

Laborator 5

1. Modele de clasificare

Preluati 3 seturi de date pentru probleme de clasificare, stocate in repository-urile date in cursul 4 - sau altele. Pentru fiecare din acestea:

1. Daca e necesar, aplicati una din metodele de missing value imputation implementate in laboratorul precedent;
2. Alegeti pentru fiecare problema o metrica pe care vreti sa o optimizati - acuratete, f1 score etc. - a se vedea [sklearn.metrics](http://scikit-learn.org/stable/modules/classes.html#module-sklearn.metrics) (<http://scikit-learn.org/stable/modules/classes.html#module-sklearn.metrics>);
3. Alegeti 4 modele de clasificare din biblioteca sklearn; documentati-i in Jupyter notebook;
4. Pentru fiecare model: efectuati o cautare a hiperparametrilor optimi - de exemplu prin [GridSearchCV](http://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html) (http://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html);
5. Comparati rezultatele date de catre fiecare model (pentru hiperparametrii optimali gasiti la punctul anterior) in forma tabelara. Performanta se va determina prin 10-fold cross validation, se vor compara mediile.

2. Modele de regresie

Efectuati aceiasi pasi ca mai sus pentru problema de regresie.

Predare: laboratoarele 4 si 5 vor fi prezentate in 20 aprilie.