

ΠΡΟΤΕΙΝΟΜΕΝΑ ΘΕΜΑΤΑ ΔΙΠΛΩΜΑΤΙΚΩΝ ΕΡΓΑΣΙΩΝ ΓΙΑ ΤΟ ΕΑΡΙΝΟ ΕΞΑΜΗΝΟ 2018-2019

Εισηγητής: Νίκος Πλόσκας (nploskas@uowm.gr), Επίκουρος Καθηγητής ΤΜΠΤ

1. Τεχνικές μηχανικής μάθησης για τον εντοπισμό προβλημάτων σχεδίασης σε συστήματα λογισμικού (Machine learning methods for detecting bad smells in software systems)

Περιγραφή: Η κακή ποιότητα του κώδικα εμφανίζεται σε ορισμένα κομμάτια κώδικα που ονομάζονται «κακές οσμές» (bad smells). Η εύρεση και η διόρθωση των κακών οσμών βελτιώνει σημαντικά την ποιότητα του κώδικα. Ωστόσο, η εύρεση τέτοιων προβλημάτων σχεδίασης σε μεγάλα συστήματα λογισμικού είναι δύσκολο να επιτευχθεί χειροκίνητα. Έχουν αναπτυχθεί εργαλεία για την αυτόματη εύρεση κακών οσμών στον κώδικα. Ο στόχος της εργασίας αυτής είναι η χρησιμοποίηση τεχνικών μηχανικής μάθησης για την αυτοματοποιημένη εύρεση προβλημάτων σχεδίασης σε συστήματα λογισμικού.

Ενδεικτική βιβλιογραφία:

1. Fontana, F. A., Mäntylä, M. V., Zanoni, M., & Marino, A. (2016). Comparing and experimenting machine learning techniques for code smell detection. *Empirical Software Engineering*, 21(3), 1143-1191.
2. Khomh, F., Vaucher, S., Guéhéneuc, Y. G., & Sahraoui, H. (2009). A bayesian approach for the detection of code and design smells. In *2009 Ninth International Conference on Quality Software* (pp. 305-314). IEEE.
3. Maneerat, N., & Muenchaisri, P. (2011). Bad-smell prediction from software design model using machine learning techniques. In *2011 Eighth International Joint Conference on Computer Science and Software Engineering (JCSSE)* (pp. 331-336). IEEE.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Τεχνητή Νοημοσύνη, Τεχνολογία Λογισμικού, Προγραμματισμός, Στατιστική.

2. Χρονοπρογραμματισμός ωρολογίου προγράμματος μαθημάτων (Scheduling class timetables)

Περιγραφή: Ο χρονοπρογραμματισμός (scheduling) είναι ένα πρόβλημα που αντιμετωπίζεται συχνά σε εκπαιδευτικά ιδρύματα και αναφέρεται στην εύρεση ενός εβδομαδιαίου προγράμματος μαθημάτων. Ακόμη και σήμερα, η δημιουργία των ωρολογίων προγραμμάτων γίνεται πολλές φορές με παραδοσιακές μεθόδους (με το χέρι) και όχι αυτόματα μέσω κάποιου προγράμματος. Στόχος της εργασίας αυτής είναι η δημιουργία μιας εφαρμογής που να κατασκευάζει αυτόματα το πρόγραμμα μαθημάτων ενός τμήματος. Θα διερευνηθούν διάφοροι αλγόριθμοι χρονοπρογραμματισμού.

Ενδεικτική βιβλιογραφία:

1. Alvarez-Valdes, R., Crespo, E., & Tamarit, J. M. (2002). Design and implementation of a course scheduling system using tabu search. *European Journal of Operational Research*, 137(3), 512-523.

2. Loo, E. H., Goh, T. N., & Ong, H. L. (1986). A heuristic approach to scheduling university timetables. *Computers & Education*, 10(3), 379-388.
3. Assi, M., Halawi, B., & Haraty, R. A. (2018). Genetic Algorithm Analysis using the Graph Coloring Method for Solving the University Timetable Problem. *Procedia Computer Science*, 126, 899-906.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Τεχνητή Νοημοσύνη, Προγραμματισμός.

3. Χρονοπρογραμματισμός εργασιών σε περιβάλλοντα cloud computing (Job scheduling in cloud computing)

Περιγραφή: Ο χρονοπρογραμματισμός εργασιών (job scheduling) χρησιμοποιείται για να αποφασισθεί ποια από τις διεργασίες που είναι έτοιμες για εκτέλεση θα δεσμεύσει τους πόρους του συστήματος. Μερικές από αυτές τις διεργασίες μπορούν να απαιτούν πολλούς και διαφορετικούς πόρους του συστήματος. Στόχος της εργασίας είναι η μελέτη αλγορίθμων χρονοπρογραμματισμού εργασιών που θα αποφέρουν βελτίωση στον χρόνο εκτέλεσης των διεργασιών.

Ενδεικτική βιβλιογραφία:

1. Singh, S., & Chana, I. (2016). A survey on resource scheduling in cloud computing: Issues and challenges. *Journal of Grid Computing*, 14(2), 217-264.
2. Arunarani, A. R., Manjula, D., & Sugumaran, V. (2019). Task scheduling techniques in cloud computing: A literature survey. *Future Generation Computer Systems*, 91, 407-415.
3. Agarwal, M., & Srivastava, G. M. S. (2018). A cuckoo search algorithm-based task scheduling in cloud computing. In *Advances in Computer and Computational Sciences* (pp. 293-299). Springer, Singapore.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Τεχνητή Νοημοσύνη, Προγραμματισμός.

4. Αυτοματοποιημένη εύρεση βέλτιστων παραμέτρων λογισμικού (Auto-tuning software parameters)

Περιγραφή: Τα συστήματα λογισμικού για επιστημονικούς υπολογισμούς διαθέτουν διάφορες παραμέτρους για την ταχύτερη εκτέλεσή τους ή/και για την εύρεση καλύτερων λύσεων. Οι δημιουργοί αυτών των λογισμικών έχουν διεξάγει πολλά πειράματα για την εύρεση κατάλληλων προκαθορισμένων τιμών για αυτές τις παραμέτρους. Ωστόσο, τα συστήματα λογισμικού δεν έχουν πάντα την επιθυμητή απόδοση σε ορισμένες κλάσεις προβλημάτων. Στόχος της εργασίας αυτής είναι η υλοποίηση και υπολογιστική σύγκριση αλγορίθμων για την εύρεση βέλτιστων παραμέτρων λογισμικού.

Ενδεικτική βιβλιογραφία:

1. Arcuri, A., & Fraser, G. (2013). Parameter tuning or default values? An empirical investigation in search-based software engineering. *Empirical Software Engineering*, 18(3), 594-623.
2. Elkhodary, A., Esfahani, N., & Malek, S. (2010). FUSION: a framework for engineering self-tuning self-adaptive software systems. In *Proceedings of the eighteenth ACM SIGSOFT international symposium on Foundations of software engineering* (pp. 7-16). ACM.

3. Katagiri, T., Kise, K., Honda, H., & Yuba, T. (2003). Fiber: A generalized framework for auto-tuning software. In *International Symposium on High Performance Computing* (pp. 146-159). Springer, Berlin, Heidelberg.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Προγραμματισμός.

5. Υλοποίηση και υπολογιστική σύγκριση αλγορίθμων για προβλήματα χωροθέτησης μονάδων (Development and computational comparison of algorithms for the solution of the facility location problem)

Περιγραφή: Η επιλογή των καταλληλότερων χώρων εγκατάστασης μονάδων είναι ένα ιδιαίτερα σύνθετο πρόβλημα μιας και εξαρτάται από μια σειρά παραγόντων και από πολλαπλά κριτήρια. Στόχος της εργασίας είναι η μελέτη αλγορίθμων για προβλήματα χωροθέτησης μονάδων και η υπολογιστική τους σύγκριση. Στην μελέτη των αλγορίθμων θα ληφθούν υπόψιν διάφορα κριτήρια επιλογής των χώρων εγκατάστασης, όπως η μέγιστη κάλυψη και η μείωση του κόστους.

Ενδεικτική βιβλιογραφία:

1. Chudak, F. A., & Shmoys, D. B. (2003). Improved approximation algorithms for the uncapacitated facility location problem. *SIAM Journal on Computing*, 33(1), 1-25.
2. Ghosh, D. (2003). Neighborhood search heuristics for the uncapacitated facility location problem. *European Journal of Operational Research*, 150(1), 150-162.
3. Al-Sultan, K. S., & Al-Fawzan, M. A. (1999). A tabu search approach to the uncapacitated facility location problem. *Annals of Operations Research*, 86, 91-103.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Προγραμματισμός.

6. Υλοποίηση και υπολογιστική σύγκριση αλγορίθμων για προβλήματα ανάθεσης πόρων (Development and computational comparison of algorithms for the solution of the assignment problem)

Περιγραφή: Το πρόβλημα ανάθεσης πόρων αφορά στην κατανομή m πόρων σε n δραστηριότητες, έτσι ώστε να ελαχιστοποιείται το κόστος ή να μεγιστοποιείται το όφελος με περιορισμούς (i) ένας πόρος να χρησιμοποιηθεί σε μια μόνο δραστηριότητα και (ii) σε κάθε δραστηριότητα να χρησιμοποιείται ένας μόνο πόρος. Στόχος της εργασίας είναι η μελέτη και η υπολογιστική σύγκριση αλγορίθμων για την επίλυση του προβλήματος ανάθεσης πόρων. Στην μελέτη των αλγορίθμων θα ληφθούν υπόψιν διάφορα κριτήρια επιλογής των αναθέσεων, όπως η ελαχιστοποίηση του κόστους ή η μεγιστοποίηση του οφέλους.

Ενδεικτική βιβλιογραφία:

1. Chu, P. C., & Beasley, J. E. (1997). A genetic algorithm for the generalised assignment problem. *Computers & Operations Research*, 24(1), 17-23.
2. Shmoys, D. B., & Tardos, É. (1993). An approximation algorithm for the generalized assignment problem. *Mathematical Programming*, 62(1-3), 461-474.
3. Özbakir, L., Baykasoğlu, A., & Tapkan, P. (2010). Bees algorithm for generalized assignment problem. *Applied Mathematics and Computation*, 215(11), 3782-3795.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Προγραμματισμός.

7. Ανάπτυξη συστήματος συστάσεων για το ταίριασμα κριτών και επιστημονικών άρθρων (Recommender system for matching reviewers with scientific papers)

Περιγραφή: Ένα συχνό πρόβλημα κατά την διαδικασία αξιολόγησης ενός επιστημονικού άρθρου είναι η εύρεση των κατάλληλων κριτών. Η επιλογή των κριτών συνήθως γίνεται χειροκίνητα από τους εκδότες του επιστημονικού περιοδικού ή συνεδρίου. Στόχος της εργασίας είναι η μελέτη και η υπολογιστική σύγκριση αλγορίθμων για το ταίριασμα κριτών και επιστημονικών άρθρων, δηλαδή την εύρεση των κατάλληλων κριτών για κάποιο άρθρο. Θα αναπτυχθεί ένα σύστημα συστάσεων που δοθέντος ενός άρθρου θα βρίσκει αυτοματοποιημένα κριτές από το Google Scholar και θα τους κατατάσσει ως προς την συνάφειά τους.

Ενδεικτική βιβλιογραφία:

1. Protasiewicz, J., Pedrycz, W., Kozłowski, M., Dadas, S., Stanisławek, T., Kopacz, A., & Gałęzewska, M. (2016). A recommender system of reviewers and experts in reviewing problems. *Knowledge-Based Systems, 106*, 164-178.
2. Liu, X., Suel, T., & Memon, N. (2014). A robust model for paper reviewer assignment. In *Proceedings of the 8th ACM Conference on Recommender systems* (pp. 25-32). ACM.
3. Wang, F., Chen, B., & Miao, Z. (2008). A survey on reviewer assignment problem. In *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems* (pp. 718-727). Springer, Berlin, Heidelberg.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Τεχνητή Νοημοσύνη, Προγραμματισμός, Στατιστική.

8. Σύγκριση τεχνικών προεπεξεργασίας δεδομένων πριν την εκτέλεση αλγορίθμων μηχανικής μάθησης (Comparison of data preprocessing techniques prior to the execution of machine learning techniques)

Περιγραφή: Ένα συχνό πρόβλημα κατά την εφαρμογή μιας μεθόδου μηχανικής μάθησης είναι η προεπεξεργασία των δεδομένων. Υπάρχουν διάφορες μέθοδοι για την προεπεξεργασία των δεδομένων και η κατάλληλη χρήση τους έχει σημαντικό αντίκτυπο στην ακρίβεια των εκτιμήσεων των τεχνικών μηχανικής μάθησης. Στόχος της εργασίας είναι η σύγκριση τεχνικών προεπεξεργασίας δεδομένων πριν την εκτέλεση αλγορίθμων μηχανικής μάθησης.

Ενδεικτική βιβλιογραφία:

1. Huang, J., Li, Y. F., & Xie, M. (2015). An empirical analysis of data preprocessing for machine learning-based software cost estimation. *Information and Software Technology, 67*, 108-127.
2. Uysal, A. K., & Gunal, S. (2014). The impact of preprocessing on text classification. *Information Processing & Management, 50*(1), 104-112.
3. Srividhya, V., & Anitha, R. (2010). Evaluating preprocessing techniques in text categorization. *International journal of computer science and application, 47*(11), 49-51.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Τεχνητή Νοημοσύνη, Προγραμματισμός, Στατιστική.