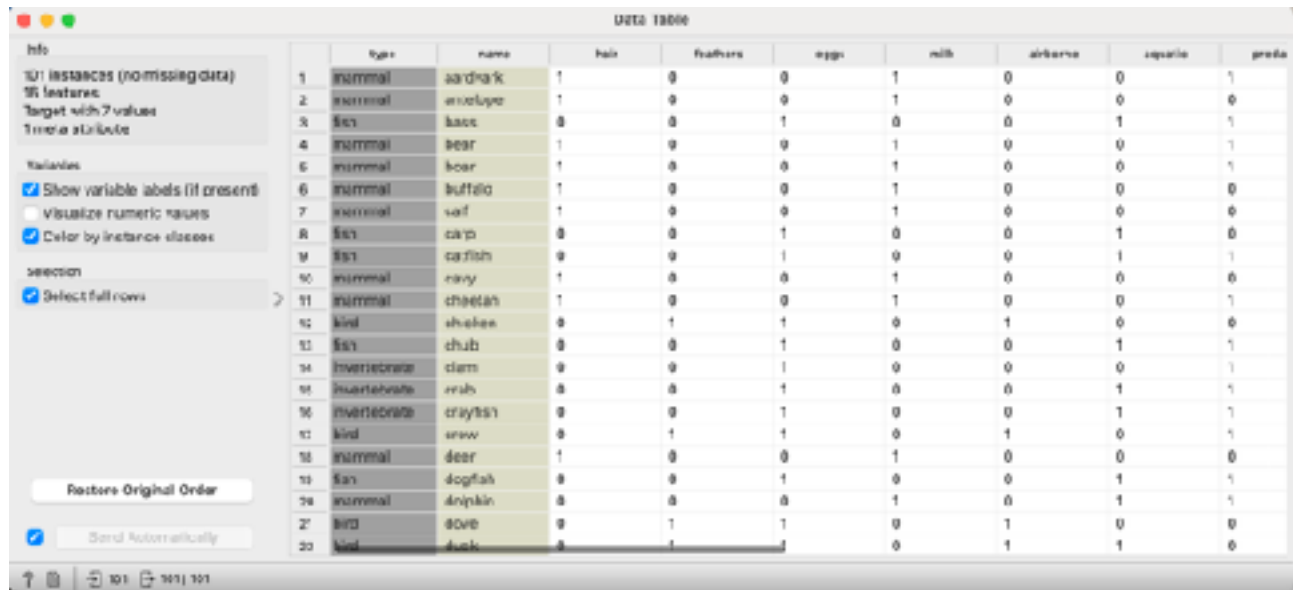


Dataset

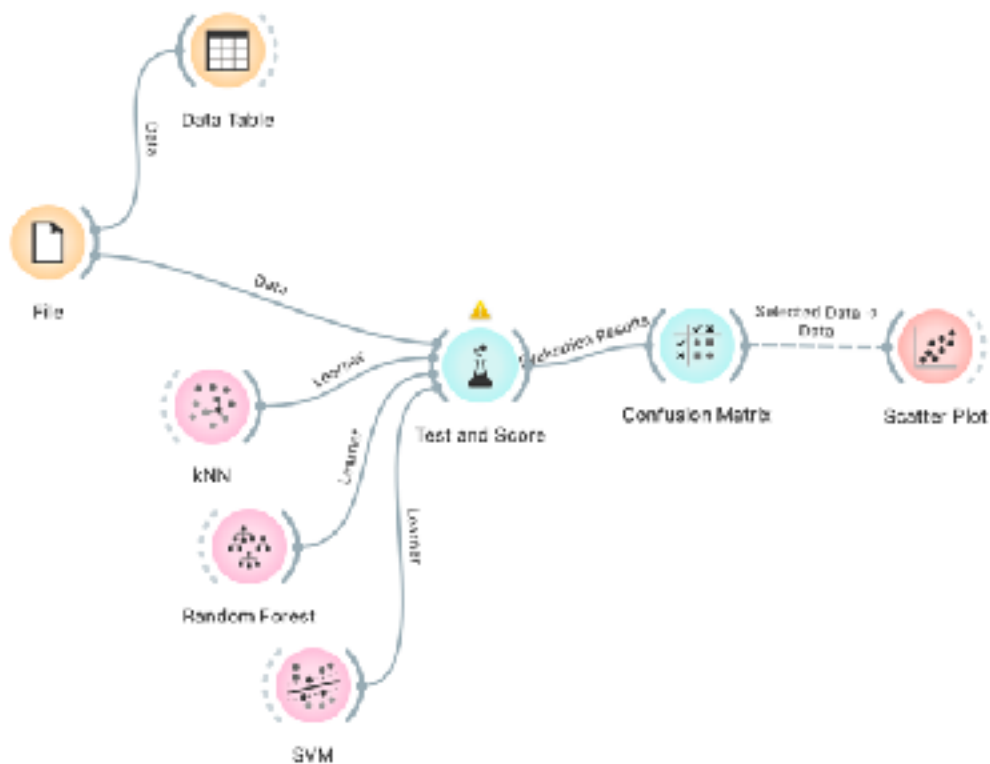
使用動物園的資料集，裡面包含不同動物的各種特徵ex:腳的數量、有無尾巴等等



	type	name	hair	feathers	eggs	milk	airborne	aquatic	predator
1	mammal	sardak	1	0	0	1	0	0	1
2	mammal	weasel	1	0	0	1	0	0	0
3	fish	hake	0	0	1	0	0	1	1
4	mammal	beaver	1	0	0	1	0	0	1
5	mammal	boar	1	0	0	1	0	0	1
6	mammal	buffalo	1	0	0	1	0	0	0
7	mammal	valet	1	0	0	1	0	0	0
8	fish	carp	0	0	1	0	0	1	0
9	fish	catfish	0	0	1	0	0	1	1
10	mammal	navy	1	0	0	1	0	0	0
11	mammal	cheetah	1	0	0	1	0	0	1
12	bird	shrike	0	1	1	0	1	0	0
13	fish	chub	0	0	1	0	0	1	1
14	invertebrate	clam	0	0	1	0	0	0	1
15	invertebrate	crab	0	0	1	0	0	1	1
16	invertebrate	crayfish	0	0	1	0	0	1	1
17	bird	crow	0	1	1	0	1	0	1
18	mammal	deer	1	0	0	1	0	0	0
19	fish	dogfish	0	0	1	0	0	1	1
20	mammal	dingo	0	0	0	1	0	1	1
21	bird	goose	0	1	1	0	1	0	0
22	bird	duck	0	1	1	0	1	1	0

Model

目標是分類，所以使用KNN、SVM、Random Forest當模型



Test and Score

三種模型表現都很好，精確率、召回率都高於90%

Model	AUC	CA	F1	Prec	Recall	MCC
kNN	0.997	0.931	0.920	0.925	0.931	0.910
Random Forest	0.990	0.931	0.928	0.936	0.931	0.910
SVM	0.991	0.941	0.937	0.945	0.941	0.922

Confusion Matrix

透過混淆矩陣可以看出只有KNN能完美分類amphibian，只有SVM能完美分類invertebrate，但KNN誤認成reptile的數量比其他兩個模型多，SVM對insect的分類能力沒有其他兩個模型好，每個模型各有優勢。

		Predicted							Σ
		amphibian	bird	fish	insect	invertebrate	mammal	reptile	
Actual	amphibian	4	0	0	0	0	0	0	4
	bird	0	20	0	0	0	0	0	20
	fish	0	0	13	0	0	0	0	13
	insect	0	0	0	8	0	0	0	8
	invertebrate	0	0	0	2	7	0	1	10
	mammal	0	0	0	0	0	41	0	41
	reptile	1	1	2	0	0	0	1	5
Σ		5	21	15	10	7	41	2	101

KNN

		Predicted							Σ
		amphibian	bird	fish	insect	invertebrate	mammal	reptile	
Actual	amphibian	2	0	0	0	0	0	2	4
	bird	0	20	0	0	0	0	0	20
	fish	0	0	13	0	0	0	0	13
	insect	0	0	0	8	0	0	0	8
	invertebrate	0	0	0	3	7	0	0	10
	mammal	0	0	0	0	0	41	0	41
	reptile	1	0	1	0	0	0	3	5
Σ		3	20	14	11	7	41	5	101

Random Forest

		Predicted							Σ
		amphibian	bird	fish	insect	invertebrate	mammal	reptile	
Actual	amphibian	2	0	0	0	0	0	2	4
	bird	0	20	0	0	0	0	0	20
	fish	0	0	13	0	0	0	0	13
	insect	0	0	0	6	2	0	0	8
	invertebrate	0	0	0	0	10	0	0	10
	mammal	0	0	0	0	0	41	0	41
	reptile	0	0	1	0	0	1	3	5
Σ		2	20	14	6	12	42	5	101

SVM