IML Term project

Julia Palorinne, Pyry Silomaa, Sara Sippola

03.12.2022

Classifiers

Testing classifiers as they are

Naive Bayes

Accuracy			Perplexity		
train	test	CV	train	test	CV
0.8448276	0.7801724	0.8146552	30.86972	82.21584	152.8545

LDA

Accuracy			Perplexity		
train	test	CV	train	test	CV
0.987069	0.4267241	0.875	1.03937	1082.212	1.683712

QDA

Accuracy			Perplexity		
train	test	CV	train	test	CV
1	0.4396552	0.8448276	1.004092	1.724053e+22	2979.474

Preprocessing data

Correlation

Measurement heights

Correlation between measurements at different heights

Table 1: Eri korkeuksissa tehdyt mittaukset: Mitä muuttujia löytyy millekin korkeudelle?

m.42	m.84	m.168	m.336	m.504	m.672
CO242	NA	CO2168	CO2336	CO2504	NA
H2O42	H2O84	H2O168	H2O336	H2O504	NA
NO42	NO84	NO168	NO336	NO504	NO672
NOx42	NOx84	NOx168	NOx336	NOx504	NOx672
O342	O384	O3168	NA	O3504	O3672
RHIRGA42	RHIRGA84	RHIRGA168	RHIRGA336	RHIRGA504	RHIRGA672
NA	NA	SO2168	NA	NA	NA
T42	T84	T168	NA	T504	T672

Table 2: Correlation (CO2)

	CO2168	CO2336	CO242	CO2504
CO2168	1.0000000	0.9996891	0.9939228	0.9988868
CO2336	0.9996891	1.0000000	0.9928568	0.9996424
CO242	0.9939228	0.9928568	1.0000000	0.9910311
CO2504	0.9988868	0.9996424	0.9910311	1.0000000

Table 3: Correlation (H20)

	H2O168	H2O336	H2O42	H2O504	H2O672	H2O84
H2O168	1.0000000	0.9998966	0.9997062	0.9997158	0.9994631	0.9998894
H2O336	0.9998966	1.0000000	0.9993506	0.9999302	0.9997498	0.9996330
H2O42	0.9997062	0.9993506	1.0000000	0.9990202	0.9986416	0.9999330
H2O504	0.9997158	0.9999302	0.9990202	1.0000000	0.9998589	0.9993631
H2O672	0.9994631	0.9997498	0.9986416	0.9998589	1.0000000	0.9990316
H2O84	0.9998894	0.9996330	0.9999330	0.9993631	0.9990316	1.0000000

Table 4: Correlation (NO)

	NO168	NO336	NO42	NO504	NO672	NO84
NO168	1.0000000	0.9942106	0.9729234	0.9888394	0.9787201	0.9955792
NO336	0.9942106	1.0000000	0.9681760	0.9962470	0.9889000	0.9912906
NO42	0.9729234	0.9681760	1.0000000	0.9648354	0.9580581	0.9766436
NO504	0.9888394	0.9962470	0.9648354	1.0000000	0.9947569	0.9859003
NO672	0.9787201	0.9889000	0.9580581	0.9947569	1.0000000	0.9766069
NO84	0.9955792	0.9912906	0.9766436	0.9859003	0.9766069	1.0000000

Table 5: Correlation (NOx)

	NOx168	NOx336	NOx42	NOx504	NOx672	NOx84
NOx168	1.0000000	0.9988830	0.9966528	0.9966839	0.9946312	0.9996773
NOx336	0.9988830	1.0000000	0.9953287	0.9986298	0.9973042	0.9984147
NOx42	0.9966528	0.9953287	1.0000000	0.9931515	0.9913512	0.9968910
NOx504	0.9966839	0.9986298	0.9931515	1.0000000	0.9987762	0.9961637
NOx672	0.9946312	0.9973042	0.9913512	0.9987762	1.0000000	0.9940603
NOx84	0.9996773	0.9984147	0.9968910	0.9961637	0.9940603	1.0000000

Table 6: Correlation (O3)

	O3168	O342	O3504	O3672	O384
O3168	1.0000000	0.9955292	0.9954429	0.9918052	0.9986470
O342	0.9955292	1.0000000	0.9843062	0.9789921	0.9987926
O3504	0.9954429	0.9843062	1.0000000	0.9989553	0.9904563
O3672	0.9918052	0.9789921	0.9989553	1.0000000	0.9857437
O384	0.9986470	0.9987926	0.9904563	0.9857437	1.0000000

Table 7: Correlation (RHIRGA)

	RHIRGA168	RHIRGA336	RHIRGA42	RHIRGA504	RHIRGA672	RHIRGA84
RHIRGA168	1.0000000	0.9992464	0.9972212	0.9981986	0.9957488	0.9989197
RHIRGA336	0.9992464	1.0000000	0.9950372	0.9994052	0.9973122	0.9970893
RHIRGA42	0.9972212	0.9950372	1.0000000	0.9931951	0.9902415	0.9991006
RHIRGA504	0.9981986	0.9994052	0.9931951	1.0000000	0.9987544	0.9955233
RHIRGA672	0.9957488	0.9973122	0.9902415	0.9987544	1.0000000	0.9927423
RHIRGA84	0.9989197	0.9970893	0.9991006	0.9955233	0.9927423	1.0000000

Table 8: Correlation (T)

	T168	T42	T504	T672	T84
T168	1.0000000	0.9997396	0.9997239	0.9993168	0.9999172
T42	0.9997396	1.0000000	0.9993129	0.9987827	0.9998980
T504	0.9997239	0.9993129	1.0000000	0.9998644	0.9995274
T672	0.9993168	0.9987827	0.9998644	1.0000000	0.9990504
T84	0.9999172	0.9998980	0.9995274	0.9990504	1.0000000