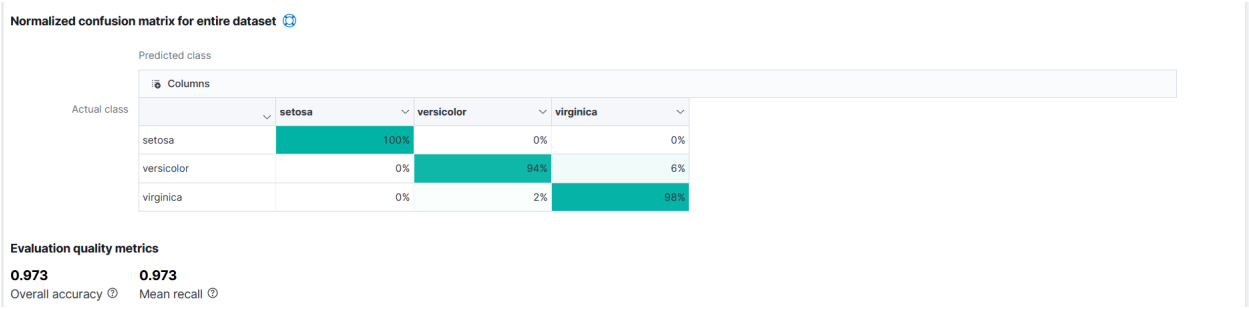
Because we were unable to use the project VM this week, we have developed the following plan to progress the project.

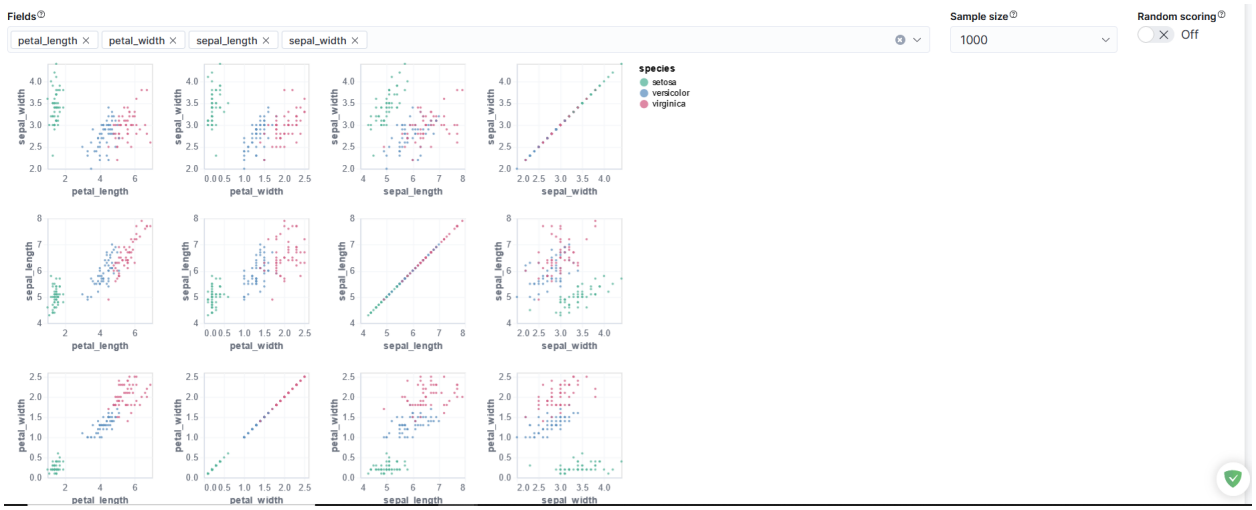
What we would have done:

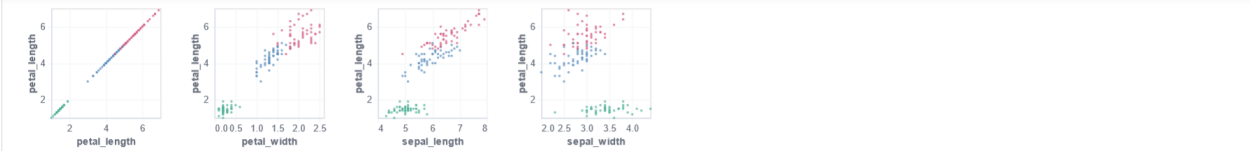
1. Use Elasticsearch to further develop our machine learning algorithms (specifically decision tree) with data sets to analyze points within the VM.
    a. Gain further in depth knowledge of the applications and properties of Elasticsearch.
2. Use Kibana to visualize the results.

What we have done:

1. Used Elasticsearch on free trial of elastic to analyze Iris data set







Results

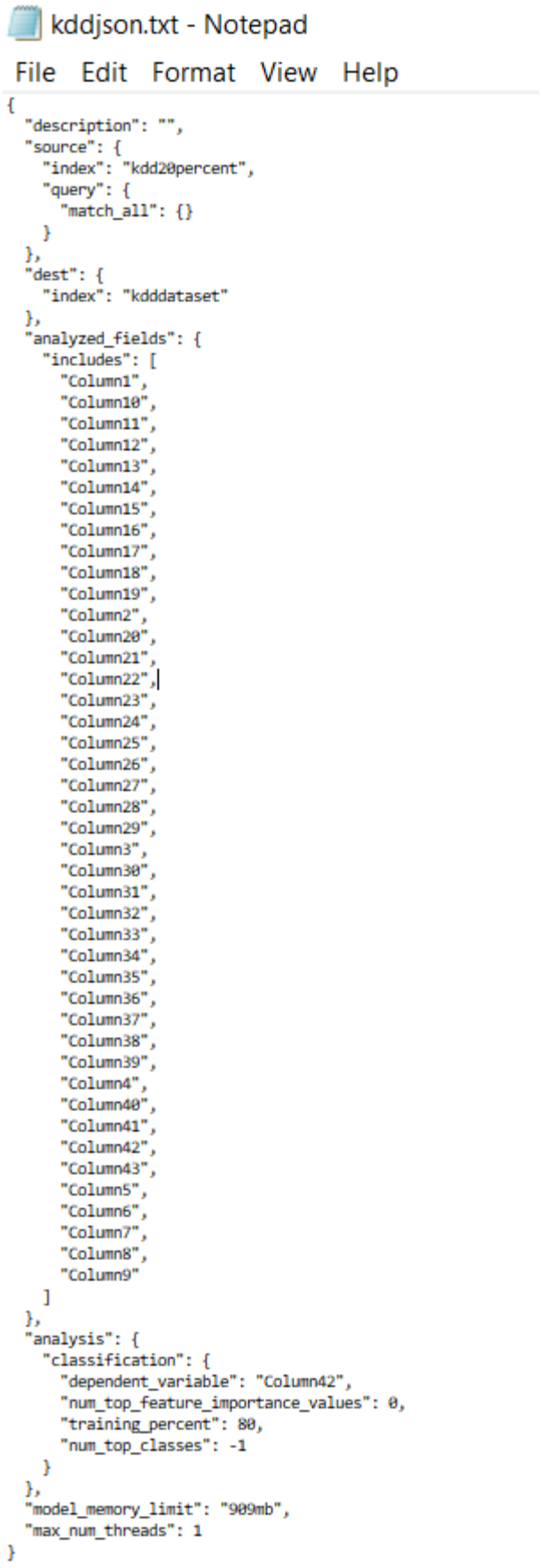
Total docs

150

Showing documents for which predictions exist

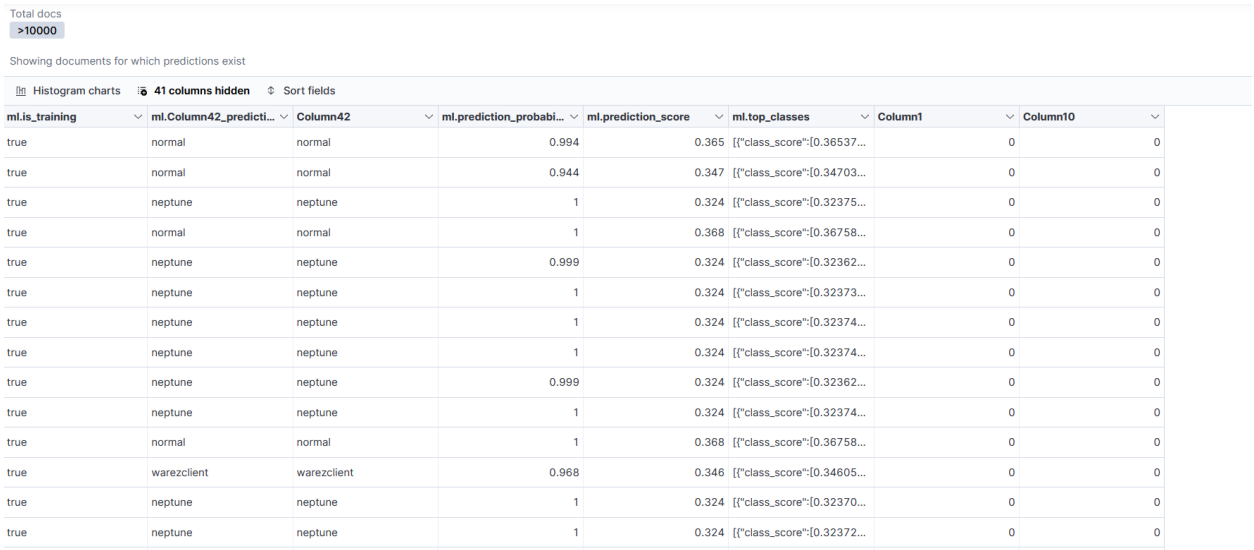
Histogram charts | 3 columns hidden | Sort fields

ml.is_training *true false*	ml.species_prediction *3 categories*	species *3 categories*	ml.prediction_probability *0.51 - 0.96*	ml.prediction_score *0.3 - 0.93*	ml.top_classes *Chart not supported.*	ml__incremental_id *0 - 149*	petal_length *1 - 6.9*
true	setosa	setosa	0.956	0.314	[{"class_score":[0.31427...	0	1.4
true	setosa	setosa	0.956	0.314	[{"class_score":[0.31427...	1	1.4
true	setosa	setosa	0.956	0.314	[{"class_score":[0.31427...	3	1.5
true	setosa	setosa	0.956	0.314	[{"class_score":[0.31427...	4	1.4
true	setosa	setosa	0.956	0.314	[{"class_score":[0.31427...	7	1.5



JSON code above:

Results



Total docs
>10000

Showing documents for which predictions exist

Histogram charts | 41 columns hidden | Sort fields

ml.is_training	ml.Column42_predicti...	Column42	ml.prediction_probabi...	ml.prediction_score	ml.top_classes	Column1	Column10
true	normal	normal	0.994	0.365	[{"class_score":[0.36537...	0	0
true	normal	normal	0.944	0.347	[{"class_score":[0.34703...	0	0
true	neptune	neptune	1	0.324	[{"class_score":[0.32375...	0	0
true	normal	normal	1	0.368	[{"class_score":[0.36758...	0	0
true	neptune	neptune	0.999	0.324	[{"class_score":[0.32362...	0	0
true	neptune	neptune	1	0.324	[{"class_score":[0.32373...	0	0
true	neptune	neptune	1	0.324	[{"class_score":[0.32374...	0	0
true	neptune	neptune	1	0.324	[{"class_score":[0.32374...	0	0
true	neptune	neptune	0.999	0.324	[{"class_score":[0.32362...	0	0
true	neptune	neptune	1	0.324	[{"class_score":[0.32374...	0	0
true	normal	normal	1	0.368	[{"class_score":[0.36758...	0	0
true	warezclient	warezclient	0.968	0.346	[{"class_score":[0.34605...	0	0
true	neptune	neptune	1	0.324	[{"class_score":[0.32370...	0	0
true	neptune	neptune	1	0.324	[{"class_score":[0.32372...	0	0

Model evaluation

Classification evaluation docs

Job status: stopped
docs evaluated: 25192

Normalized confusion matrix for entire dataset

Predicted class

Columns

Actual class

	back	ipsweep	neptune	nmap	normal	portsweep	4 more
back	100%	0%	0%	0%	0%	0%	
ipsweep	0%	100%	0%	0%	0%	0%	
neptune	0%	0%	100%	0%	0%	0%	
nmap	0%	0%	0%	100%	0%	0%	
normal	0%	0%	0%	0%	100%	0%	
portsweep	0%	0%	0%	0%	0%	100%	

Evaluation quality metrics

0.999
Overall accuracy

0.948
Mean recall

Next Step: Using code in elastic to analyze IRIS data set using python