

ΠΡΟΤΕΙΝΟΜΕΝΑ ΘΕΜΑΤΑ ΔΙΠΛΩΜΑΤΙΚΩΝ ΕΡΓΑΣΙΩΝ ΓΙΑ ΤΟ ΕΑΡΙΝΟ ΕΞΑΜΗΝΟ 2019-2020

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

1. 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, Αριθμητική Ανάλυση.

2. 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.

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

Περιγραφή: Ένα συχνό πρόβλημα κατά τη 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., & Gałężewska, 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.

4. 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, Αριθμητική Ανάλυση, Στατιστική.

5. Σύγκριση μεθόδων μείωσης διαστάσεων των δεδομένων (Comparison of dimension reduction techniques)

Περιγραφή: Πολλές εφαρμογές χρησιμοποιούν διάφορους αισθητήρες για την συλλογή δεδομένων. Τα δεδομένα που συλλέγονται είναι μεγάλα και ανομοιόμορφα, οπότε είναι δύσκολη η επεξεργασία τους και η λήψη αποφάσεων με αυτά. Μια τεχνική που χρησιμοποιείται πριν εφαρμοστεί κάποιος αλγόριθμος μηχανικής μάθησης είναι η μείωση των διαστάσεων των δεδομένων (dimension reduction). Στόχος της εργασίας είναι η σύγκριση τεχνικών μείωσης των διαστάσεων δεδομένων πριν την εκτέλεση αλγορίθμων μηχανικής μάθησης.

Ενδεικτική βιβλιογραφία:

1. Papageorgiou, A., Cheng, B., & Kovacs, E. (2015). Real-time data reduction at the network edge of Internet-of-Things systems. In: *11th International Conference on Network and Service Management (CNSM)*, 284-291.
2. Li, L. (2010). Dimension Reduction for High-Dimensional Data. In: Bang, H., Zhou, X., van Epps, H., & Mazumdar, M. (eds.) *Statistical Methods in Molecular Biology. Methods in Molecular Biology (Methods and Protocols)*, 620, Humana Press, Totowa, NJ.
3. Van Der Maaten, L., Postma, E., & Van den Herik, J. (2009). Dimensionality reduction: a comparative. *Journal of Machine Learning Research*, 10(66-71), 13.

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

6. Σύγκριση μεθόδων χειρισμού ελλιπών δεδομένων στη μηχανική μάθηση (Comparison of methods for handling missing data in machine learning)

Περιγραφή: Τα δεδομένα από πραγματικές εφαρμογές είναι σπάνια πλήρη και ομογενοποιημένα. Συνήθως περιέχουν θόρυβο και είναι ημιτελή και ασυνεπή. Είναι σημαντικό να γίνει προεπεξεργασία των δεδομένων έτσι ώστε η χρήση μεθόδων μηχανικής μάθησης να οδηγήσει σε καλά αποτελέσματα. Στόχος της εργασίας είναι η σύγκριση μεθόδων χειρισμού ελλιπών δεδομένων στη μηχανική μάθηση.

Ενδεικτική βιβλιογραφία:

1. Farhangfar, A., Kurgan, L., & Dy, J. (2008). Impact of imputation of missing values on classification error for discrete data. *Pattern Recognition*, 41(12), 3692-3705.
2. Zhang, S., Zhang, J., Zhu, X., Qin, Y., & Zhang, C. (2008). Missing value imputation based on data clustering. In: Gavrilova, M. L., & Tan, C. J. K. (eds.) *Transactions on Computational Science I. Lecture Notes in Computer Science*, 4750. Springer, Berlin, Heidelberg.

3. Kotsiantis, S. B., Kanellopoulos, D., & Pintelas, P. E. (2006). Data preprocessing for supervised learning. *International Journal of Computer Science*, 1(2), 111-117.

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

7. Αλγόριθμοι μηχανικής μάθησης για δημιουργία λεζάντας σε εικόνες (Machine learning algorithms for labelling images)

Περιγραφή: Οι αλγόριθμοι βαθιάς μάθησης (deep learning) έχουν χρησιμοποιηθεί σε ένα μεγάλο εύρος εφαρμογών. Ένας από αυτούς είναι η εξαγωγή πληροφοριών από εικόνες. Στόχος της εργασίας είναι να χρησιμοποιηθούν αλγόριθμοι βαθιάς μάθησης για την αυτόματη δημιουργία λεζάντας σε εικόνες.

Ενδεικτική βιβλιογραφία:

1. Maggiori, E., Tarabalka, Y., Charpiat, G., & Alliez, P. (2017). Can semantic labeling methods generalize to any city? the inria aerial image labeling benchmark. In: *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 3226-3229.
2. Paisitkriangkrai, S., Sherrah, J., Janney, P., & Van Den Hengel, A. (2016). Semantic labeling of aerial and satellite imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(7), 2868-2881.
3. Mnih, V., & Hinton, G. E. (2012). Learning to label aerial images from noisy data. In: *Proceedings of the 29th International conference on machine learning (ICML-12)*, 567-574.

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

8. Σύγκριση αλγορίθμων δημιουργίας υποκατάστατων μοντέλων (Comparison of surrogate-based methods)

Περιγραφή: Σε αρκετά προβλήματα βελτιστοποίησης δεν υπάρχει η αναλυτική μορφή της αντικειμενικής συνάρτησης, οπότε χειριζόμαστε το πρόβλημα ως ένα πρόβλημα μαύρου κουτιού (black box optimization). Υπάρχουν διάφοροι αλγόριθμοι που δημιουργούν υποκατάστατα μοντέλα (surrogate models) για να προσεγγίσουν τις αντικειμενικές συναρτήσεις σε αυτά τα προβλήματα. Στόχος της εργασίας είναι η σύγκριση αλγορίθμων δημιουργίας υποκατάστατων μοντέλων.

Ενδεικτική βιβλιογραφία:

1. Owen, N. E., Challenor, P., Menon, P. P., & Bennani, S. (2017). Comparison of surrogate-based uncertainty quantification methods for computationally expensive simulators. *SIAM/ASA Journal on Uncertainty Quantification*, 5(1), 403-435.
2. Qian, Z., Seepersad, C. C., Joseph, V. R., Allen, J. K., & Wu, C. J. (2006). Building surrogate models based on detailed and approximate simulations. *Journal of Mechanical Design*, 128(4), 668-677.
3. Tabatabaei, M., Hakanen, J., Hartikainen, M., Miettinen, K., & Sindhya, K. (2015). A survey on handling computationally expensive multiobjective optimization problems using surrogates: non-nature inspired methods. *Structural and Multidisciplinary Optimization*, 52(1), 1-25.

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

9. Προσδιορισμός βέλτιστης θέσης αναμεταδοτών σε ασύρματα δίκτυα (Optimal placement of relays in wireless networks)

Περιγραφή: Για τη δημιουργία ασύρματων δικτύων απαιτείται η εύρεση της θέσης των αναμεταδοτών. Οι στόχοι για την τοποθέτησή τους μπορεί να είναι η κάλυψη των περισσότερων χρηστών, η μείωση του κόστους, κ.α. Στόχος της εργασίας είναι η υλοποίηση αλγορίθμων για τον προσδιορισμό της βέλτιστης θέσης αναμεταδοτών σε ασύρματα δίκτυα.

Ενδεικτική βιβλιογραφία:

1. Hashim, H. A., Ayinde, B. O., & Abido, M. A. (2016). Optimal placement of relay nodes in wireless sensor network using artificial bee colony algorithm. *Journal of Network and Computer Applications*, 64, 239-248.
2. Lin, B., Ho, P. H., Xie, L. L., Shen, X., & Tapolcai, J. (2009). Optimal relay station placement in broadband wireless access networks. *IEEE Transactions on Mobile Computing*, 9(2), 259-269.
3. So, A., & Liang, B. (2009). Optimal placement and channel assignment of relay stations in heterogeneous wireless mesh networks by modified bender's decomposition. *Ad Hoc Networks*, 7(1), 118-135.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Προγραμματισμός.

10. Υλοποίηση και υπολογιστική σύγκριση αλγορίθμων για προβλήματα βέλτιστης κοπής (Development and computational comparison of algorithms for the solution of the cutting stock problem)

Περιγραφή: Το πρόβλημα της βέλτιστης κοπής αφορά την κοπή ενός υλικού σε κομμάτια σταθερού μεγέθους, έτσι ώστε να ελαχιστοποιείται η φύρα (σε άλλες παραλλαγές του προβλήματος υπάρχουν και άλλοι στόχοι). Στόχος της εργασίας είναι η μελέτη και η υπολογιστική σύγκριση αλγορίθμων για την επίλυση του προβλήματος βέλτιστης κοπής. Στη μελέτη των αλγορίθμων θα ληφθούν υπόψιν διάφοροι στόχοι.

Ενδεικτική βιβλιογραφία:

1. Muter, İ., & Sezer, Z. (2018). Algorithms for the one-dimensional two-stage cutting stock problem. *European Journal of Operational Research*, 271(1), 20-32.
2. Delorme, M., Iori, M., & Martello, S. (2016). Bin packing and cutting stock problems: Mathematical models and exact algorithms. *European Journal of Operational Research*, 255(1), 1-20.
3. Delorme, M., Iori, M., & Martello, S. (2018). BPPLIB: a library for bin packing and cutting stock problems. *Optimization Letters*, 12(2), 235-250.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Προγραμματισμός.

11. Υλοποίηση αλγορίθμων γραμμικού προγραμματισμού (Implementation of linear programming algorithms)

Περιγραφή: Τα προβλήματα γραμμικού προγραμματισμού εμφανίζονται σε διάφορες εφαρμογές και η βελτίωση των ήδη υπαρχόντων αλγορίθμων είναι μεγάλης σημασίας. Στόχος της εργασίας είναι η μελέτη και η υπολογιστική σύγκριση αλγορίθμων γραμμικού προγραμματισμού τύπου simplex. Θα υλοποιηθούν και θα συγκριθούν διάφορες μέθοδοι σε κάθε βήμα των αλγορίθμων τύπου simplex.

Ενδεικτική βιβλιογραφία:

1. Dantzig, G. B. (1998). *Linear programming and extensions* (Vol. 48). Princeton University Press.
2. Vanderbei, R. J. (2015). *Linear programming*. Heidelberg: Springer.
3. Ploskas, N., & Samaras, N. (2017). *Linear programming using MATLAB®* (Vol. 127). Switzerland: Springer.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Προγραμματισμός, Επιχειρησιακή Έρευνα, Αριθμητική Ανάλυση.

12. Υλοποίηση αλγορίθμων δρομολόγησης σε περιβάλλοντα έξυπνων πόλεων (Implementation of algorithms for route planning in smart cities)

Περιγραφή: Το πρόβλημα δρομολόγησης οχημάτων είναι ένα πολύ σημαντικό πρόβλημα για το οποίο έχουν αναπτυχθεί πολλοί αλγόριθμοι. Σε περιβάλλοντα έξυπνων πόλεων χρησιμοποιούνται αισθητήρες για τη μετάδοση πληροφοριών με απώτερο σκοπό τη λήψη πληροφοριών που μπορούν να χρησιμοποιηθούν για τη βελτίωση της ποιότητας ζωής. Ένα τέτοιο παράδειγμα είναι η δυναμική δρομολόγηση οχημάτων με βάσει δεδομένα που συλλέγονται από αισθητήρες. Στόχος της εργασίας είναι η μελέτη και η υπολογιστική σύγκριση αλγορίθμων δρομολόγησης σε περιβάλλοντα έξυπνων πόλεων. Θα υλοποιηθούν και θα συγκριθούν διάφοροι αλγόριθμοι για την επίλυση αυτού του προβλήματος.

Ενδεικτική βιβλιογραφία:

1. Nha, V. T. N., Djahel, S., & Murphy, J. (2012). A comparative study of vehicles' routing algorithms for route planning in smart cities. In *IEEE 2012 First International Workshop on Vehicular Traffic Management for Smart Cities (VTM)*, 1-6.
2. Anagnostopoulos, T., Zaslavsky, A., & Medvedev, A. (2015). Robust waste collection exploiting cost efficiency of IoT potentiality in Smart Cities. In *IEEE 2015 International conference on recent advances in Internet of Things (RIoT)*, 1-6.
3. Medvedev, A., Fedchenkov, P., Zaslavsky, A., Anagnostopoulos, T., & Khoruzhnikov, S. (2015). Waste management as an IoT-enabled service in smart cities. In *Internet of Things, Smart Spaces, and Next Generation Networks and Systems*, Springer, 104-115.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Προγραμματισμός.