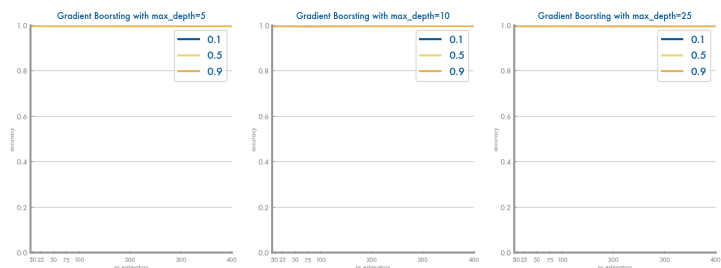


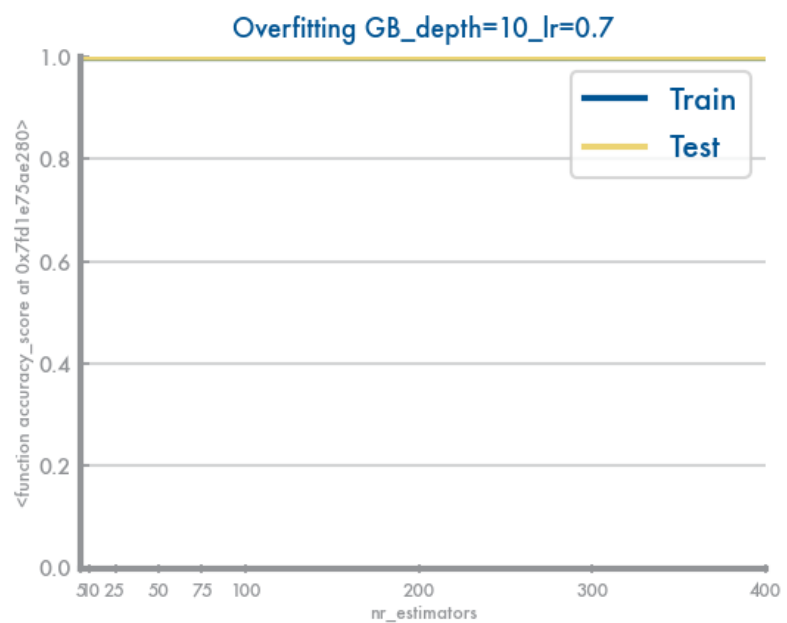
Dataset1:

Gradient Boosting set1:

Max depth
Nr of estimators
Learning rate variation

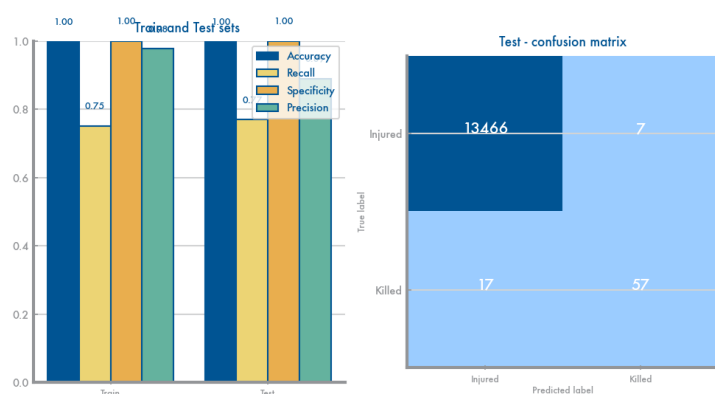


Overfitting



Feature Selection
(imagem do melhor modelo
com feature selection nas
variáveis!)

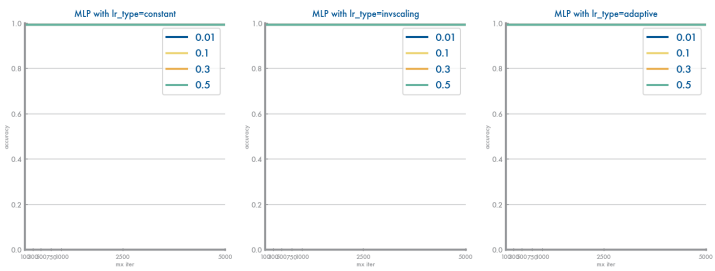
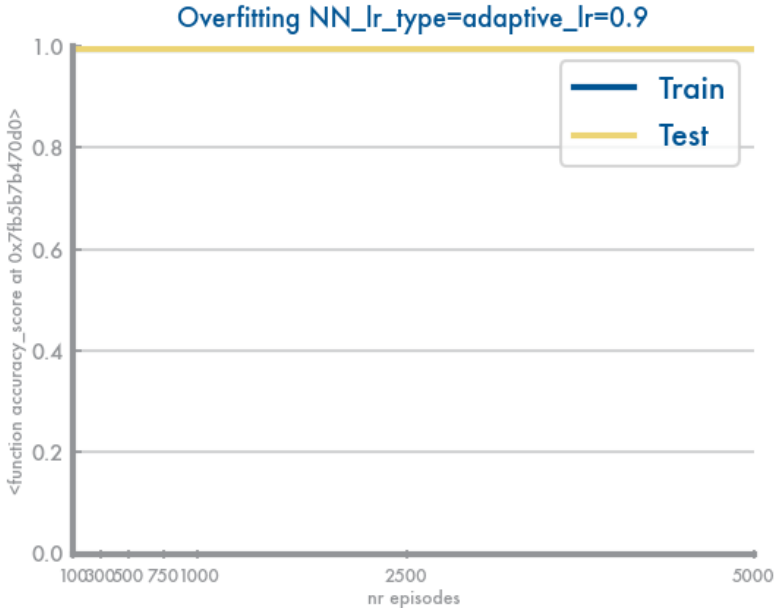
Sem feature selection os
resultados foram iguais.



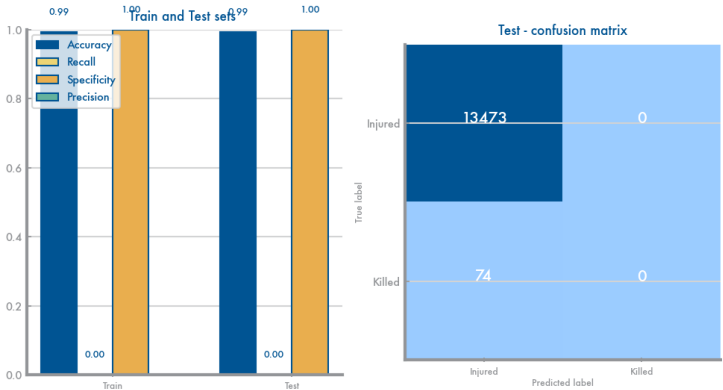
Gradient Boosting best model:

Best model evaluation	Best results with depth=5, learning rate=0.10 and 5 estimators, with accuracy=1.00																																
Confusion Matrix	<div><div><div>Train and Test sets</div><table><tr><th>Metric</th><th>Train</th><th>Test</th></tr><tr><td>Accuracy</td><td>1.00</td><td>1.00</td></tr><tr><td>Recall</td><td>1.00</td><td>1.00</td></tr><tr><td>Specificity</td><td>1.00</td><td>1.00</td></tr><tr><td>Precision</td><td>1.00</td><td>1.00</td></tr></table></div><div><div>Test - confusion matrix</div><table><tr><th></th><th>Injured</th><th>Killed</th></tr><tr><th>Injured</th><td>13466</td><td>7</td></tr><tr><th>Killed</th><td>17</td><td>57</td></tr></table></div></div>	Metric	Train	Test	Accuracy	1.00	1.00	Recall	1.00	1.00	Specificity	1.00	1.00	Precision	1.00	1.00		Injured	Killed	Injured	13466	7	Killed	17	57								
Metric	Train	Test																															
Accuracy	1.00	1.00																															
Recall	1.00	1.00																															
Specificity	1.00	1.00																															
Precision	1.00	1.00																															
	Injured	Killed																															
Injured	13466	7																															
Killed	17	57																															
Features Importance	<div><div>Gradient Boosting Features importance</div><table><tr><th>Feature</th><th>Importance</th></tr><tr><td>IL_STATUS</td><td>0.75</td></tr><tr><td>SON_AGE</td><td>0.10</td></tr><tr><td>NIQUE_ID</td><td>0.05</td></tr><tr><td>OMPLAINT</td><td>0.05</td></tr><tr><td>OCATION</td><td>0.05</td></tr><tr><td>QUIPMENT</td><td>0.05</td></tr><tr><td>.Y_INJURY</td><td>0.05</td></tr><tr><td>L_VEHICLE</td><td>0.05</td></tr><tr><td>ION_TYPE</td><td>0.05</td></tr><tr><td>FACTOR_2</td><td>0.05</td></tr><tr><td>FACTOR_1</td><td>0.05</td></tr><tr><td>Y_ACTION</td><td>0.05</td></tr><tr><td>EJECTION</td><td>0.05</td></tr><tr><td>SON_SEX</td><td>0.05</td></tr><tr><td>PED_ROLE</td><td>0.05</td></tr></table></div>	Feature	Importance	IL_STATUS	0.75	SON_AGE	0.10	NIQUE_ID	0.05	OMPLAINT	0.05	OCATION	0.05	QUIPMENT	0.05	.Y_INJURY	0.05	L_VEHICLE	0.05	ION_TYPE	0.05	FACTOR_2	0.05	FACTOR_1	0.05	Y_ACTION	0.05	EJECTION	0.05	SON_SEX	0.05	PED_ROLE	0.05
Feature	Importance																																
IL_STATUS	0.75																																
SON_AGE	0.10																																
NIQUE_ID	0.05																																
OMPLAINT	0.05																																
OCATION	0.05																																
QUIPMENT	0.05																																
.Y_INJURY	0.05																																
L_VEHICLE	0.05																																
ION_TYPE	0.05																																
FACTOR_2	0.05																																
FACTOR_1	0.05																																
Y_ACTION	0.05																																
EJECTION	0.05																																
SON_SEX	0.05																																
PED_ROLE	0.05																																

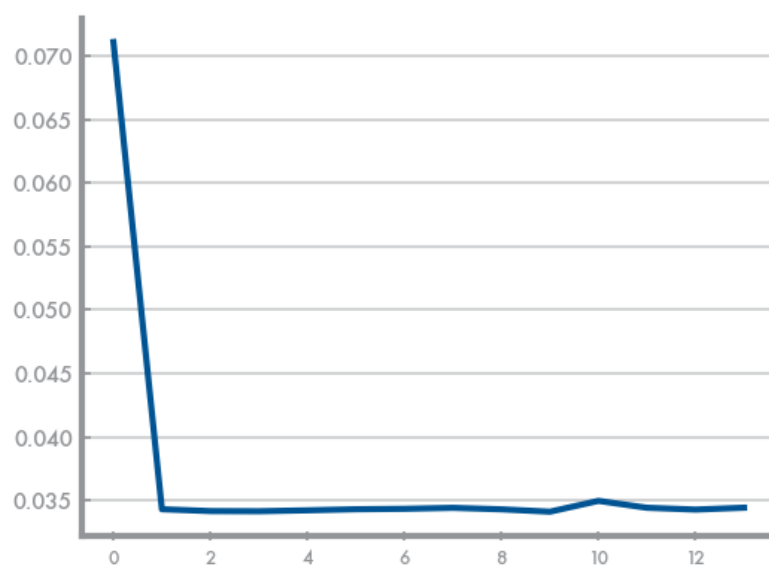
Multi-Layer Perceptrons study:

<p>Nr of iterations</p> <p>Learning rate variation</p> <p>Learning rate values</p>	
<p>Overfitting</p>	
<p>Feature Selection</p>	

Multi-Layer Perceptrons best model:

<p>Best model evaluation</p>	<p>Best results with lr_type=constant, learning rate=0.01 and 100 max iter, with accuracy=0.9945375359858272</p>
<p>Confusion Matrix</p>	

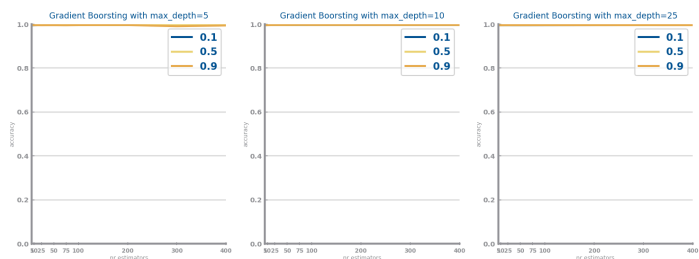
Loss evolution



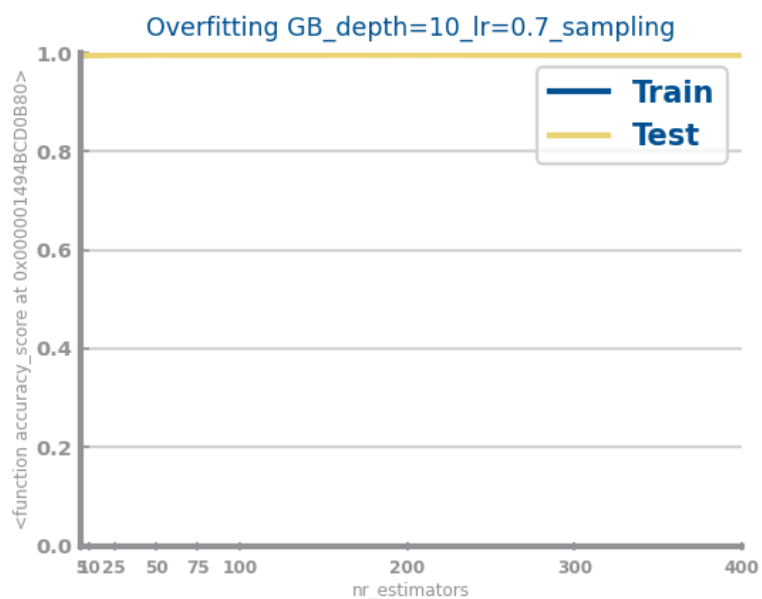
Dataset2:

Gradient Boosting set2:

Max depth
Nr of estimators
Learning rate variation

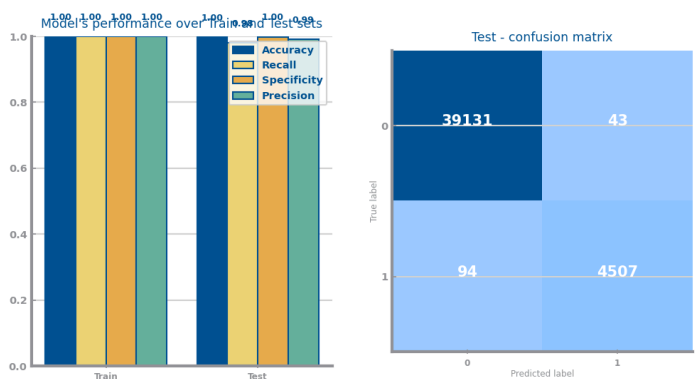


Overfitting



Feature Selection
(imagem do melhor modelo com
feature selection nas variáveis!)

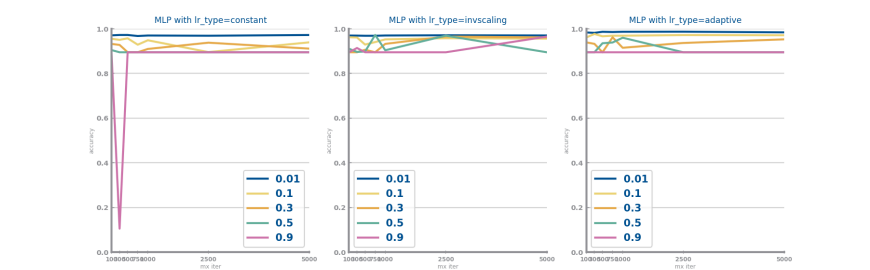
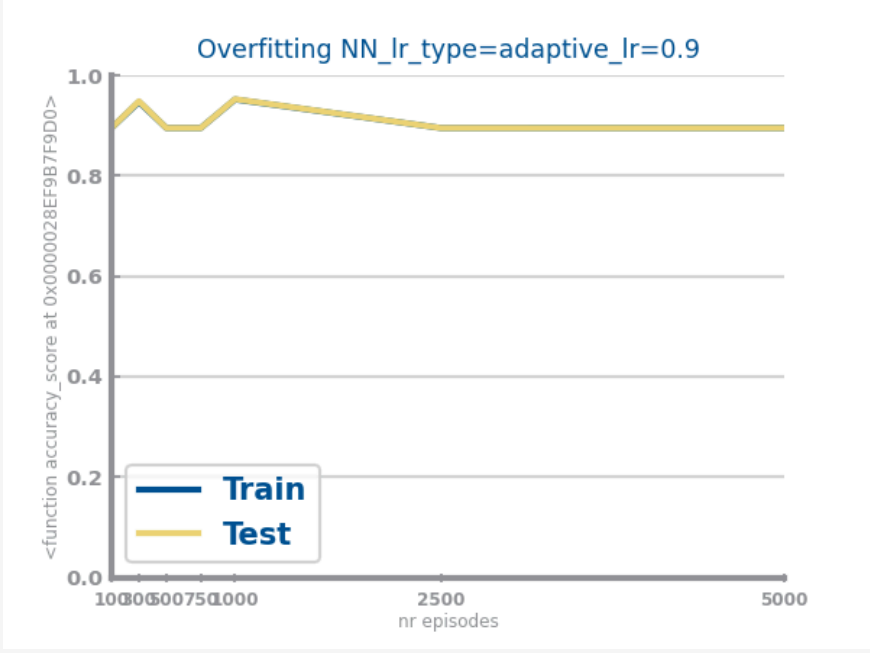
Sem feature selection os
resultados foram iguais.



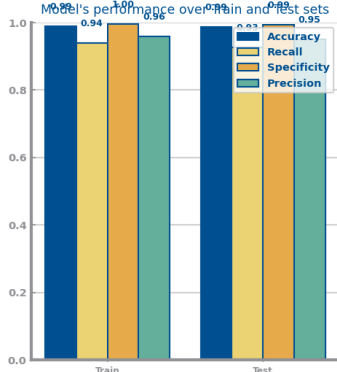
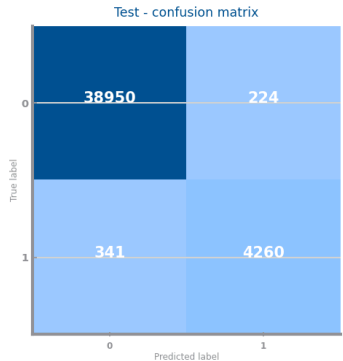
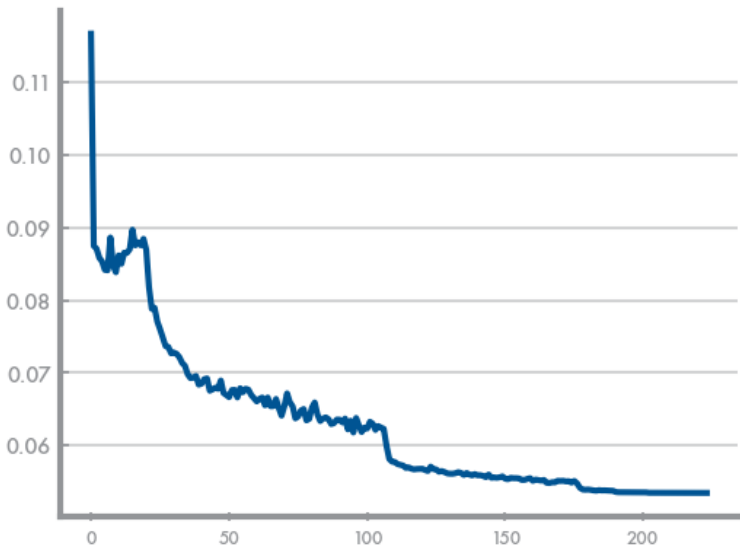
Gradient Boosting best model:

Best model evaluation	Best results with depth=5, learning rate=0.10 and 300 estimators, with accuracy=1.00																																																																
Confusion Matrix	<div><div>Model's performance over Train and Test sets</div><table><tr><th>Metric</th><th>Train</th><th>Test</th></tr><tr><td>Accuracy</td><td>1.00</td><td>1.00</td></tr><tr><td>Recall</td><td>1.00</td><td>1.00</td></tr><tr><td>Specificity</td><td>1.00</td><td>1.00</td></tr><tr><td>Precision</td><td>1.00</td><td>1.00</td></tr></table></div> <div><div>Test - confusion matrix</div><table><tr><th></th><th>Predicted label 0</th><th>Predicted label 1</th></tr><tr><th>True label 0</th><td>39131</td><td>43</td></tr><tr><th>True label 1</th><td>94</td><td>4507</td></tr></table></div>	Metric	Train	Test	Accuracy	1.00	1.00	Recall	1.00	1.00	Specificity	1.00	1.00	Precision	1.00	1.00		Predicted label 0	Predicted label 1	True label 0	39131	43	True label 1	94	4507																																								
Metric	Train	Test																																																															
Accuracy	1.00	1.00																																																															
Recall	1.00	1.00																																																															
Specificity	1.00	1.00																																																															
Precision	1.00	1.00																																																															
	Predicted label 0	Predicted label 1																																																															
True label 0	39131	43																																																															
True label 1	94	4507																																																															
Features Importance	<div><div>Gradient Boosting Features importance</div><table><tr><th>Feature</th><th>Importance</th></tr><tr><td>Mean_5</td><td>0.80</td></tr><tr><td>Mean_10</td><td>0.18</td></tr><tr><td>Max_10</td><td>0.15</td></tr><tr><td>Std_12.5</td><td>0.12</td></tr><tr><td>Max_12.5</td><td>0.10</td></tr><tr><td>Max_03</td><td>0.08</td></tr><tr><td>Std_M10</td><td>0.07</td></tr><tr><td>Mean_03</td><td>0.06</td></tr><tr><td>Std_03</td><td>0.05</td></tr><tr><td>Min_12.5</td><td>0.04</td></tr><tr><td>Min_M10</td><td>0.03</td></tr><tr><td>date</td><td>0.02</td></tr><tr><td>Std_NO2</td><td>0.01</td></tr><tr><td>Min_03</td><td>0.01</td></tr><tr><td>Std_SO2</td><td>0.01</td></tr><tr><td>City_Gb</td><td>0.01</td></tr><tr><td>Min_NO2</td><td>0.01</td></tr><tr><td>Max_NO2</td><td>0.01</td></tr><tr><td>Mean_02</td><td>0.01</td></tr><tr><td>Max_02</td><td>0.01</td></tr><tr><td>Field_1</td><td>0.01</td></tr><tr><td>Mean_0</td><td>0.01</td></tr><tr><td>Mean_02</td><td>0.01</td></tr><tr><td>Min_SO2</td><td>0.01</td></tr><tr><td>City_EN</td><td>0.01</td></tr><tr><td>Min_CO</td><td>0.01</td></tr><tr><td>Std_CO</td><td>0.01</td></tr><tr><td>Max_CO</td><td>0.01</td></tr><tr><td>FPD</td><td>0.01</td></tr><tr><td>Prov_EN</td><td>0.01</td></tr><tr><td>GbProv</td><td>0.01</td></tr></table></div>	Feature	Importance	Mean_5	0.80	Mean_10	0.18	Max_10	0.15	Std_12.5	0.12	Max_12.5	0.10	Max_03	0.08	Std_M10	0.07	Mean_03	0.06	Std_03	0.05	Min_12.5	0.04	Min_M10	0.03	date	0.02	Std_NO2	0.01	Min_03	0.01	Std_SO2	0.01	City_Gb	0.01	Min_NO2	0.01	Max_NO2	0.01	Mean_02	0.01	Max_02	0.01	Field_1	0.01	Mean_0	0.01	Mean_02	0.01	Min_SO2	0.01	City_EN	0.01	Min_CO	0.01	Std_CO	0.01	Max_CO	0.01	FPD	0.01	Prov_EN	0.01	GbProv	0.01
Feature	Importance																																																																
Mean_5	0.80																																																																
Mean_10	0.18																																																																
Max_10	0.15																																																																
Std_12.5	0.12																																																																
Max_12.5	0.10																																																																
Max_03	0.08																																																																
Std_M10	0.07																																																																
Mean_03	0.06																																																																
Std_03	0.05																																																																
Min_12.5	0.04																																																																
Min_M10	0.03																																																																
date	0.02																																																																
Std_NO2	0.01																																																																
Min_03	0.01																																																																
Std_SO2	0.01																																																																
City_Gb	0.01																																																																
Min_NO2	0.01																																																																
Max_NO2	0.01																																																																
Mean_02	0.01																																																																
Max_02	0.01																																																																
Field_1	0.01																																																																
Mean_0	0.01																																																																
Mean_02	0.01																																																																
Min_SO2	0.01																																																																
City_EN	0.01																																																																
Min_CO	0.01																																																																
Std_CO	0.01																																																																
Max_CO	0.01																																																																
FPD	0.01																																																																
Prov_EN	0.01																																																																
GbProv	0.01																																																																

Multi-Layer Perceptrons study:

<p>Nr of iterations</p> <p>Learning rate variation</p> <p>Learning rate values</p>	
<p>Overfitting</p>	
<p>Feature Selection</p>	

Multi-Layer Perceptrons best model:

Best model evaluation	Best results with lr_type=adaptive, learning rate=0.01 and 300 max iter, with accuracy=0.9729333333333333																								
Confusion Matrix	<div><div><p>Model's performance over train and test sets</p><table><thead><tr><th>Metric</th><th>Train</th><th>Test</th></tr></thead><tbody><tr><td>Accuracy</td><td>0.98</td><td>0.95</td></tr><tr><td>Recall</td><td>0.94</td><td>0.93</td></tr><tr><td>Specificity</td><td>0.98</td><td>0.95</td></tr><tr><td>Precision</td><td>0.96</td><td>0.95</td></tr></tbody></table></div><div><p>Test - confusion matrix</p><table><thead><tr><th>True label \ Predicted label</th><th>0</th><th>1</th></tr></thead><tbody><tr><th>0</th><td>389</td><td>50</td></tr><tr><th>1</th><td>341</td><td>4260</td></tr></tbody></table></div></div>	Metric	Train	Test	Accuracy	0.98	0.95	Recall	0.94	0.93	Specificity	0.98	0.95	Precision	0.96	0.95	True label \ Predicted label	0	1	0	389	50	1	341	4260
Metric	Train	Test																							
Accuracy	0.98	0.95																							
Recall	0.94	0.93																							
Specificity	0.98	0.95																							
Precision	0.96	0.95																							
True label \ Predicted label	0	1																							
0	389	50																							
1	341	4260																							
Loss evolution	 <table><thead><tr><th>Iteration</th><th>Loss</th></tr></thead><tbody><tr><td>0</td><td>0.115</td></tr><tr><td>10</td><td>0.085</td></tr><tr><td>50</td><td>0.070</td></tr><tr><td>100</td><td>0.062</td></tr><tr><td>150</td><td>0.057</td></tr><tr><td>200</td><td>0.055</td></tr></tbody></table>	Iteration	Loss	0	0.115	10	0.085	50	0.070	100	0.062	150	0.057	200	0.055										
Iteration	Loss																								
0	0.115																								
10	0.085																								
50	0.070																								
100	0.062																								
150	0.057																								
200	0.055																								