

Αναγνώριση Προτύπων

Ακαδημαϊκό έτος 2011-2012

Χειμερινό Εξάμηνο

3^ο έτος

- **Ονομασία μαθήματος:**
Αναγνώριση Προτύπων
- **Διδάσκοντες:** Αγγελος Πικράκης, Λέκτορας,
Γεώργιος Τσιχριντζής, Καθηγητής
- **Ονομασία μαθήματος (στα Αγγλικά):** Pattern Recognition
- **Εξάμηνο:** 5ο, Υποχρεωτικό Μάθημα

- **Προτεινόμενα συγγράμματα** για το ακαδημαϊκό έτος 2011-2012:

- [1] "*Εισαγωγή στην Αναγνώριση Προτύπων με MATLAB*", S. Theodoridis, A. Pikrakis, K. Koutroumbas, D. Cavouras, 2011, Εκδόσεις Πασχαλίδης. Κωδικός στο σύστημα συγγραμμάτων Εύδοξος:45214. Μετάφραση του: "Introduction to Pattern Recognition: a MATLAB approach", S. Theodoridis, A. Pikrakis, K. Koutroumbas, D. Cavouras, Academic Press (imprint of Elsevier Science), 2010.
- [2] "*Αναγνώριση Προτύπων*", S. Theodoridis, K. Koutroumbas, 2011, Εκδόσεις Πασχαλίδης. Κωδικός στο σύστημα συγγραμμάτων Εύδοξος:12537236. Πρόκειται για μετάφραση του: "Pattern Recognition, 4th edition", S. Theodoridis, K. Koutroumbas, Academic Press, 2009.

- **Βαθμολόγηση:** 70% γραπτές εξετάσεις + 30 % εργασία. Η εργασία παραδίδεται τον Ιανουάριο και έχει ισχύ έως και τις εξετάσεις του Σεπτεμβρίου. Η εργασία εκπονείται σε ομάδες ενός/δύο φοιτητών και έχει ως θέμα την ανάπτυξη συστήματος ταξινόμησης και το έλεγχο της απόδοσής του. Για την εκπόνηση της εργασίας προτείνεται η χρήση GNU Octave (ή Matlab), αλλά είναι αποδεκτή και οποιαδήποτε έγκυρη γλώσσα προγραμματισμού.
- **Διεξαγωγή μαθήματος:** Η διδασκαλία του μαθήματος περιλαμβάνει, σε εβδομαδιαία βάση, 3 ώρες θεωρίας (σε αμφιθέατρο) και 1 ώρα εφαρμογής της θεωρίας (σε αμφιθέατρο). Για το εφαρμοσμένο σκέλος χρησιμοποιείται GNU Octave (ή Matlab όπου διατίθεται). Στην ενότητα εγγράφων του μαθήματος θα βρείτε σχετικά παραδείγματα.

- **Ώρες Γραφείου για φοιτητές για το μάθημα Αναγνώρισης Προτύπων: Δευτέρα 10:00-12:00, Τρίτη 12:00-14:00**
- **Χρήσιμες γνώσεις για επιτυχή παρακολούθηση του μαθήματος: Βασική Γραμμική Άλγεβρα και Στατιστική, γλώσσα προγραμματισμού με δυνατότητα διαχείρισης γραφικών (προτείνεται GNU Octave/Matlab/Visual γλώσσες).**

- **Διάλεξη 1-3:** Εισαγωγή στην Αναγνώριση Προτύπων, Ταξινομητές που βασίζονται στη θεωρία αποφάσεων του Bayes
- **Διάλεξη 4-5:** Ταξινομητές που βασίζονται στη Βελτιστοποίηση Συνάρτησης Κόστους
- **Διαλέξεις 6-7:** Μετασχηματισμοί Δεδομένων: Γένεση Χαρακτηριστικών και Μείωση Αριθμού Διαστάσεων
- **Διάλεξη 8:** Επιλογή Χαρακτηριστικών
- **Διάλεξη 9:** Σύγκριση με Πρότυπα Αναφοράς
- **Διάλεξη 10:** Hidden Markov Μοντέλα
- **Διαλέξεις 11-13:** Ομαδοποίηση

Ταξινομητές που βασίζονται στη θεωρία αποφάσεων του Bayes

- Καλούμαστε να ταξινομήσουμε ένα πρότυπο (*pattern*) σε μία από c δυνατές κλάσεις.
- Κάθε πρότυπο αναπαρίσταται με ένα σύνολο τιμών, $x(i)$, $i=1,2,\dots,l$, που συνιστούν ένα l -διάστατο διάνυσμα χαρακτηριστικών

$$x = [x(1), x(2), \dots, x(l)]^T \in \mathcal{R}^l$$

- Υποθέτουμε ότι κάθε πρότυπο αντιπροσωπεύεται, με μοναδικό τρόπο, από ένα και μόνο διάνυσμα χαρακτηριστικών και μπορεί να ανήκει σε μία μόνο κλάση

- Δοθέντος ενός διανύσματος $x \in \mathcal{R}^l$ και ενός συνόλου c κλάσεων, σύμφωνα με το θεώρημα Bayes

$$P(\omega_i|x)p(x) = p(x|\omega_i)P(\omega_i)$$

και

$$p(x) = \sum_{i=1}^c p(x|\omega_i)P(\omega_i)$$

- $P(\omega_i)$ είναι η εκ των προτέρων (a priori) πιθανότητα της κλάσης
- $P(\omega_i|x)$ είναι η εκ των υστέρων (a posteriori) πιθανότητα της κλάσης δοθέντος του x

- $p(x)$ είναι η συνάρτηση πυκνότητας πιθανότητας του x
- είναι η συνάρτηση της υπό-συνθήκη πυκνότητας πιθανότητας του x δοθείσης της κλάσης ω_i (πιθανοφάνεια).

ΘΕΩΡΙΑ ΑΠΟΦΑΣΕΩΝ ΤΟΥ BAYES

- Έστω $x \equiv [x(1), x(2), \dots, x(l)]^T \in \mathcal{R}^l$
- Το x ταξινομείται στην ω_i σύμφωνα με την

$$P(\omega_i|x) > P(\omega_j|x), \quad \forall j \neq i$$

$$p(x|\omega_i)P(\omega_i) > p(x|\omega_j)P(\omega_j), \quad \forall j \neq i$$

- Ο Bayesian ταξινομητής είναι βέλτιστος υπό την έννοια ότι ελαχιστοποιεί την πιθανότητα σφάλματος