

ΕΜΠ ΔΠΜΣ

Εφαρμοσμένες Μαθηματικές
Επιστήμες

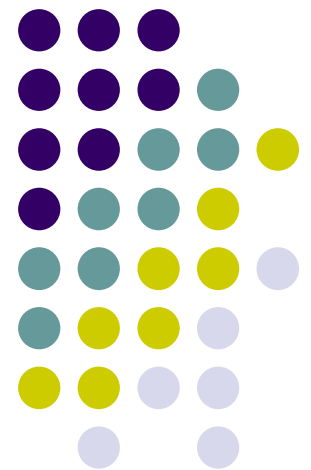
Αλγόριθμοι Εξόρυξης Πληροφορίας

Διάλεξη 07:

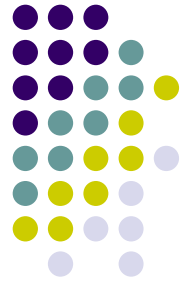
Αλγόριθμοι Εκμάθησης

Συναρτήσεις και μετα-μαθησιακοί

Αλγόριθμοι

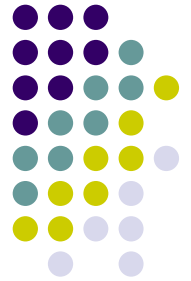


Γραμμικά Μοντέλα



- Λειτουργούν εκ φύσεως με αριθμητικά χαρακτηριστικά
- Τυπική τεχνική αριθμητικής πρόβλεψης: γραμμική παλινδρόμηση
-Το αποτέλεσμα είναι γραμμικός συνδυασμός των χαρακτηριστικών
- Οποιαδήποτε τεχνική παλινδρόμησης μπορεί να χρησιμοποιηθεί, εκτός της αριθμητικής πρόβλεψης, για ταξινόμηση (classification)
- Το πρόβλημα που δημιουργείται είναι η παραβίαση κάποιων υποθέσεων κατά την υιοθέτηση γραμμικής παλινδρόμησης σε προβλήματα ταξινόμησης
- Γενικά τα γραμμικά μοντέλα είναι ακατάλληλα στην περίπτωση που τα δεδομένα παρουσιάζουν μη γραμμικές εξαρτήσεις
- Ωστόσο μπορούν να αποτελέσουν πρώτη ύλη για την κατασκευή περισσότερο πολύπλοκων σχημάτων (π.χ. model trees). Μια εναλλακτική προσέγγιση είναι η εκμάθηση υπερεπιπέδου (perceptron) που διαχωρίζει τα υποδείγματα διαφορετικών τάξεων

Μηχανές Διανυσμάτων Υποστήριξης



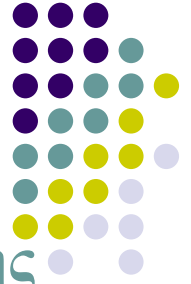
- Οι μηχανές διανυσμάτων υποστήριξης αποτελούν αλγορίθμους που επιτυγχάνουν εκμάθηση γραμμικών ταξινομητών
- Έχουν χαμηλό υπολογιστικό κόστος ακόμη και στην περίπτωση μη γραμμικότητας
- Οι αλγόριθμοι μηχανών διανυσμάτων υποστήριξης επιταχύνονται στην περίπτωση που τα δεδομένα είναι αραιά (sparse, δηλαδή πολλές τιμές είναι μηδενικές) και αυτό γιατί υπολογίζονται πολλά διανυσματικά γινόμενα

Νευρωνικά Δίκτυα



- Σύνδεση πολλών απλών όμοιων προς το perceptron μοντέλο σε μία ιεραρχική δομή : νευρωνικά δίκτυα (neural networks)
- Είναι ικανά να απεικονίσουν μη γραμμικά όρια αποφάσεων
- Ένα perceptron περιγράφεται συχνά ως ένας τεχνικός νευρώνας
- Η ακόλουθη παρατήρηση ενέπνευσε τη δημιουργία τεχνητών νευρώνων: οι νευρώνες του ανθρώπινου εγκεφάλου διασυνδέονται σε μαζική κλίμακα, επιτρέποντας την αποσύνθεση ενός προβλήματος σε υποπροβλήματα που μπορούν να επιλυθούν σε επίπεδο νευρώνα.

Μετα-μαθησιακά Σχήματα



- Θεμελιώδης ιδέα: κατασκευή πολλών «εμπειρογνωμόνων», ανάδειξη πλειοψηφούσας γνώμης
 - «εμπειρογνώμονας» είναι ένα μοντέλο που δημιουργήθηκε με τεχνικές μηχανικής μάθησης
- Πλεονεκτήματα: συχνά βελτιώνει σημαντικά την προβλεπτική ικανότητα
- Μειονεκτήματα: τα εξαγόμενα είναι πολύ δύσκολο να αναλυθούν
- Σχήματα (εφαρμόζονται τόσο για ταξινόμηση όσο και για αριθμητική πρόβλεψη)
 1. Εμφωλίαση (bagging)
 2. Ενδυνάμωση (boosting)
 3. Συσσώρευση (stacking)
 4. Κώδικες διόρθωσης σφαλμάτων εξόδου (error-correcting output codes)