

Cyclone Infrared Image Classification using Support Vector Machines

Όνομ/νυμο: Σταμάτιος Ορφανός
Αριθμός Μητρώου: mtn2211
Μάθημα: Μηχανική Μάθηση



Περιεχόμενα

1. Δεδομένα
2. Επεξεργασία Δεδομένων
3. Support Vector Machines
4. Εκπαίδευση
5. Accuracy Metrics
6. Demo
7. Συμπεράσματα



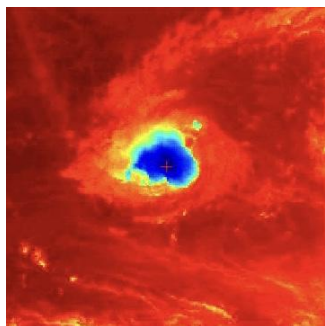
Δεδομένα

Τα δεδομένα για την ανάπτυξη του μοντέλου προέρχονται από την σελίδα Kaggle και συγκεκριμένα είναι το παρακάτω data set:

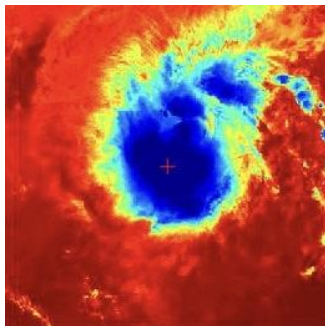
<https://www.kaggle.com/datasets/sshubam/insat3d-infrared-raw-cyclone-images-20132021>

Δεδομένα

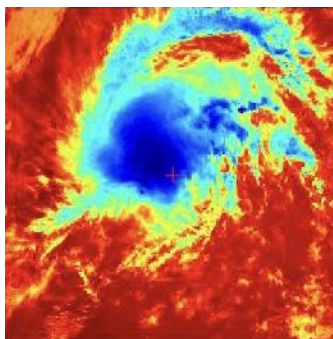
Τα δεδομένα σε αυτή την εργασία είναι Infrared εικόνες από Κυκλώνες στον Ινδικό ωκεανό και χωρίζονται σε 5 κατηγορίες με βάση την ισχύ τους:



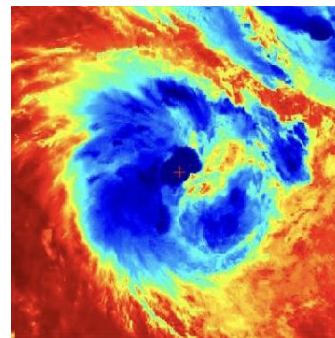
Κατηγορία 1



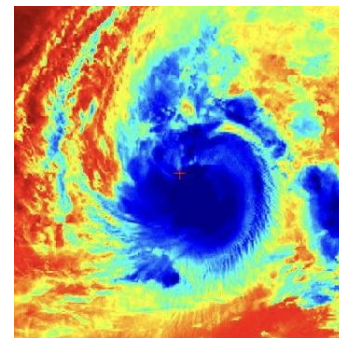
Κατηγορία 2



Κατηγορία 3



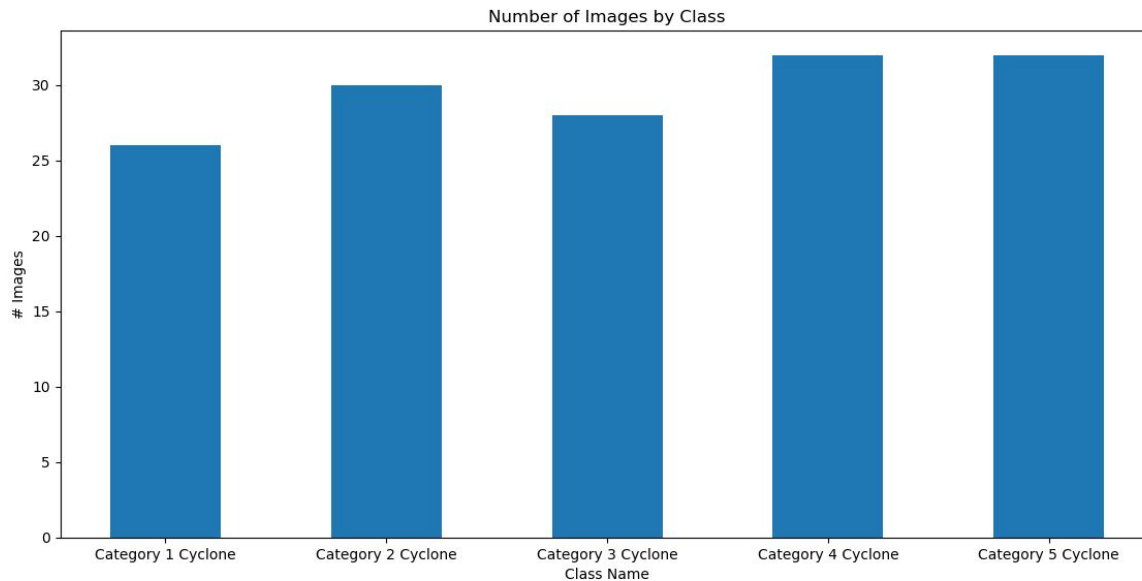
Κατηγορία 4



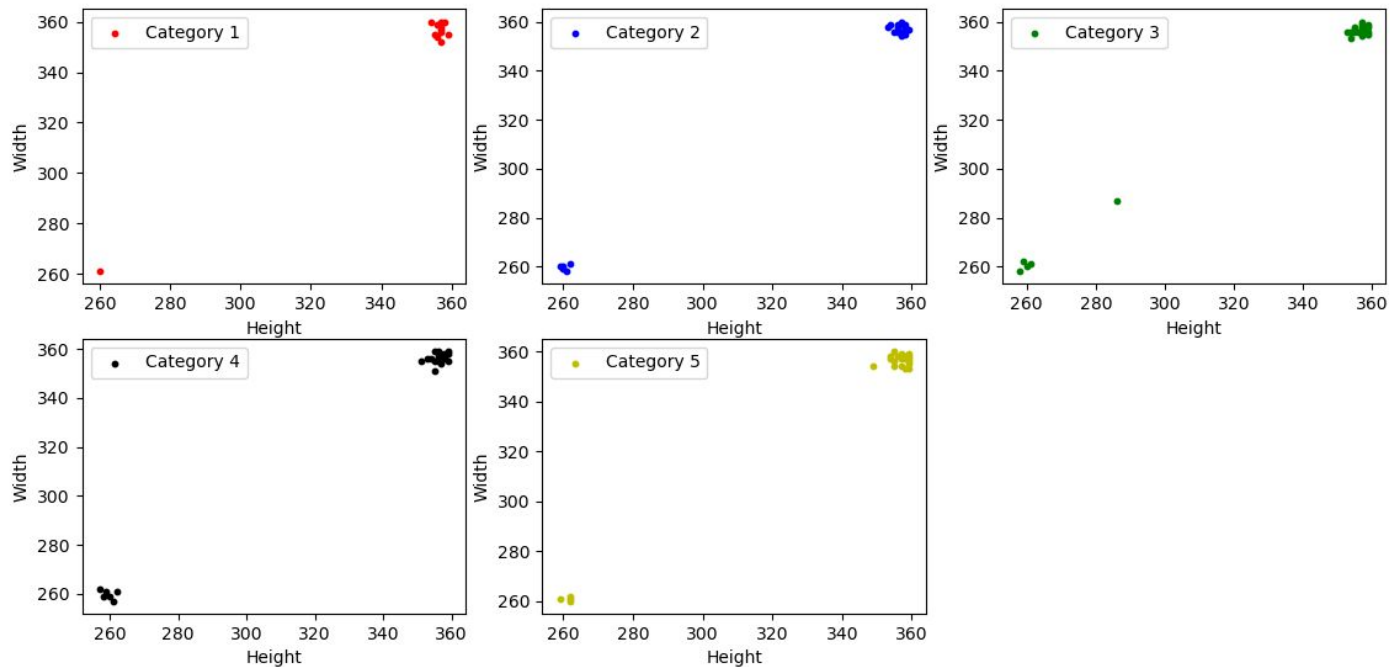
Κατηγορία 5



Επεξεργασία Δεδομένων

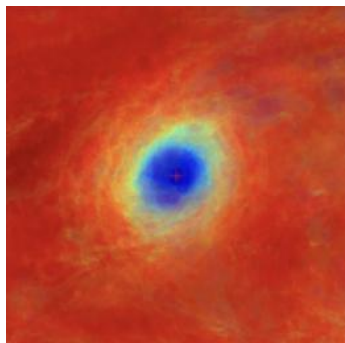


Επεξεργασία Δεδομένων

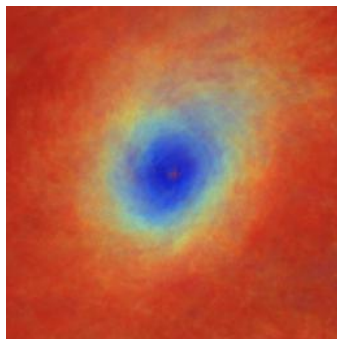


Επεξεργασία Δεδομένων

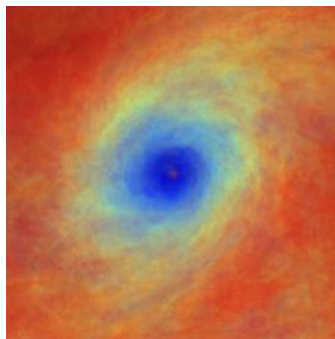
Αποτελέσματα μέσης φωτογραφίας για κάθε κατηγορία



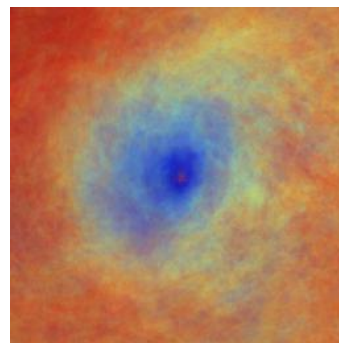
Κατηγορία 1



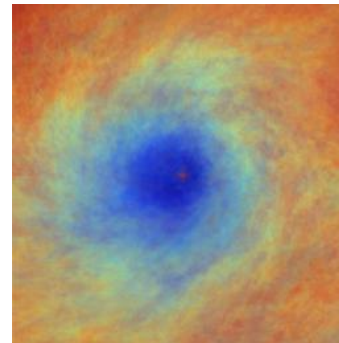
Κατηγορία 2



Κατηγορία 3



Κατηγορία 4



Κατηγορία 5



Support Vector Machines

Ο SVM είναι ένας supervised machine learning αλγόριθμος, που βοηθά σε προβλήματα classification ή regression. Στοχεύει στην εύρεση ενός βέλτιστου ορίου μεταξύ των πιθανών εξόδων.

Με απλά λόγια, το SVM κάνει σύνθετους μετασχηματισμούς δεδομένων ανάλογα με την επιλεγμένη συνάρτηση kernel και με βάση αυτούς τους μετασχηματισμούς, προσπαθεί να μεγιστοποιήσει τα όρια διαχωρισμού μεταξύ των σημείων δεδομένων σας ανάλογα με τις ετικέτες ή τις κλάσεις που έχουμε ορίσει.

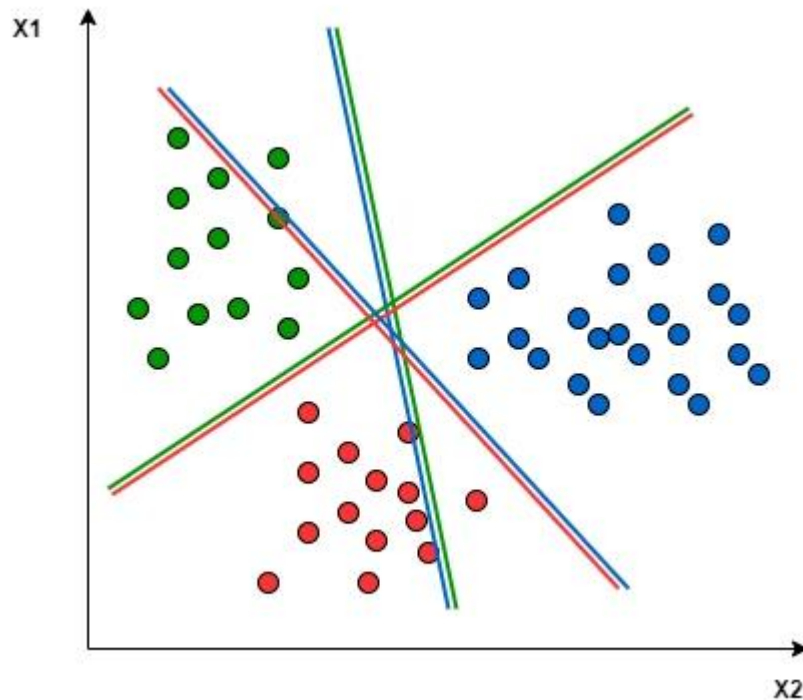
Support Vector Machines

One-to-One Προσέγγιση:

Τα SVMs δεν υποστηρίζουν nativally την multiclass classification αλλά binary classification

Η προσέγγιση One-to-One αναλύει το πρόβλημα πολλαπλών κλάσεων σε πολλαπλά προβλήματα δυαδικής ταξινόμησης. Ένας δυαδικός ταξινομητής για κάθε ζεύγος κλάσεων.

Έχουμε δηλαδή: $m(m-1)/2$ ταξινομητές.



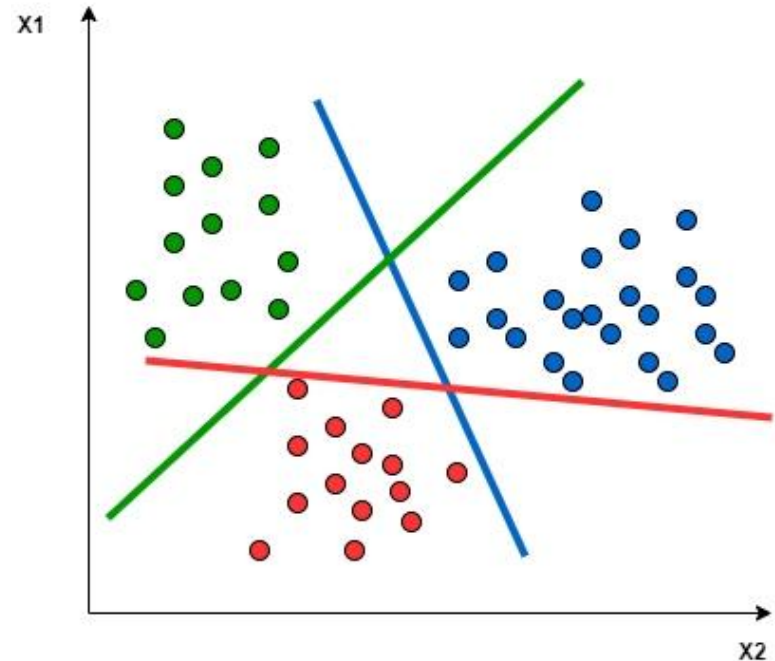
Support Vector Machines

One-to-Rest Προσέγγιση:

Σε αυτήν την προσέγγιση, η ανάλυση ορίζεται σε έναν δυαδικό ταξινομητή ανά κλάση.

Ένα μόνο SVM κάνει δυαδική ταξινόμηση και μπορεί να διαφοροποιήσει δύο κλάσεις. Έτσι ώστε, σύμφωνα με τις δύο προσεγγίσεις ανάλυσης, να ταξινομηθούν τα σημεία δεδομένων από το σύνολο δεδομένων m classes:

Έχουμε δηλαδή: m ταξινομητές.





Εκπαίδευση

Regularisation parameter C : 0.85

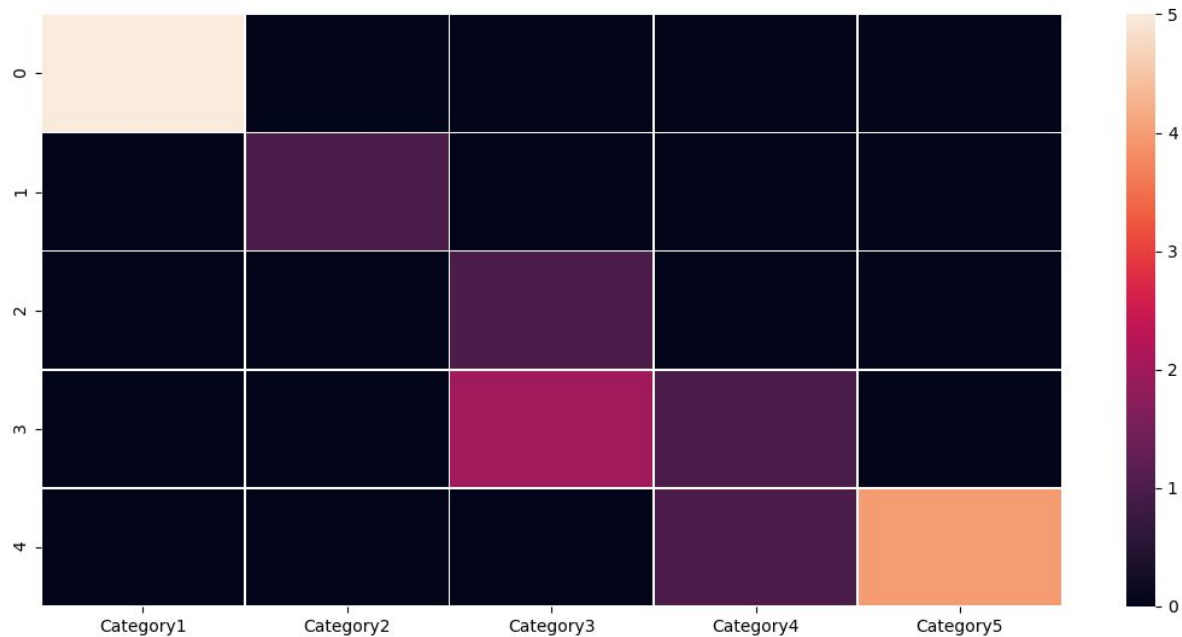
Kernel: poly – Polynomial

Kernel coefficient: γ : auto

Accuracy Metrics

One-versus-One SVM

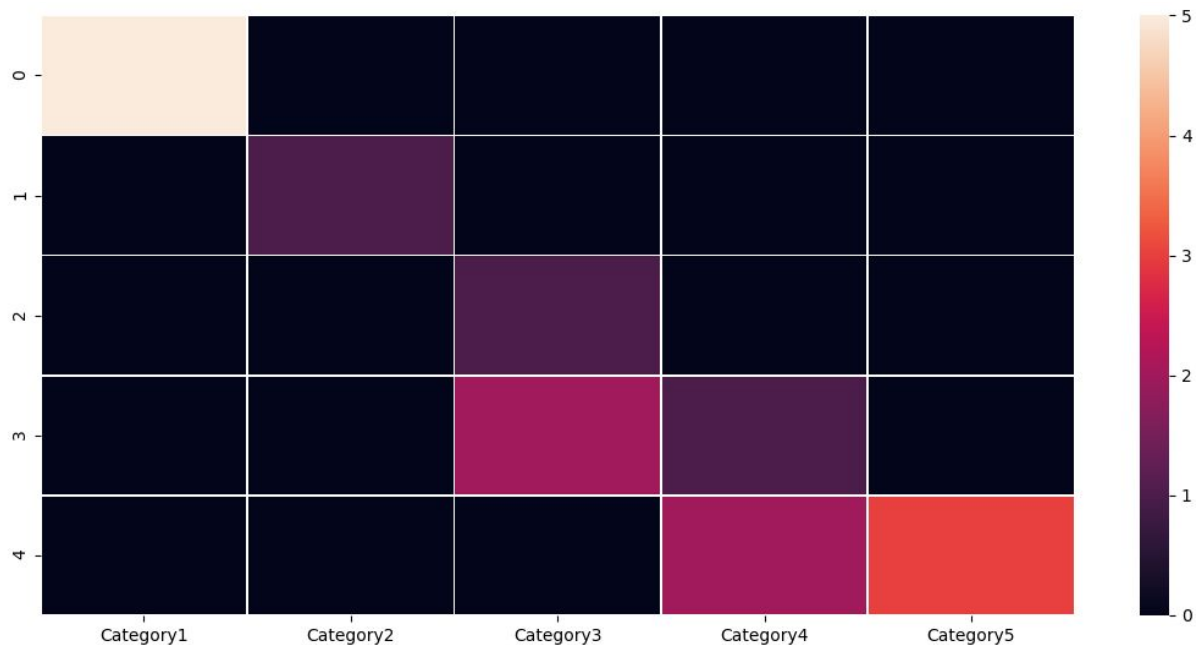
1. Accuracy: 73%
2. F1 score: 75%
3. Confusion Matrix



Accuracy Metrics

One-versus-Rest SVM

1. Accuracy: 73%
2. F1 score: 73%
3. Confusion Matrix





Συμπεράσματα

- + Γρήγορη εκτέλεση αλγορίθμου
- + Καλά αποτελέσματα με λίγα multidimensional δεδομένα
- Εξαρτάται αρκετά στο shuffling των δεδομένων για σωστά αποτελέσματα



Ευχαριστώ για την προσοχή σας