



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Αλγοριθμική Επιστήμη Δεδομένων 2022-23

Διδάσκοντες: Α. Παγουρτζής, Θ. Σούλιου

1η Σειρά Ασκήσεων

Άσκηση 1. Να λύσετε τις ασκήσεις 6.3.1 και 6.4.1 του βιβλίου MMDS.

Άσκηση 2.

(α) Εξηγήστε την ορθότητα της μεθόδου A-priori για το πρόβλημα της εξόρυξης συχνών συνόλων στοιχείων (frequent itemset mining) από δεδομένα τύπου ‘καλαθιού αγορών’ (market basket data). Πόσες διασχίσεις (passes) πραγματοποιούνται στην βάση δεδομένων;

(β) Συζητήστε αν και γιατί η χρονική πολυπλοκότητα του αλγορίθμου A-priori είναι *πολυωνυμική ως προς την είσοδο ή πολυωνυμική ως προς την έξοδο*.

(γ) Συζητήστε τα ίδια ερωτήματα με το (β) παραπάνω σε σχέση με τον αλγόριθμο FP-Growth.

(δ) Εκτελέστε τους αλγορίθμους A-priori και FP-Growth στο παρακάτω παράδειγμα. Υποθέστε ότι το κατώφλι στήριξης είναι $s = 5$.

{a b c d}	{ b d}
{a d}	{a d}
{ b c d}	{a c }
{ b d}	{a b d}
{a c }	{ c d}
{ b c }	{a c d}
{a b c }	{ b c d}

(ε) Εκτελέστε επίσης τον αλγόριθμο του Τοϊνονεν στο ίδιο παράδειγμα, θεωρώντας ότι το δείγμα είναι οι 4 πρώτες εγγραφές (της αριστερής στήλης). Χρησιμοποιήστε ως κατώφλι στο δείγμα $s' = 2$ και $s'' = 1$ (θα κάνετε δύο διαφορετικές εκτελέσεις του αλγορίθμου). Τι παρατηρείτε;

Άσκηση 3. (*) Μελετήστε τις υλοποιήσεις αλγορίθμων εξόρυξης συχνών συνόλων στοιχείων που υπάρχουν στο αποθετήριο <http://fimi.uantwerpen.be/> και προσπαθήστε να επαναλάβετε ορισμένα από τα πειράματα που περιγράφονται εκεί (τουλάχιστον 2 αλγορίθμους με τουλάχιστον 3 dataset, της επιλογής σας). Περιγράψτε συνοπτικά τι κάνατε και παρουσιάστε με λεπτομέρειες τα αποτελέσματα που πήρατε.

Άσκηση 4. Να λύσετε τις ασκήσεις 5.2.2, 5.2.3, 5.2.4 και 5.3.1 του βιβλίου MMDS.

Άσκηση 5. Να λύσετε τις ασκήσεις 7.2.2, 7.2.3, 7.3.2 και 7.4.1 του βιβλίου MMDS.

Προθεσμία υποβολής και οδηγίες. Οι απαντήσεις θα πρέπει να υποβληθούν έως τις 28/5/2023, σε ηλεκτρονική μορφή. Τα ερωτήματα με (*) είναι προαιρετικά. Εφ'όσον τα επιλύσετε μπορούν να προσμετρηθούν στη θέση ερωτημάτων που δεν απαντήσατε. Για απορίες / διευκρινίσεις: στείλτε μήνυμα στη διεύθυνση ads@corelab.ntua.gr.