

Comentarios Sesión 4: Redes Neuronales

Jairo Flores (BCRP)

Encuentro de Economistas 2024

21 de octubre de 2024

Neural Architecture and Hyperparameter Optimization of Time Series Forecasting Models with Tree Parzen Estimators

Andrew Garcia y Marco Vega (2024)

Resumen

- Presentan un modelo AutoML para la optimización de Neural Architecture Search (NAS) e hyperparametros basado en el algoritmo TPE.
- Se aplica para el pronóstico de la inflación de Perú para dos casos: univariado y por componentes.

Resultado

- Los mejores modelos son NBEATS and NHITS.
- K-fold CV para el caso univariado y 1 solo conjunto de validación para el modelo basado en los componentes de la inflación.
- El análisis de perturbación basado en gradiente ayuda a entender el modelo capturando dinámicas no lineales a lo largo de períodos específicos.

Comentarios/Preguntas

- Es posible que Time2Vec ayude en parte debido a que MLP (Multilayer Perceptron) no incorpora noción del tiempo. Alternativas: Recurrent Neural Networks (RNNs), Long Short-Term Memory Networks (LSTMs).
- Resultaría interesante analizar cuáles fueron los efectos de la pandemia sobre la estabilidad de los hiperparámetros y/o parámetros.
- ¿Cómo se construiría un intervalo de confianza alrededor de la estimación puntual bajo este esquema?.
- ¿Por qué K-fold CV es impráctico en el modelo basado en los componentes de la inflación?.
- ¿Es posible agregar predictores? ¿Cuál es el costo en términos de tiempo de realizar este procedimiento?

Uncovering inflation nonlinearities with Kolmogorov-Arnold Networks

Andrew Garcia y Marco Vega (2024)

Resumen

- En el paper se aplica el modelo Kolmogorov-Arnold Networks (KANs) para descubrir e interpretar formas funcionales de la Curva de Phillips.

Resultados

- La curva Phillips parece presentar no linealidad entre el tipo de cambio y la inflación.

Comentarios/Preguntas

- En la primera etapa IV, existe valores del estadístico F por debajo de 10, sobre todo en la versión trimestral (valores alrededor de 4). Esto puede ser una señal de instrumentos débiles. Tal vez sea conveniente aplicar algún test para verificar que no sea el caso.
- También se podría aplicar algún test de sobreidentificación (J-test de Hansen) para descartar problemas adicionales de identificación.
- Se podría considerar pasar de una metodología de 2 etapas (IV+KAN) a una sola etapa con GMM no lineal.
- ¿Cómo se construiría un intervalo de confianza alrededor de la estimación puntual de los coeficientes en un modelo KAN?.
- ¿Existe algún benchmark para comparar los resultados de obtenidos por KAN?