

Report

Contents

- Original Data
- Filter Data
- Elimination Data
- Slided Data (*window size=2*)
- Optimization Linear Regression (*window size=2*)
- Best hyperparameters Linear Regression (*window size=2*)
- Metrics evaluation Linear Regression (*window size=2*)
- Slided Data (*window size=3*)
- Optimization Linear Regression (*window size=3*)
- Best hyperparameters Linear Regression (*window size=3*)
- Metrics evaluation Linear Regression (*window size=3*)
- Slided Data (*window size=4*)
- Optimization Linear Regression (*window size=4*)
- Best hyperparameters Linear Regression (*window size=4*)
- Metrics evaluation Linear Regression (*window size=4*)
- Distribution Data
- PCA analysis
- ICA Analysis
- Correlation matrix
- Validation curve Linear Regression (*fit_intercept, window size=2*)
- Validation curve Linear Regression (*normalize, window size=2*)
- Prediction Linear Regression (*UCI, window size=2*)
- Validation curve Linear Regression (*fit_intercept, window size=3*)
- Validation curve Linear Regression (*normalize, window size=3*)
- Prediction Linear Regression (*UCI, window size=3*)
- Validation curve Linear Regression (*fit_intercept, window size=4*)
- Validation curve Linear Regression (*normalize, window size=4*)
- Prediction Linear Regression (*UCI, window size=4*)
- Window size comparison Linear Regression

Original Data

Fecha	cod_in	provincia	Casos	Fallecidos	Hospitalizados	UCI
2020-01-01	0	No consta	0	0	0	0
2020-01-01	1	Araba/Álava	0	0	0	0
2020-01-01	2	Albacete	0	0	0	0
2020-01-01	3	Alicante/Alacant	0	0	0	0
2020-01-01	4	Almería	0	0	0	0

Filter Data

Fecha	cod_ine	Casos	Fallecidos	Hospitalizados	UCI
2020-01-01	110	0	0	1	0
2020-01-02	110	0	0	0	0
2020-01-03	110	0	0	0	0
2020-01-04	110	0	0	0	0
2020-01-05	110	0	0	0	0

Elimination Data

Fecha	Casos	Fallecidos	Hospitalizados	UCI
2020-01-01	0	0	1	0
2020-01-02	0	0	0	0
2020-01-03	0	0	0	0
2020-01-04	0	0	0	0
2020-01-05	0	0	0	0

Slided Data (*window size=2*)

Fecha	Casos_t-1	Casos	Fallecidos_t-1	Fallecidos	Hospitalizados_t-1	Hospitalizados
2020-01-02	0	0	0	0	1	0
2020-01-03	0	0	0	0	0	0
2020-01-04	0	0	0	0	0	0
2020-01-05	0	0	0	0	0	0
2020-01-06	0	0	0	0	0	0

UCI_t-1	UCI
0	0
0	0
0	0
0	0
0	0

Optimization Linear Regression (*window size=2*)

mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_normalize	param_fit_intercept	params
0.0032	0.0004	0.0022	0.0004	True	True	{'normalize': 'True', 'fit_intercept': 'True'}
0.0032	0.0004	0.002	0.0	False	True	{'normalize': 'False', 'fit_intercept': 'True'}
0.003	0.0	0.0016	0.0005	True	False	{'normalize': 'True', 'fit_intercept': 'False'}
0.0028	0.0007	0.0024	0.0008	False	False	{'normalize': 'False', 'fit_intercept': 'False'}

split0_test_score	split1_test_score	split2_test_score	split3_test_score	split4_test_score	mean_test_score	std_test_score
0.7807	0.8211	0.0649	0.4123	0.6399	0.5438	0.2789
0.7807	0.8211	0.0649	0.4123	0.6399	0.5438	0.2789
0.7807	0.8211	0.0649	0.4123	0.6399	0.5438	0.2789
0.7807	0.8211	0.0649	0.4123	0.6399	0.5438	0.2789

rank_test_score	split0_train_score	split1_train_score	split2_train_score	split3_train_score	split4_train_score	mean_train_score
1	0.843	0.851	0.8429	0.8846	0.8512	0.8545
1	0.843	0.851	0.8429	0.8846	0.8512	0.8545
1	0.843	0.851	0.8429	0.8846	0.8512	0.8545
1	0.843	0.851	0.8429	0.8846	0.8512	0.8545

std_train_score
0.0155
0.0155
0.0155
0.0155

Best hyperparameters Linear Regression (*window size=2*)

Model	normalize	fit_intercept
Linear Regression	True	True

Metrics evaluation Linear Regression (*window size=2*)

Model	MAE	MSE
Linear Regression	2.5634	14.2279

Slided Data (*window size=3*)

Fecha	Casos_t-2	Casos_t-1	Casos	Fallecidos_t-2	Fallecidos_t-1	Fallecidos
2020-01-03	0	0	0	0	0	0
2020-01-04	0	0	0	0	0	0
2020-01-05	0	0	0	0	0	0
2020-01-06	0	0	0	0	0	0
2020-01-07	0	0	0	0	0	0

Hospitalizados_t-2	Hospitalizados_t-1	Hospitalizados	UCI_t-2	UCI_t-1	UCI
1	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0

Optimization Linear Regression (*window size=3*)

mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_normalize	param_fit_intercept	params
0.0034	0.0005	0.002	0.0	True	True	{'normalize': 'True', 'fit_intercept': 'True'}
0.003	0.0006	0.002	0.0	False	True	{'normalize': 'False', 'fit_intercept': 'True'}
0.0048	0.0016	0.0034	0.0008	True	False	{'normalize': 'True', 'fit_intercept': 'False'}
0.0036	0.0008	0.0026	0.0008	False	False	{'normalize': 'False', 'fit_intercept': 'False'}

split0_test_score	split1_test_score	split2_test_score	split3_test_score	split4_test_score	mean_test_score	std_test_score
0.7497	0.7528	0.0087	0.2808	0.617	0.4818	0.2925
0.7497	0.7528	0.0087	0.2808	0.617	0.4818	0.2925
0.7497	0.7528	0.0087	0.2808	0.617	0.4818	0.2925
0.7497	0.7528	0.0087	0.2808	0.617	0.4818	0.2925

rank_test_score	split0_train_score	split1_train_score	split2_train_score	split3_train_score	split4_train_score	mean_train_score
1	0.8458	0.8542	0.8436	0.8897	0.8538	0.8574
1	0.8458	0.8542	0.8436	0.8897	0.8538	0.8574
1	0.8458	0.8542	0.8436	0.8897	0.8538	0.8574
1	0.8458	0.8542	0.8436	0.8897	0.8538	0.8574

std_train_score
0.0167
0.0167
0.0167
0.0167

Best hyperparameters Linear Regression (*window size=3*)

Model	normalize	fit_intercept
Linear Regression	True	True

Metrics evaluation Linear Regression (*window size=3*)

Model	MAE	MSE
Linear Regression	2.5579	14.2907

Slided Data (*window size=4*)

Fecha	Casos_t-3	Casos_t-2	Casos_t-1	Casos	Fallecidos_t-3	Fallecidos_t-2
2020-01-04	0	0	0	0	0	0
2020-01-05	0	0	0	0	0	0
2020-01-06	0	0	0	0	0	0
2020-01-07	0	0	0	0	0	0
2020-01-08	0	0	0	0	0	0

Fallecidos_t-1	Fallecidos	Hospitalizados_t-3	Hospitalizados_t-2	Hospitalizados_t-1	Hospitalizados	UCI_t-3
0	0	1	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	1	0

UCI_t-2	UCI_t-1	UCI
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0

Optimization Linear Regression (*window size=4*)

mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_normalize	param_fit_intercept	params
0.0034	0.0005	0.002	0.0	True	True	{'normalize': 'True', 'fit_intercept': 'True'}
0.0032	0.0004	0.0018	0.0004	False	True	{'normalize': 'False', 'fit_intercept': 'True'}
0.0032	0.0004	0.002	0.0	True	False	{'normalize': 'True', 'fit_intercept': 'False'}
0.003	0.0	0.002	0.0	False	False	{'normalize': 'False', 'fit_intercept': 'False'}

split0_test_score	split1_test_score	split2_test_score	split3_test_score	split4_test_score	mean_test_score	std_test_score
0.7277	0.7939	0.0643	0.2496	0.593	0.4857	0.2823
0.7277	0.7939	0.0643	0.2496	0.593	0.4857	0.2823
0.7277	0.7939	0.0643	0.2496	0.593	0.4857	0.2823
0.7277	0.7939	0.0643	0.2496	0.593	0.4857	0.2823

rank_test_score	split0_train_score	split1_train_score	split2_train_score	split3_train_score	split4_train_score	mean_train_score
1	0.8531	0.8572	0.8497	0.8956	0.8608	0.8633
1	0.8531	0.8572	0.8497	0.8956	0.8608	0.8633
1	0.8531	0.8572	0.8497	0.8956	0.8608	0.8633
1	0.8531	0.8572	0.8497	0.8956	0.8608	0.8633

std_train_score
0.0166
0.0166
0.0166
0.0166

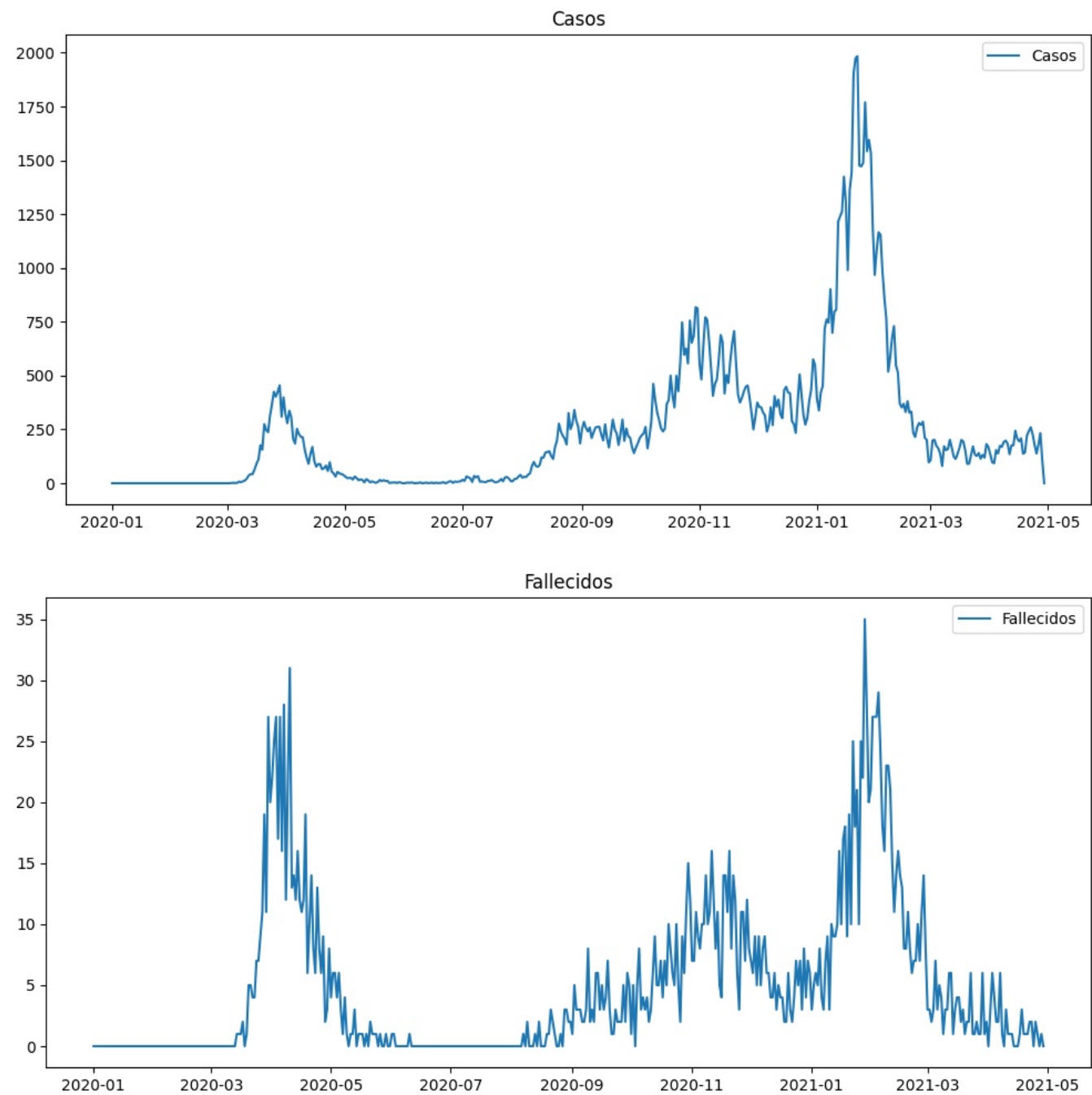
Best hyperparameters Linear Regression (*window size=4*)

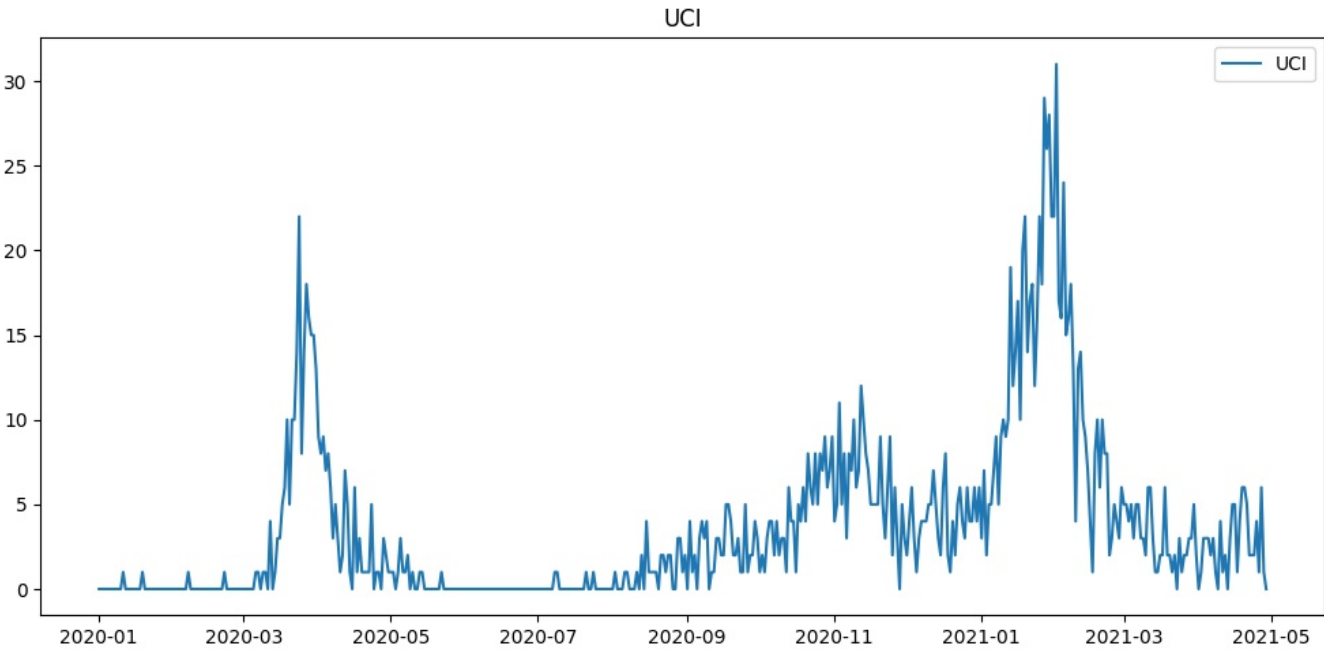
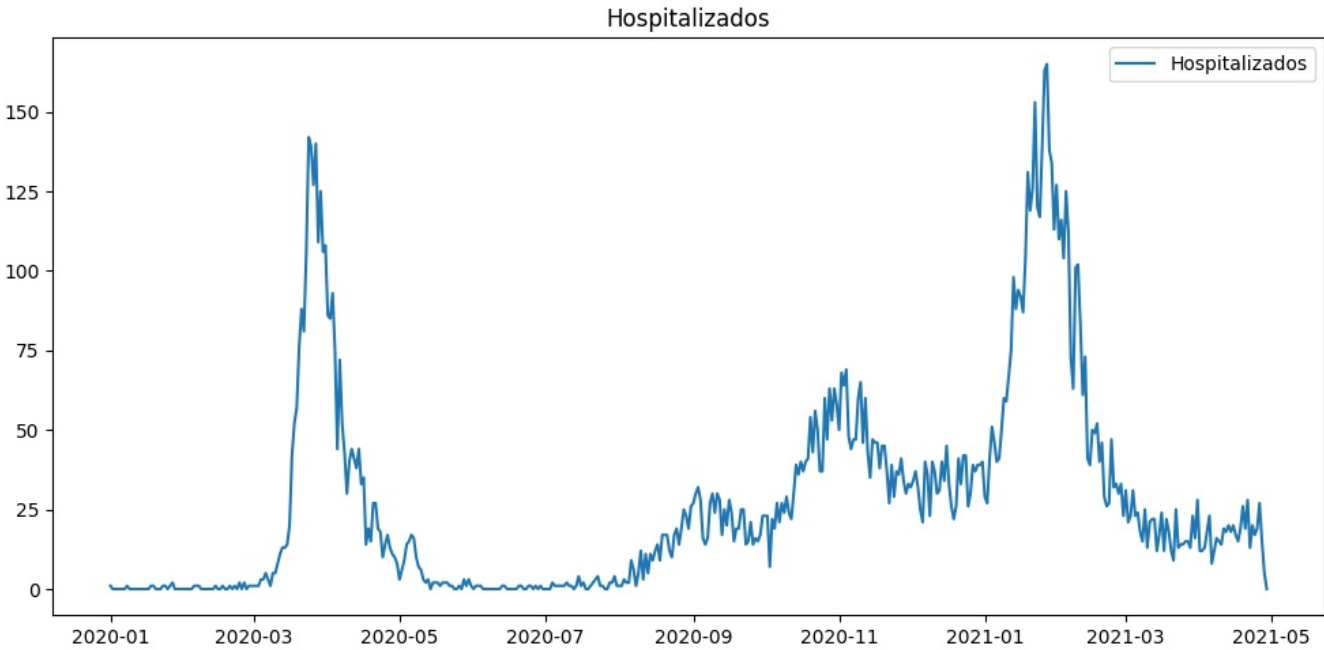
Model	normalize	fit_intercept
Linear Regression	True	True

Metrics evaluation Linear Regression (*window size=4*)

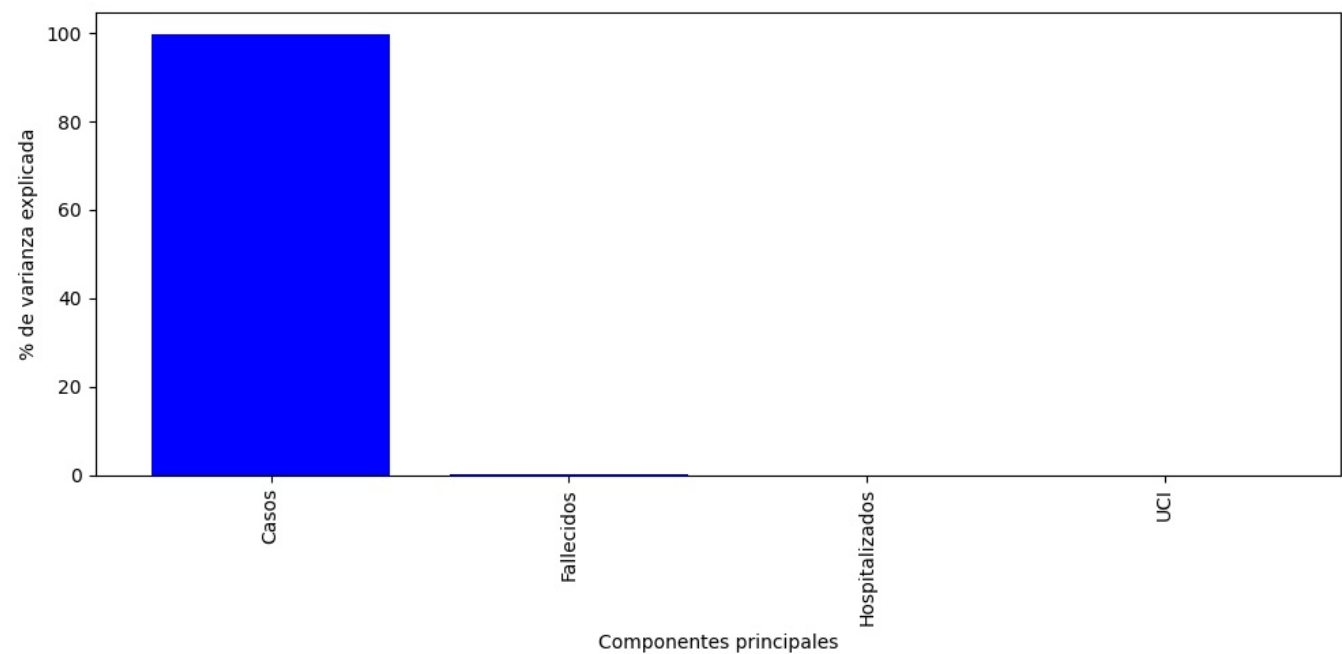
Model	MAE	MSE
Linear Regression	2.4524	12.3391

Distribution Data

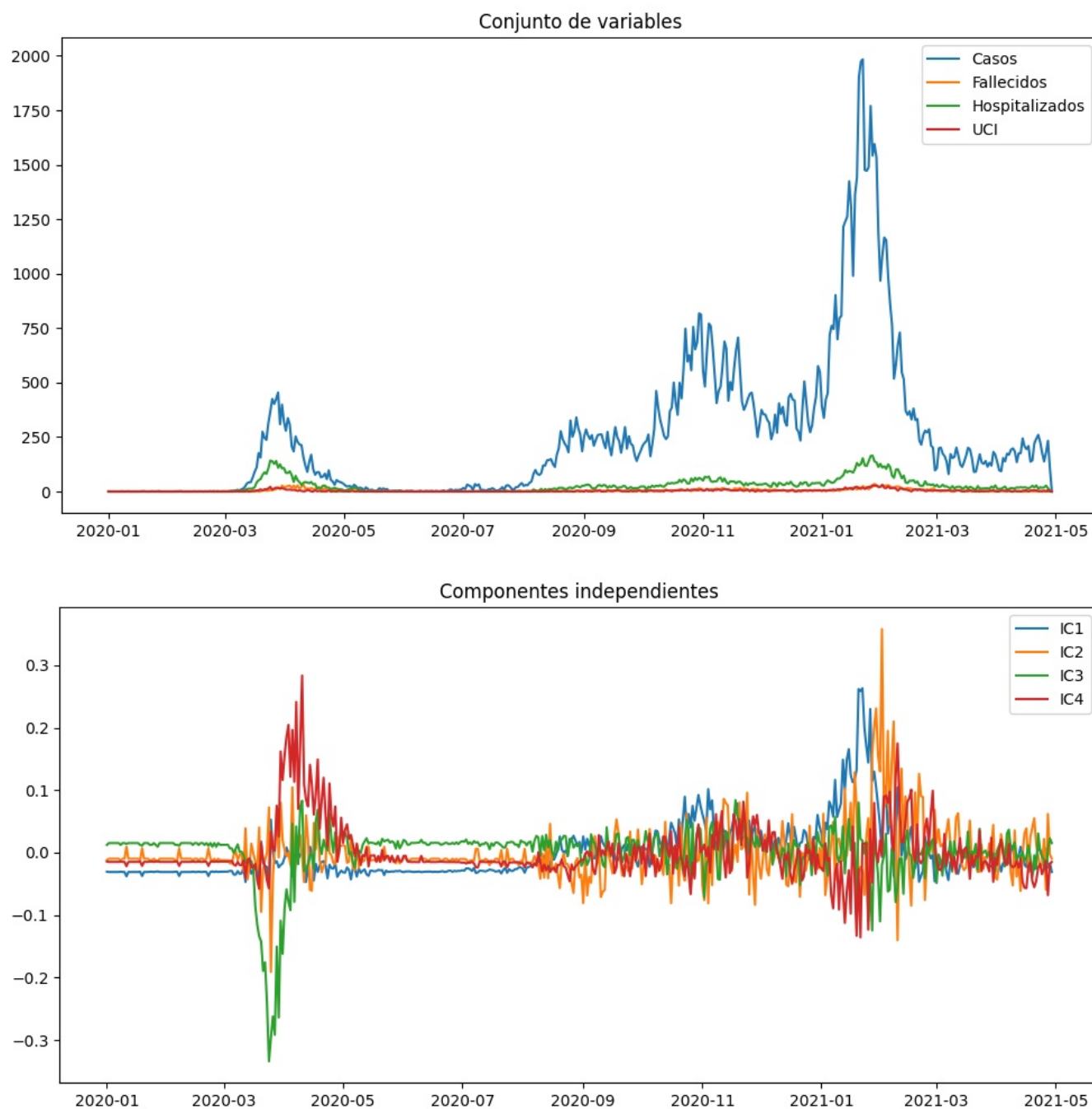




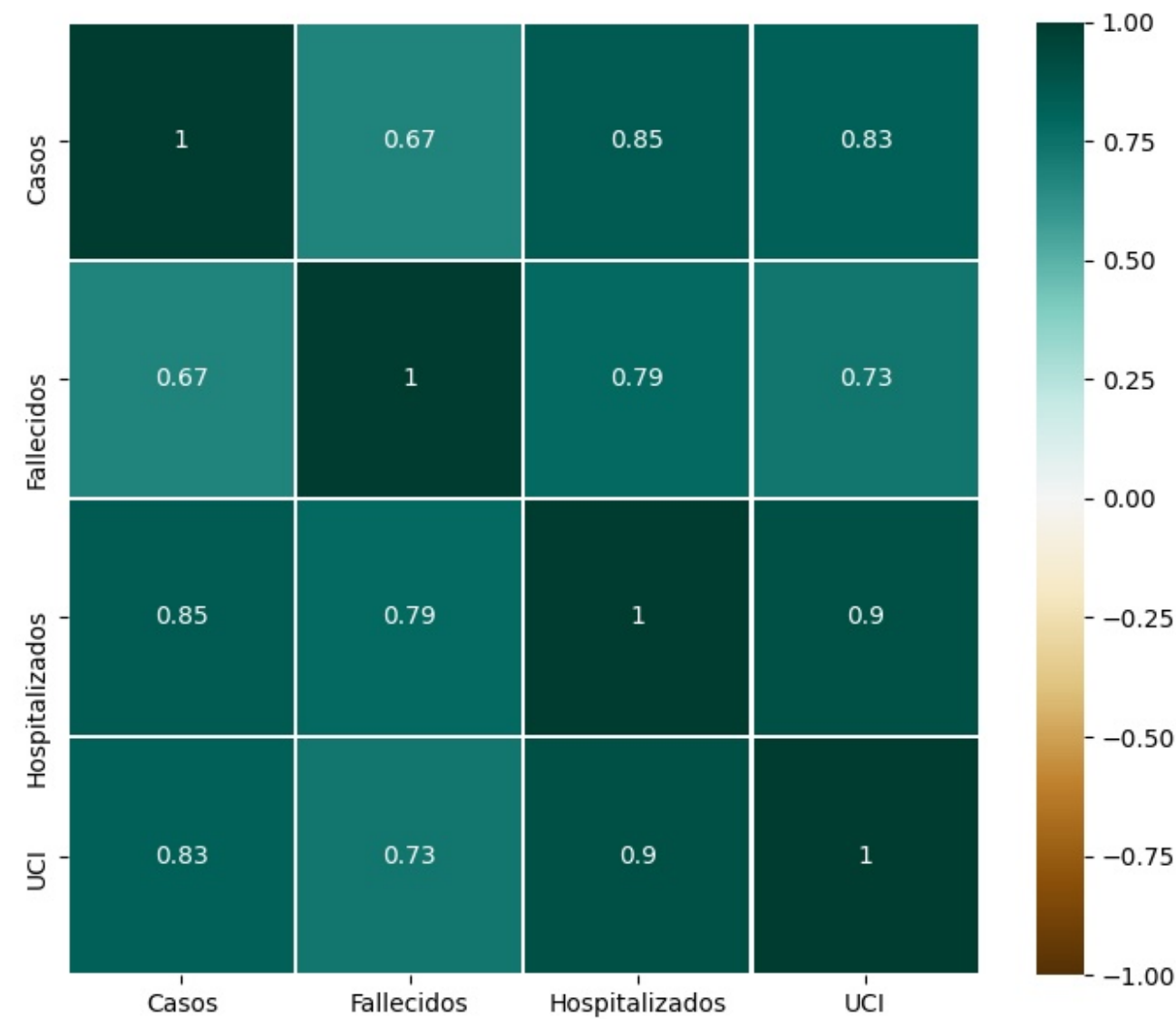
PCA analysys



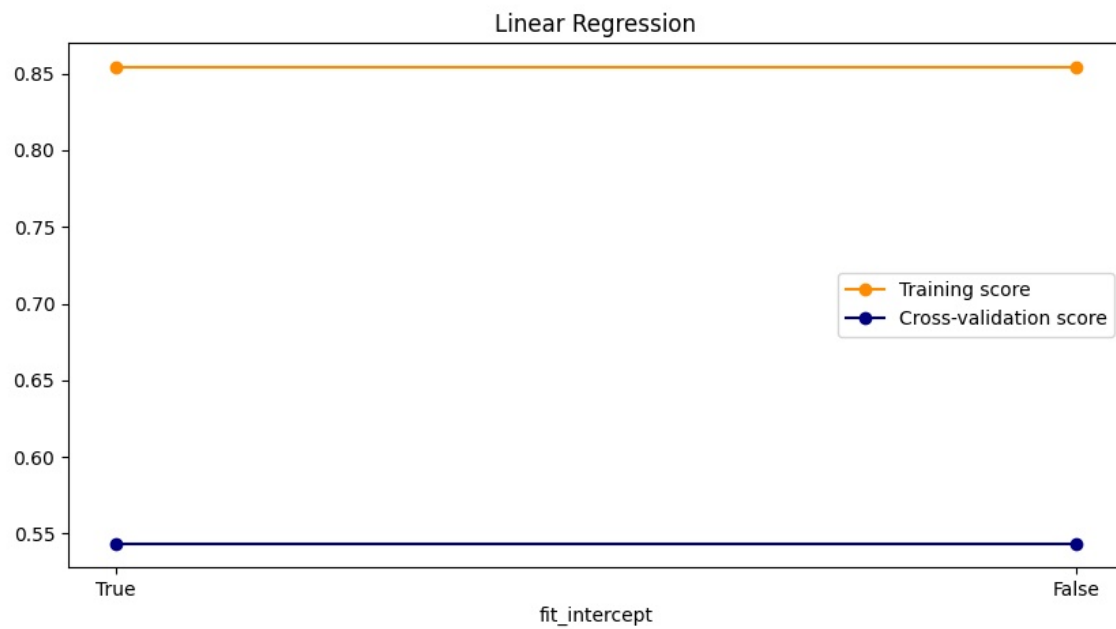
ICA Analysis



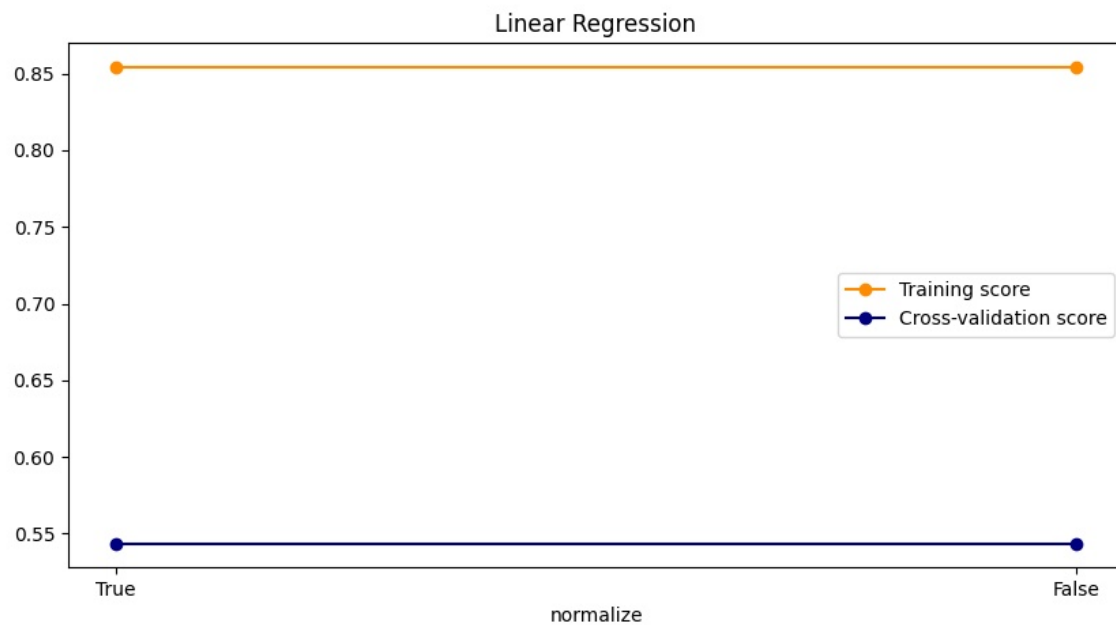
Correlation matrix



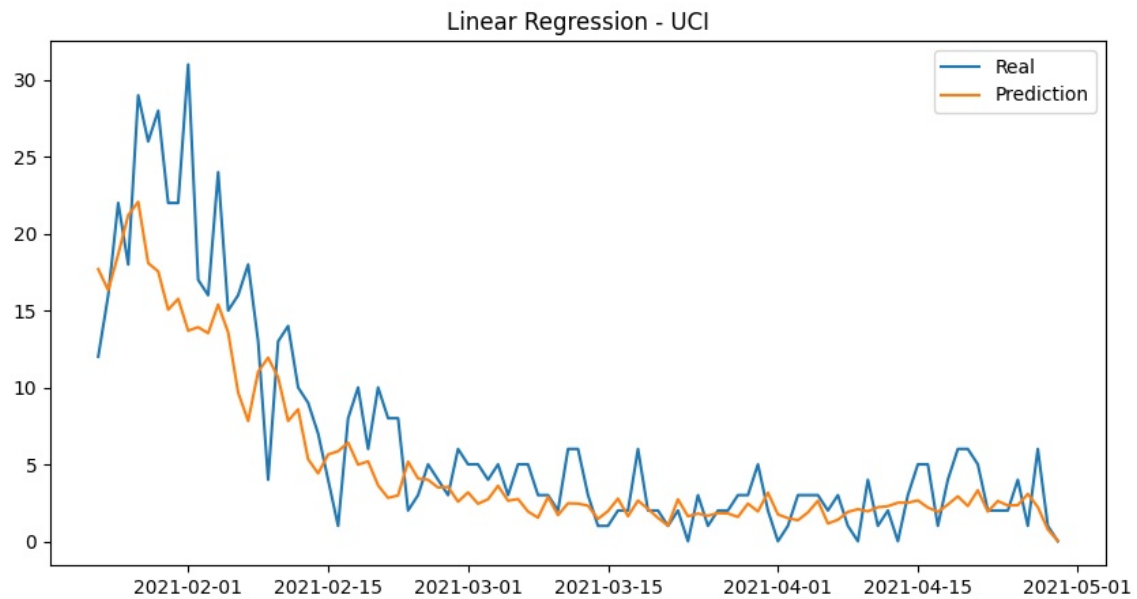
Validation curve Linear Regression (*fit_intercept*, *window_size=2*)



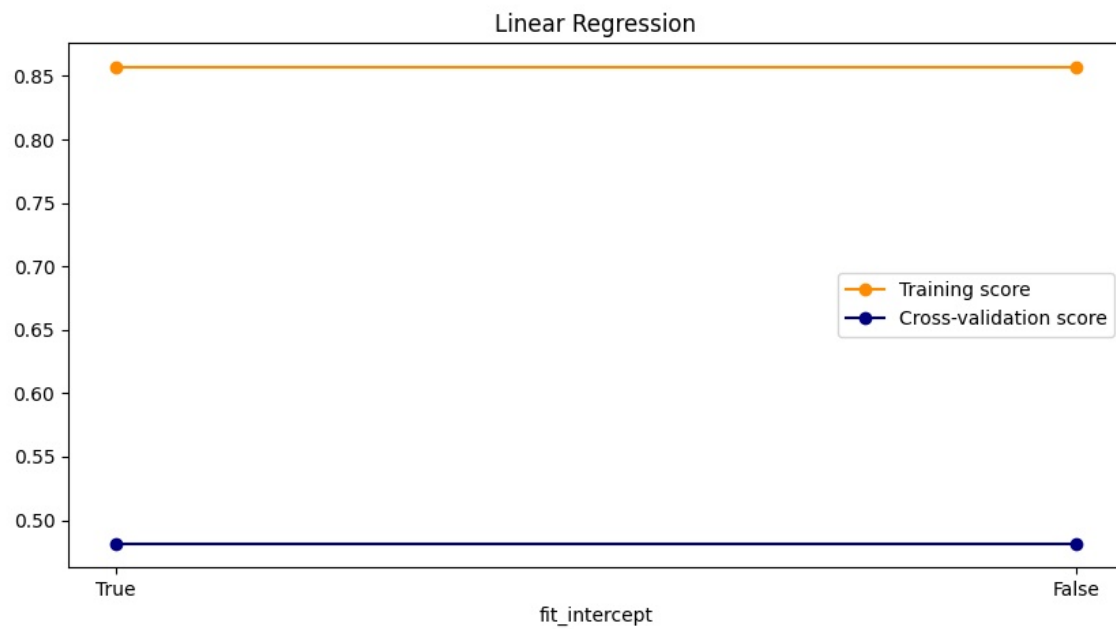
Validation curve Linear Regression (*normalize*, *window_size=2*)



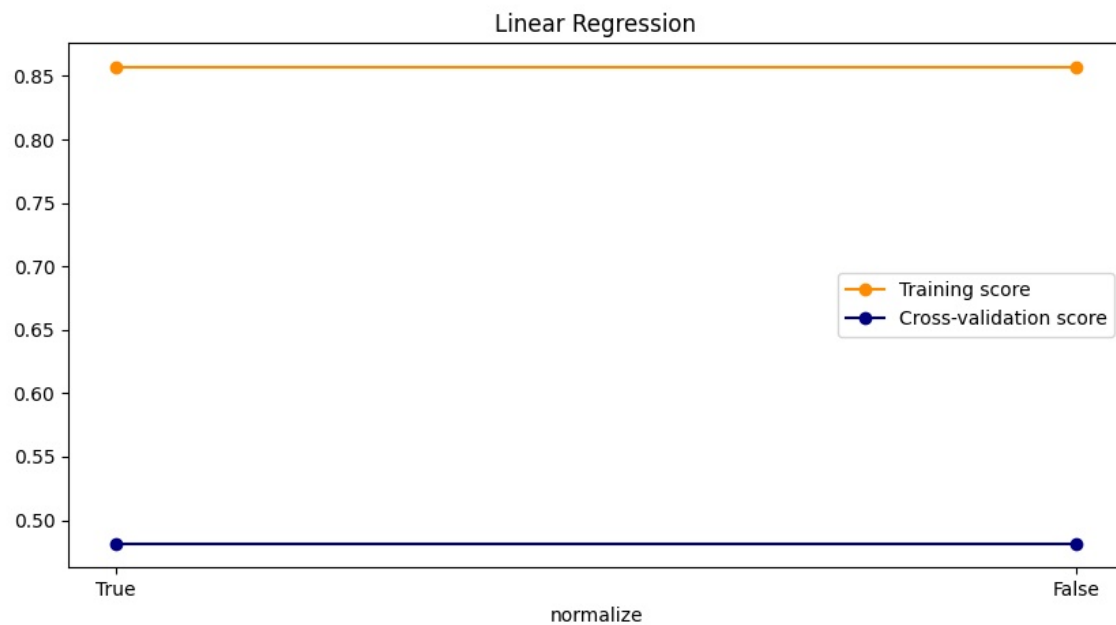
Prediction Linear Regression (*UCI*, *window size=2*)



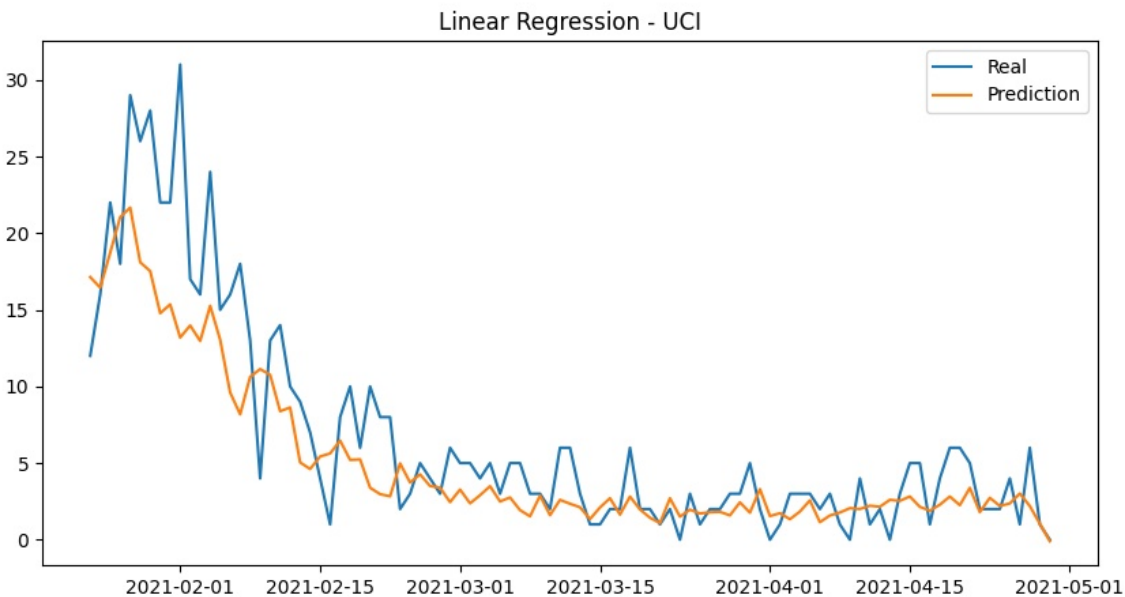
Validation curve Linear Regression (*fit_intercept*, *window_size=3*)



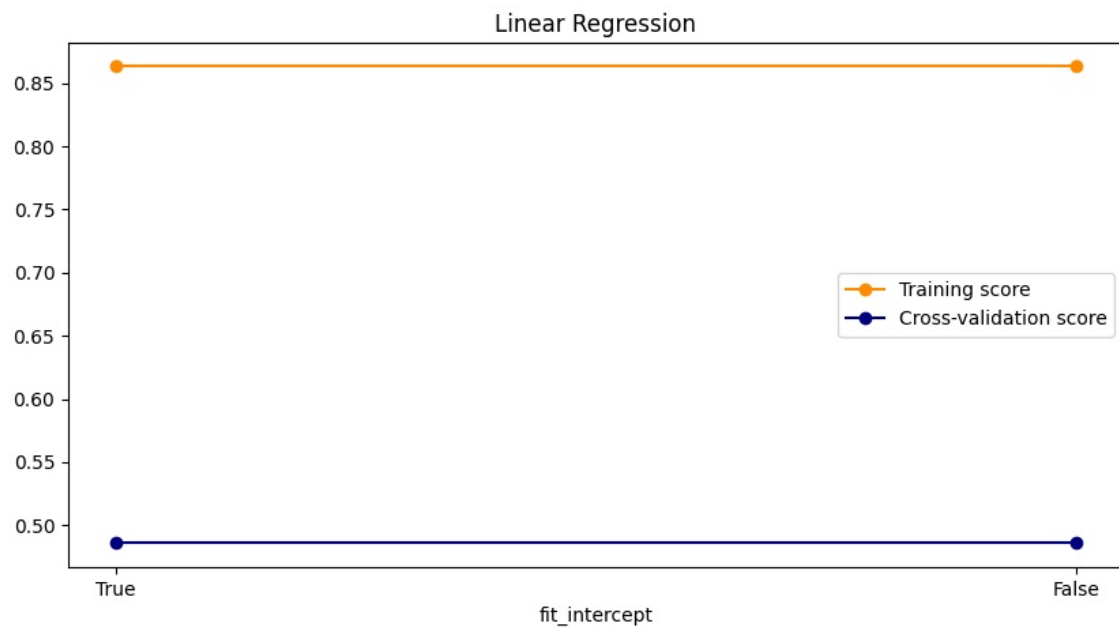
Validation curve Linear Regression (*normalize*, *window_size=3*)



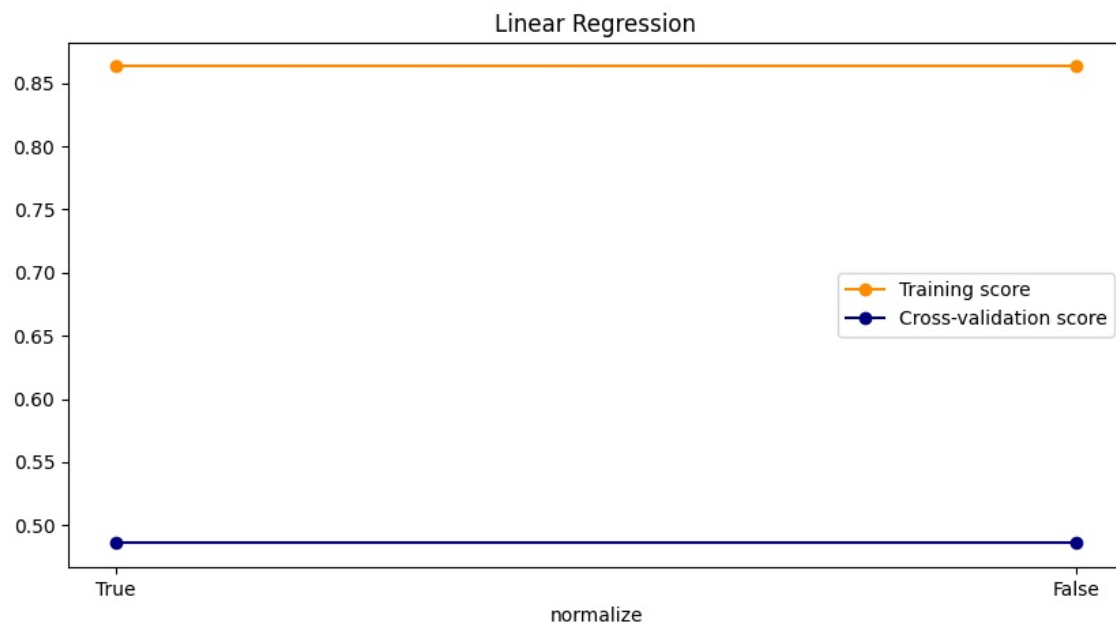
Prediction Linear Regression (*UCI*, *window*=3)



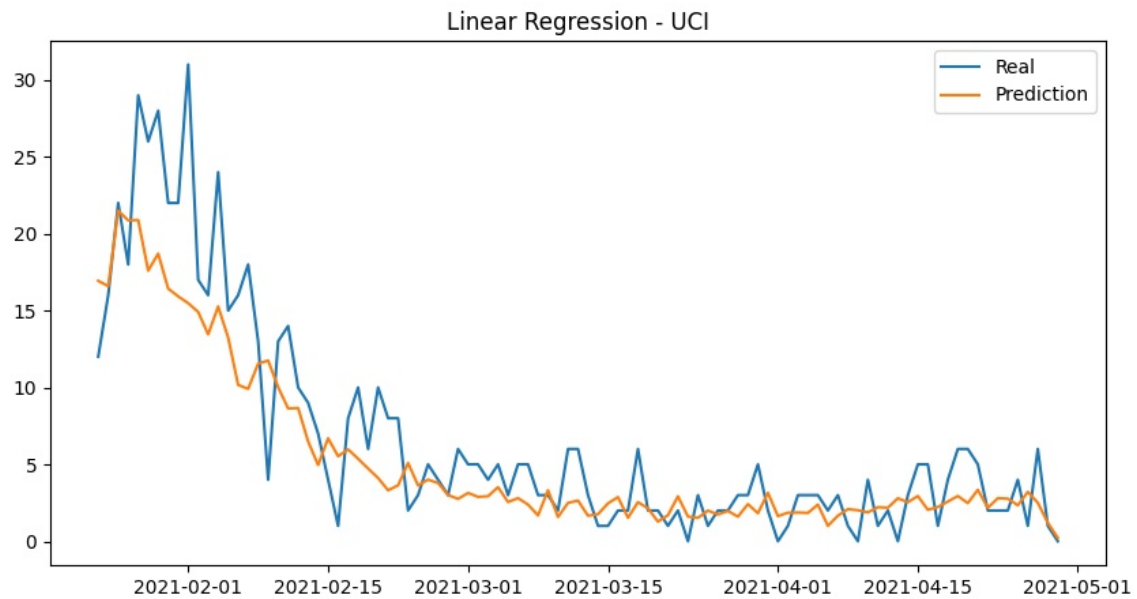
Validation curve Linear Regression (*fit_intercept*, *window_size=4*)



Validation curve Linear Regression (*normalize*, *window size=4*)



Prediction Linear Regression (*UCI*, *window size=4*)



Window size comparison Linear Regression

