1. Record Nr. TD21003131

Autore CURIA, FRANCESCO

Titolo Explainable clinical decision support system: opening black-box meta-learner algorithm expert's based [Tesi di dottorato]

Lingua di pubblicazione Inglese

Formato Tesi di dottorato

Livello bibliografico Monografia

Note diritti: info:eu-repo/semantics/openAccess

In relazione con info:eu-repo/semantics/altIdentifier/hdl/11573/1538472

Sommario
Mathematical optimization methods are the basic mathematical tools of all artificial intelligence theory. In the field of machine learning and deep learning the examples with which algorithms learn (training data) are used by sophisticated cost functions which can have solutions in closed form or through approximations. The interpretability of the models used and the relative transparency, opposed to the opacity of the black-boxes, is related to how the algorithm learns and this occurs through the optimization and minimization of the errors that the machine makes in the learning process. In particular in the present work is introduced a new method for the determination of the weights in an ensemble model, supervised and unsupervised, based on the well known Analytic Hierarchy Process method (AHP). This method is based on the concept that behind the choice of different and possible algorithms to be used in a machine learning problem, there is an expert who controls the decisionmaking process. The expert assigns a complexity score to each algorithm (based on the concept of complexity-interpretability trade-off) through which the weight with which each model contributes to the training and prediction phase is determined. In addition, different methods are presented to evaluate the performance of these algorithms and explain how each feature in

the model contributes to the prediction of the outputs. The interpretability techniques used in machine learning are also combined with the method introduced based on AHP in the context of clinical decision support systems in order to make the algorithms (black-box) and the results interpretable and explainable, so that clinical-decision-makers can take controlled decisions together with the concept of "right to explanation" introduced by the legislator, because the decision-makers have a civil and legal responsibility of their choices in the clinical field based on systems that make use of artificial intelligence. No less, the central point is the interaction between the expert who controls the algorithm construction process and the domain expert, in this case the clinical one. Three applications on real data are implemented with the methods known in the literature and with those proposed in this work: one application concerns cervical cancer, another the problem related to diabetes and the last one focuses on a specific pathology developed by HIV-infected individuals. All applications are supported by plots, tables and explanations of the results, implemented through Python libraries. The main case study of this thesis regarding HIV-infected individuals concerns an unsupervised ensemble-type problem, in which a series of clustering algorithms are used on a set of features and which in turn produce an output used again as a set of meta-features to provide a set of labels for each given cluster. The meta-features and labels obtained by choosing the best algorithm are used to train a Logistic regression meta-learner, which in turn is used through some explainability methods to provide the value of the contribution that each algorithm has had in the training phase. The use of Logistic regression as a meta-learner classifier is motivated by the fact that it provides appreciable results and also because of the easy explainability of the estimated coefficients.

Localizzazioni e accesso	http://memoria.depositolegale.it/*/http://hdl.handle.net/11573/1538472