

ΑΝΑΓΝΩΡΙΣΗ ΠΡΟΤΥΠΩΝ ΚΑΙ ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ

Εργασία 2

ΠΑΝΤΕΛΕΗΜΩΝ ΠΡΩΙΟΣ

ice18390023

9ο Εξάμηνο

ice18390023@uniwa.gr

Τμήμα Α 15:00-17:00



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ
UNIVERSITY OF WEST ATTICA

Υπεύθυνοι καθηγητές

Νικόλαος Βασιλάς

Σωτήριος Δαλιάνης

Τμήμα Μηχανικών και Πληροφορικής Υπολογιστών
10 Ιανουαρίου 2023

Περιεχόμενα

1 Μοντέλα Μηχανικής Μάθησης	1
1.1 Classification Learner	1
1.2 Regression Learner	9

Κώδικες

Κατάλογος σχημάτων

1.1 Iris Fine KNN Matrix	1
1.2 Iris Fine KNN Scatter	2
1.3 Iris Fine Tree Matrix	2
1.4 Iris Fine Tree Scatter	3
1.5 Iris Gaussian Naive Bayes Matrix	3
1.6 Iris Gaussian Naive Bayes Scatter	4
1.7 Iris Linear Discrimination Matrix	4
1.8 Iris Linear Discrimination Scatter	5
1.9 Cancer Gaussian Naive Bayes Matrix	6
1.10 Cancer Gaussian Naive Bayes Scatter	6
1.11 Cancer Linear Discriminant Matrix	7
1.12 Cancer Linear Discriminant Scatter	7
1.13 Cancer Medium KNN Matrix	8
1.14 Cancer Quadratic SVM Matrix	8
1.15 Cancer Quadratic SVM Scatter	9
1.16 Iris Fine Gaussian SVM Predictions	10
1.17 Iris Fine Gaussian SVM Response	10
1.18 Iris Fine Tree Predictions	11
1.19 Iris Fine Tree Response	11
1.20 Iris Linear Predictions	12
1.21 Iris Linear Response	12
1.22 Iris Linear SVM Predictions	13
1.23 Iris Linear SVM Response	13
1.24 Cancer Cubic SVM Predictions	14
1.25 Cancer Cubic SVM Response	14
1.26 Cancer Fine Tree Predictions	15
1.27 Cancer Fine Tree Response	15
1.28 Cancer Linear SVM Predictions	16
1.29 Cancer Linear SVM Response	16
1.30 Cancer Rational Quadratic GPR Predictions	17
1.31 Cancer Rational Quadratic GPR Response	17

1 Μοντέλα Μηχανικής Μάθησης

Τα dataset τα οποία επιλέχθηκαν για τα μοντέλα μηχανικής μάθησης είναι:

- Fisher iris dataset
- ovariancancer dataset

1.1 Classification Learner

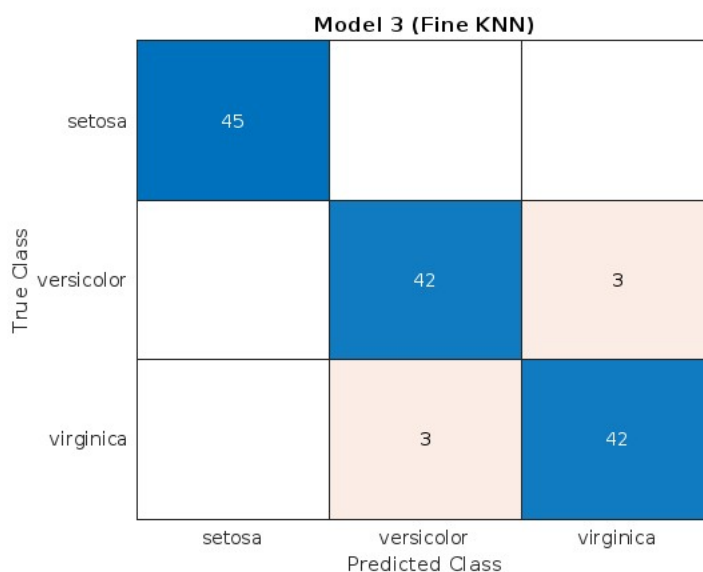
Τα μοντέλα που χρησιμοποιήθηκαν για το iris dataset είναι:

1. Fine KNN
2. Fine Tree
3. Gaussian Naive Bayes
4. Linear Discrimination

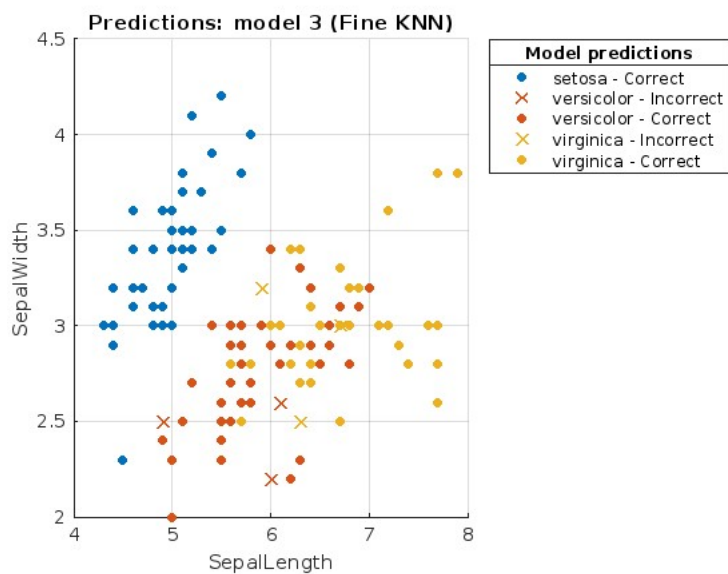
Από τον πίνακα 1.1 μπορούμε να διακρίνουμε πως ενώ το καλύτερο μοντέλο είναι το Linear Discrimination στο validation, με 99.26%, στο test έχει την χειρότερη απόδοση με 86.66%. Ενώ τα καλύτερα μοντέλα στο test είναι αυτά του Fine Tree και του Gaussian Naive Bayes με 93.33%, θα επιλέξουμε του Gaussian Naive Bayes ως το καλύτερο, διότι είχε καλύτερα αποτελέσματα στο validation.

Πίνακας 1.1: Iris classification learner

Model Type	Accuracy % (Validation)	Accuracy % (Test)
Tree	95.55555556	93.33333333
Discriminant	99.25925926	86.66666667
KNN	95.55555556	86.66666667
Naive Bayes	96.2962963	93.33333333



Σχήμα 1.1: Iris Fine KNN Matrix

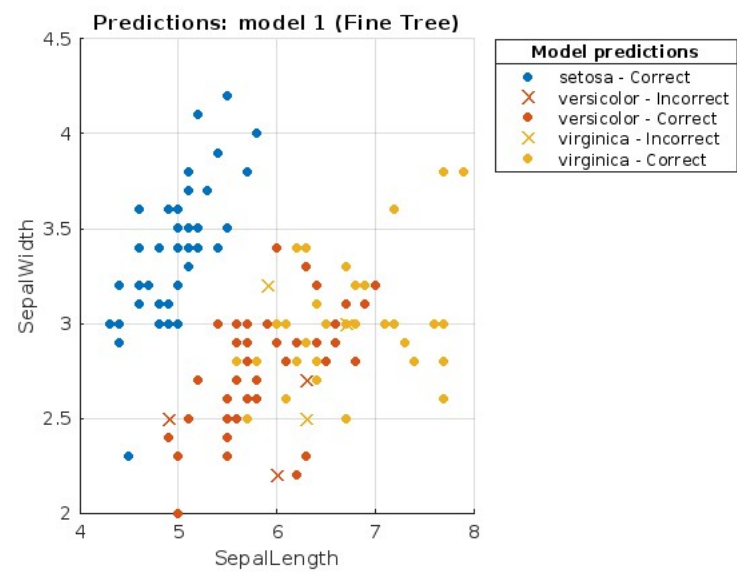


Σχήμα 1.2: Iris Fine KNN Scatter

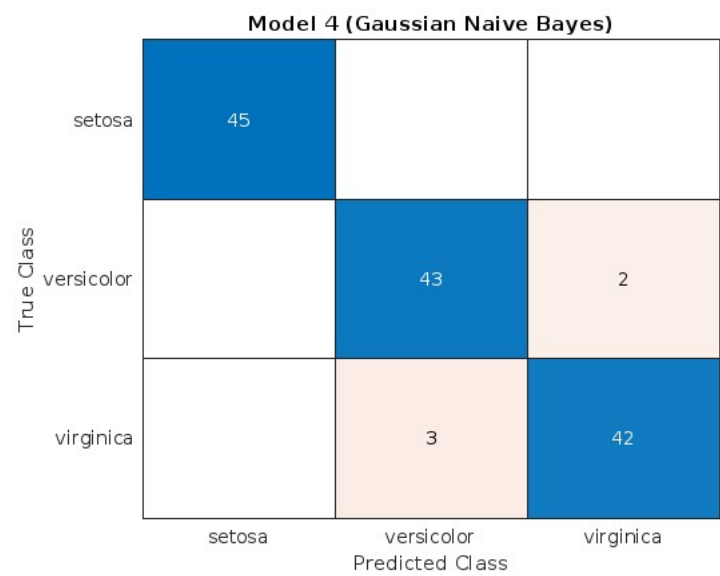
Model 1 (Fine Tree)

True Class	setosa	versicolor	virginica
setosa	45		
versicolor		42	3
virginica		3	42
	setosa	versicolor	virginica
	Predicted Class		

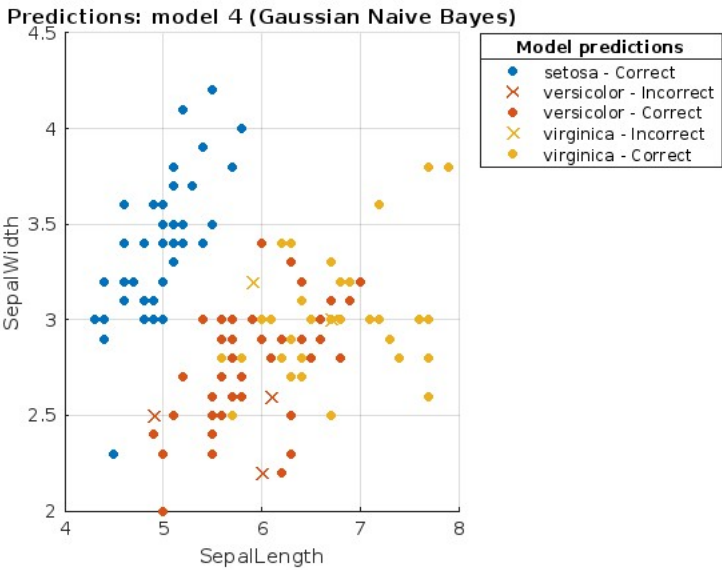
Σχήμα 1.3: Iris Fine Tree Matrix



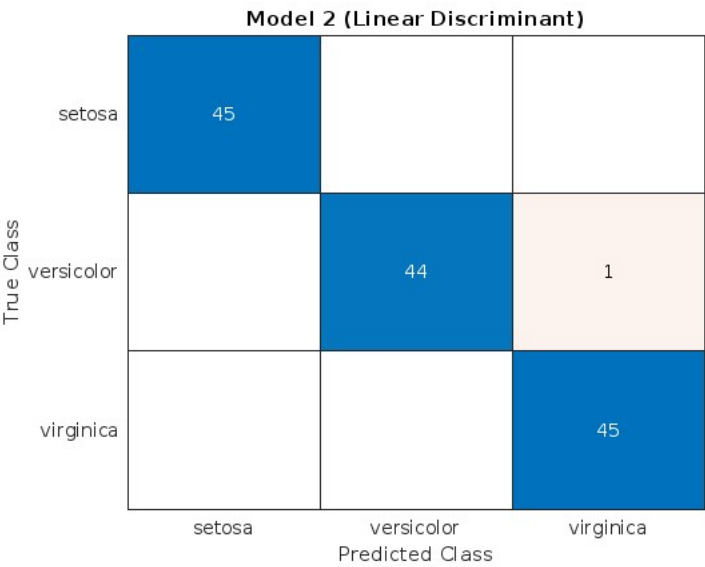
Σχήμα 1.4: Iris Fine Tree Scatter



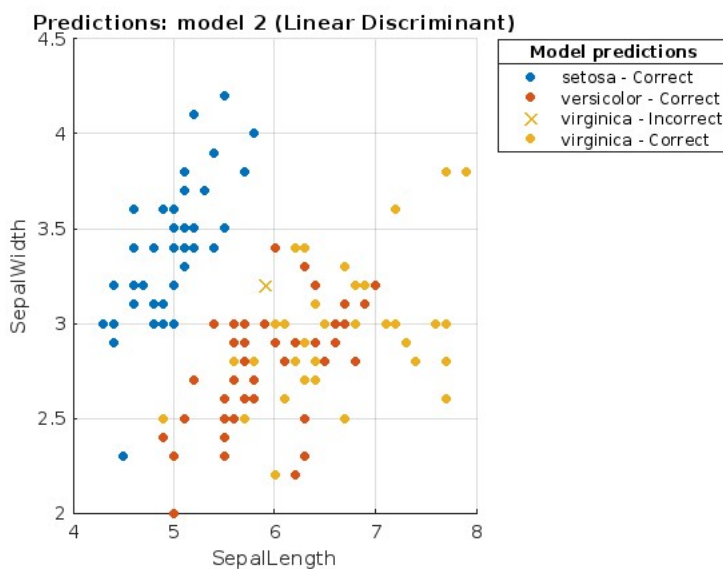
Σχήμα 1.5: Iris Gaussian Naive Bayes Matrix



Σχήμα 1.6: Iris Gaussian Naive Bayes Scatter



Σχήμα 1.7: Iris Linear Discrimination Matrix



Σχήμα 1.8: Iris Linear Discrimination Scatter

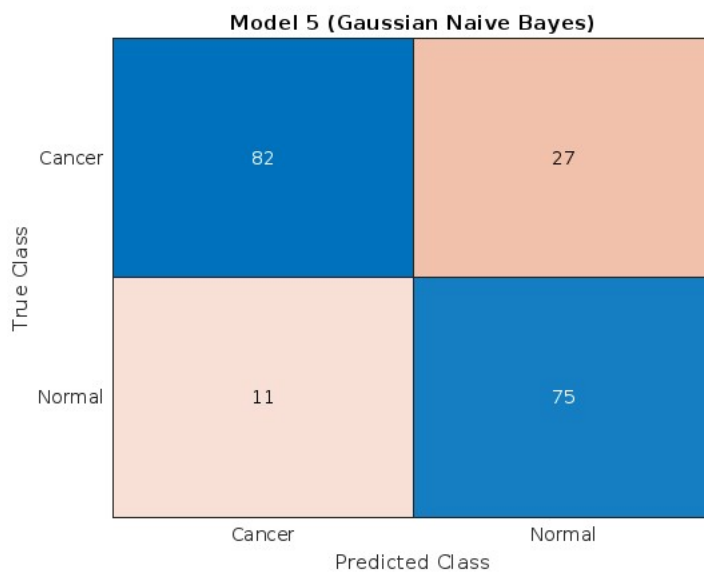
Τα μοντέλα που χρησιμοποιήθηκαν για το cancer dataset είναι:

1. Gaussian Naive Bayes
2. Medium KNN
3. Quadratic SVM
4. Linear Discrimination

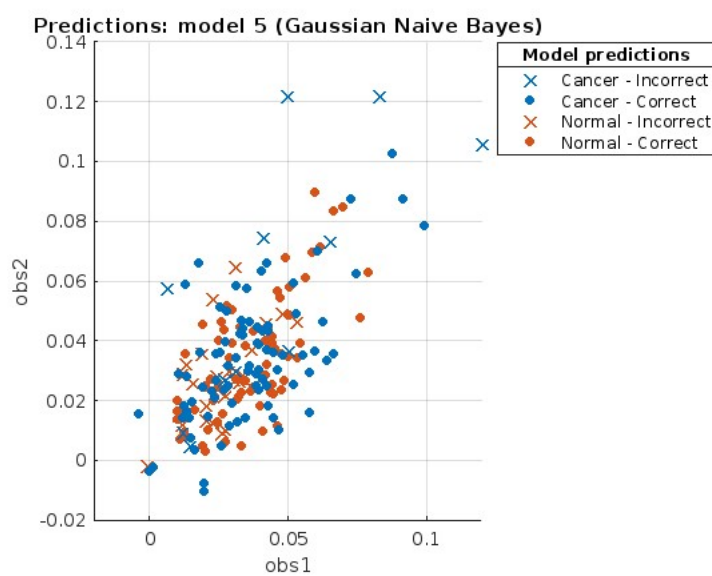
Πίνακας 1.2: Cancer classification learner

Model Type	Accuracy % (Validation)	Accuracy % (Test)
Discriminant	98.46153846	100
SVM	98.46153846	100
Naive Bayes	80.51282051	95.23809524
KNN	94.35897436	100

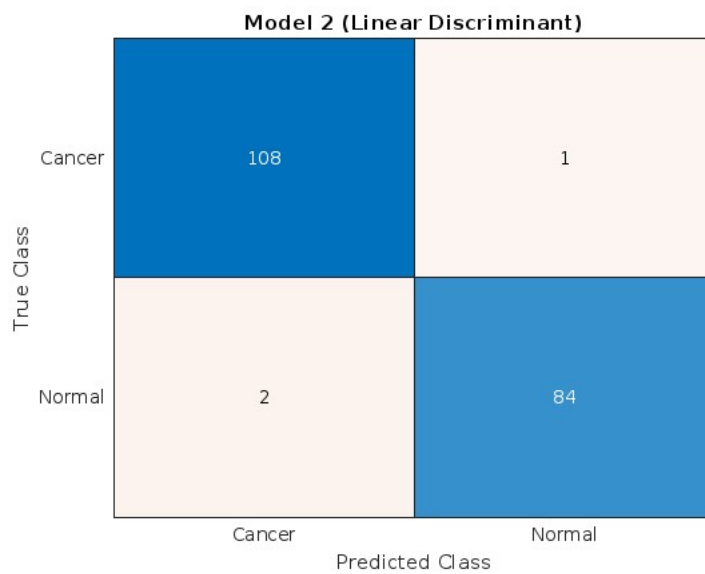
Στον πίνακα 1.2 διακρίνουμε πως το μόνο μοντέλο που δεν είχε 100% accuracy στο test, είναι αυτό του Gaussian Naive Bayes. Από τα τρία μοντέλα που τα είχαν 100% επιτυχία στο test, αυτό που είχε την χαμηλότερη επιτυχία στο validation, είναι αυτό του medium KNN, ενώ τα δύο άλλα είχαν τα ίδια ακριβώς αποτελέσματα.



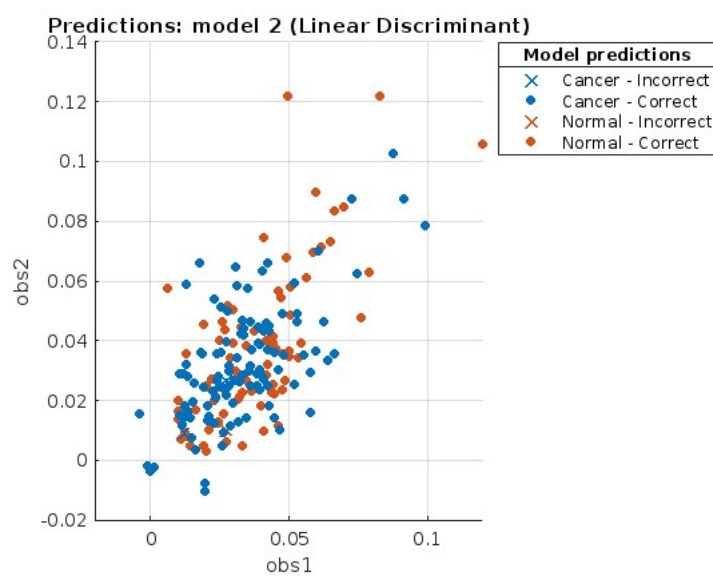
Σχήμα 1.9: Cancer Gaussian Naive Bayes Matrix



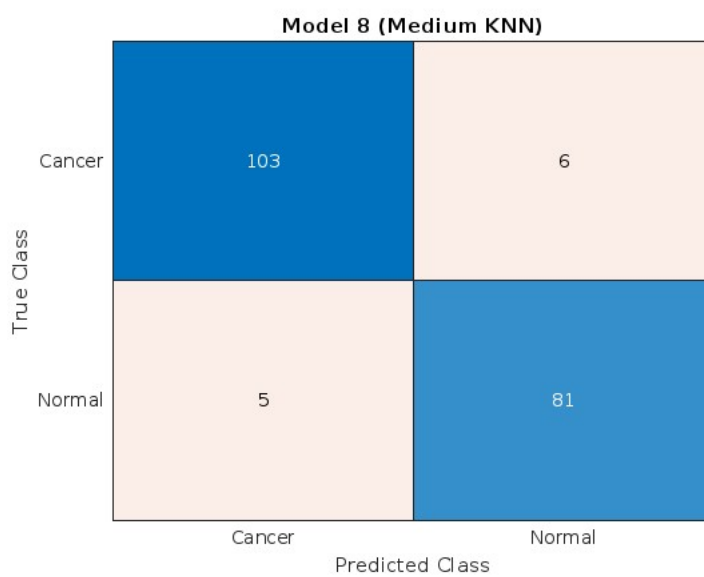
Σχήμα 1.10: Cancer Gaussian Naive Bayes Scatter



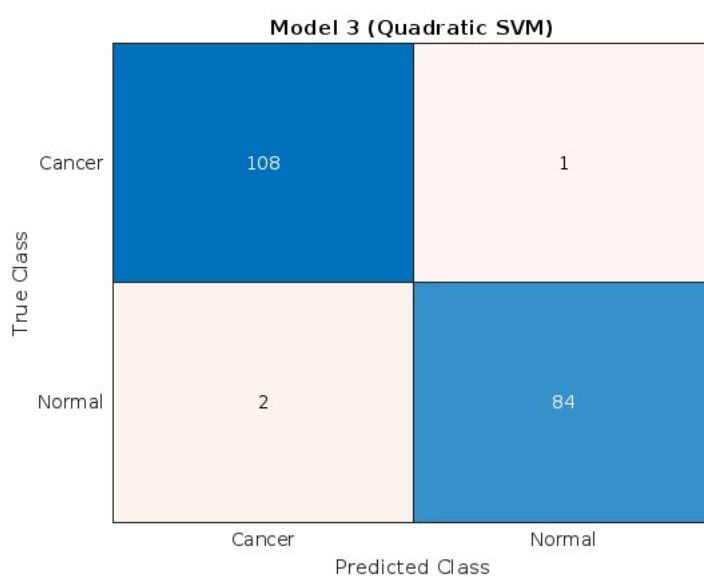
Σχήμα 1.11: Cancer Linear Discriminant Matrix



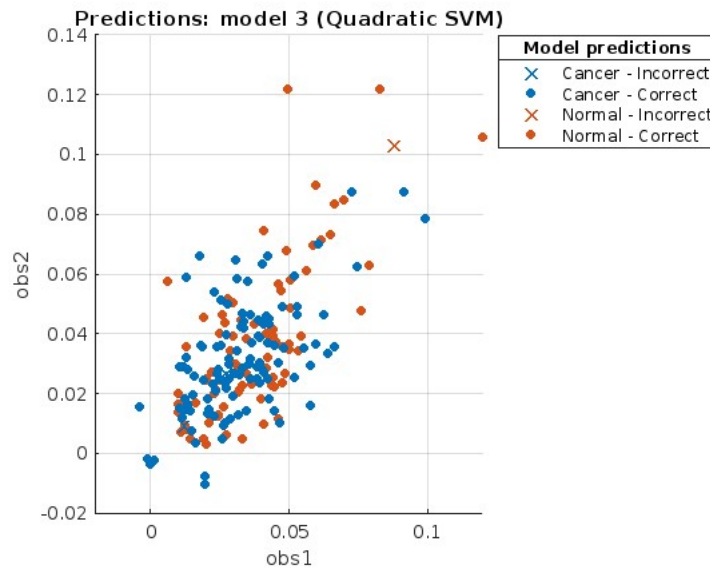
Σχήμα 1.12: Cancer Linear Discriminant Scatter



Σχήμα 1.13: Cancer Medium KNN Matrix



Σχήμα 1.14: Cancer Quadratic SVM Matrix



Σχήμα 1.15: Cancer Quadratic SVM Scatter

1.2 Regression Learner

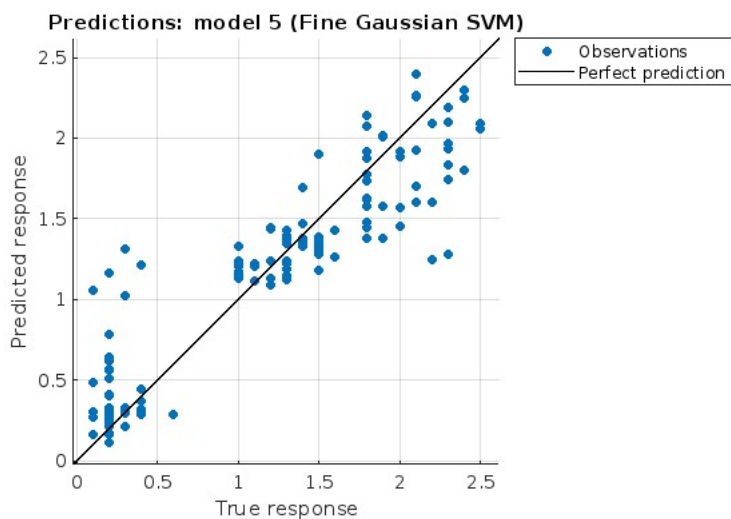
Τα μοντέλα που χρησιμοποιήθηκαν για το iris dataset είναι:

1. Fine Gaussian SVM
2. Fine Tree
3. Linear
4. Linear SVM

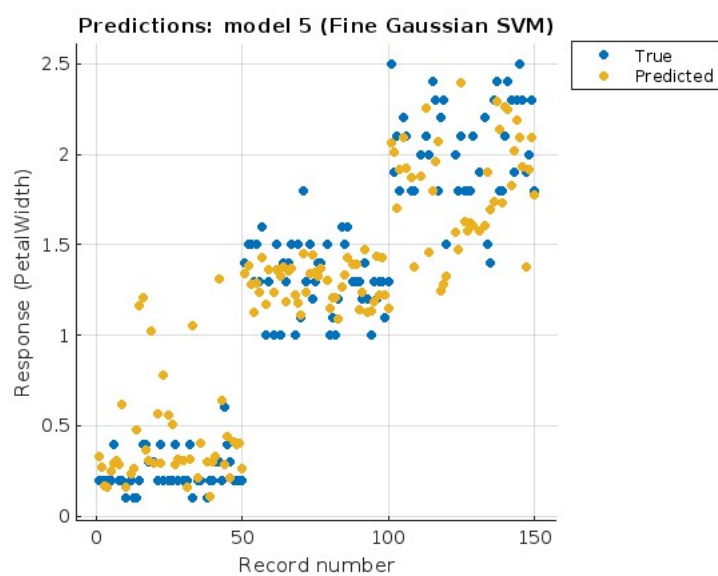
Στον πίνακα 1.3, παρατηρούμε πως ενώ το καλύτερο RMSE του validation είναι του μοντέλου Linear Regression, το καλύτερο RMSE στο test, είναι αυτό του Linear SVM. Οπότε, το μοντέλο που θα χρησιμοποιούσαμε, έτσι ώστε να προβλέψη με μεγάλη επιτυχία το μέγεθος του petal width (σε cm) είναι αυτό Linear SVM.

Πίνακας 1.3: Iris Regression Learner

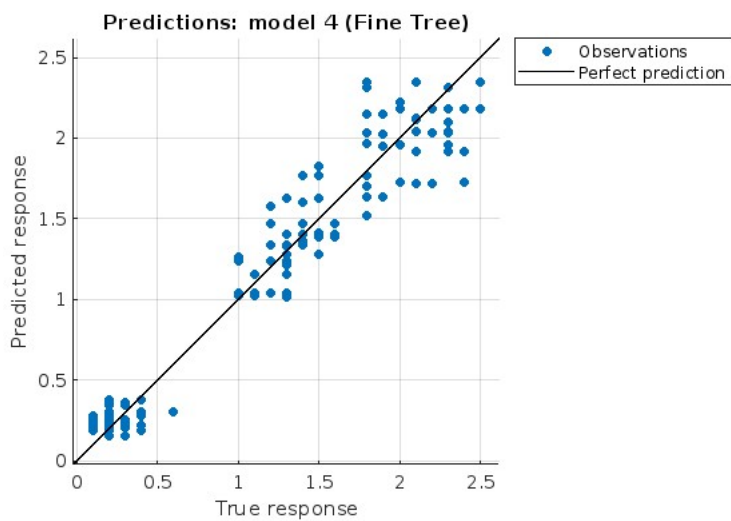
Model Type	RMSE (Validation)	RMSE (Test)
Linear Regression	0.165623108	0.199447886
Linear SVM	0.166582018	0.1969296
Fine Tree	0.193901175	0.223268336
Fine Gaussian SVM	0.313766514	0.380063468



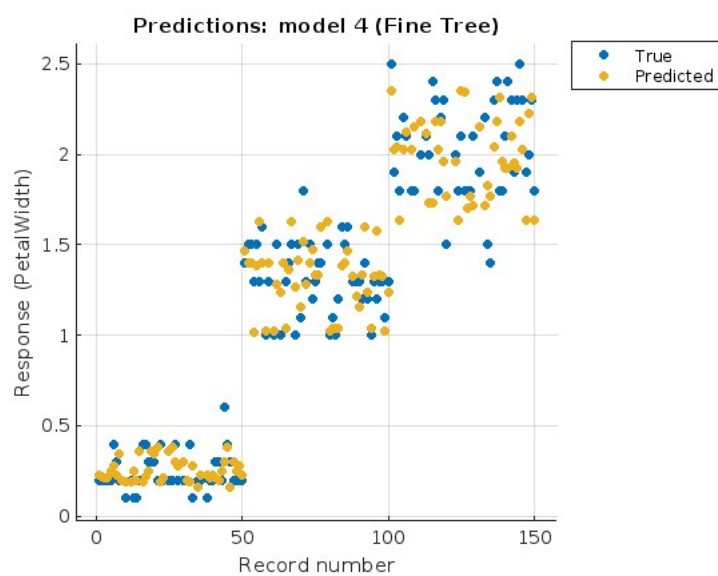
Σχήμα 1.16: Iris Fine Gaussian SVM Predictions



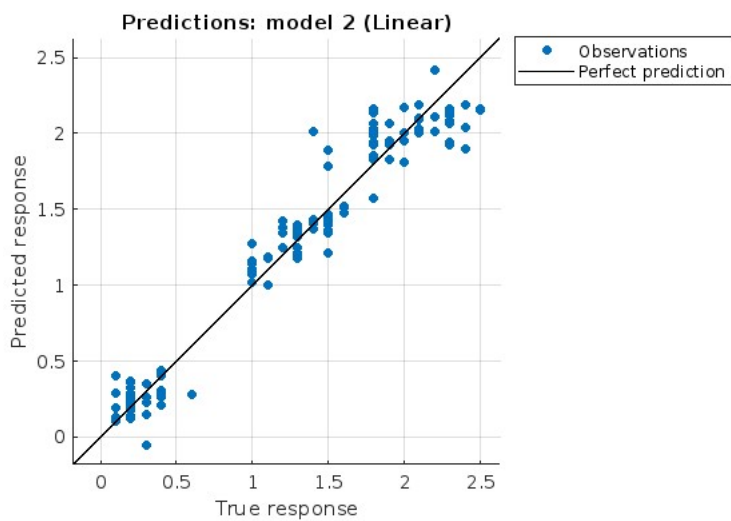
Σχήμα 1.17: Iris Fine Gaussian SVM Response



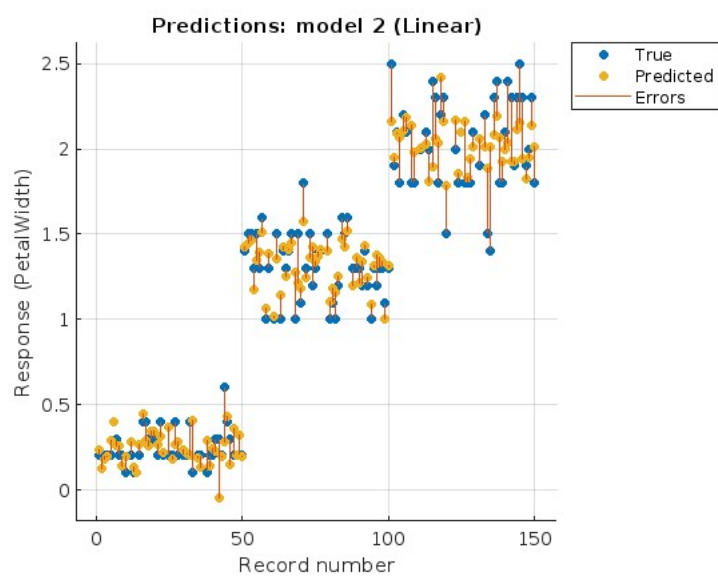
Σχήμα 1.18: Iris Fine Tree Predictions



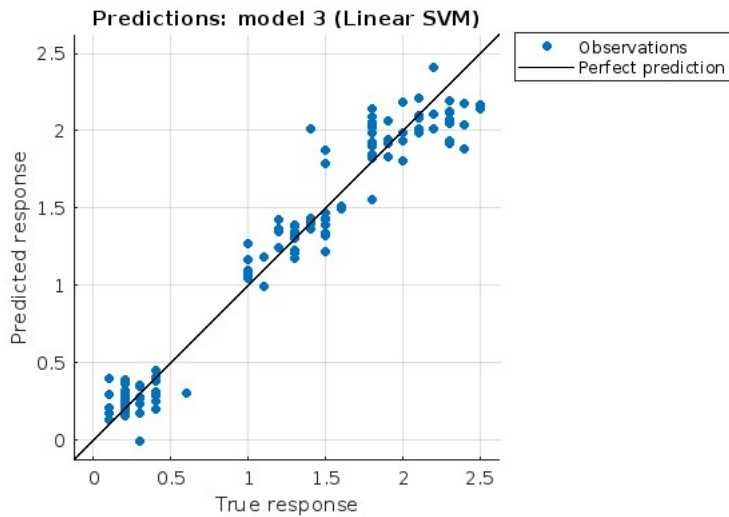
Σχήμα 1.19: Iris Fine Tree Response



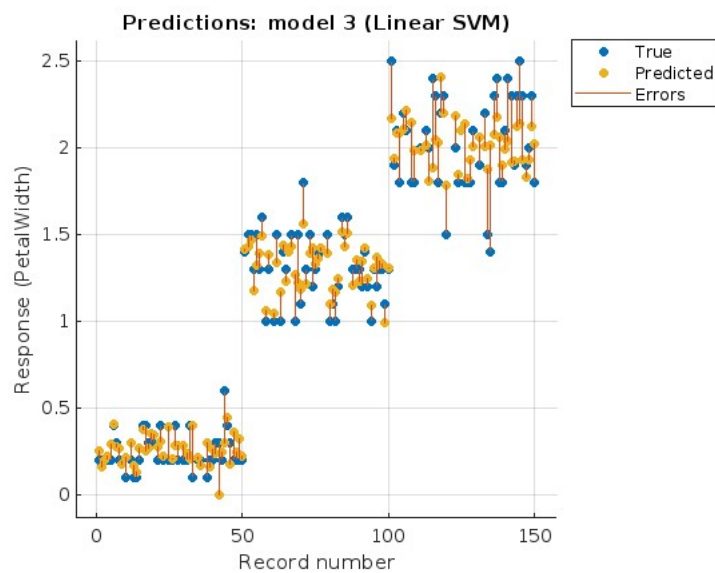
Σχήμα 1.20: Iris Linear Predictions



Σχήμα 1.21: Iris Linear Response



Σχήμα 1.22: Iris Linear SVM Predictions



Σχήμα 1.23: Iris Linear SVM Response

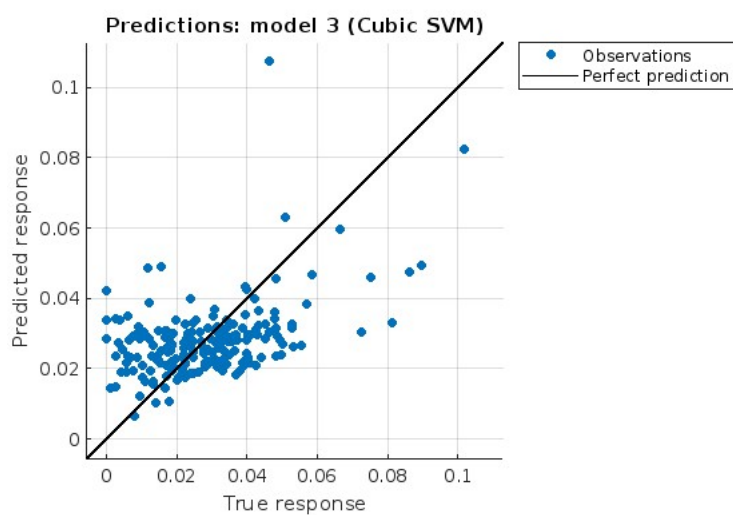
Τα μοντέλα που χρησιμοποιήθηκαν για το cancer dataset είναι:

1. Cubic SVM
2. Fine Tree
3. Linear SVM
4. Rational Quadratic GPR

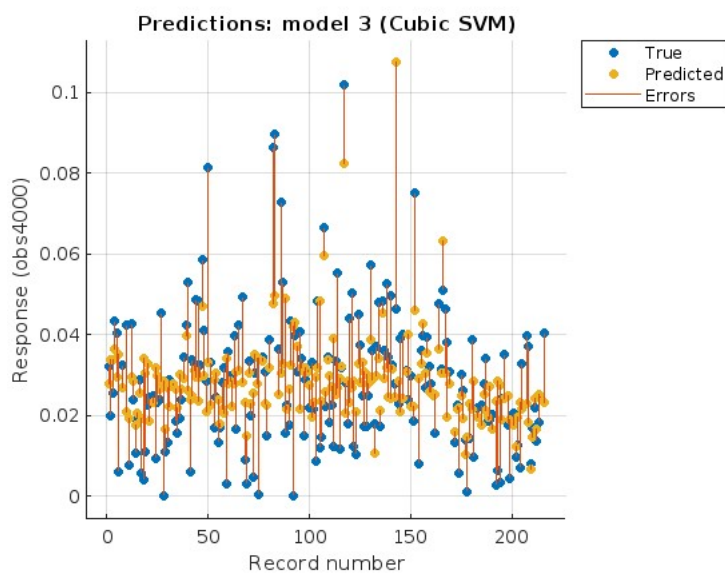
Το μοντέλο το οποίο μπορούσε να προβλέψει με μεγαλύτερη επιτυχία την τετρακοσιστή παρατήρηση στο test, όπως μπορούμε να παρατηρήσουμε στον πίνακα 1.4, είναι αυτό linear SVM, το οποίο είναι αντίστοιχα και αυτό που είχε τα καλύτερα αποτελέσματα στο Validation.

Πίνακας 1.4: Cancer Regression Learner

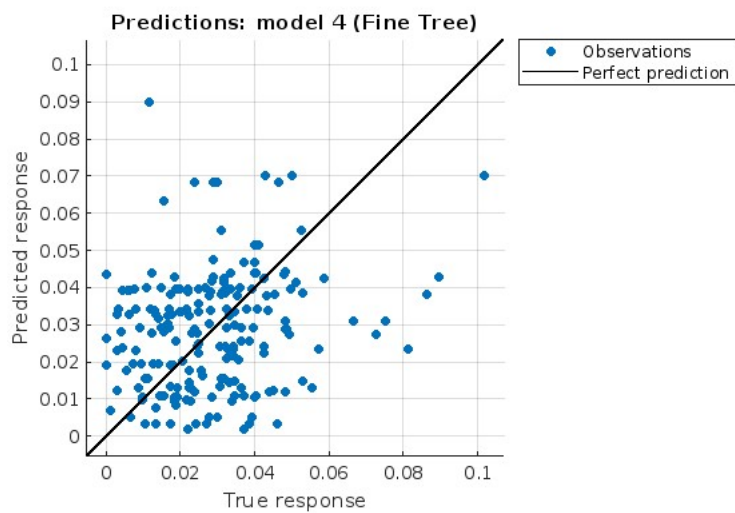
Model Type	RMSE (Validation)	RMSE (Test)
Cubic SVM	0.015317128974378	0.00871646079188
Linear SVM	0.015296372270678	0.007809000102565
Fine Tree	0.02071776380951	0.020012283163472
Gaussian Process Regression	0.015426885233592	0.008659863260083



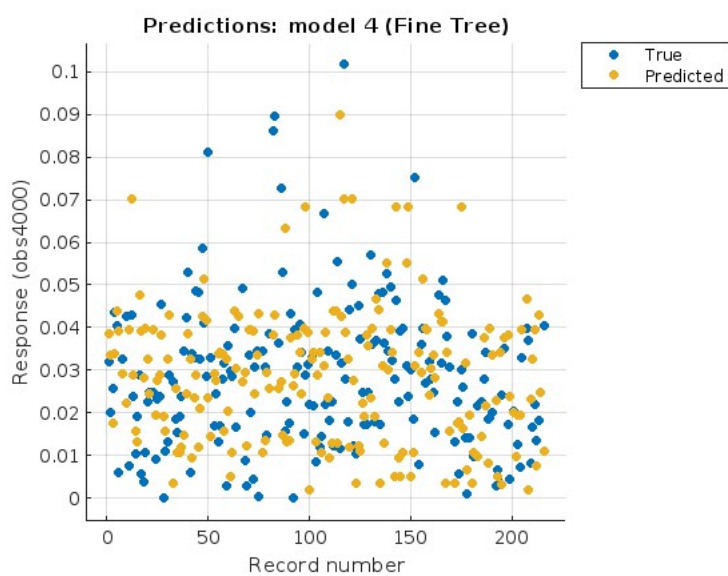
Σχήμα 1.24: Cancer Cubic SVM Predictions



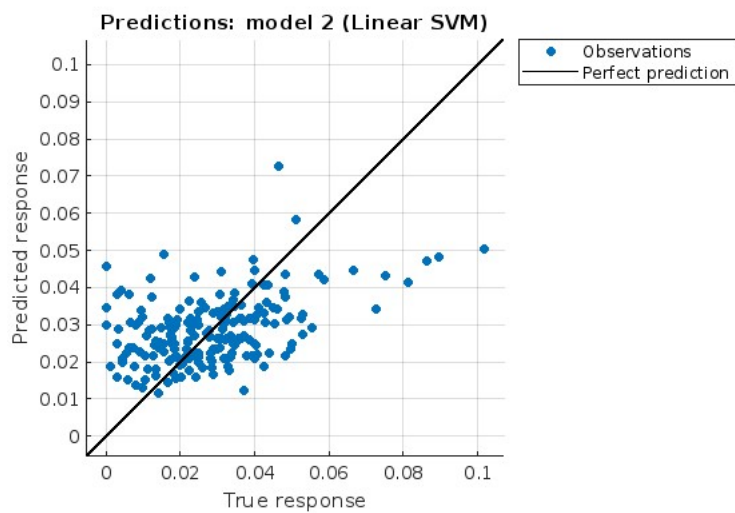
Σχήμα 1.25: Cancer Cubic SVM Response



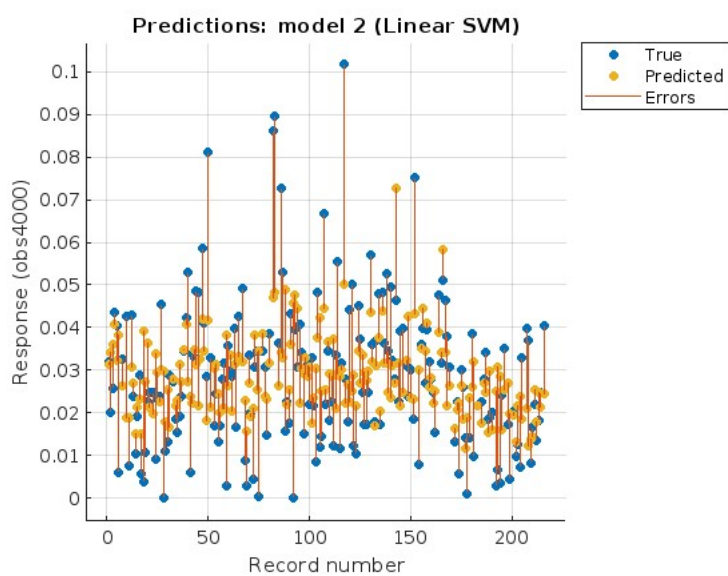
Σχήμα 1.26: Cancer Fine Tree Predictions



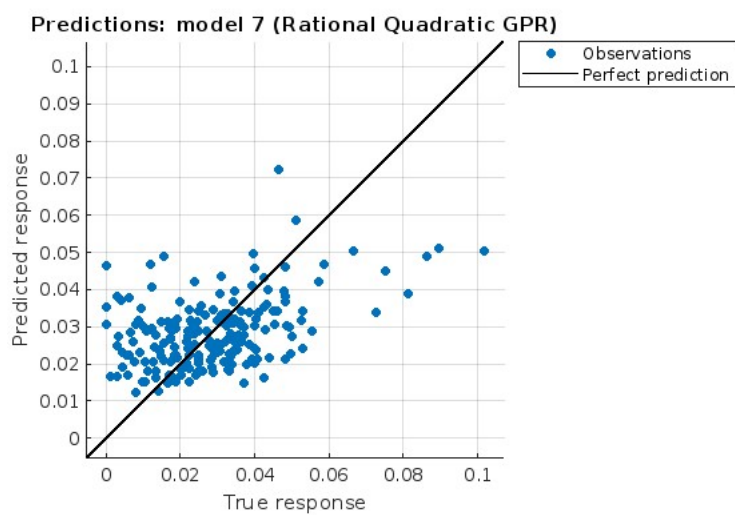
Σχήμα 1.27: Cancer Fine Tree Response



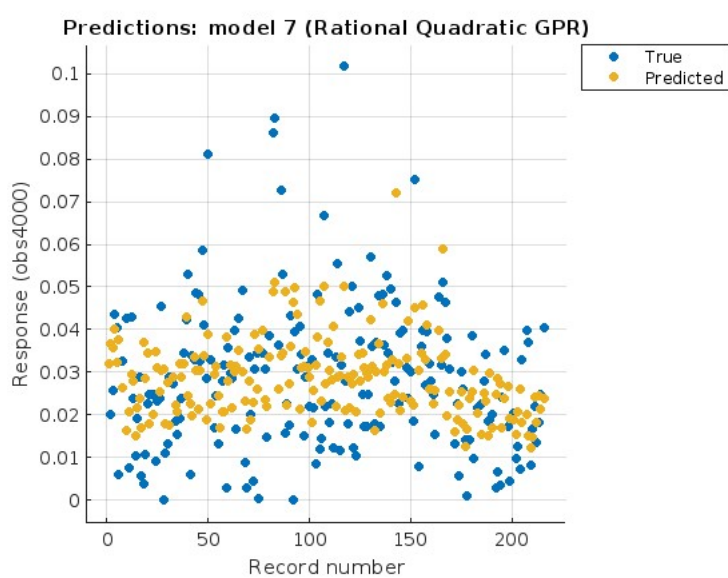
Σχήμα 1.28: Cancer Linear SVM Predictions



Σχήμα 1.29: Cancer Linear SVM Response



Σχήμα 1.30: Cancer Rational Quadratic GPR Predictions



Σχήμα 1.31: Cancer Rational Quadratic GPR Response