

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ**

**ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ**

**ΠΑΡΟΥΣΙΑΣΗ / ΕΞΕΤΑΣΗ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ**

**Ξανθόπουλος Ιορδάνης  
Μεταπτυχιακός Φοιτητής**

**Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης**

**Επόπτης Μεταπτυχιακής Εργασίας: Καθηγητής, Ι. Τσαμαρδίνος**

**Πέμπτη 3 Σεπτεμβρίου 2020 ,ώρα 18:00 μ.μ.**

**Τηλεδιάσκεψη ( μέσω του συστήματος e:Presence), Τμήμα Επιστήμης Υπολογιστών,  
Πανεπιστήμιο Κρήτης**

Διεύθυνση μετάδοσης (url): <http://video.ucnet.uoc.gr/live/show/303>

Κανάλι YouTube του Τμήματος

[https://www.youtube.com/channel/UC7uE3QiMTQjkrpByB\\_Gnt6Q/live](https://www.youtube.com/channel/UC7uE3QiMTQjkrpByB_Gnt6Q/live)

**“Μια Ποιοτική, ποσοτική και επικεντρωμένη στον Χρήστη Μεθοδολογία αξιολόγησης  
Συστημάτων Αυτοματοποιημένης Μηχανικής Μάθησης ”**

### **Περίληψη**

Η Αυτοματοποιημένη Μηχανική Μάθηση (AutoML) είναι ένα ταχέως αναπτυσσόμενο υπο-πεδίο της Μηχανικής Μάθησης (MM). Το AutoML στοχεύει στην πλήρη αυτοματοποίηση της διαδικασίας της MM από άκρο σε άκρο, εκδημοκρατώντας τη Μηχανική Μάθηση σε μη ειδικούς και αυξάνοντας δραστικά την παραγωγικότητα των τελευταίων. Μέχρι στιγμής, οι περισσότερες συγκρίσεις των συστημάτων AutoML εστιάζονται αποκλειστικά σε ποσοτικά κριτήρια όπως η πραγματική προβλεπτική απόδοση και ο χρόνος εκτέλεσης. Σε αυτή τη διατριβή, παρουσιάζουμε

μια μεθοδολογία πολλαπλών επιπέδων για την επαρκή αξιολόγηση αυτών των πολύπλοκων συστημάτων. Ξεκινάμε εξετάζοντας τις υπηρεσίες AutoML για τη δημιουργία μοντέλων MM από την οπτική γωνία ενός χρήστη, μη εξετάζοντας την προγνωστική απόδοση. Παρουσιάζουμε μια μεγάλη γκάμα κριτηρίων για την αξιολόγηση και σύγκριση αυτών των υπηρεσιών από την πλευρά του χρήστη. Η σύγκριση δείχνει τα πλεονεκτήματα και τις αδυναμίες κάθε υπηρεσίας, τις ανάγκες που καλύπτει, τη μερίδα χρηστών που είναι καταλληλότερο και τις δυνατότητες βελτίωσης. Για να αξιολογήσουμε περαιτέρω την εμπειρία του χρήστη, δημιουργούμε και διεξάγουμε μια μελέτη, εστιάζοντας στη χρηστικότητα των συστημάτων AutoML. Σε αυτήν τη μελέτη, οι χρήστες καλούνται να πραγματοποιήσουν μια ανάλυση ML χρησιμοποιώντας 3 συστήματα και να βαθμολογήσουν την ευκολία χρήσης τους. Οι απαντήσεις τους μπορούν να παρέχουν πληροφορίες στις ομάδες ανάπτυξης των συστημάτων AutoML σχετικά με τα σημεία συμφόρησης του UX και τις λανθασμένες αποφάσεις σχεδιασμού. Συνεχίζοντας, παρουσιάζουμε τη δική μας μεθοδολογία ποσοτικής αξιολόγησης, θέτοντας κανόνες για την επιλογή υποψηφίων συστημάτων AutoML, δεδομένων που θα αναλυθούν και μετρήσεων απόδοσης. Για τα ποσοτικά μας κριτήρια, δίνουμε έμφαση στην ακρίβεια της εκτίμησης της προγνωστικής απόδοσης, καθώς και στην πραγματική απόδοσή τους. Επιπλέον, πραγματοποιούμε μια ανάλυση με βάση πληροφορίες για τη δομή των δεδομένων μας και αξιολογούμε πώς τα διαφορετικά χαρακτηριστικά δεδομένων συσχετίζονται με το αποτέλεσμα. Τα αποτελέσματα δείχνουν ότι τα περισσότερα συστήματα υπερεκτιμούν την απόδοση του παραχθέντος προβλεπτικού μοντέλου, ενώ δεν υπάρχουν σημαντικές διαφορές μεταξύ τους όσον αφορά την κατάταξη με βάση τις επιδόσεις τους. Και στις δύο περιπτώσεις, τα αποτελέσματα έχουν στατιστικά σημαντικούς συσχετισμούς με τα επιλεγμένα χαρακτηριστικά των δεδομένων μας.

**University of Crete**

**Computer Science Department**

**M.Sc. Thesis presentation / examination**

**Iordanis Xanthopoulos**

**Master's Thesis Supervisor: Professor I. Tsamardinos**

**Thursday, 3 September 2020, 18:00 p.m**

**Teleconference (will use the e: Presence system), Computer Science Department,  
University of Crete**

(url) : <http://video.ucnet.uoc.gr/live/show/303>

YouTube channel :

[https://www.youtube.com/channel/UC7uE3QiMTQjkrpByB\\_Gnt6Q/live](https://www.youtube.com/channel/UC7uE3QiMTQjkrpByB_Gnt6Q/live)

**“A Qualitative, Quantitative and User-based Methodology of Automated Machine Learning Systems Evaluation”**

**Abstract**

Automated Machine Learning (AutoML) is a rapidly rising sub-field of Machine Learning. AutoML aims to fully automate the machine learning process end-to-end, democratizing Machine Learning to non-experts and drastically increasing the productivity of expert analysts. So far, most comparisons of AutoML systems focus solely on quantitative criteria such as predictive performance and execution time. In this thesis, we present a multi-level methodology to adequately evaluate these complex systems. We start off by examining AutoML services for predictive modeling tasks from a user's perspective, going beyond predictive performance. We present a wide palette of criteria and dimensions on which to evaluate and compare these services as a user. The comparison indicates the strengths and weaknesses of each service, the needs that it covers, the segment of users that is most appropriate for, and the possibilities for improvements. To further expand on the evaluation of the user experience, we create and conduct a custom user study, focusing on the user experience and usability of AutoML systems. In this study, the users are asked to perform a ML analysis using 3 state-of-the-art systems and grade their ease-of-use. Their responses provide useful feedback to the AutoML systems' development teams regarding UX bottlenecks and flawed design decisions. Continuing, we present our own quantitative evaluation methodology, by setting rules and preferences for selecting candidate AutoML systems, benchmark datasets and performance metrics. For our quantitative criteria, we emphasize on the accuracy of the estimation of predictive performance, as well as their hold-out performance. Additionally, we perform an analysis based on the metafeatures of our benchmark and evaluate how different data characteristics are correlated to the outcome. The results show most systems overestimate their output's performance, while there are no major differences between

them when it comes to ranking them based on hold-out performance. In both cases, the results have a trend with the selected metafeatures.