

Μηχανική μάθηση - Εξόρυξη δεδομένων

1. Μορφές επιρροής στη μηχανική μάθηση, τεχνικές εντοπισμού και αποφυγής

Επιβλέποντες: Βαρλάμης, Δίου, Μιχαήλ

Στη σημερινή εποχή όλο και περισσότερα συστήματα βασίζονται τη λειτουργία τους σε αλγόριθμους μηχανικής μάθησης. Η χρήση τέτοιων συστημάτων σε πολύ ευαίσθητους τομείς όπως λόγου χάρη της υγείας κάνει επιτακτική την ανάγκη τα συστήματα αυτά να είναι απαλλαγμένα από προκαταλήψεις κατά ορισμένων ομάδων. Σκοπός της εργασίας είναι να εξεταστεί η ύπαρξη bias στους αλγόριθμους μηχανικής μάθησης και να προταθούν λύσεις για την εξάλειψή του.

Ενδεικτική βιβλιογραφία/πηγές

- AI Fairness της IBM: <https://aif360.mybluemix.net/>
- FairML project: <https://github.com/adebayoj/fairml>
- Fairness tutorial: https://dssg.github.io/fairness_tutorial/
- Survey on bias in AI: <https://arxiv.org/abs/1908.09635>

2. Ανίχνευση εξαιρέσεων σε δεδομένα χρονοσειρών με χρήση ensemble τεχνικών

Επιβλέποντες: Βαρλάμης, Βιόλος, Δίου

Η ανίχνευση εξαιρέσεων στη συμπεριφορά μιας μέτρησης με το χρόνο μπορεί να έχει πολλαπλά επιχειρηματικά και άλλα οφέλη αν γίνει έγκαιρα και σωστά. Η ανίχνευση εξαιρέσεων ή ανωμαλιών σε δεδομένα χρονοσειρών βασίζεται αρχικά στον εντοπισμό της αναμενόμενης συμπεριφοράς με βάση κάποιο προγνωστικό μοντέλο και στη συνέχεια στην αξιολόγηση μεμονωμένων αποκλίσεων από αυτήν. Στόχος της εργασίας είναι να αναπτύξει ένα αλγόριθμο που θα εντοπίζει ανωμαλίες σε δεδομένα πολλαπλών χρονοσειρών σε δύο βήματα:

- Αρχικά θα αναλύει κάθε χρονοσειρά και θα επιλέγει το καλύτερο (ή τα καλύτερα) μαθηματικά μοντέλα που μπορούν να χρησιμοποιηθούν για την περιγραφή της κανονικής συμπεριφοράς τους.
- Στη συνέχεια θα καθορίζει την αναμενόμενη τιμή στα δεδομένα αυτά σε επόμενες χρονικές στιγμές, καθώς και το ρυθμό δειγματοληψίας
- Τέλος θα αξιολογεί τις ερχόμενες τιμές στην εκάστοτε χρονοσειρά και θα εκτιμά αν πρόκειται για μη αναμενόμενες τιμές.

Η χρήση πολλαπλών μοντέλων (ensemble) θα συγκριθεί σε σχέση με την αντίστοιχη χρήση μεμονωμένων μοντέλων.

Ενδεικτική βιβλιογραφία/πηγές:

- Braei, M., & Wagner, S. (2020). Anomaly detection in univariate time-series: A survey on the state-of-the-art. *arXiv preprint arXiv:2004.00433*.
- Cook, A. A., Misirlı, G., & Fan, Z. (2019). Anomaly detection for IoT time-series data: A survey. *IEEE Internet of Things Journal*, 7(7), 6481-6494.

- I. Tinawi. Machine Learning for Time Series Anomaly Detection
<https://dspace.mit.edu/bitstream/handle/1721.1/123129/1128282917-MIT.pdf?sequence=1&isAllowed=y>
- <https://github.com/rob-med/awesome-TS-anomaly-detection>

3. Κατανεμημένη Μηχανική Μάθηση και Μεταφορά Γνώσης και αξιολόγηση “Distributed ML and Knowledge Transfer and evaluation”

Επιβλέποντες: Βαρλάμης, Τσερπές, Δίου

Στόχος: Ξεκινώντας από ένα κεντρικό μοντέλο ML που υποθέτουμε ότι διαμοιράζεται σε πολλούς χρήστες που συνεχίζουν να το εκπαιδεύουν με δικά τους δεδομένα, παράγουμε αρκετά Delta models. Η εργασία θα μελετήσει πως μπορούμε να επικοινωνήσουμε αποδοτικά τα delta models μεταξύ των χρηστών και να τα συνδυάζουμε σε ένα νέο ενιαίο μοντέλο.

Βήματα:

- Εκπαίδευση ενός classification model (**shared model**) σε κλάσεις **A** και **B** με δεδομένα από πολλούς χρήστες
- Επανεκπαίδευση του μοντέλου στο χρήστη user 1 (**model1**) και ανίχνευση αλλαγών (**delta1**) με χρήση instances των κλάσεων **A, B**, και μιας νέας κλάσης **C**
- Το ίδιο για τον user 2 (**model2**) με αλλαγές (**delta2**) και εκπαίδευση σε κλάσεις **A, B**, και μια νέα **D**
- Συγκώνευση των delta1 και model2 και αξιολόγηση του νέου μοντέλου model2_1 vs model2 σε όλες τις κλάσεις (A,B,C,D)
- Περισσότεροι συνδυασμοί.

Δεδομένα:

<https://archive.ics.uci.edu/ml/datasets/WISDM+Smartphone+and+Smartwatch+Activity+and+Biometrics+Dataset+>

Περιέχει δεδομένα από επιταχυνσιόμετρα και γυροσκόπια χρονολογικών σειρών που συλλέγονται από ένα smartphone και ένα έξυπνο ρολόι, σε 51 άτομα δοκιμής που εκτελούν 18 δραστηριότητες για 3 λεπτά το καθένα.

Ενδεικτική βιβλιογραφία:

- Li, Q., He, B., & Song, D. (2020). Model-Agnostic Round-Optimal Federated Learning via Knowledge Transfer. *arXiv preprint arXiv:2010.01017*.
- Guha, N., Talwalkar, A., & Smith, V. (2019). One-shot federated learning. *arXiv preprint arXiv:1902.11175*.
- Li, Q., He, B., & Song, D. (2020). Model-Agnostic Round-Optimal Federated Learning via Knowledge Transfer. *arXiv preprint arXiv:2010.01017*.
- Shoham, N., Avidor, T., Keren, A., Israel, N., Benditkis, D., Mor-Yosef, L., & Zeitak, I. (2019). Overcoming Forgetting in Federated Learning on Non-IID Data. *arXiv preprint arXiv:1910.07796*.

4. «Αξιολόγηση πολυπλοκότητας επιχειρηματικών διαδικασιών»

Επιβλέποντες: Βαρλάμης, Μιχαήλ, Τσαδήμας

Η διαχείριση επιχειρηματικών διαδικασιών (Business Process Modelling) θεωρείται ως ένα σύνολο μεθόδων που στοχεύουν στην υποστήριξη του σχεδιασμού, της ανάλυσης και της βελτιστοποίησης των διαδικασιών. Βασική πτυχή του BPM είναι η χρήση τεχνικών μοντελοποίησης για την αναπαράσταση επιχειρηματικών διαδικασιών με τρόπο που διευκολύνει την εφαρμογή πρωτοβουλιών επανασχεδιασμού. Στόχος της εργασίας είναι να εντοπίσει στη βιβλιογραφία μετρικές για την αξιολόγηση της πολυπλοκότητας επιχειρηματικών διαδικασιών. Στη συνέχεια θα προτείνει ένα κατάλληλα προσαρμοσμένο σύνολο μετρικών για την αξιολόγηση ενός συνόλου 1000 περίπου διαδικασιών δημοσίων φορέων που έχουν καταγραφεί με συστηματικό τρόπο και θα αναπτύξει ένα λογισμικό που θα μπορεί να αξιολογήσει μαζικά το σύνολο των διαδικασιών και να εντοπίσει διαδικασίες υψηλής πολυπλοκότητας.

Ενδεικτική βιβλιογραφία:

- Latva-Koivisto, A. M. (2001). Finding a complexity measure for business process models. Technical report, Helsinki University of Technology.
- Sánchez-González, L., Ruiz, F., García, F., & Cardoso, J. (2011, March). Towards thresholds of control flow complexity measures for BPMN models. In *Proceedings of the 2011 ACM symposium on Applied computing* (pp. 1445-1450).
- Fotoglou, C., Tsakalidis, G., Vergidis, K., & Chatzigeorgiou, A. (2020, May). Complexity Clustering of BPMN Models: Initial Experiments with the K-means Algorithm. In *International Conference on Decision Support System Technology* (pp. 57-69). Springer, Cham.

Διαχείριση οχημάτων σε περιβάλλον προσομοίωσης

5. «Ανίχνευση και αποφυγή εμποδίων από δεδομένα αισθητήρων για μη επανδρωμένα εναέρια οχήματα UAVs»

Επιβλέποντες: Βαρλάμης, Δημητράκοπουλος, Μιχαήλ

Τα μη επανδρωμένα εναέρια οχήματα (UAVs) είναι οχήματα που μπορούν να εκτελέσουν επιθυμητές αποστολές χωρίς συνεχή ανθρώπινη καθοδήγηση. Η εργασία θα εγκαταστήσει την εφαρμογή προσομοίωσης πτήσης AirSim (<https://github.com/microsoft/AirSim>) και θα αναπτύξει δοκιμαστικά σενάρια πλοήγησης μη επανδρωμένου UAV με ύπαρξη εμποδίων. Στόχος της εργασίας είναι να αξιοποιήσει τους αισθητήρες της πλατφόρμας (<https://microsoft.github.io/AirSim/sensors/>) και να αναπαράγει σενάρια με σταθερά ή κινούμενα εμπόδια. Θα πρέπει παράλληλα να αναπτύξει τεχνικές κατηγοριοποίησης των μοντέλων σε κινούμενα ή σταθερά και να αξιοποιήσει ή να αναπτύξει βασικά πλάνα πλοήγησης που θα αποφεύγουν τα εμπόδια.

Ενδεικτική βιβλιογραφία:

- Vemprala, S., & Saripalli, S. (2018, May). Vision based collaborative path planning for micro aerial vehicles. In *2018 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 1-7). IEEE.
- Shah, S., Dey, D., Lovett, C., & Kapoor, A. (2018). Airsim: High-fidelity visual and physical simulation for autonomous vehicles. In *Field and service robotics* (pp. 621-635). Springer, Cham.
- Julian, K., Mern, J., & Tompa, R. (2017). Uav depth perception from visual images using a deep convolutional neural network. In *Tech. Rep.*.
- Liu, Y. C., Tian, J., Ma, C. Y., Glaser, N., Kuo, C. W., & Kira, Z. (2020). Who2com: Collaborative perception via learnable handshake communication. *arXiv preprint arXiv:2003.09575*.

6. “«Ανίχνευση και αποφυγή εμποδίων από δεδομένα αισθητήρων για οχήματα σε περιβάλλον προσομοίωσης»

Επιβλέποντες: Βαρλάμης, Τσερπές, Μιχαήλ

Η εργασία θα εγκαταστήσει το περιβάλλον προσομοίωσης οδήγησης CARLA (<https://carla.org/>) και θα αναπτύξει δοκιμαστικά σενάρια πλοήγησης με ύπαρξη εμποδίων. Στόχος της εργασίας είναι να αξιοποιήσει τους αισθητήρες της πλατφόρμας και να αναπαράγει σενάρια με σταθερά ή κινούμενα εμπόδια. Θα πρέπει παράλληλα να αναπτύξει τεχνικές κατηγοριοποίησης των μοντέλων σε κινούμενα ή σταθερά και να αξιοποιήσει ή να αναπτύξει βασικά πλάνα πλοήγησης που θα αποφεύγουν τα εμπόδια.

Ενδεικτική βιβλιογραφία:

- Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A., & Koltun, V. (2017). CARLA: An open urban driving simulator. *arXiv preprint arXiv:1711.03938*.
- Dharmawan, W., & Nambo, H. (2019, December). End-to-End Xception Model Implementation on Carla Self Driving Car in Moderate Dense Environment. In *Proceedings of the 2019 2nd Artificial Intelligence and Cloud Computing Conference* (pp. 139-143).
- Zapridou, E., Bartocci, E., & Katsaros, P. (2020, October). Runtime Verification of Autonomous Driving Systems in CARLA. In *International Conference on Runtime Verification* (pp. 172-183). Springer, Cham.