



3TS: ML Engineering II

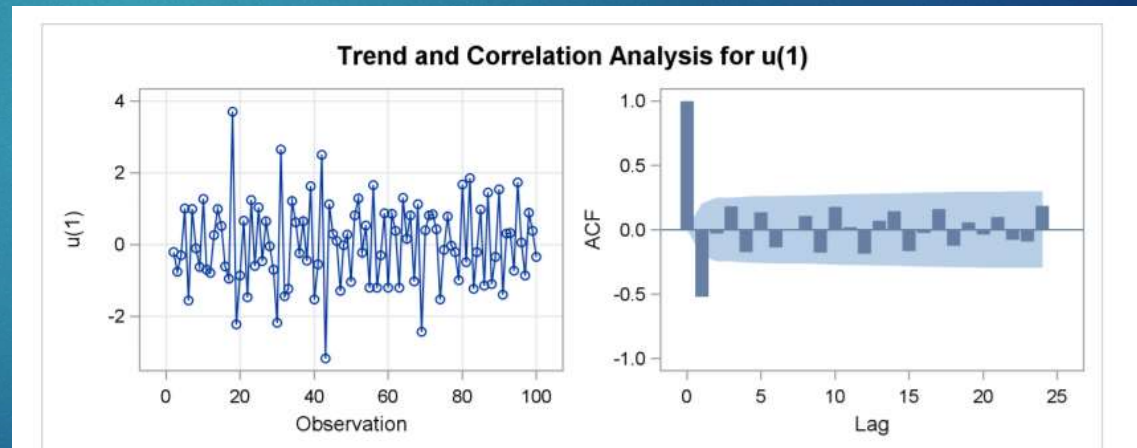
MACHINE LEARNING APPLIED TO FORECASTING

Objetivos

- ▶ Generación Energía Eléctrica en España
- ▶ Solucionar problemas de Forecasting con modelos de Machine Learning en lugar de con modelos estadísticos:
 - ▶ ARIMA
 - ▶ GARCH
 - ▶ VAR

Objetivos

- ▶ Debatir, discutir y compartir experiencias y prácticas.
- ▶ Preguntarnos por qué



REPO: https://github.com/manualrg/DSLAB_Python/ML_Engineering

Índice de la sesión

- ▶ Introducción
- ▶ EDA (Exploratory Data Analysis) de series temporales
- ▶ Métodos de Feature Engineering específicos para problemas de forecasting
- ▶ Predicción de la demanda energética en España con modelos de Machine Learning

Introducción

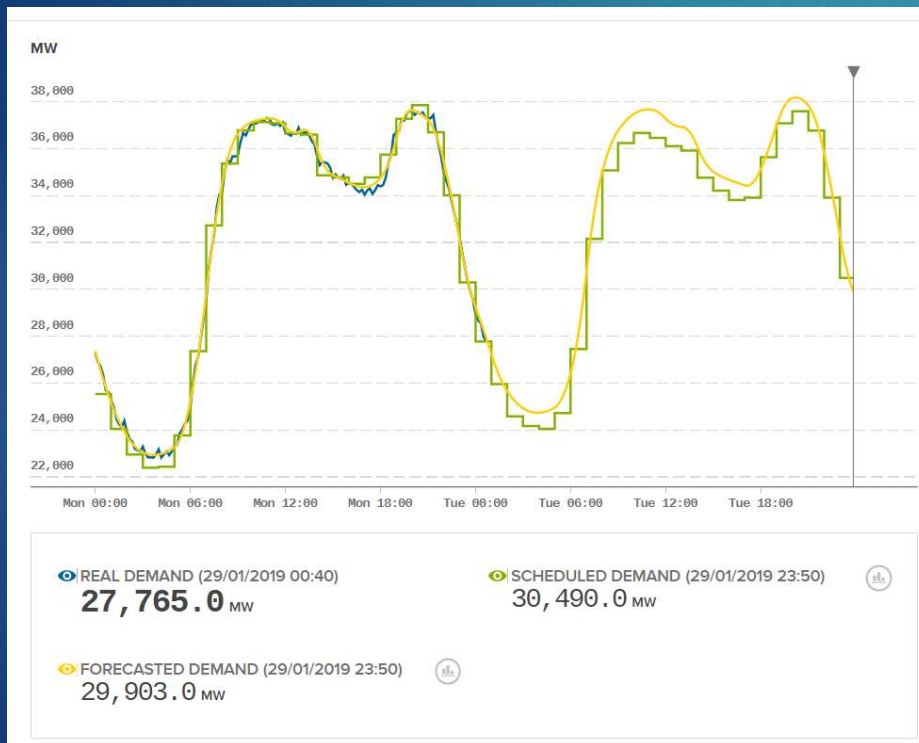
- ▶ Sistema Eléctrico:
 - ▶ Generadores
 - ▶ Mercado mayorista (bilateral, pool)
 - ▶ Transportista
 - ▶ Distribuidoras
 - ▶ Comercializadoras
 - ▶ Consumidores
 - ▶ Mercado minorista (PVPC, libre)

Introducción

- ▶ REE (Red Eléctrica de España)
 - ▶ Operador del sistema Eléctrico
 - ▶ Transportista
- ▶ OMIE (Operador del Mercado Ibérico de Electricidad)
 - ▶ Pool: Subasta diaria marginalista

Introducción

► Fuente: <https://www.esios.ree.es/es>



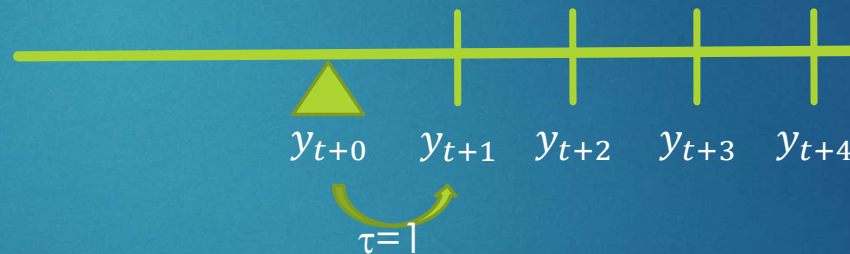
Introducción

► Datos:

► Demanda de electricidad semanal: 2015-01-11 al 2018-08-26

► Predicción de 4 periodos

t	$y(t)$
31-12-18	90
07-01-19	130
14-01-19	120
21-01-19	100
28-01-19	105

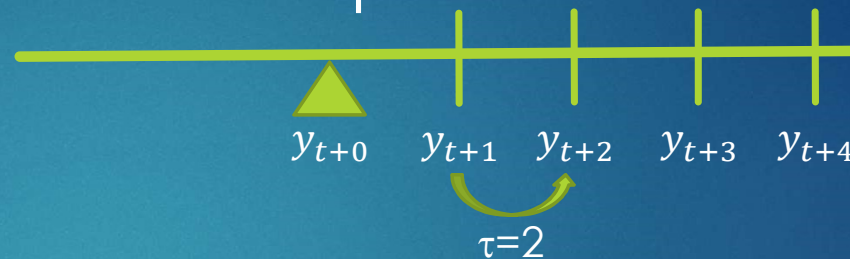


t	y_lag1	y_lag2	y_t
31-12-18	80	75	90
07-01-19	90	80	130
14-01-19	130	90	120

Introducción

► Forecasting > Regresión multi-periodo

t	y(t)
31-12-18	90
07-01-19	130
14-01-19	120
21-01-19	100
28-01-19	105



t	y_lag1	y_lag2	y_t	y_t1
31-12-18	80	75	90	130
07-01-19	90	80	130	120
14-01-19	130	90	120	100

sklearn.multioutput.MultiOutputRegressor

```
class sklearn.multioutput. MultiOutputRegressor (estimator, n_jobs=None)
```

[\[source\]](#)

Multi target regression

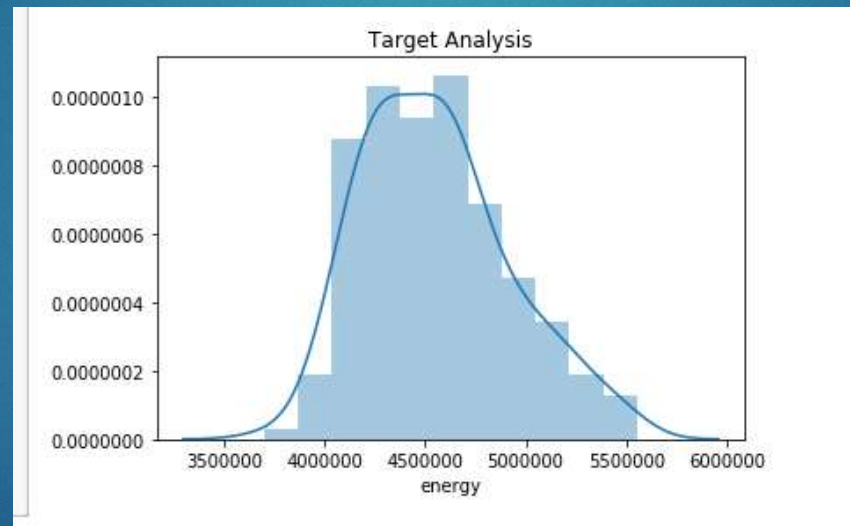
This strategy consists of fitting one regressor per target. This is a simple strategy for extending regressors that do not natively support multi-target regression.

Time Series EDA

- ▶ ¿Valores nulos?
- ▶ Análisis de la señal
- ▶ Patrones estacionales
- ▶ Tendencia

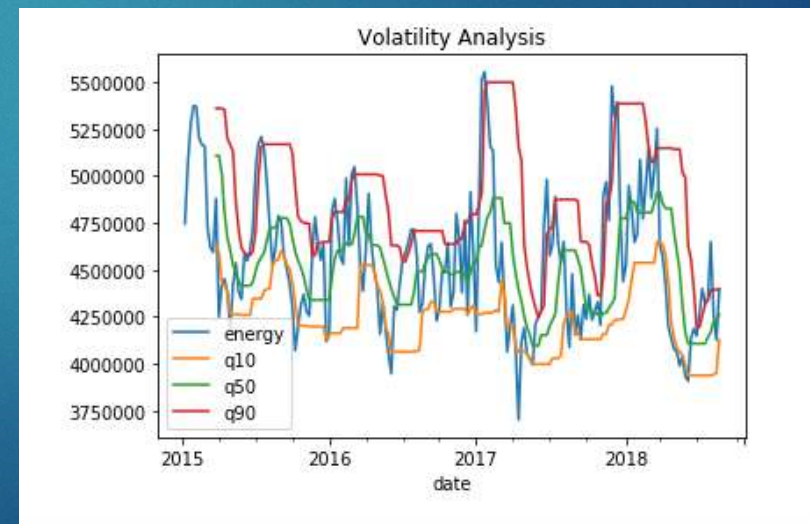
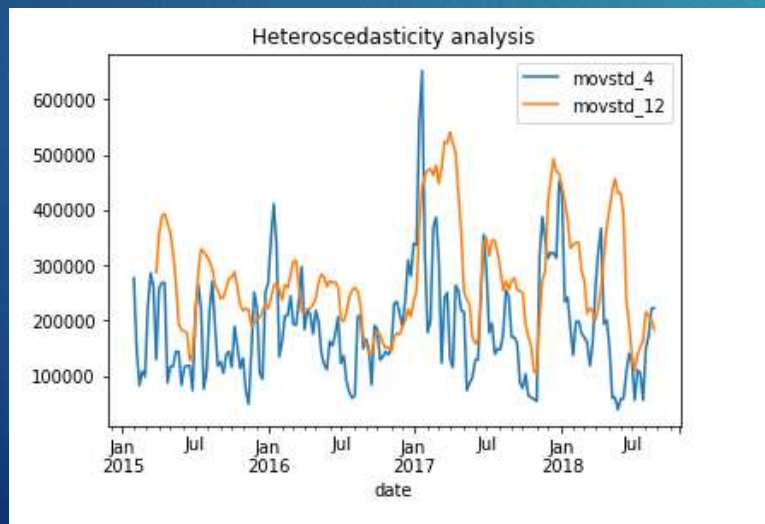
Time Series EDA

- ▶ Análisis de la señal
 - ▶ Normalidad y outliers



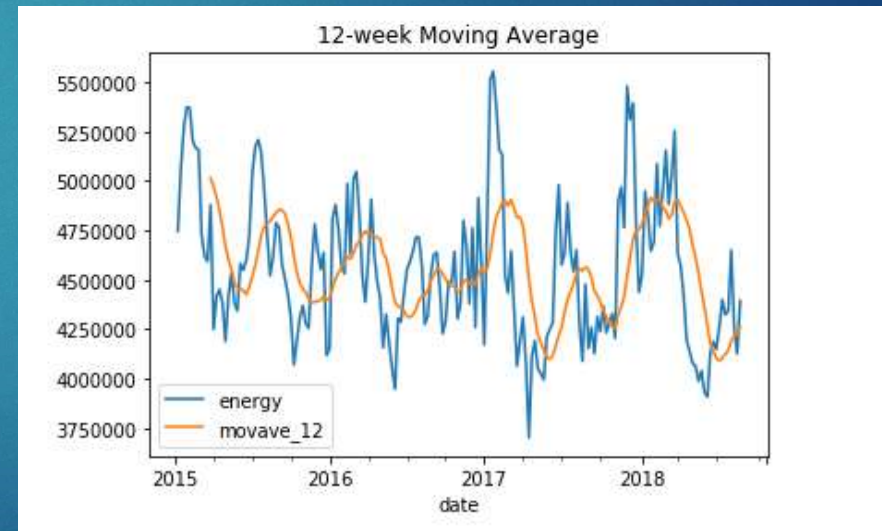
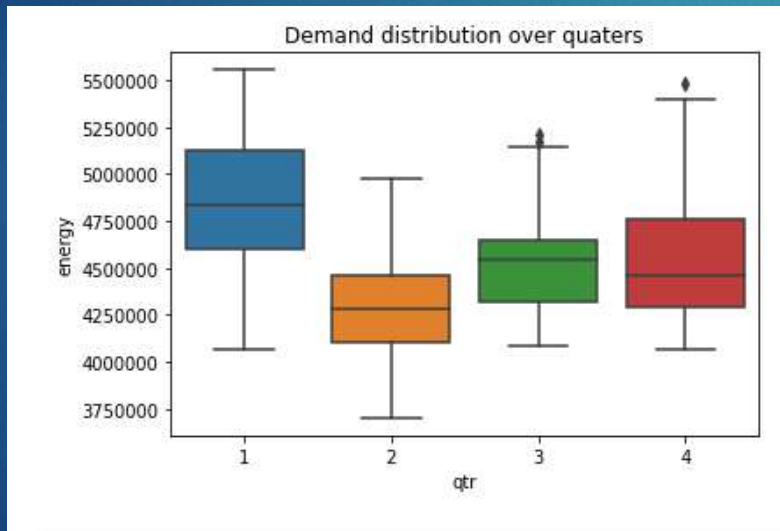
Time Series EDA

- ▶ Análisis de la señal
 - ▶ Volatilidad (heterocedasticidad)
 - ▶ La desviación típica cambia con t
 - ▶ ¿Existen patrones o causas en la volatilidad?



Time Series EDA

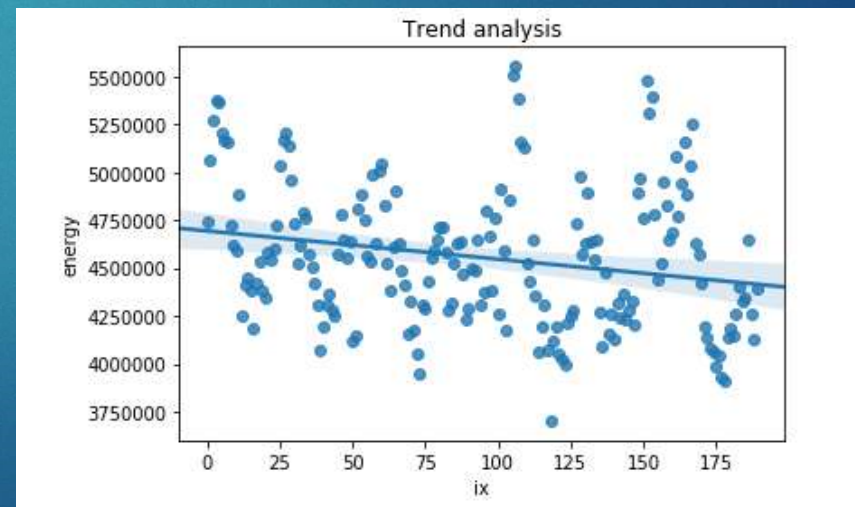
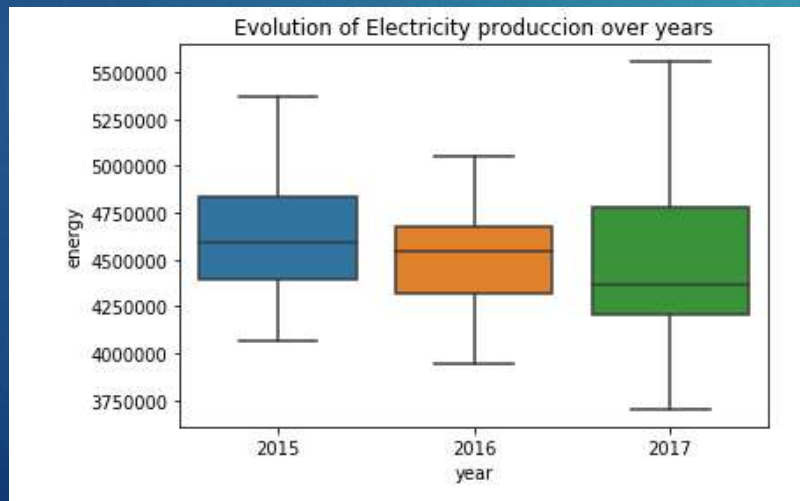
► Patrones estacionales



Time Series EDA

► Tendencia

- Lineal, amortiguada, exponencial
- Cambios bruscos

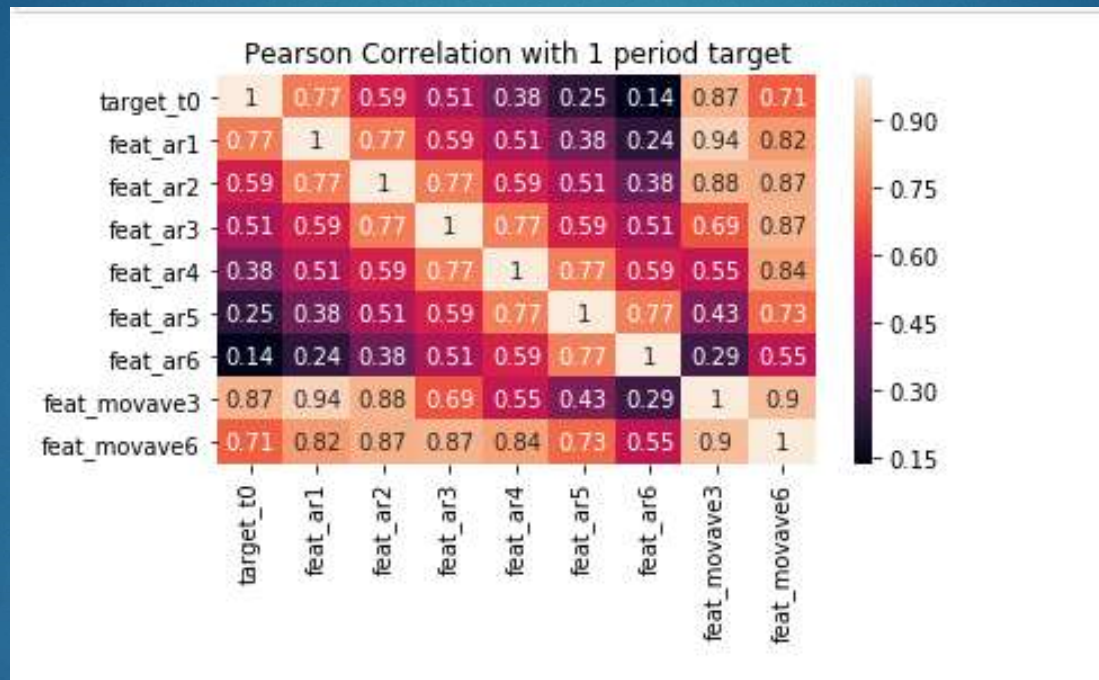


Feature Engineering in Forecasting

- ▶ Automático y sin asunciones estadísticas
- ▶ Lags
- ▶ Estadísticas de ventana móvil
- ▶ Avanzados:
 - ▶ RSI (Relative Strength Index)
 - ▶ Momentum
 - ▶ Mean Reversion
 - ▶ Sequence mining

$$RSI = 100 - \frac{100}{1 + RS}$$
$$RS = \text{Average Gain} / \text{Average Loss}$$

Feature Engineering in Forecasting

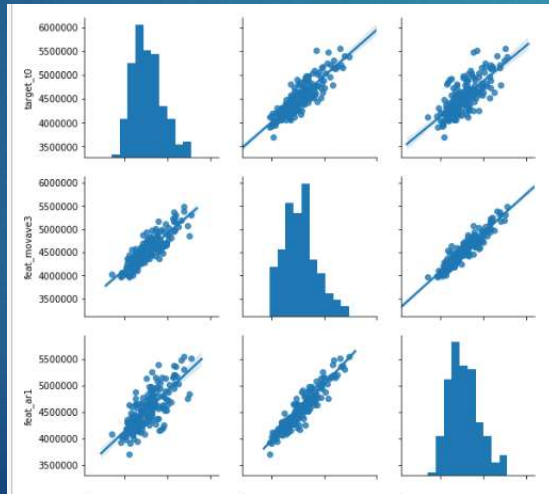


Machine Learning and Multi-Period Forecasting

► Regresión lineal

► Existe relación lineal entre los predictores y la señal

► Inconvenientes?



Machine Learning and Multi- Period Forecasting

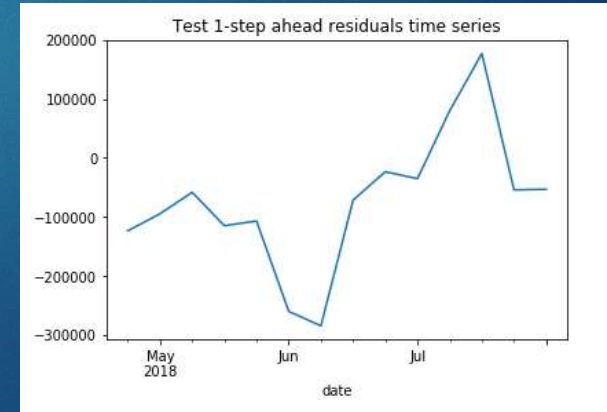
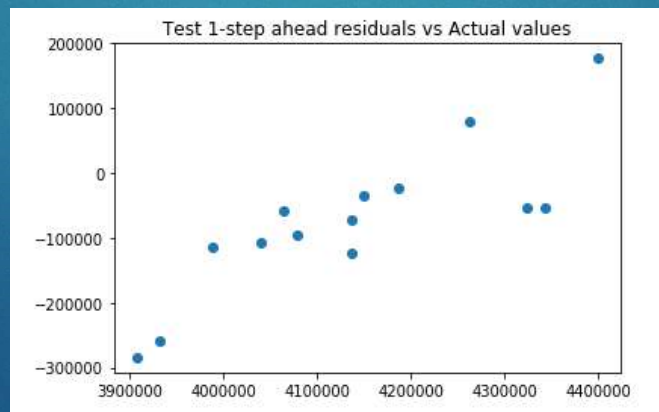
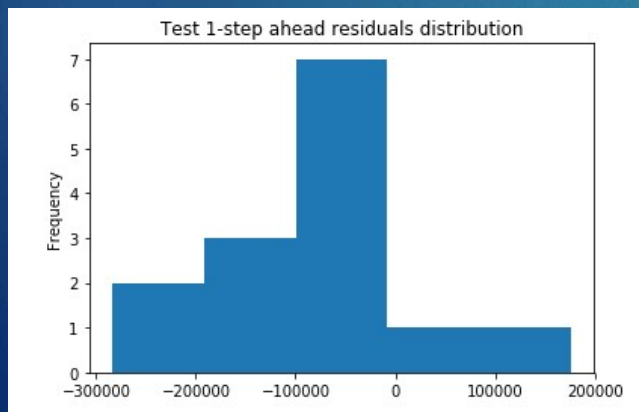
- ▶ Árboles y Ensembles de árboles
 - ▶ Modelar cualquier relación
 - ▶ Herramientas de regularización
- ▶ Inconvenientes?

Machine Learning and Multi- Period Forecasting

- ▶ Entrenamiento y ajuste de hiperparámetros
 - ▶ Train (2-4 ciclos estacionales)
 - ▶ Validation: Ajuste hiperparámetros
 - ▶ Test: Evaluación y análisis residual

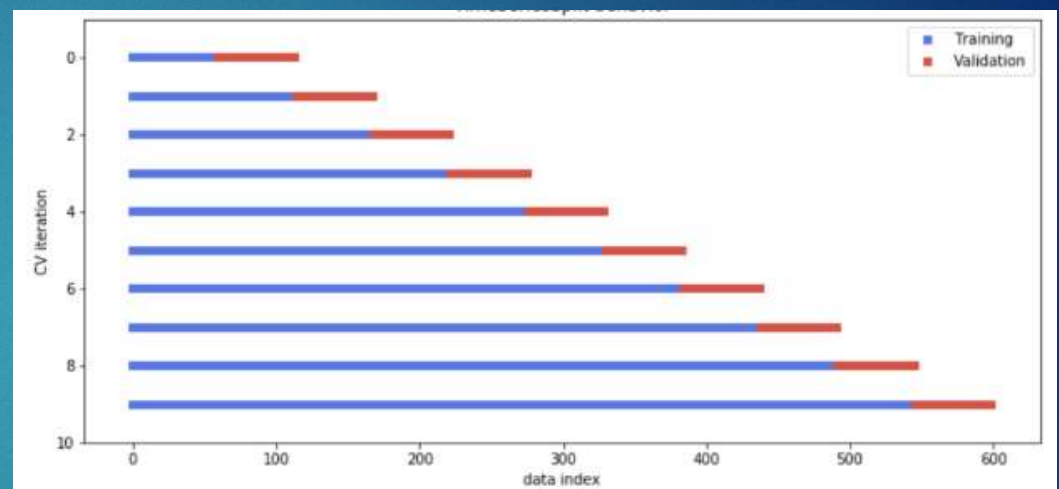
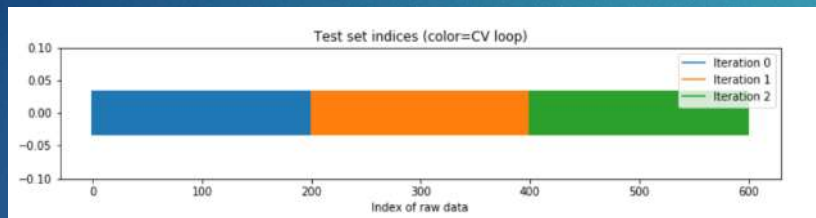
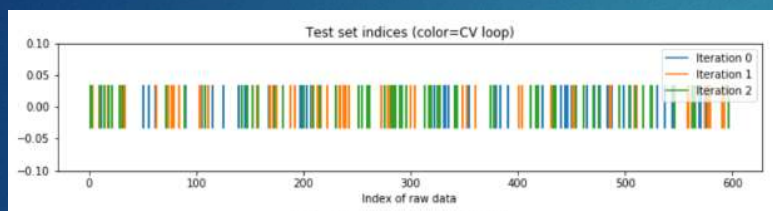
Machine Learning and Multi-Period Forecasting

- Evaluación: Problema de regresión: RMSE, MAPE...
- Análisis residual (para cada modelo)
 - Distribución de los residuos
 - Forecasted vs actual
 - Residuals over time

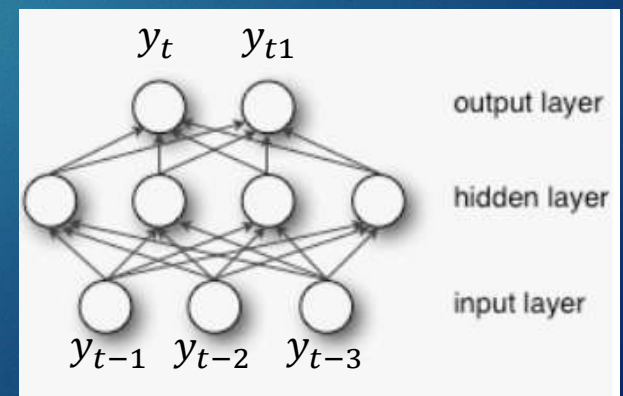


More ideas

► Hiperparameter tuning with Time CV



► AR-MLP



More ideas

- ▶ Real problems:
 - ▶ +250k SKUs
 - ▶ interrupted data
 - ▶ new products
 - ▶ Hierarchical forecasting

GRACIAS !!!

- ▶ BIBLIOGRAFÍA:
- ▶ Forecasting: Principles and Practice (R)
 - ▶ <https://otexts.com/fpp2/>
- ▶ Practical Time Series Analysis (Python)
 - ▶ <https://www.packtpub.com/big-data-and-business-intelligence/practical-time-series-analysis>