

# Δέντρα απόφασης .

## ( Decision Trees )

---

### Λίγα λόγια σχετικά με τα δένδρα απόφασης .

---

Ένα δένδρο απόφασης είναι μία τεχνική μοντελοποίησης πρόβλεψης η οποία χρησιμοποιείται στην κατηγοριοποίηση , στην ομαδοποίηση και στις εργασίες πρόβλεψης . Τα δένδρα αποφάσεων χρησιμοποιούν την τεχνική που << διαιρεί και βασίλευε >>για να χωρίσουν το χώρο αναζήτησης του προβλήματος σε υποσύνολα .

### Ένας αλγόριθμος κατασκευής δέντρων απόφασης

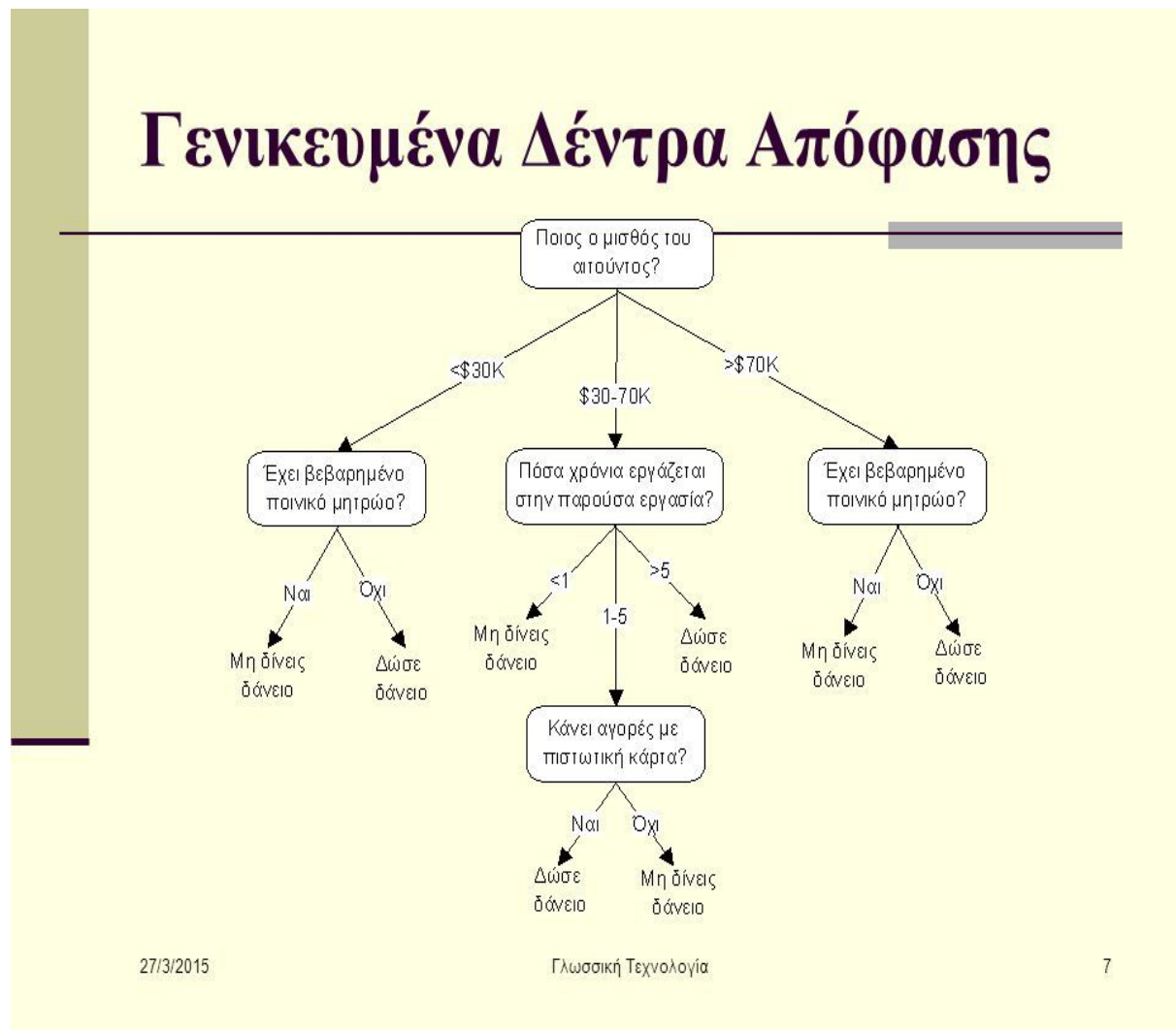
---

- Τα δέντρα απόφασης κατασκευάζονται χρησιμοποιώντας μόνο εκείνα τα γνωρίσματα που είναι σε θέση να διακρίνουν τις έννοιες προς εκμάθηση.
- Για να χτίσουμε ένα δέντρο απόφασης, πρέπει αρχικά να επιλέξουμε ένα υποσύνολο περιπτώσεων από το σύνολο των δεδομένων που θα χρησιμοποιηθούν στην εκπαίδευση (υποσύνολο δεδομένων εκπαίδευσης - training set).
- Αυτό το υποσύνολο (δεδομένα ελέγχου - test set) χρησιμοποιείται έπειτα από τον αλγόριθμο για να κατασκευάσει το δέντρο απόφασης.
- Τα υπόλοιπα δεδομένα, τα δεδομένα training set, χρησιμοποιούνται στην εξέταση της ακρίβειας του κατασκευασμένου δέντρου. Εάν το δέντρο απόφασης ταξινομεί τις περιπτώσεις σωστά, η διαδικασία ολοκληρώνεται. Εάν μια περίπτωση είναι ανακριβώς ταξινομημένη, η περίπτωση προστίθεται στο επιλεγμένο υποσύνολο των training set και ένα νέο δέντρο κατασκευάζεται
- Τα βήματα του αλγορίθμου είναι τα ακόλουθα:
  1. Έστω T είναι το σύνολο στιγμιότυπων εκπαίδευσης, το training set
  2. Επιλέγουμε ένα χαρακτηριστικό που διαφοροποιεί καλύτερα τις περιπτώσεις που περιλαμβάνονται στο T.
  3. Δημιουργούμε έναν κόμβο στο δέντρο του οποίου η αξία είναι το επιλεγμένο χαρακτηριστικό. Δημιουργούμε θυγατρικούς δεσμούς από αυτόν τον κόμβο, όπου κάθε σύνδεση αντιπροσωπεύει μια μοναδική αξία για τα επιλεγμένα χαρακτηριστικά. Χρησιμοποιούμε τις τιμές των θυγατρικών δεσμών για να υποδιαιρέσουμε περαιτέρω τα στιγμιότυπα σε δευτερεύουσες κλάσεις.
  4. Για κάθε δευτερεύουσα κλάση που δημιουργήθηκε στο βήμα 3: - Εάν τα στιγμιότυπα στη δευτερεύουσα κλάση ικανοποιούν προκαθορισμένα κριτήρια ή εάν το σύνολο των υπολοίπων επιλογών γνωρισμάτων γι' αυτή τη διαδρομή του δέντρου είναι μηδέν, καθορίζουμε την κατηγοριοποίηση των καινούργιων στιγμιότυπων που ακολουθούν αυτή τη διαδρομή αποφάσεων. - Εάν η δευτερεύουσα κλάση δεν ικανοποιεί τα προκαθορισμένα κριτήρια, και υπάρχει τουλάχιστον ένα γνώρισμα για να υποδιαιρέσει περαιτέρω τη διαδρομή του δέντρου, αφήστε το T να είναι το τρέχον σύνολο των στιγμιότυπων της δευτερεύουσας κλάσης και επιστρέψουμε στο βήμα 2.

# Δέντρα απόφασης . ( Decision Trees )

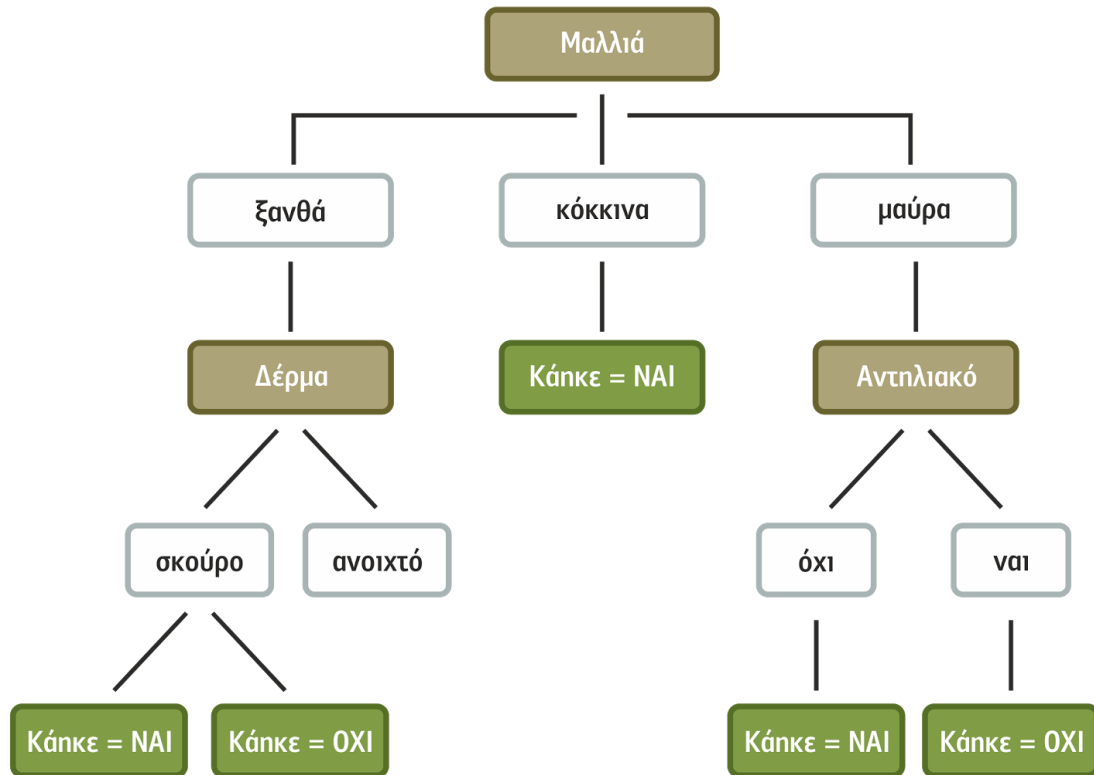
Παράδειγμα δένδρων απόφασης .

1)



# Δέντρα απόφασης . ( Decision Trees )

2)



$R_1$ : IF (Μαλλιά μαύρα) AND (Δέρμα σκούρο) THEN Κάπκε= ΟΧΙ  
 $R_2$ : IF (Μαλλιά μαύρα) AND (Δέρμα ανοιχτό) THEN Κάπκε= ΝΑΙ  
 $R_3$ : IF (Μαλλιά κόκκινα) THEN Κάπκε= ΝΑΙ  
 $R_4$ : IF (Μαλλιά ξανθά) AND (Αντιπηλικό ναι ) THEN Κάπκε= ΟΧΙ  
 $R_5$ : IF (Μαλλιά ξανθά) AND (Αντιπηλικό όχι ) THEN Κάπκε= ΝΑΙ