Developing Partial Least Squares Discriminant Analysis (PLS-DA) model in The UnscramblerX



Created by: Edwin Caballero-Agosto

Manager: Samuel Hernandez-Rivera

SOP-01	Edwin Caballero- Agosto	University of Puerto Rico at Mayagüez
Effectivity: Nov/12/2021	Developing Partial Least Squares Discriminant Analysis model in The UnscramblerX	Revised by:
Revised:		Approved by:

DEVELOPING PLS-DA MODEL IN THE UNSCRAMBLERX

This SOP uses the "Tutorial H" data set from The UnscramblerX program PLS-DA tutorial.

Create Excel file (.xlsx).
Write data with samples in rows and variables in columns.
Save spreadsheet file as "PLS-DA Data.xlsx" in desktop by going to File \rightarrow Save As on the Desktop area.
Open The UnscramblerX program.
Rename project with right click \rightarrow Rename to "PLS-DA Development".
Import data with File \rightarrow Import Data \rightarrow Excel, select the created file in the Desktop area, click "Open".
Select "Sheet 1" and press "OK" on the "Excel Preview" window.
Click "OK" on the pop-up window if it appears.
Change the "Class" column from text to category by left click on the "4" below the "Class " text, right click \rightarrow Change Data Type \rightarrow Category
Click "OK" on the "Category Converter" window.
Create sample and variable sets with Edit $ ightarrow$ Define Range
Create "TRAIN" sample set by click on the first row, going to the last row with A/B values, press "Shift" tab, click on the last row, rename set to "TRAIN", and click "Create".
Repeat step 12 with samples that do not have A/B values.
Create "X" column set with column 5 to 14.
Create "Y" column set with column 1.
Confirm ranges by clicking "OK". DO NOT PRESS ENTER, it does not save the ranges.
Develop PCA model with Tasks \rightarrow Analyze \rightarrow Partial Least Squares
Select "PLS-DA Data_Sheet 1" matrix, select "TRAIN" sample set, for the Predictors select "X" variable set, for the Responses select "Y" variable set, write "7" maximum component, and check "Mean center data" box in the "Model Inputs" tab.
Select "All", select "A/(SDev+B)" in the "X Weights" tab.
Select "Setup" button in "Validation" tab.
Select "Full" cross validation method in the "Cross Validation Setup" window, and press "OK".
Select "NIPALS" circle in the "Algorithm" tab and click "Finish".
Copy data for a plot by right clicking on the graph \rightarrow View \rightarrow Numerical.
Save file with File \rightarrow Save As

Appendix Section

Category columns describing data, Response matrix, and Predictor matrix with both TRAIN and TEST sample sets.

	A/B	А	В	Class	Fe	Ti	Ва	Ca	К	Mn	Rb	Sr	Y	Zr
A#1	1	1	0	А	1100	390	55	920	460	45	120	57	58	142
A#2	1	1	0	А	1173	417	54	961	441	47	135	55	60	145
A#3	1	1	0	А	1164	404	56	916	446	42	120	58	45	148
A#4	1	1	0	А	1030	373	59	920	487	38	128	53	58	138
A#5	1	1	0	А	1077	373	55	888	455	38	97	51	54	145
A#6	1	1	0	А	1080	403	53	919	442	41	133	60	45	155
A#7	1	1	0	А	1020	360	59	883	473	43	119	40	50	134
A#8	1	1	0	А	1050	396	56	924	482	48	140	74	71	157
A#9	1	1	0	А	1100	373	53	910	477	51	137	61	58	152
A#10	1	1	0	А	1069	375	51	958	429	42	100	51	47	128
B#1	-1	0	1	В	863	183	8	626	452	34	121	15	58	70
B#2	-1	0	1	В	1108	289	7	783	426	41	109	15	57	67
B#3	-1	0	1	В	1210	276	10	966	430	44	117	20	44	73
B#4	-1	0	1	В	1205	291	10	975	420	43	115	25	58	73
B#5	-1	0	1	В	1100	267	10	910	500	40	145	25	65	95
B#6	-1	0	1	В	1100	280	10	872	515	49	145	38	60	65
B#7	-1	0	1	В	689	114	9	534	404	26	110	25	50	55
B#8	-1	0	1	В	1186	257	10	940	431	40	121	20	53	73
B#9	-1	0	1	В	860	182	7	722	418	33	115	20	55	53
E#1				E	1050	195	46	865	400	36	97	10	50	160
E#2				E	1010	357	65	900	455	42	112	65	55	145
E#3				E	920	320	60	830	440	33	110	43	60	143
E#4				E	1000	194	58	965	460	42	125	30	60	145
F#1				F	920	215	6	650	435	37	122	15	65	75
F#2				F	780	140	12	605	415	35	130	15	50	70
F#3				F	680	105	10	460	400	31	127	15	57	87

Advisor Signature

Co-Advisor Signature