model approach is to change hyperparameters in algorithms, but not the samples
  - If we let hyperparameters chosen randomly, then we can have sort of "confidence intervals"
  - So we can measure the performance uncertainty by iterating training an algorithm with randomly chosen hyperparameters

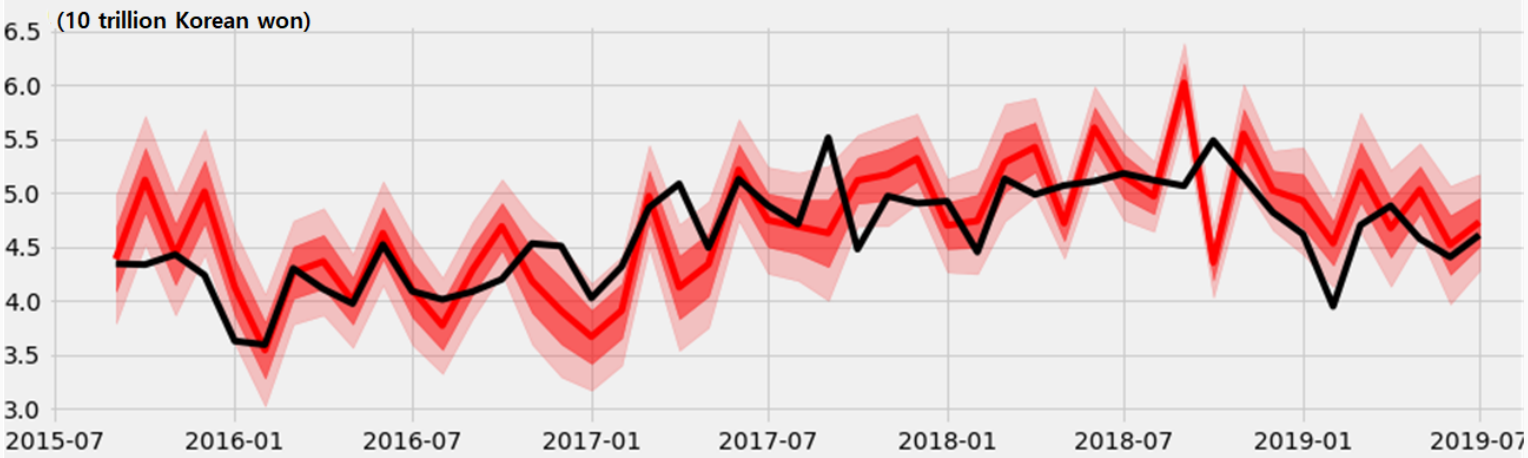
Model Selection Depending on Usages

Korean monthly nominal export: VAR versus MLP

[VAR with confidence intervals]



[Predictions from 100 DL: nodes and layers are randomly chosen]



Conclusion

- We need to keep asking that under which criteria we should select a model
- This study turns a classical way of model selection into a brilliant application to select a model and a ML algorithm
- If we have to choose between ML algorithms, we can just pool the results to compare performance uncertainty
    - After all, we are never going to have perfect data set for any purpose
    - It's a risk to select only a hyperparameter if we cannot fully control uncertainty of data