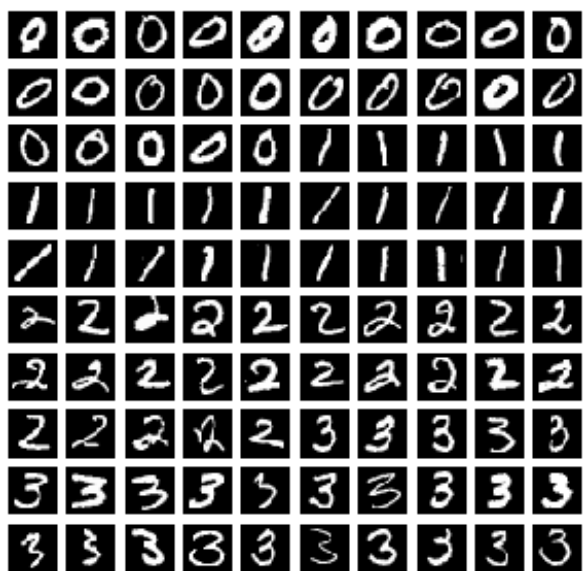


Test de laborator la Inteligență Artificială, seria 35

- 1 februarie 2022 -

1. Clasificarea cifrelor pare și impare (4.5 puncte)

În această lucrare de laborator veți avea de construit un clasificator binar care distinge între imagini cu cifre pare de imagini cu cifre impare. Prima clasă (*clasa pară* – *clasa 0*) conține cifrele pare 0, 2, 4, 6, 8 iar a doua clasă (*clasa impară* – *clasa 1*) conține cifrele impare 1, 3, 5, 7, 9. Vă veți construi acest clasificator binar folosind datele furnizate, fișierele *0.txt*, *1.txt*,... *9.txt* conțin fiecare 100 de exemple (imagini 28×28 pixeli) din fiecare cifră. Codul Python de la care porniți (fișierul *problema_1.py*) afișează primele 25 de exemple pentru cifrele 0, 1, 2, 3.



Primele 25 de exemple din cifrele 0, 1, 2, 3.

Realizați următoarele:

- Împărțiți cele 100 de exemple pentru fiecare cifră astfel: primele 75 de exemple pentru fiecare cifră intră în mulțimea de antrenare, următoarele 25 intră în mulțimea de testare. Veți avea astfel o mulțime de antrenare cu 750 de exemple de antrenare și o mulțime de testare cu 250 de exemple de cifre. **(1 punct)**.
- Construiți-vă clasificatorul binar astfel: antrenați un SVM liniar pe datele de antrenare (750 exemple) pe clasele inițiale ale cifrelor, adică pe 10 clase (un exemplu de cifra 0 are clasa 0, un exemplu de cifra 1 are clasa 1, un exemplu de cifra 2 are clasa 2, etc.). Folosind acest clasificator pe 10 clase obțineți clasificatorul final pe cele 2 clase și anume pe *clasa pară* vs *clasa impară*. Care este acuratețea clasificatorului binar obținut pe mulțimea de testare de la punctul a? Afișați matricea de confuzie a clasificatorului binar. **(2 puncte)**
- Diferit de punctul b acum vă antrenați direct clasificatorul binar din cele 750 de exemple de antrenare, considerând etichete binare: eticheta *clasa pară* (puteți pune eticheta 0) pentru exemple cu cifrele 0, 2, 4, 6, 8 și eticheta *clasa impară* (puteți pune eticheta 1) pentru exemplele cu cifrele 1, 3, 5, 7, 9. Care este acuratețea clasificatorului binar obținut pe mulțimea de testare de la punctul a? Afișați matricea de confuzie a clasificatorului binar. **(1 punct)**
- Există o diferență între performanța clasificatorului binar de la punctul b și cel de la punctul c? Justificați răspunsul. **(0.5 puncte)**