

Rezolvarea unor probleme prin metode de învățare automată



Obiective

Introducere în dezvoltarea sistemelor care învață singure. Tipuri de probleme rezolvabile cu metode de învățare automată (regresie și clasificare). Măsurile de evaluare a performanței acestor metode.



Aspecte teoretice

Clasificarea problemelor care necesită metode de învățare automată.

Proiectarea sistemelor care învață singure.

Evaluarea sistemelor care învață singure. Metrice de performanță:

- Eroare
- Acuratețe, Precizie, Rapel,
- Funcție de cost

Termen de predare și evaluare



Laborator 7

Punctajele acordate:

- Determinarea erorii de predicție - regresie multi-target – 100 puncte
- Determinarea acurateții, preciziei, rapelului - clasificare multi-class – 100 puncte
- Determinarea loss-ului - probleme de regresie și clasificare multi-class și multi-label – 200 puncte

Cerințe

Specificați, proiectați și implementați rutine de evaluare a performanței unui algoritm de ML.



Live:

- performanța predicției în cazul unei probleme de regresie (cu un singur output)
- performanța clasificării (acuratețe, precizie, rapel) în cazul unei clasificări binare (cu outputuri de tip etichetă) - cazul unui set de date echilibrat și cazul unui set de date neechilibrat
- performanța clasificării (acuratețe, precizie, rapel) în cazul unei clasificări binare (cu outputuri de tip probabilități - matrice cu `noSamples` x `noClasses` elemente)

Temă:

- să se specifice procedura de evaluare a unui algoritm de ML care a rezolvat o problema de regresie multi-target și să se determine eroarea de predicție în cazul unei regresii multi-target
- să se specifice procedura de evaluare a unui algoritm de ML care a rezolvat o problema de clasificare multi-clasă și să se determine acuratețea, precizia, rapelul în cazul unei clasificări multi-class

Temă opțională

- Determinarea loss-ului (funcție de cost) în cazul problemelor de regresie
- Determinarea loss-ului (funcție de cost) în cazul problemelor de clasificare binară (outputul clasificatorului este reprezentat ca o matrice cu `noSamples` x 2 valori reale subunitare, fiecare linie având suma 1)
- Determinarea loss-ului (funcție de cost) în cazul problemelor de clasificare multi-clasă (outputul clasificatorului este reprezentat ca o matrice cu `noSamples` x `noClasses` valori reale)
- Determinarea loss-ului (funcție de cost) în cazul problemelor de clasificare multi-label (outputul clasificatorului este reprezentat ca o matrice cu `noSamples` x `noClasses` valori reale)