

ΘΕΩΡΙΑ ΑΠΟΦΑΣΕΩΝ

ΕΡΓΑΣΙΑ 2: Παραχάραξη χαρτονομισμάτων

Χρησιμοποιείτε τα δεδομένα Data.xlsx φορτώνοντας τα με τη βιβλιοθήκη pandas.

Τα δεδομένα εξήχθησαν από εικόνες που λήφθηκαν από γνήσια και πλαστά δείγματα από χαρτονομίσματα. Για την ψηφιοποίηση χρησιμοποιήθηκε μια βιομηχανική κάμερα. Αποκτήθηκαν εικόνες σε κλίμακα του γκρι. Για την εξαγωγή χαρακτηριστικών από τις εικόνες χρησιμοποιήθηκε η μέθοδος του μετασχηματισμού κυματιδίων (Wavelet Transform). Τα χαρακτηριστικά που εξάχθηκαν είναι:

Διασπορά, ασυμμετρία, κύρτωση, της μετασχηματισμένης εικόνας, εντροπία της εικόνας.

Χρησιμοποιήστε το 60% από κάθε κατηγορία (0-1) για εκπαίδευση, το 20% για validation, 20% για τεστ.

Κάνετε κανονικοποίηση των χαρακτηριστικών αν χρειάζεται.

1. Εφαρμόστε μέθοδο PCA για μείωση χαρακτηριστικών στα 2 και κάντε παράθυρα parzen για ταξινόμηση για διαφορετικά παράθυρα Gauss.
2. Εφαρμόστε k-NN για διαφορετικές τιμές του k.
3. Χρησιμοποιείτε κατάλληλο γραμμικό και μη γραμμικό SVM με κατάλληλες υπερπαραμέτρους. Δείξτε τα αποτελέσματα των δοκιμών σας.

Συγκρίνετε τις μεθόδους και σχολιάστε. Ποιο το F1 score σε κάθε περίπτωση;

Προαιρετικά:

4. Χρησιμοποιήστε μικτή κατανομή gauss για να μοντελοποιήσετε τα θετικά δείγματα. Δώστε την καμπύλη ROC για διαφορετικό κατώφλι πιθανότητας και διαφορετικό αριθμό στοιχείων της κατανομής.