



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

Τεχνητή Νοημοσύνη II

Εργαστηριακή Άσκηση 6

Μουστάκας Κωνσταντίνος

Τμήμα Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών

ΑΝΟΙΚΤΑ ακαδημαϊκά **ΠΠ**
μαθήματα

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ
ΥΠΟΛΟΓΙΣΤΩΝ

ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ II

Ακαδημαϊκό Έτος 2014-15

6η Εργαστηριακή Άσκηση¹

Καταληκτική Ημερομηνία Υποβολής: 17/6/2015

Αλγόριθμοι Επιβλεπόμενης και Μη Επιβλεπόμενης Μάθησης

ΜΕΡΟΣ Α'

Φορτώστε στο WEKA το αρχείο zoo.arff. Το αρχείο περιέχει 18 χαρακτηριστικά (features) από 100 είδη ζώων. Το χαρακτηριστικό "animal" περιέχει το όνομα του ζώου. Το χαρακτηριστικό "type" περιέχει την κατηγορία του (θηλαστικό, πτηνό, κλπ, συνολικά 7 κατηγορίες).

Κατασκευάστε δένδρο αποφάσεων που θα αποφασίζει την κατηγορία ενός ζώου από τα υπόλοιπα χαρακτηριστικά του.

Στην έκθεσή σας:

1. Γράψτε ποιόν αλγόριθμο χρησιμοποιήσατε και με ποιές παραμέτρους τον τρέξατε. Αιτιολογήστε τις επιλογές σας.
2. Προσθέστε ένα screen capture του διαγράμματος του δένδρου όπως το έβγαλε το WEKA.
3. Γράψτε περιληπτικά, από τη στιγμή που έχετε το δένδρο, με ποιά διαδικασία αποφασίζετε σε ποιά κατηγορία ανήκει ένα νέο ζώο.
4. Γράψτε ποιά χαρακτηριστικά δεν χρειάζεται να λάβετε καθόλου υπόψη σας για την απόφασή σας. Στη συνέχεια αφαιρέστε τα από το σύνολο των δεδομένων σας και ξανατρέξτε τον αλγόριθμο. Έβγαλε το ίδιο δέντρο; Εξηγήστε γιατί.

ΜΕΡΟΣ Β'

Κατασκευάστε δίκτυο Bayes από τα δεδομένα του αρχείου zoo.arff με στόχο επίσης τον προσδιορισμό της κατηγορίας ενός ζώου από τα χαρακτηριστικά του. Στην έκθεσή σας:

1. Γράψτε ποιόν αλγόριθμο χρησιμοποιήσατε και με ποιές παραμέτρους τον τρέξατε. Αιτιολογήστε τις επιλογές σας.
2. Βάλτε ένα screen capture του διαγράμματος του δικτύου όπως το έβγαλε το WEKA. Προσθέστε τους πίνακες πιθανοτήτων σε κάθε κόμβο.
3. Γράψτε περιληπτικά, από τη στιγμή που έχετε το διάγραμμα και τους πίνακες πιθανοτήτων, με ποιά διαδικασία αποφασίζετε σε ποιά κατηγορία ανήκει ένα νέο ζώο.
4. Γράψτε ποιά χαρακτηριστικά δεν χρειάζεται να λάβετε καθόλου υπόψη σας για την

¹ Έκδοση άσκησης 1.02 (11/6/15), Κ.Σγάμπας. Αρχεία δεδομένων: R.S.Forsyth.

απόφασή σας. Στη συνέχεια αφαιρέστε τα από το σύνολο των δεδομένων σας και ξανατρέξτε τον αλγόριθμο. Έβγαλε το ίδιο διάγραμμα; Με τους ίδιους ακριβώς πίνακες πιθανοτήτων; Εξηγήστε γιατί.

ΜΕΡΟΣ Γ'

Τώρα αφαιρέστε το χαρακτηριστικό type από τα δεδομένα του αρχείου zoo.arff και προσπαθήστε να ομαδοποιήσετε τα ζώα σε 7 κατηγορίες χρησιμοποιώντας διαδοχικά τον αλγόριθμο K-Means και τον αλγόριθμο EM.

Στην έκθεσή σας:

1. Γράψτε κάθε φορά ποιόν αλγόριθμο χρησιμοποιήσατε και με ποιές παραμέτρους τον τρέξατε. Αιτιολογήστε τις επιλογές σας.
2. Γράψτε τα αποτελέσματα που έβγαλε το WEKA για κάθε αλγόριθμο.
3. Συγκρίνετε τις κατηγορίες που έβγαλε ο κάθε αλγόριθμος με τις σωστές (αυτές που αφαιρέσατε).
4. Γράψτε ποιός από τους δύο (K-Means ή EM) έδωσε καλύτερα αποτελέσματα.

ΜΕΡΟΣ Δ'

Ελέγξτε αν μπορείτε να βοηθήσετε τους αλγορίθμους K-Means και EM να κάνουν καλύτερη ομαδοποίηση αφαιρώντας μερικά χαρακτηριστικά. Δοκιμάστε διαδοχικά να αφαιρέσετε:

- (α) το σύνολο των χαρακτηριστικών που δεν χρειάστηκε το δέντρο αποφάσεων στο Μέρος Α' της άσκησης,
- (β) το σύνολο των χαρακτηριστικών που δεν χρειάστηκε το δίκτυο Bayes στο Μέρος Β' της άσκησης,
- (γ) την τομή των δύο συνόλων χαρακτηριστικών,
- (δ) την ένωση των δύο συνόλων χαρακτηριστικών.

Στην έκθεσή σας:

1. Γράψτε ποιιά χαρακτηριστικά αφαιρέσατε σε κάθε περίπτωση.
2. Γράψτε τί αποτελέσματα έβγαλε το WEKA σε κάθε περίπτωση.
3. Συγκρίνετε τα αποτελέσματα και μεταξύ τους και με αυτά του Μέρους Γ'.

Βιβλιογραφία

- [1] Stuart Russell, Peter Norvig, "Τεχνητή Νοημοσύνη - Μια Σύγχρονη Προσέγγιση", 2η Αμερικανική Έκδοση, Κλειδάριθμος, 2004.
- [2] Ιωάννης Βλαχάβας, Πέτρος Κεφαλάς, Νικόλαος Βασιλειάδης, Φώτης Κόκκορας, Ηλίας Σακελλαρίου, "Τεχνητή Νοημοσύνη" Γ' Έκδοση, Β. Γκιούρδας Εκδοτική, 2006.
- [3] Remco R. Bouckaert, Eibe Frank, Mark Hall, Richard Kirkby, Peter Reutemann, Alex Seewald, David Scuse, "WEKA Manual", 2013.
- [4] Ian H. Witten, Eibe Frank, Mark A. Hall, "Data Mining - Practical Machine Learning Tools and Techniques", 3rd Edition, Morgan Kaufmann/Elsevier, 2011.

Παραδοτέα

Για την άσκηση αυτή υποβάλετε στο eclass ένα αρχείο zip ή rar που θα περιέχει:

- την έκθεσή σας (κατά προτίμηση σε pdf format)
- τυχόν αρχεία .arff (μόνο στην περίπτωση που τροποποιήσατε το αρχικό).

Προθεσμία υποβολής: 17/6/2015.

ΚΑΛΗ ΕΠΙΤΥΧΙΑ

Σημειώματα

Σημείωμα Ιστορικού Εκδόσεων Έργου

Το παρόν έργο αποτελεί την έκδοση **1.0**

- Έκδοση **1.0** διαθέσιμη [εδώ](#).

Σημείωμα Αναφοράς

Copyright Πανεπιστήμιο Πατρων, Φακωτάκης Νικόλαος, Σγάρμπας Κυριάκος, Πέππας Παύλος, Μουστάκας Κωνσταντίνος. «Τεχνητή Νοημοσύνη II, Εργαστήριο 6». Έκδοση: 1.0. Πάτρα 2015.

Διαθέσιμο από τη δικτυακή διεύθυνση:

https://eclass.upatras.gr/modules/course_metadata/opencourses.php?fc=15

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.

Σημείωμα Χρήσης Έργων Τρίτων

Το Έργο αυτό κάνει χρήση των ακόλουθων έργων:

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 τη 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.

