

Για την ταξινόμηση χρησιμοποιήθηκαν οι ακόλουθες ταξινομητές:

1. k-πλησιέστερων γειτόνων
2. ελάχιστης ευκλείδειας απόστασης
3. αφελής ταξινομητής κατά Bayes.

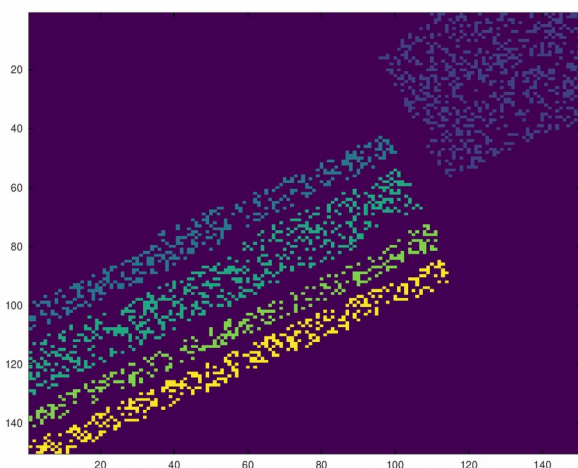
Οι ταξινομητές εκπαιδεύτηκαν με το Σύνολο Εκπαίδευσης (Training Set). Ακολουθούν τα μητρώα σύγκρισης και η οπτικοποίηση της ταξινόμησής του Συνόλου Αξιολόγησης (Test Set):

	1	2	3	4	5
1	652				
2		337	5		
3			549		
4				271	4
5				9	338
Ταξινομητής k-Πλησιέστερων Γειτόνων (k = 1)					

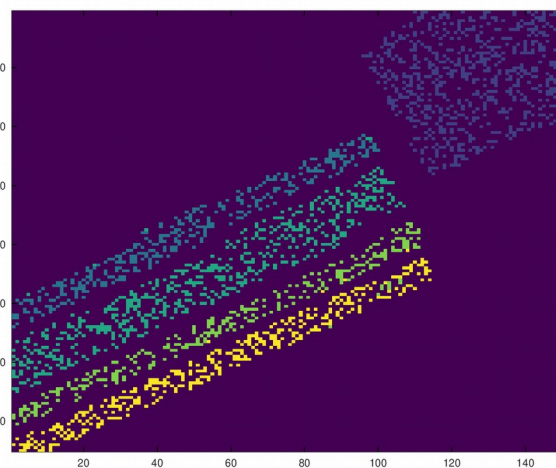
	1	2	3	4	5
1	652				
2		321	21		
3			549		
4				270	5
5		2		25	320
Ταξινομητής Ελάχιστης Ευκλείδειας Απόστασης					

	1	2	3	4	5
1	652				
2	2	333	7		
3		4	545		
4				267	8
5	3	1		12	331
Αφελής Ταξινομητής κατά Bayes					

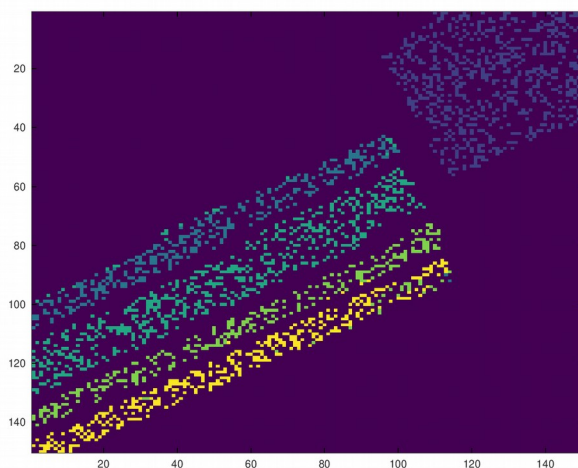
Πίνακας 1: Μητρώα Σύγκρισης για το Σύνολο Αξιολόγησης (Test Set)



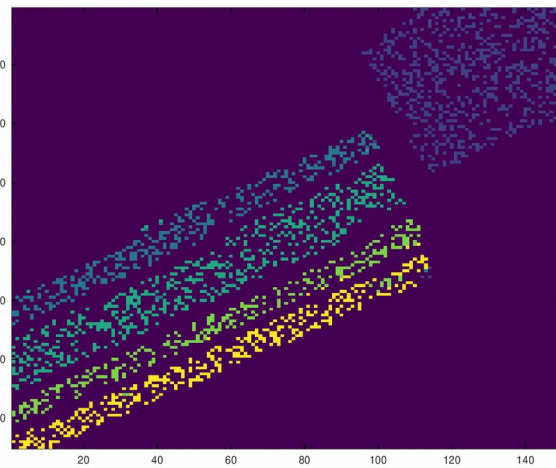
Εικόνα 1.1: Ετικέτες Συνόλου Αξιολόγησης



Εικόνα 1.2: Ταξινομητής k-Πλησιέστερων Γειτόνων (k=2)



Εικόνα 1.3: Ταξινομητής Ελάχιστης Ευκλείδειας Απόστασης



Εικόνα 1.4: Αφελής Ταξινομητής κατά Bayes

Ενώ για τις αποδόσεις των ταξινομητών, αθροίζοντας τα στοιχεία της κύριας διαγώνιου των μητρώων σύγχυσης έχουμε:

1. Ταξινομητής k-πλησιέστερω γειτόνων: 99.169 %
2. Ταξινομητής ελάχιστης ευκλείδειας απόστασης: 97.552 %
3. Αφελής Ταξινομητής κατά Bayes: 98.291 %

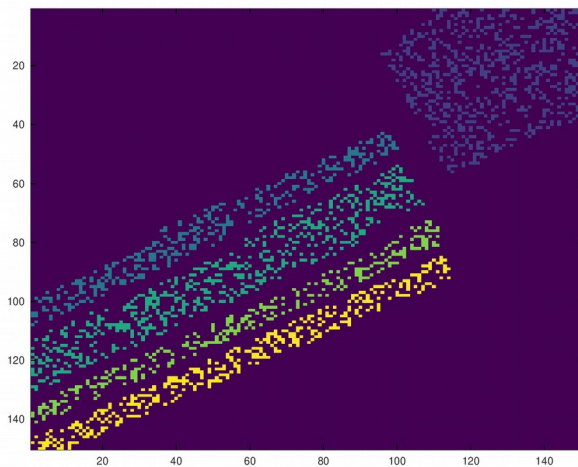
Στη συνέχεια παρουσιάζονται τα μητρώα σύγχυσης και οι αντίστοιχες εικόνες για το Σύνολο Λειτουργίας (Operational Set):

	1	2	3	4	5
1	1157				
2		530	4		
3			958		
4				444	7
5				19	490
Ταξινομητής k-Πλησιέστερων Γειτόνων (k = 1)					

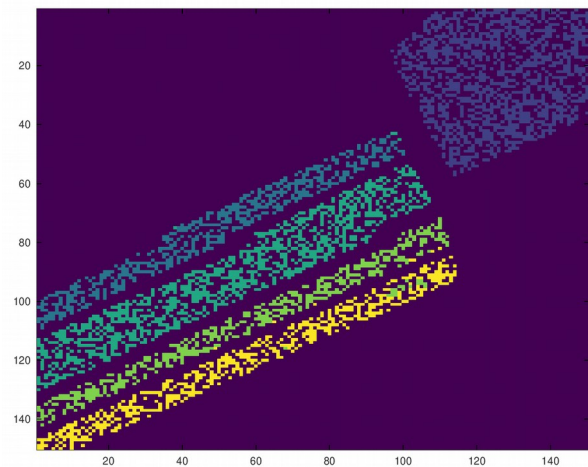
	1	2	3	4	5
1	1155			1	1
2		505	29		
3			958		
4				446	5
5		4		61	444
Ταξινομητής Ελάχιστης Ευκλείδειας Απόστασης					

	1	2	3	4	5
1	1157				
2	4	520	10		
3		12	946		
4				438	13
5	9	1		20	479
Αφελής Ταξινομητής κατά Bayes					

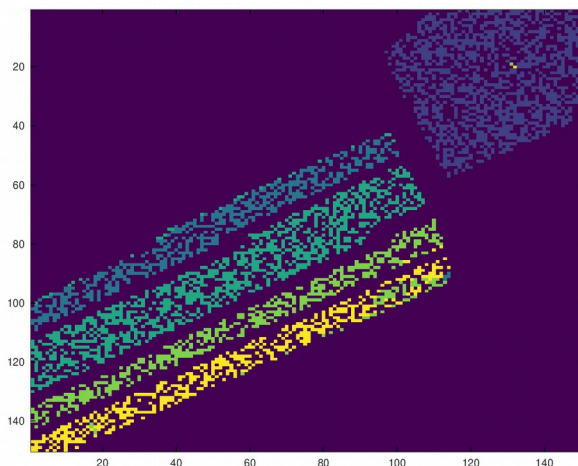
Πίνακας 2: Μητρώα Σύγχυσης για το Σύνολο Λειτουργίας (Operational Set)



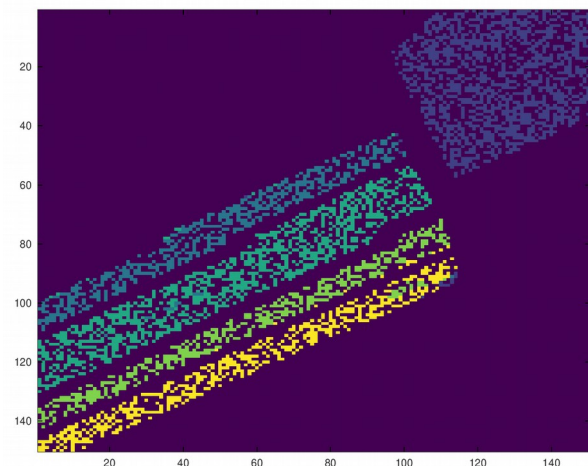
Εικόνα 2.1: Ετικέτες Συνόλου Λειτουργίας



Εικόνα 2.2: Ταξινομητής k-Πλησιέστερων Γειτόνων



Εικόνα 2.3: Ταξινομητής Ελάχιστης Ευκλείδειας Απόστασης



Εικόνα 2.4: Αφελής Ταξινομητής κατά Bayes

Για τις αποδόσεις των ταξινομητών στο Σύνολο Λειτουργίας, έχουμε:

1. Ταξινομητής k-πλησιέστερω γειτόνων: 99.169 %
2. Ταξινομητής ελάχιστης ευκλείδειας απόστασης: 97.201 %
3. Αφελής Ταξινομητής κατά Bayes: 98.088 %

### **Για τον Ταξινομητή κατά Bayes**

Στη γενική περίπτωση, λόγω την “κατάρας των μεγάλων διαστάσεων” (curse of dimensionality), αν  $N$  το πλήθος στοιχεία θα αρκούσαν για την εκπαίδευση ενός ταξινομητή στη μία διάσταση θα χρειαζόμασταν  $N^d$  σημεία για να εκπαιδεύσουμε έναν ταξινομητή στον  $d$ -διάστατο χώρο.

Παρ' όλα αυτά οι υποθέσεις που έχουμε κάνει παραπάνω για την χρήση πιο αφελών ταξινομητών, φαίνεται να λειτουργούν στην συγκεκριμένη περίπτωση. Συγκεκριμένα:

1. Το γεγονός πως ο Αφελής Ταξινομητής κατά Bayes δουλεύει πολύ ικανοποιητικά, μας δείχνει πως πράγματι οι T.M. των χαρακτηριστικών των σημείων της κάθε κλάσης είναι ανεξάρτητες (σε μεγάλο βαθμό), ο  $\Sigma_i$  είναι διαγώνιος για κάθε κλάση  $i$ .
2. Το γεγονός πως ο Ταξινομητής Ελάχιστης Ευκλείδειας απόστασης δουλεύει ικανοποιητικά, μας δείχνει πως  $\Sigma_i \approx I_d$ . Συνεπώς στην συγκεκριμένη περίπτωση θα μπορούσαμε απλώς να υποθέσουμε πως  $\Sigma_i = I_d$  για κάθε κλάση  $i$ .
3. Για το  $\mu_i$ , το γεγονός πως ο Ταξινομητής Ελάχιστης Ευκλείδειας Απόστασης δουλεύει ικανοποιητικά δείχνει ότι μάλλον το πλήθος των στοιχείων αρκεί για να εκτιμήσουμε το  $\mu_i$ .

### **Παρατηρήσεις:**

- Για την δημιουργία της εφαρμογής χρησιμοποιήθηκε Octave 4.2.2.
- Το πρόγραμμα τρέχει με την εντολή `octave ./project` στον φάκελο της εφαρμογής