



ΠΑΝΕΠΙΣΤΗΜΙΟ  
ΠΑΤΡΩΝ  
UNIVERSITY OF PATRAS

ΑΝΟΙΚΤΑ ακαδημαϊκά  
μαθήματα ΠΠ

# Ευφυής Προγραμματισμός

Ενότητα 13: Δημιουργία Βάσεων Κανόνων Από  
Δεδομένα- Αξιολόγηση Βάσης Κανόνων

Ιωάννης Χατζηλυγερούδης

Πολυτεχνική Σχολή

Τμήμα Μηχανικών Η/Υ & Πληροφορικής

Δημιουργία Βάσεων Κανόνων  
Από Δεδομένα- Αξιολόγηση  
Βάσης Κανόνων

# Δημιουργία Βάσης Κανόνων (1)

- Οι κανόνες που έχουν εξαχθεί μέσω της μεθόδου ΔΑ
  - Μπορεί να έχουν μη ικανοποιητική απόδοση ή η απόδοσή τους να μπορεί να βελτιωθεί.
  - Μπορεί κάποιοι από αυτούς να μην ανταποκρίνονται στην πραγματικότητα.
- Οπότε απαιτείται αναθεώρηση των κανόνων

# Δημιουργία Βάσης Κανόνων (2)

- Διαδικασία αναθεώρησης κανόνων:
  1. Υλοποίηση των κανόνων μέσω ενός εργαλείου ανάπτυξης συστημάτων κανόνων.
  2. Αναθεώρηση των κανόνων με τη βοήθεια εμπειρογνώμονα του πεδίου και του συνόλου δεδομένων.
  3. (Επαν) αξιολόγηση των κανόνων.
  4. Αν τα αποτελέσματα είναι ικανοποιητικά, τότε σταμάτα. Αλλιώς, πήγαινε στο βήμα 2.

# Εργαλεία Υποστήριξης Εξαγωγής και Υλοποίησης Κανόνων

## WEKA

- Προσφέρει σύνοψη δεδομένων
- Οπτικοποίηση κατανομής παραδειγμάτων
- Χειρισμός ελλιπών τιμών
- Χειρισμός θορύβου δεδομένων
- Κάνει επιλογή χαρακτηριστικών
- Κάνει διακριτοποίηση
- Παράγει κανόνες

# Εργαλεία Υποστήριξης Εξαγωγής και Υλοποίησης Κανόνων

## CLIPS (JESS)

- Κέλυφος ανάπτυξης συστημάτων βασισμένων σε κανόνες.
- Υλοποίηση προτασιακών κανόνων και κανόνων πρώτης τάξεως.
- Ελεύθερο λογισμικό ανοικτού κώδικα.
- Διαθέτει εργαλεία εκσφαλμάτωσης.
- Επιλογή στρατηγικών συλλογισμού και επίλυσης συγκρούσεων.

# Μετρικές Αξιολόγησης-Προβλήματα Ταξινόμησης (1)

## Μετρικές

- **Accuracy** =  $(TP+TN)/(TP+FN+FP+TN)$  (Ορθότητα)
- **Sensitivity** =  $TP/(TP+FN)$  (Ευαισθησία)
- **Specificity** =  $TN/(TN+FP)$  (Εξειδίκευση).
- **Precision** =  $TP/(TP+FP)$  (Ακρίβεια)

Εναλλακτικά της «ορθότητας» χρησιμοποιείται και ο:

**Error rate** =  $(FN+FP)/(TP+FN+FP+TN)$  (λόγος λάθους)

Σημαντικότερη η «ορθότητα», αφού εξασφαλίσουμε ότι «ευαισθησία» και «εξειδίκευση» είναι ισορροπημένες.

# Μετρικές Αξιολόγησης-Προβλήματα Ταξινόμησης (2)

- Οι παραπάνω μετρικές χρησιμοποιούνται ως έχουν σε περιπτώσεις δυαδικής εξόδου (ΝΑΙ-ΌΧΙ, ΑΛΗΘΗΣ-ΨΕΥΔΗΣ).
- Σε περιπτώσεις εξόδων πολλαπλών κλάσεων, η μέτρηση TN και οι μετρικές που την χρησιμοποιούν (ορθότητα, εξειδίκευση) χάνουν την αξιοπιστία τους. Τότε χρησιμοποιούνται οι μετρικές:
  - **Recall** =  $TP/(TP+FN)$  (Ανάκληση)
  - **Precision** =  $TP/(TP+FP)$  (Ακρίβεια)
  - **F\_measure** =  $(2 * precision * recall) / (precision + recall)$
  - Για κάθε κλάση υπολογίζονται οι μετρικές και στη συνέχεια κάποιος M.O., ενδεχομένως με χρήση βαρών



# Μετρικές Αξιολόγησης-Προβλήματα Ταξινόμησης (3)

Αν έχουμε  $n$  Κλάσεις και το σύνολο ελέγχου έχει  $k$  στιγμιότυπα με  $k_j$  τα στιγμιότυπα της κλάσης  $j$  τότε ο μέσος όρος της Ανάκλησης:

$$\textit{Weight\_Average\_Recall} = \sum_{j=1}^n \frac{k_j}{k} \times \textit{Recall}_j$$

# Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στο πλαίσιο του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Αθηνών**» έχει χρηματοδοτήσει μόνο την αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



# Σημείωμα Ιστορικού Εκδόσεων Έργου

Το παρόν έργο αποτελεί την έκδοση 1.0.



# Σημείωμα Αναφοράς

Copyright Πανεπιστήμιο Πατρών, Ιωάννης Χατζηλυγερούδης 2015.  
«Ευφυής Προγραμματισμός». Έκδοση: 1.0. Πάτρα 2015. Διαθέσιμο από τη  
δικτυακή διεύθυνση:

<https://eclass.upatras.gr/courses/CEID1095/>



# Σημείωμα Αδειοδότησης

Το παρόν υλικό διατίθεται με τους όρους της άδειας χρήσης Creative Commons Αναφορά, Μη Εμπορική Χρήση Παρόμοια Διανομή 4.0 [1] ή μεταγενέστερη, Διεθνής Έκδοση. Εξαιρούνται τα αυτοτελή έργα τρίτων π.χ. φωτογραφίες, διαγράμματα κ.λ.π., τα οποία εμπεριέχονται σε αυτό και τα οποία αναφέρονται μαζί με τους όρους χρήσης τους στο «Σημείωμα Χρήσης Έργων Τρίτων».



[1] <http://creativecommons.org/licenses/by-nc-sa/4.0/>

Ως **Μη Εμπορική** ορίζεται η χρήση:

- που δεν περιλαμβάνει άμεσο ή έμμεσο οικονομικό όφελος από την χρήση του έργου, για το διανομέα του έργου και αδειοδόχο
- που δεν περιλαμβάνει οικονομική συναλλαγή ως προϋπόθεση για τη χρήση ή πρόσβαση στο έργο
- που δεν προσπορίζει στο διανομέα του έργου και αδειοδόχο έμμεσο οικονομικό όφελος (π.χ. διαφημίσεις) από την προβολή του έργου σε διαδικτυακό τόπο

Ο δικαιούχος μπορεί να παρέχει στον αδειοδόχο ξεχωριστή άδεια να χρησιμοποιεί το έργο για εμπορική χρήση, εφόσον αυτό του ζητηθεί.

# Διατήρηση Σημειωμάτων

Οποιαδήποτε αναπαραγωγή ή διασκευή του υλικού θα πρέπει να συμπεριλαμβάνει:

- το Σημείωμα Αναφοράς
- το Σημείωμα Αδειοδότησης
- τη δήλωση Διατήρησης Σημειωμάτων
- το Σημείωμα Χρήσης Έργων Τρίτων (εφόσον υπάρχει)

μαζί με τους συνοδευόμενους υπερσυνδέσμους.

