

TRABAJO FINAL

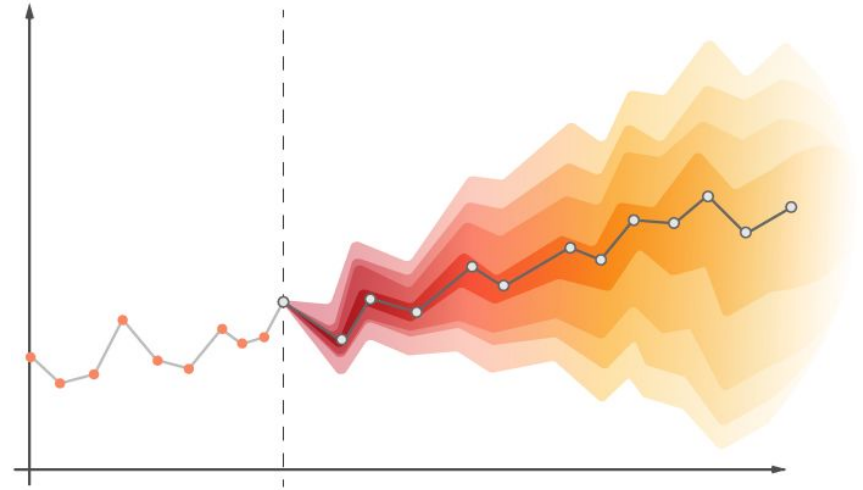
ANÁLISIS DE SERIES DE TIEMPO

ALUMNO: ALEX BARRIA

Modelado de extractor en planta productora
de aceite de girasol y soja

TEMAS ABORDADOS

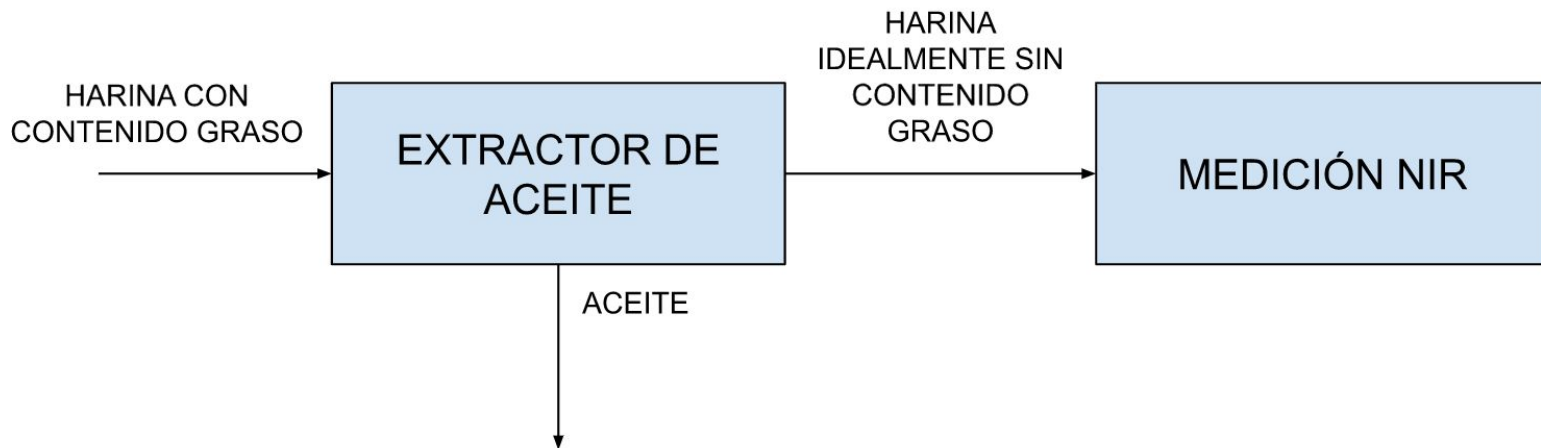
- ANÁLISIS EXPLORATORIO DE DATOS
PARA SERIES TEMPORALES
- TEST ESTADÍSTICOS -
ESTACIONARIEDAD
- MODELOS ESTADÍSTICOS
- MODELOS DE DEEP LEARNING



DESCRIPCIÓN DEL PROBLEMA

Predicción de contenido graso en harina a la salida del extractor, en una planta productora de aceite de soja.

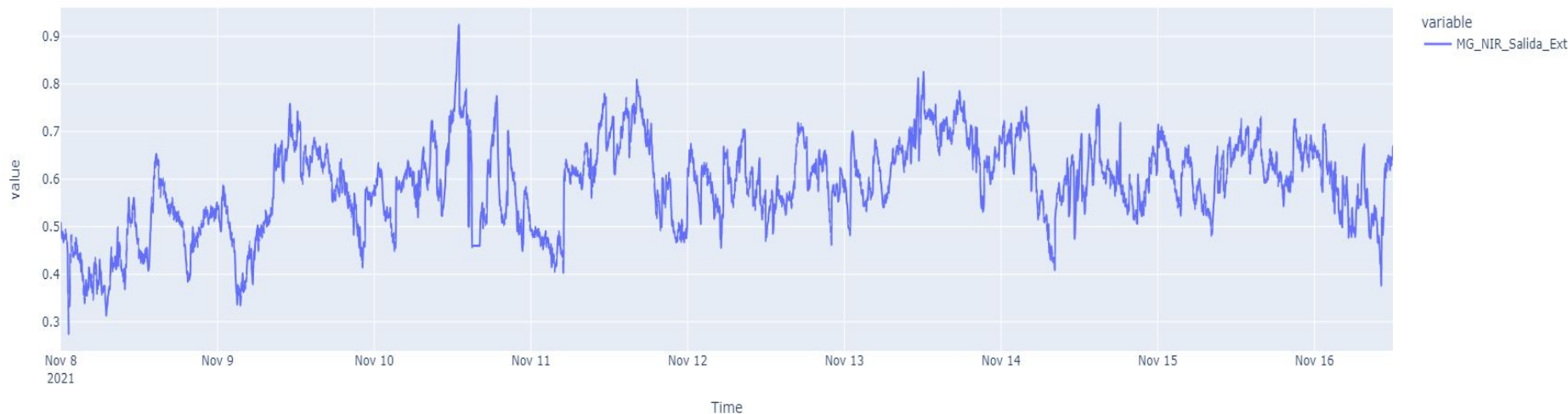
- 70% extracción en forma **mecánica**.
- 30% extracción en forma **química**.
- **Contras** de una mala extracción.
- Método de **medición**.

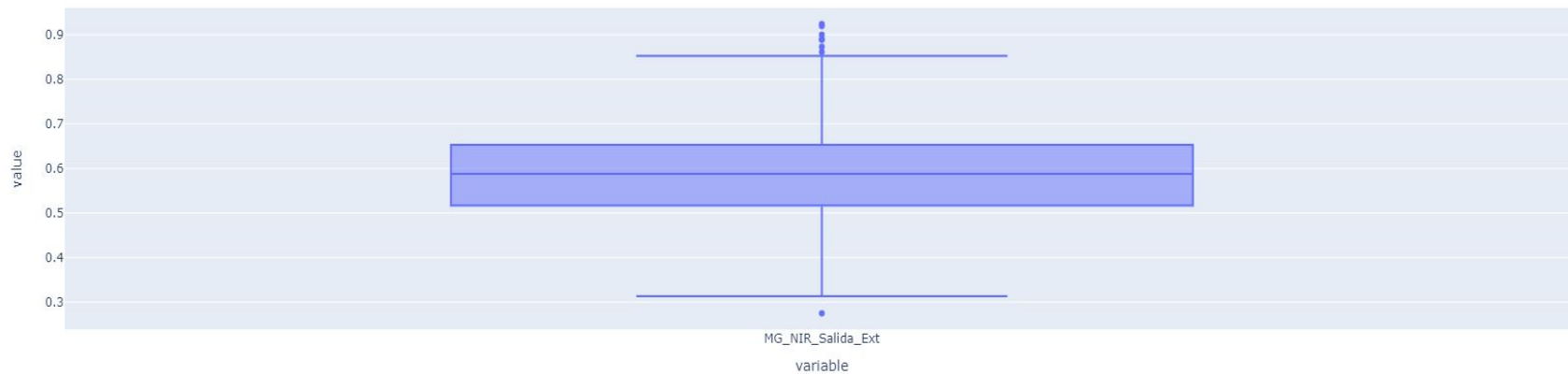
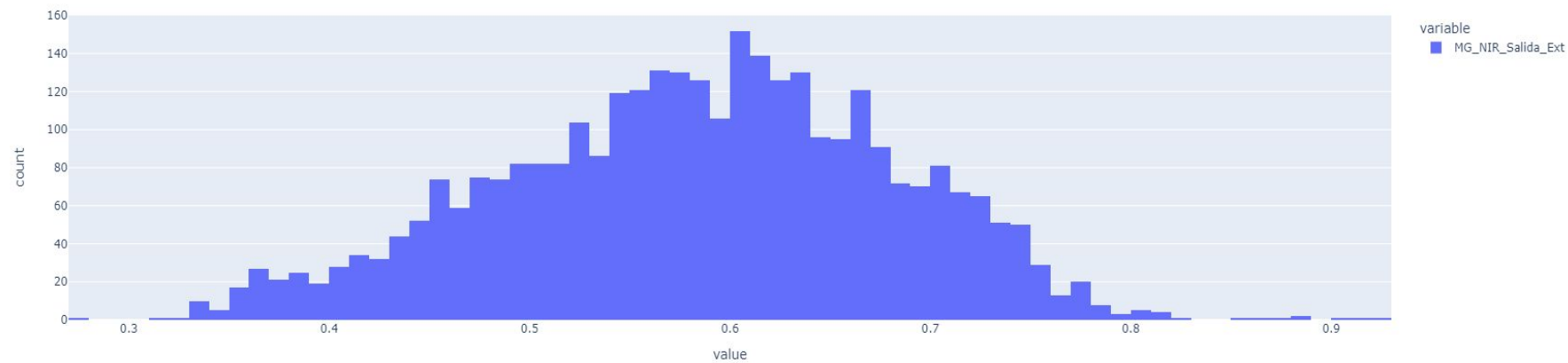


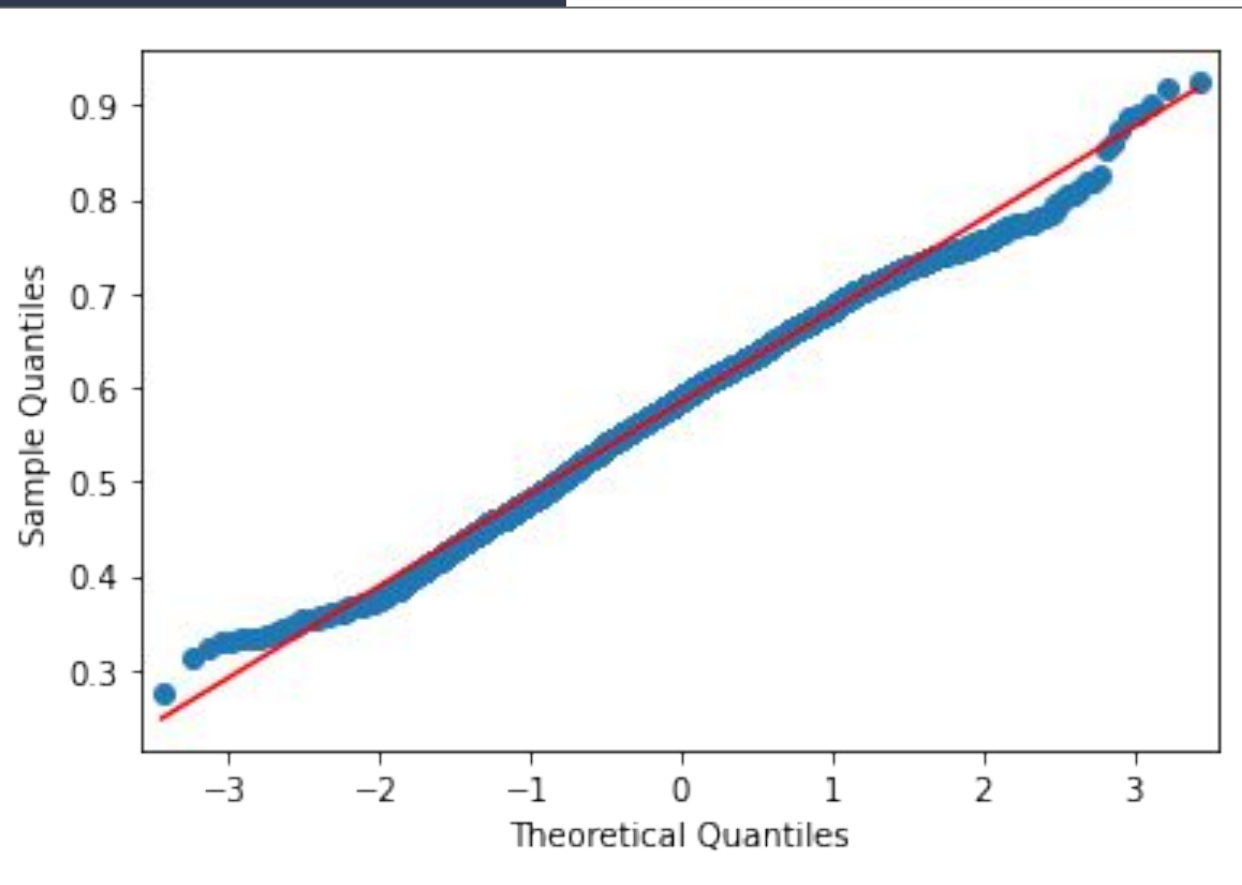
DESCRIPCIÓN DE LA SERIE

Medición entregada por la instrumentación NIR a la salida del extractor.

- 9 días de medición.
- $T_s = 120$ s.
- Por inspección visual parece **estacionaria**.
- **Interpretación.**

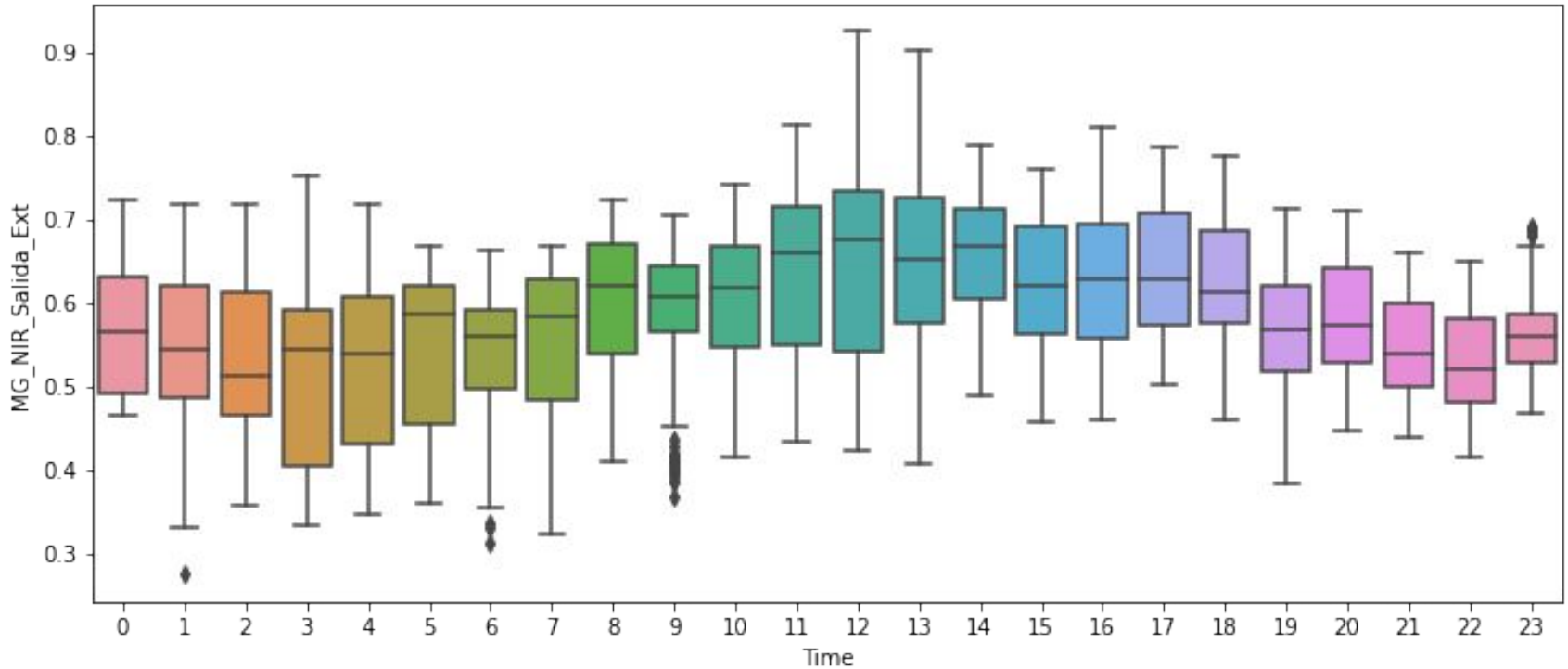




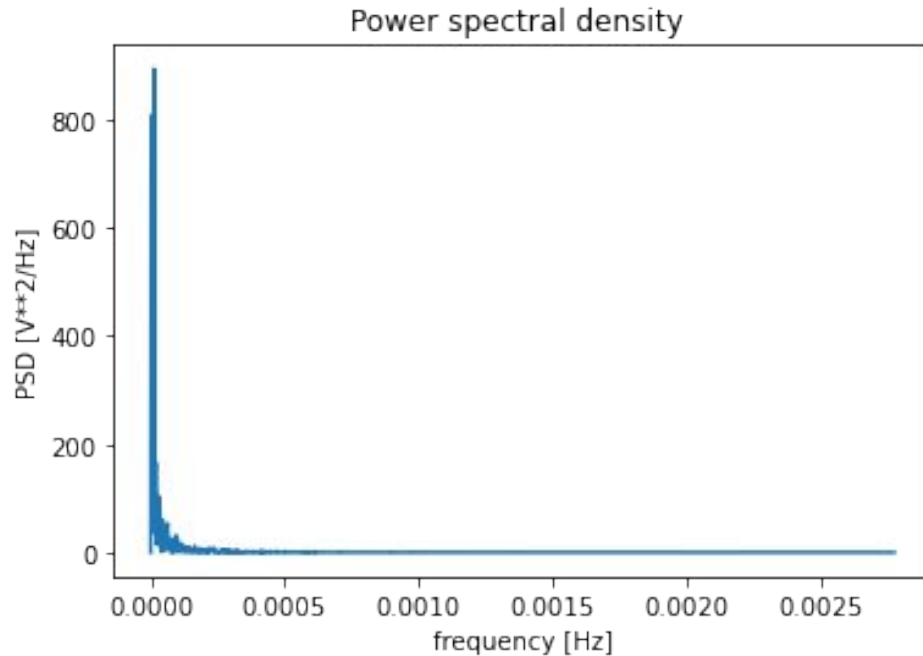


ANÁLISIS POR HORA

- Aparentemente podría haber una componente estacional.
- Se observa gran desvío en el análisis individual por cada hora.

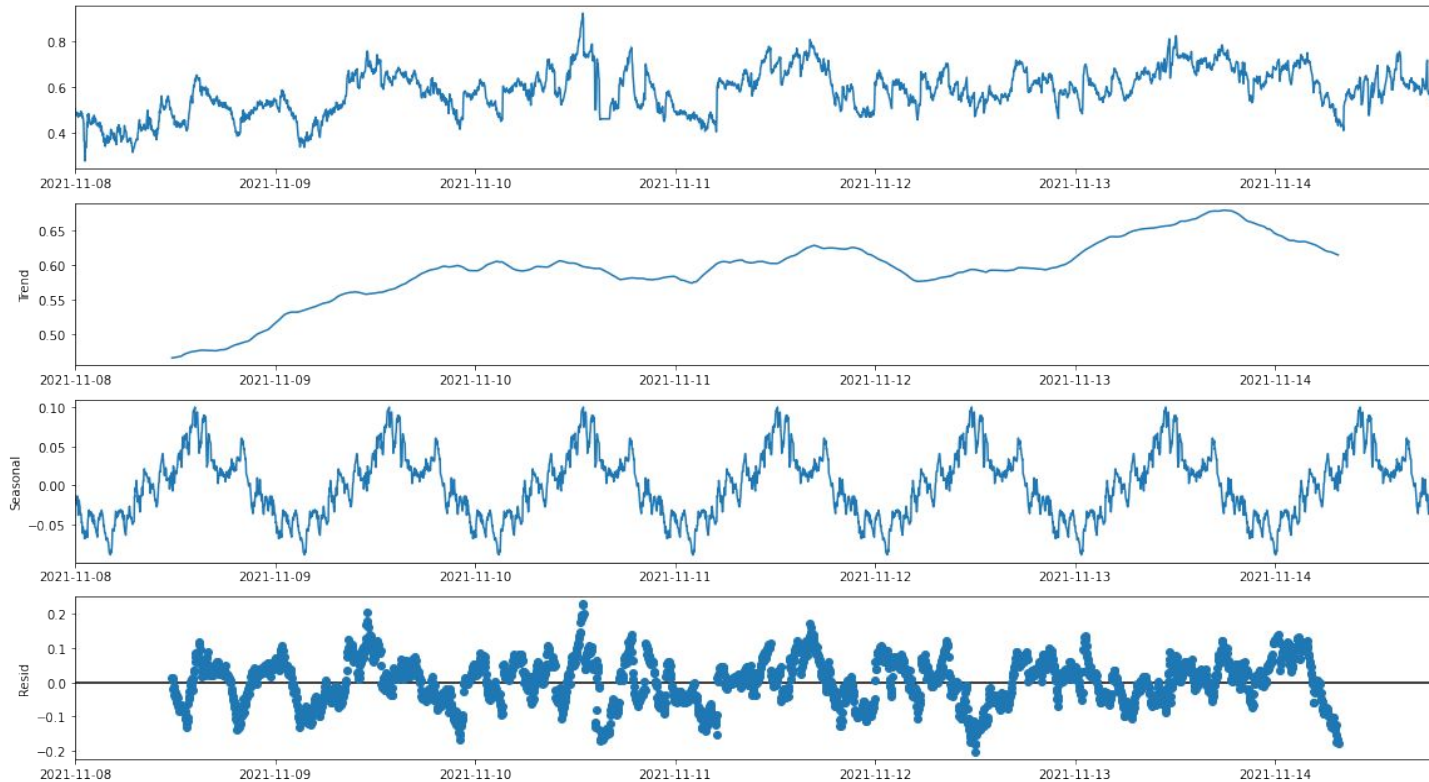


ANÁLISIS ESPECTRAL DE POTENCIA



La frecuencia de la
fundamental es: $[1.1910\text{e-}05]$
[Hz.] que representa un periodo
de: 23.321 hs.

DESCOMPOSICIÓN DE LA SEÑAL



Test de Dickey-Fuller

H0 : la serie tiene una raíz unitaria (serie no estacionaria).

H1 : la serie no tiene una raíz unitaria (serie estacionaria).

Si la prueba falla al rechazar H0 (**p-value > 0.05**), este test puede ser prueba de que la serie no es estacionaria.

```
adfuller(y)
```

```
(-4.5042631850148656,
```

```
0.00019318132272168076,
```

```
15,
```

```
3249,
```

```
{ '1%': -3.4323643041051666,
```

```
'5%': -2.862429999065346,
```

```
'10%': -2.5672437656428793},
```

```
-16657.24865815721)
```

Rechazo H0: podría ser estacionaria,
aplicando al menos una diferenciación.

Test de KPSS

H0 : la serie no tiene una raíz unitaria (serie estacionaria).

H1 : la serie tiene una raíz unitaria (serie no estacionaria).

Si la prueba falla al rechazar H0 (**p-value > 0.05**), este test puede ser prueba de que la serie es estacionaria.

```
kpss (y.dropna())
```

```
(2.627170360737052,
```

```
0.01,
```

```
36,
```

```
{ '10%': 0.347, '5%': 0.463, '2.5%':  
0.574, '1%': 0.739})
```

Rechazo H0: no podría ser estacionaria.

Diferenciación

Test de KPSS

H0 : la serie no tiene una raíz unitaria (serie estacionaria).

H1 : la serie tiene una raíz unitaria (serie no estacionaria).

Si la prueba falla al rechazar H0 (**p-value > 0.05**), este test puede ser prueba de que la serie es estacionaria.

```
kpss(y.diff().dropna())
```

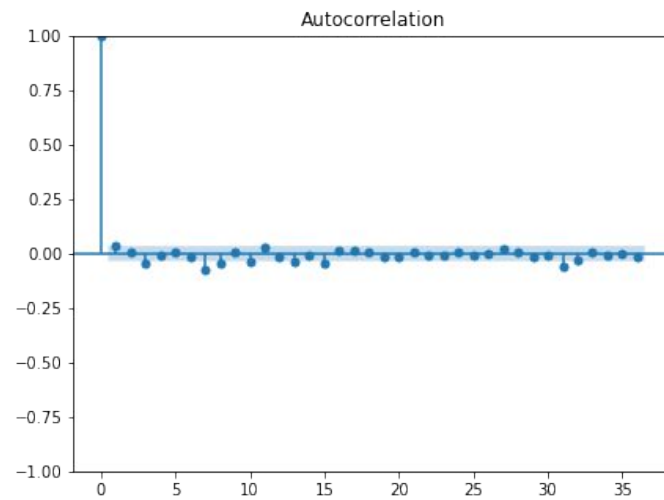
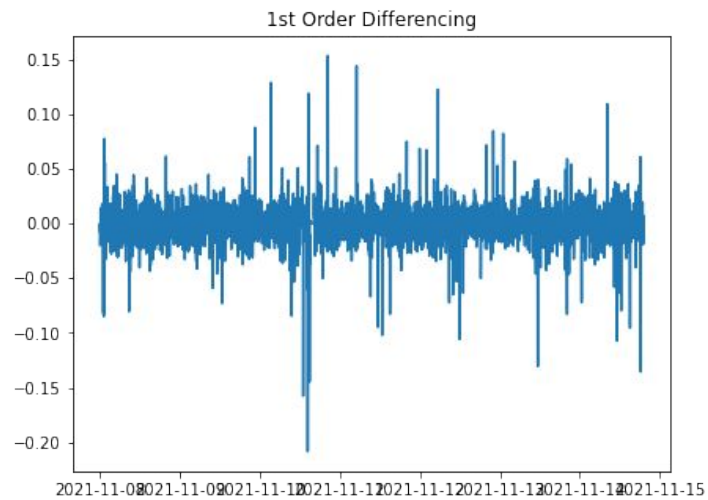
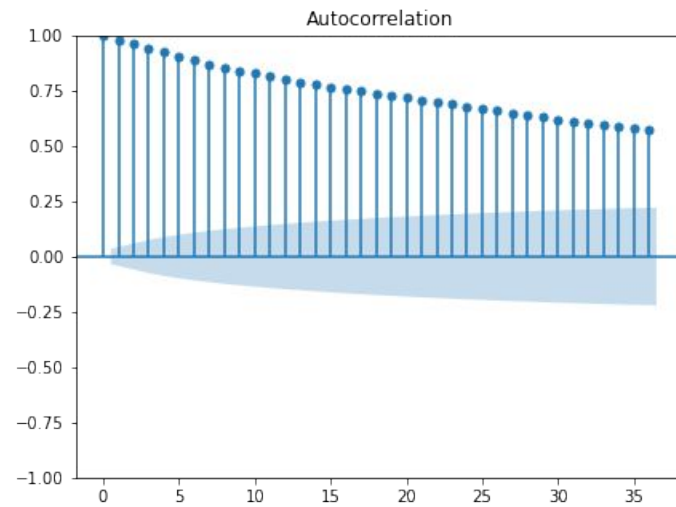
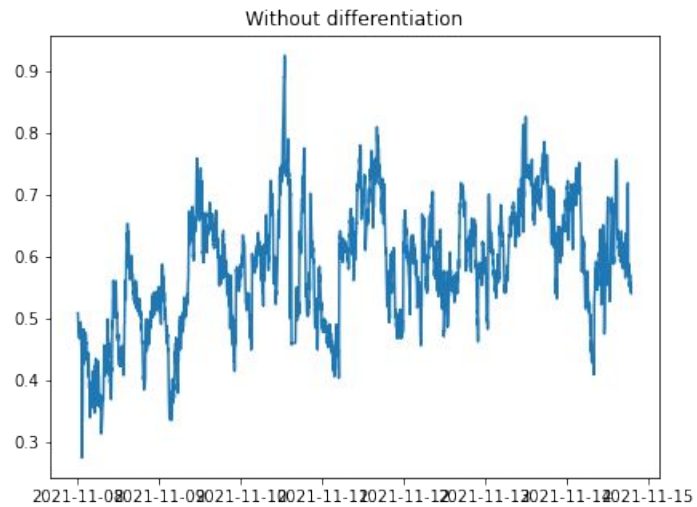
```
(0.011539242684388825,
```

```
0.1,
```

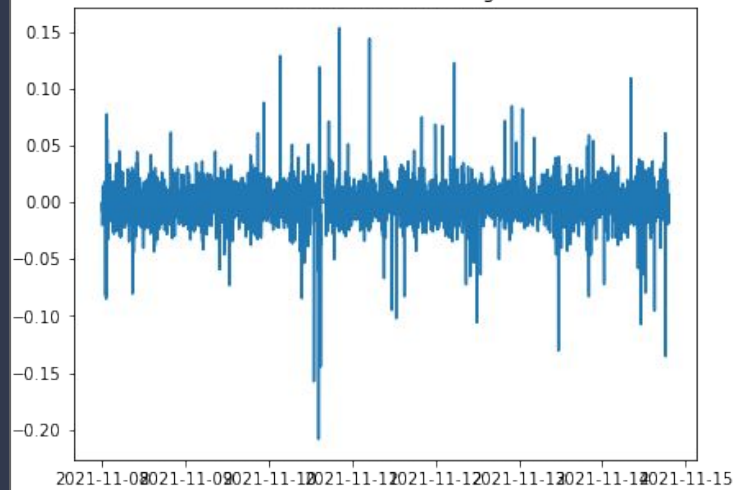
```
9,
```

```
{ '10%': 0.347, '5%': 0.463, '2.5%':  
0.574, '1%': 0.739})
```

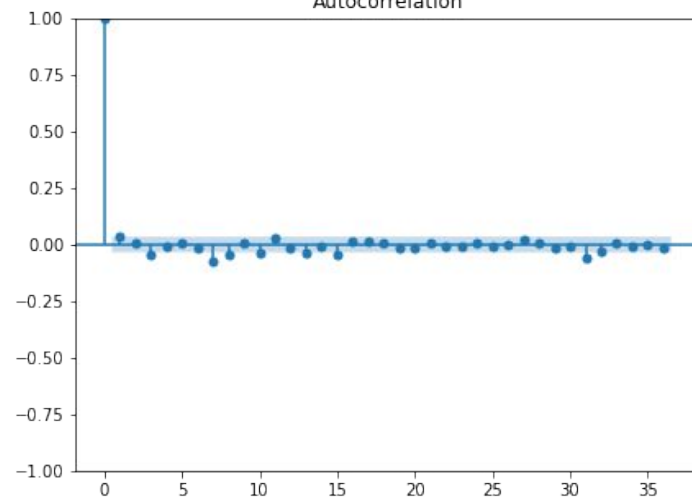
No rechazo H0: podría ser estacionaria.



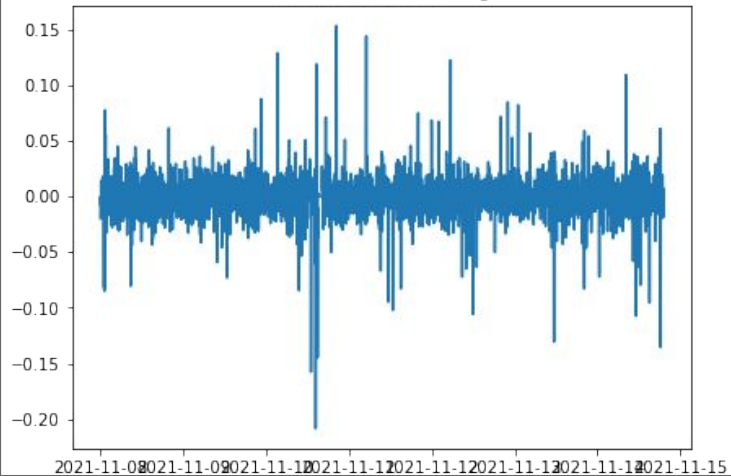
1st Order Differencing



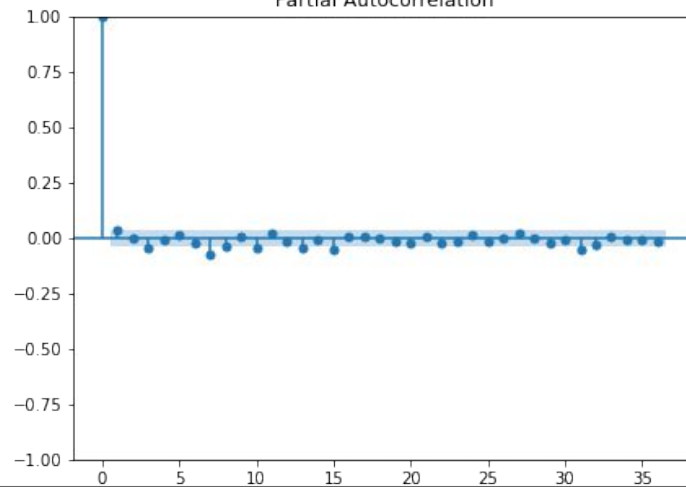
Autocorrelation



1st Order Differencing



Partial Autocorrelation



MODELO ARIMA

$p = 0 / 1$

$q = 0 / 1$

$d = 1$

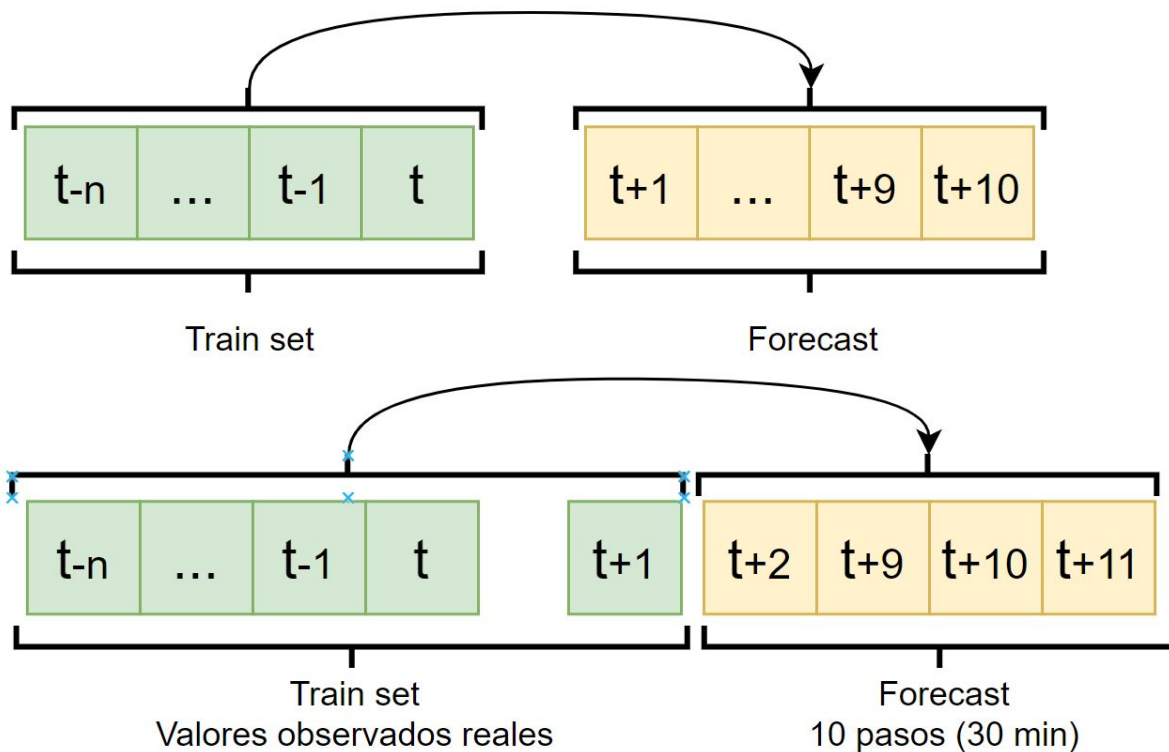
```
SARIMAX Results
=====
Dep. Variable:      MG_NIR_Salida_Ext  No. Observations:      3265
Model:              ARIMA(1, 1, 1)    Log Likelihood         8356.032
Date:               Sat, 11 Jun 2022  AIC                             -16706.064
Time:               16:30:25          BIC                             -16687.792
Sample:             11-08-2021        HQIC                            -16699.520
                  - 11-14-2021
Covariance Type:    opg
=====
              coef    std err          z      P>|z|      [0.025    0.975]
-----
ar.L1          0.0192     0.440     0.044     0.965    -0.844     0.882
ma.L1          0.0176     0.441     0.040     0.968    -0.846     0.881
sigma2         0.0003    3.23e-06   108.208     0.000     0.000     0.000
=====
Ljung-Box (L1) (Q):                0.00  Jarque-Bera (JB):                32054.09
Prob(Q):                           0.99  Prob(JB):                     0.00
Heteroskedasticity (H):              1.13  Skew:                          -0.58
Prob(H) (two-sided):                0.05  Kurtosis:                       18.31
=====
```

MÉTODO DE PREDICCIÓN

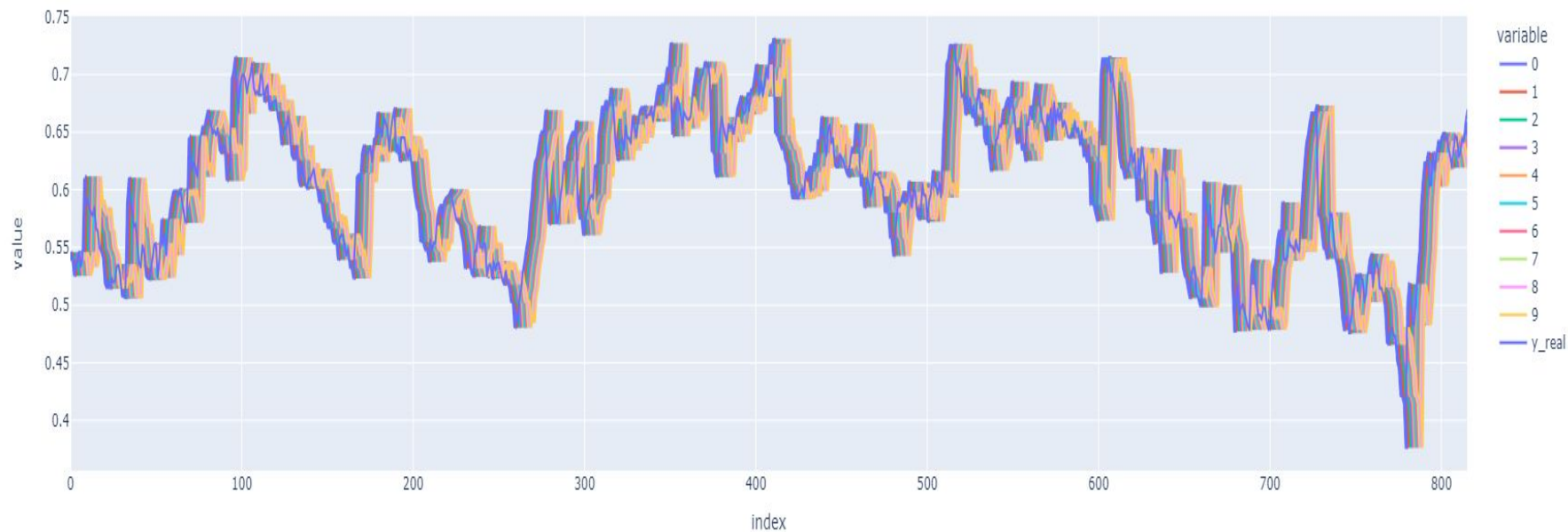
FORWARD VALIDATION

Modelos que predicen varios instantes de tiempo:

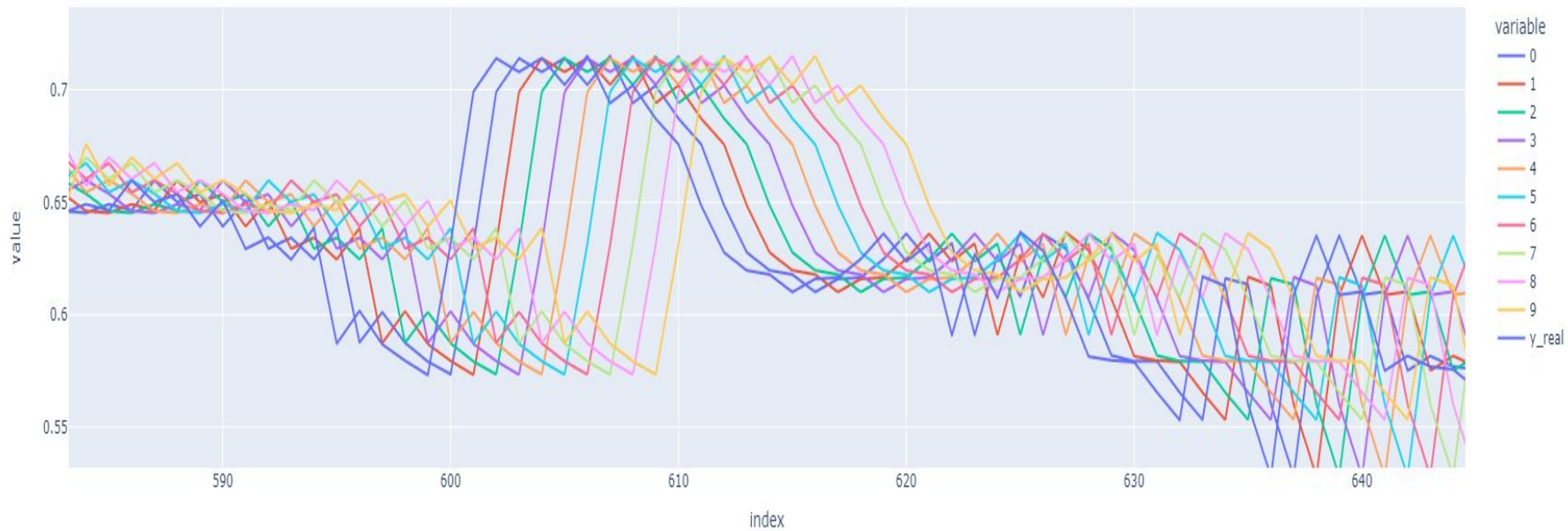
- Predicciones de disparo único
- Predicciones autorregresivas



PREDICCIÓN DEL MODELO ARIMA

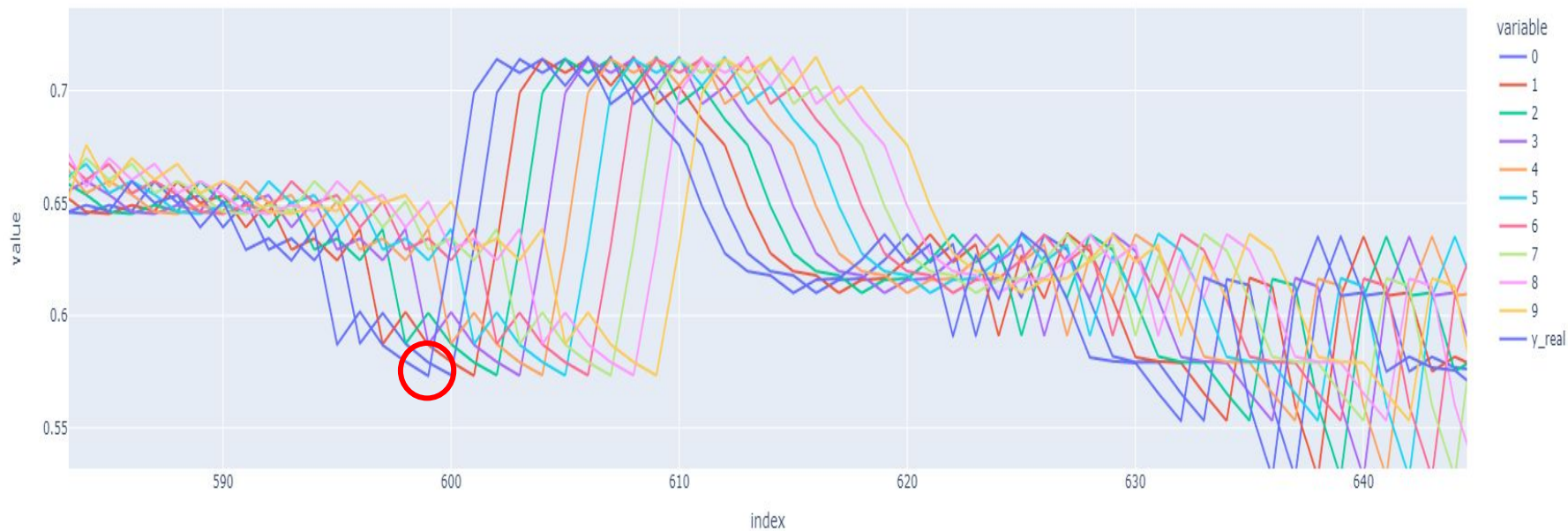


PREDICCIÓN DEL MODELO ARIMA



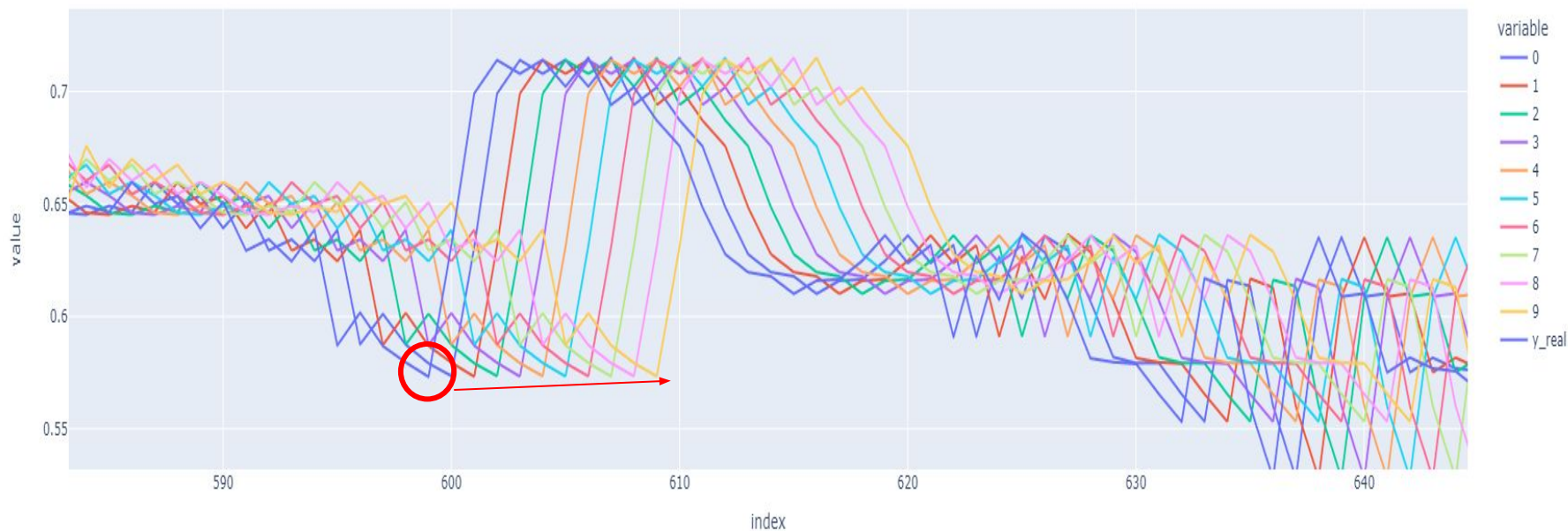
PREDICCIÓN DEL MODELO ARIMA

Las predicciones futuras poseen mucha información del último valor observado.



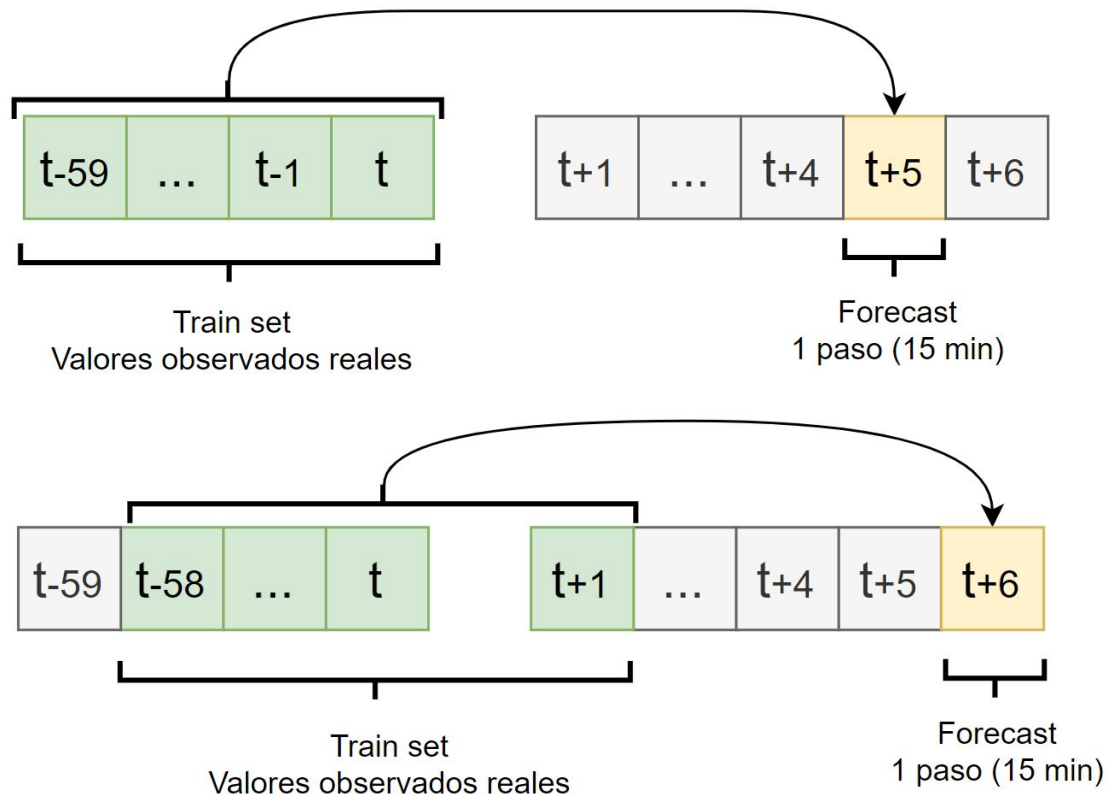
PREDICCIÓN DEL MODELO ARIMA

Las predicciones futuras poseen mucha información del último valor observado.



MODELO LSTM

SLIDING WINDOW



MODELO LSTM

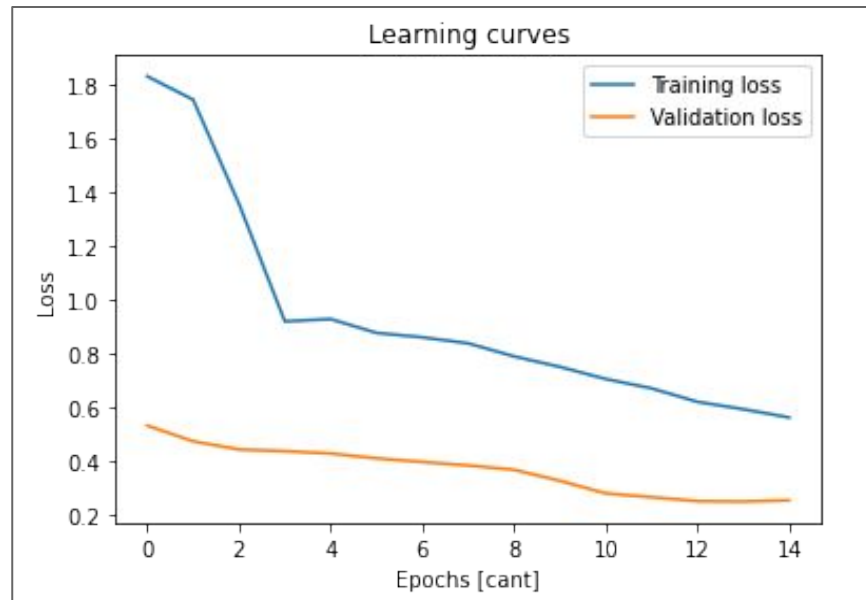
Model: "sequential_6"

Layer (type)	Output Shape	Param #
lstm_16 (LSTM)	(None, 60, 32)	4992
lstm_17 (LSTM)	(None, 60, 16)	3136
dropout_12 (Dropout)	(None, 60, 16)	0
lstm_18 (LSTM)	(None, 8)	800
dropout_13 (Dropout)	(None, 8)	0
dense_6 (Dense)	(None, 1)	9

Total params: 8,937

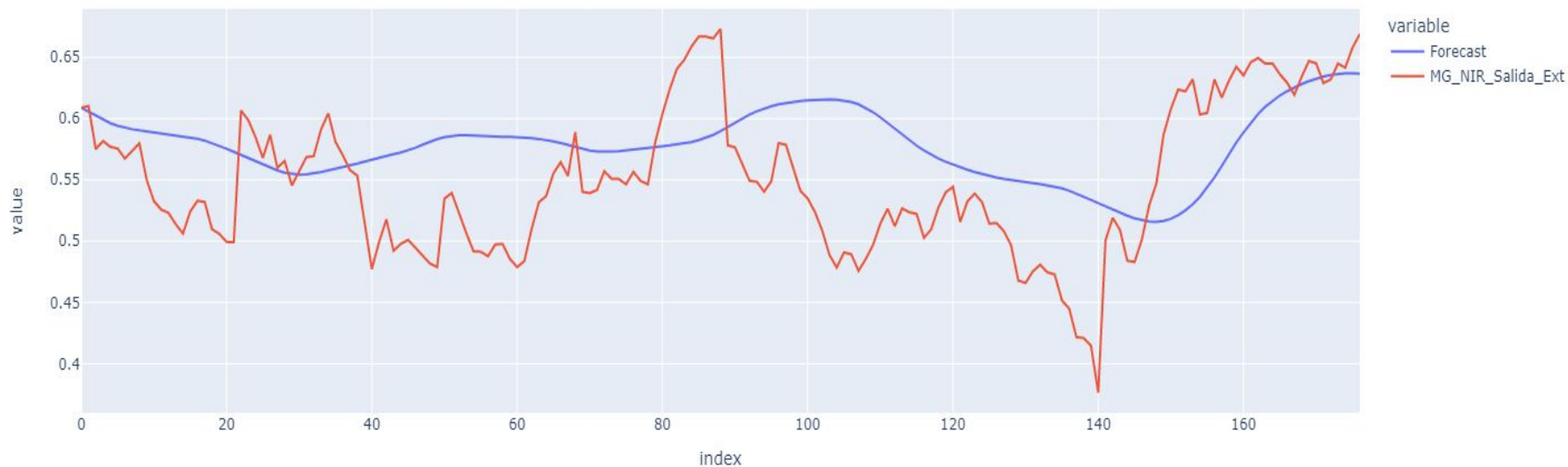
Trainable params: 8,937

Non-trainable params: 0



PREDICCIÓN DEL MODELO LSTM

El MSE obtenido con LSTM es: 0.8119624721168339
El MAE obtenido con LSTM es: 0.6795461893272343
El EVS obtenido con LSTM es: 0.09383994571872878
El R2 obtenido con LSTM es: -0.1439878986291503

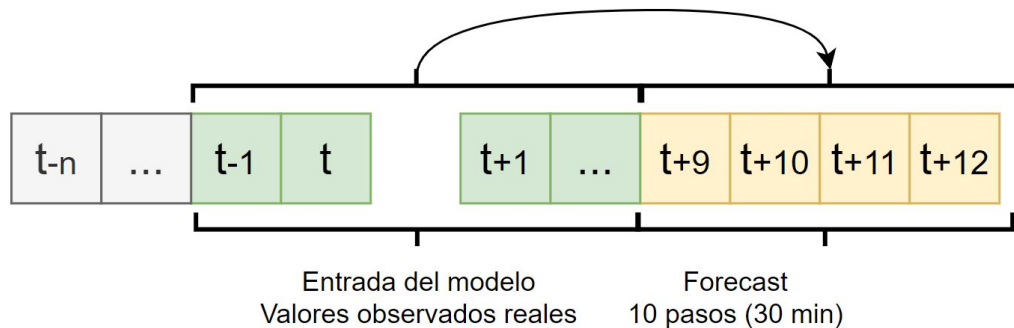
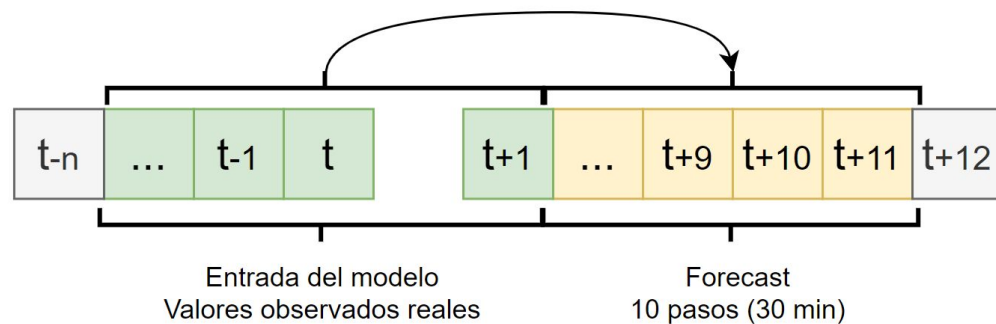
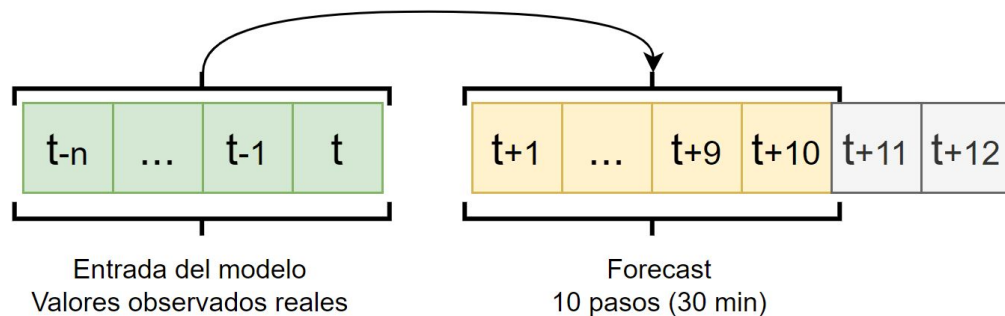


CONCLUSIONES Y PRÓXIMOS PASOS: MEJORAS AL

PREDICCIONES DE DISPARO ÚNICO

MODELOS DE DL PUEDEN CAPTURAR
TENDENCIAS

SE PODRÍAN COMBINAR AMBOS
MODELOS



PRÓXIMOS PASOS: LSTNet

Modeling Long- and Short-Term Temporal Patterns with Deep Neural Networks

SIGIR'18, July 2018, Ann Arbor, MI, USA

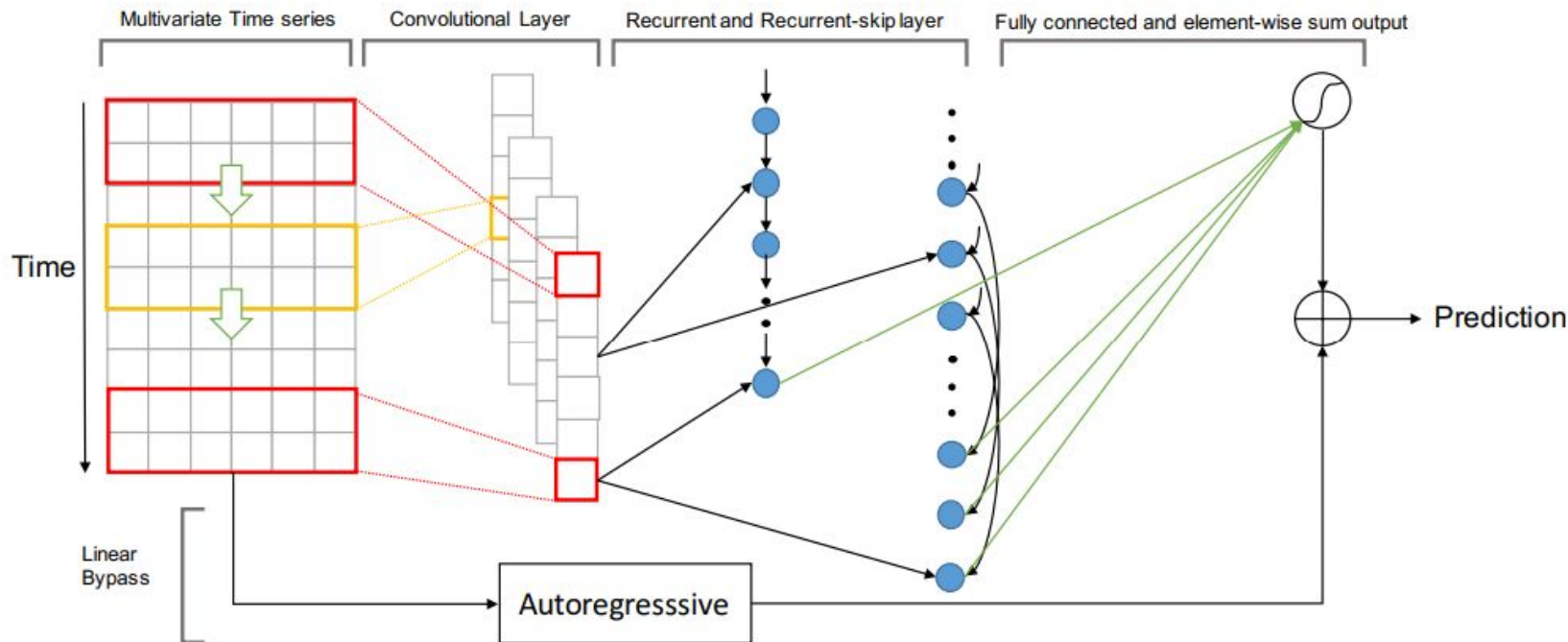


Figure 2: An overview of the Long- and Short-term Time-series network (LSTNet)