

Raport Proiect: Clasificare imaginilor

-Snowboarders vs skiers-

Scopul si descrierea proiectului

Scopul proiectului este utilizarea algoritmului k-Means pentru a obține clasificarea imaginilor in snowboarderi si schiori. Acest lucru poate fi util in dezvoltarea unui sistem automatizat pentru recunoașterea acestora într-un eveniment sportiv, sau, in cazul meu, pentru a învăța si a înțelege cum funcționează algoritmul de clasificare k-Means.

Modul de obținere si de organizare a datelor

Setul de date de antrenare l-am obținut inițial prin utilizarea unui script făcut in python pentru a descărca imagini de pe internet (un image scraper). Astfel am obținut aproximativ 400 de imagini pentru fiecare categorie (schiori si snowboarderi). Acest lucru l-am făcut pentru a scuti munca de a descărca manual fiecare poza de pe internet, însă după utilizarea scriptului, am verificat fiecare imagine si am făcut o triere a lor astfel încât ele sa fie relevante temei alese. Așadar setul de date de antrenare are 175 de poze pentru categoria „schiori” si 123 de date pentru categoria „snowboarderi”. Însă, pentru o acuratețe mai buna, am adăugat la acest set aceleași imagini, transformate (rescalate, blurate si răsturnate). Acest set de date de antrenare este organizat in subfoldere („schiori”, „snowboarderi”), necesare etichetărilor.

De asemenea am si un set de date pentru testarea modelului, având in total 12 imagini (6 pentru schiori si 6 pentru snowboarderi).

Algoritmului utilizat si parametrii acestuia

Algoritmul utilizat este k-Means, fiind unul dintre cei mai cunoscuți algoritmi de învățare nesupervizata care se aplica pe date neetichetate pentru a le împărți în mai multe clusetere (grupuri), fiecare cluster conținând un set de obiecte dintr-o anumita categorie. In cazul meu, numărul de clustere este egal cu 2, reprezentând cele 2 categorii: schiori si snowboarderi. Parametrii utilizați sunt:

- `n_cluster=2`. Acest parametru seteaza numarul de clustere in care vor fi împărțite datele
- `random_state=0`. Acest parametru controlează reproductibilitatea rezultatelor atunci când algoritmul implica aspecte aleatorii (asigura ca algoritmul va genera aceleași rezultate atunci cand este rulat pe aceleași date de intrare).

Biblioteci Python utilizate

- `os`: Pentru manipularea structurii de directoare și fișiere.
- `cv2` (OpenCV): Pentru citirea și prelucrarea imaginilor.
- `numpy`: Pentru manipularea eficientă a datelor în formă de matrice.
- `sklearn.cluster.KMeans`: Pentru implementarea algoritmului K-Means.
- `sklearn.preprocessing.StandardScaler`: Pentru scalarea datelor.
- `matplotlib.pyplot`: Pentru vizualizarea imaginilor și rezultatelor.

Rezultate obținute

Fiecare imagine conține si un text cu rezultatul obținut

Test Image 8 - ==SNOWBOARDER==



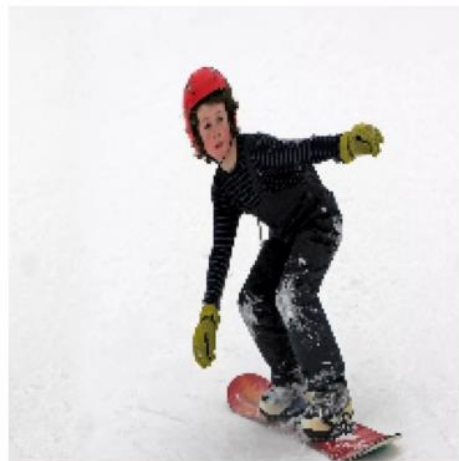
Test Image 6 - ==SNOWBOARDER==



Test Image 1 - ==SCHIOR==



Test Image 10 - ==SNOWBOARDER==



Test Image 11 - ==SNOWBOARDER==



Test Image 3 - ==SCHIOR==



Test Image 9 - ==SNOWBOARDER==



Test Image 4 - ==SCHIOR==



Test Image 5 - ==SCHIOR==



Test Image 7 - ==SCHIOR==



Test Image 12 - ==SCHIOR==



Test Image 2 - ==SCHIOR==



Interpretare model

Voi considera predictia „schior” ca fiind valoarea pozitiva, si predictia „snowboarder” ca valoare negativa. Se observa ca:

- sunt 5 imagini cu schiori avand predictia tot „schior”, deci $TP = 5$;
- sunt 4 imagini cu snowboarderi avand predictia tot „snowboarder”, deci $TN = 4$;
- este o imagine cu schiori avand predictia „snowboarder”, deci $FP = 1$;
- sunt 2 imagini cu snowboarderi avand predictia „schiori”, deci $FN = 2$;

Avand termenii matricei de confuzie:

		VALORI ACTUALE	
		Pozitive	Negative
VALORI PREZISE	Pozitive	TP	FP
	Negative	FN	TN

Putem face urmatoarele analize:

- Acuratetea este 0.75 (75%), dupa formula : $Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$
- Precizia este 0.83 (83%), dupa formula: $Precision = \frac{TP}{TP + FP}$
- Sensibilitatea este 0.71 (71%), dupa formula: $Recall = \frac{TP}{TP + FN}$
- Specificitatea este 0.8 (80%), dupa formula: $Specificitate = \frac{TN}{TN + FP}$
- Scorul F1 este 0.77 (77%), dupa formula: $Scorul\ F1 = 2 \cdot \frac{(Recall \cdot Precizie)}{(Recall + Precizie)}$

Acuratețea măsoară proporția de instanțe clasificate corect din numărul total de articole. Este raportul dintre numărul de predicții corecte și numărul total de predicții. Faptul ca aceasta este de 75% înseamnă ca modelul nu este un dezechilibru atât de mare între cele 2 clase.

Precizia este o măsură a corectitudinii care se realizează în predicția adevărată. Cu alte cuvinte, îmi spune câte predicții sunt de fapt pozitive din totalul prezis pozitiv. Este un lucru bun ca este de 0.83, fiind cat mai aproape de 1 (1 fiind ideal). Aceasta ar fi utila in cazul in care m-ar preocupa mai mult valorile fals pozitiv fata de cele fals negativ, adică daca modelul meu a prezis ca este schior când de fapt nu este (e snowboarder).

Sensibilitatea este o măsură a observațiilor reale care sunt prezise corect, adică câte observații din clasă pozitivă sunt de fapt prezise ca pozitive. Aceasta este o alegere validă a metricii de evaluare atunci când se dorește să se surprindă cât mai multe aspecte pozitive.

Specificitatea este o măsură a cât de bine identifică clasificatorul valorile negative. Adică cat bine identifica clasificatorul snowboarderii. Mă bucura faptul ca identifica mai bine snowboarderii (80%) fata de schiori (71%), deoarece si eu mă dau cu placa si mereu este o competiție între „plăcari” si schiori;

Scorul F1 menține un echilibru între precizie și recall pentru clasificatorul utilizat. Dacă precizia este scăzută, scorul F1 este scăzut și dacă recall-ul este și el scăzut scorul F1 este scăzut. Scorul F1 ar fi o măsură utilă dacă atât cazurile fals pozitive, cât și cele fals negative au consecințe grave. Un scor F1 ridicat ar indica faptul că modelul face predicții pozitive precise, minimizând în același timp falsele pozitive și falsele negative.

Concluzii

Se poate observa ca din 12 imagini, doar 3 au o etichetare greșită (test image 6, test image 7 si test image 12). Aceste greșeli sunt cauzate in principal din faptul ca cele 2 categorii sunt foarte asemănătoare între ele. Doar poziția omului, ar face diferența, însă de multe aceasta este asemănătoare in ambele cazuri. Placa sau schiurile de asemenea ar putea face diferența între cele 2 categorii, însă nu întotdeauna acestea sunt vizibile, din cauza zăpezii, care este prezenta in ambele cazuri, astfel fiind chiar greu si pentru un om sa își dea seama.