

ARIMA, SARIMA, AIC



Week 10 - Day 03

Recap

bloody complex

Forecasting is ~~easy~~

AR = Simple linear regression of lags

MA = Reactions to previous shocks

BTW what's the meaning of “**SHIOK**”?

:)

$$\text{ARMA} = \text{AR} + \text{MA}$$

Non seasonal!

Stationarity is needed!

Differencing \rightarrow Stationarity

ARIMA

AR + I + MA

I = Integrated = Differencing

ARIMA = ARMA + stationary management

ARIMA(1,1,0) = ???

$$\text{ARIMA}(1,1,0) = \text{diff}(1) + \text{AR}(1)$$

SARIMA

Seasonal ARIMA

SARIMA

=

ARIMA + seasonality

$$\text{ARIMA} = (p,d,q)$$

$$\text{SARIMA} = (p,d,q)(P,D,Q,\text{order})$$

Normal ARIMA

Seasonal
Period

$$\text{SARIMA} = (2,0,1)(0,0,1,12)$$

Seasonal part

AIC

Do you remember regularization?

Loss + Penalty

of parameters



$$\text{AIC} = \text{Log}(\text{loss}) + 2 * K$$



Based on RSS

Simple = Better

“relative metric”

=

one model against the other,

given the same dataset!

More like Accuracy

Less like AUC

<https://www.youtube.com/watch?v=QuNhTLVgV2>

Y

“Every time someone solely uses an AIC statistic for model selection, an angel loses its wings.

Every time someone thoughtlessly minimises it, an angel not only loses its wings, but is cast out of Heaven and falls in most extreme agony into the everlasting fire.”

Check your models
residuals, predictions, etc.

Do they make sense??

Residuals: strange patterns?

Model Evaluation

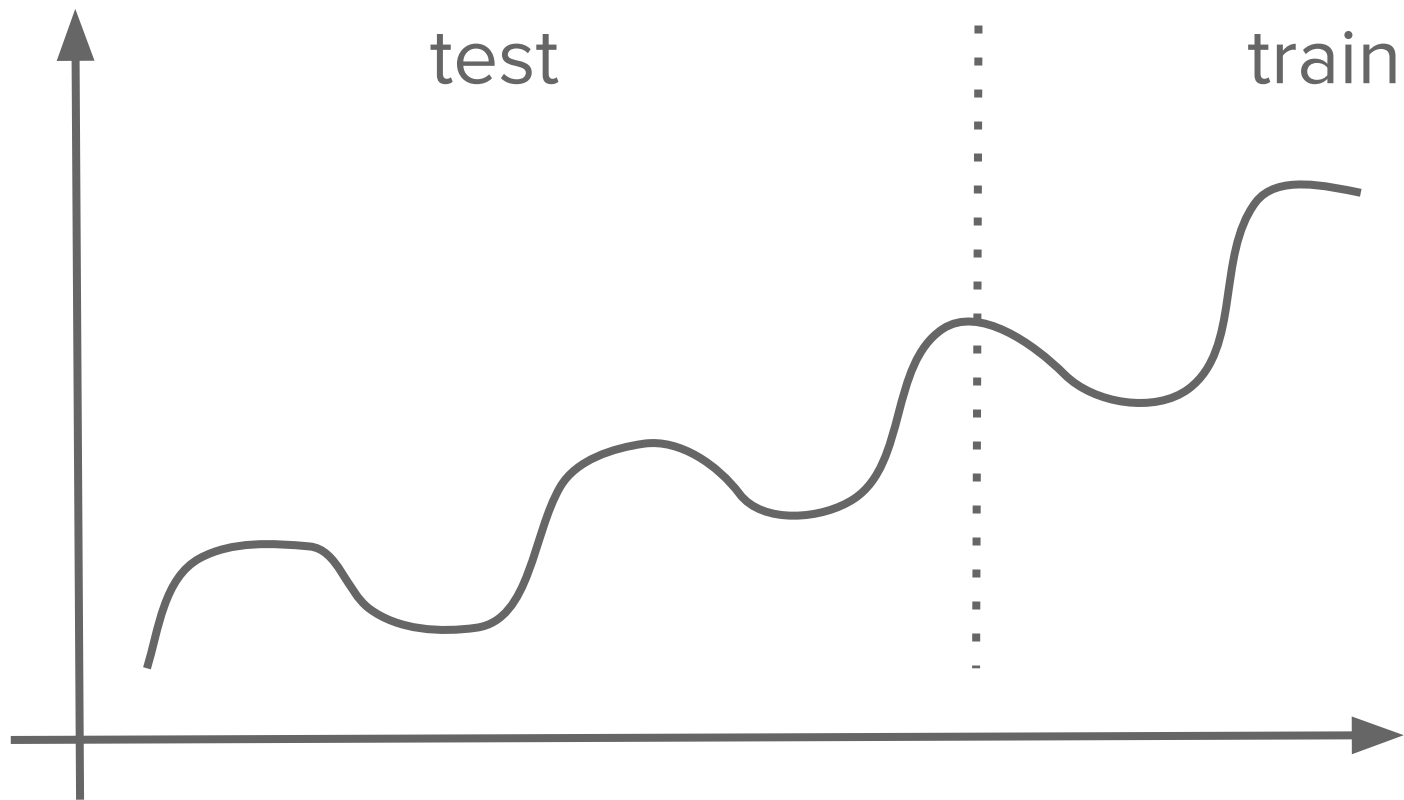
Standard Cross Validation is wrong

Why?

Standard Cross Validation is wrong

=

Peeking at the future



<u>Time</u>	<u>Measure</u>
0	32
1	34
2	27
3	22
4	34
5	35
6	31
7	29
8	27

train

test

<u>Time</u>	<u>Measure</u>
0	32
1	34
2	27
3	22
4	34
5	35
6	31
7	29
8	27



Used data



predict

<u>Time</u>	<u>Measure</u>
0	32
1	34
2	27
3	22
4	34
5	35
6	31
7	29
8	27



Used data



predict

<u>Time</u>	<u>Measure</u>
0	32
1	34
2	27
3	22
4	34
5	35
6	31
7	29
8	27



Used data



predict

```
ar1 = ARMA(data['unemp_diff'], (1, 0)).fit()  
ar1.predict(start=100, dynamic=True)
```

Multivariate Time Series

SARIMAX

=

SARIMA + eXogenous variables

VAR

=

Vector Auto Regression