

Esame 27/01/2021

Andrea Passerini
andrea.passerini@unitn.it

Informatica

Programma python

Scrivere un programma che prenda in ingresso:

- 1 un file con un elenco di predizioni di score di artefatti al polmone fatte da una rete neurale. Il file contiene una serie di pazienti e per ogni paziente una serie di predizioni, una per ogni frame di cui e' composto il video dell'ecografia polmonare
- 2 un file con lo score associato ad ogni paziente (e relativo video) da medici esperti.

Calcoli per ogni paziente lo score piu' frequente tra quelli predetti per i suoi frame, e confronti tale score con quello dei medici, riportando l'accuratezza media (frazione di predizioni corrette rispetto al numero di pazienti).

Esempio file di predizioni

```
$ cat frame_predictions.csv
patient,frame,pred
convex_202003061327250047ABD,0,2
convex_202003061327250047ABD,1,2
convex_202003061327250047ABD,2,2
[...]
convex_202003061327250047ABD,181,0
convex_202003061327250047ABD,182,2
convex_202003061327250047ABD,183,0
[...]
convex_202003101657370100ABD,137,0
convex_202003101657370100ABD,138,0
convex_202003101657370100ABD,139,0
```

Esempio file di annotazioni mediche

```
$ cat video_labels.csv
patient,label
convex_202003061327250047ABD,2
convex_202003171447150456ABD,1
convex_202003171447490457ABD,3
linear_202003171452410466SMP,3
convex_202003181411300131ABD,0
[...]
```

Esempio esecuzione

```
$ python majority_score.py  
Inserire nome file predizioni: frame_predictions.csv  
Inserire nome file annotazioni: video_labels.csv  
Majority accuracy = 0.568966
```

Programma python: suggerimento

Si possono implementare 5 funzioni separate:

- 1 una che legga il file di predizioni e restituisca una mappa da paziente a lista di predizioni per i suoi frame
- 2 una che legga il file di annotazioni mediche e restituisca una mappa da paziente a score del suo video
- 3 una che data una lista di predizioni restituisca lo score piu' frequente (opzionale: se due score sono ugualmente frequenti restituire quello piu' alto)
- 4 una che date le due mappe lette da file calcoli l'accuratezza del predire per ogni paziente lo score piu' frequente (usando la funzione sopra)
- 5 una (o un main) che realizzi il programma richiesto usando le funzioni di cui sopra

Esercizi da linea di comando

Calcolare quante sequenze nella directory `fastas` NON cominciano con una metionina (M) e finiscono con un acido aspartico (D) o glutammico (E).

Risultato atteso

8

Esercizi da linea di comando

Calcolare il numero di pazienti che hanno almeno un frame predetto con score 3 (il massimo)

Risultato atteso

22

Modalita' di esecuzione e consegna

- 1 Accendere la macchina ed aspettare (la modalità esame si avvia automaticamente)
- 2 Aprire un browser e collegarsi al sito
`examina.icts.unitn.it`
- 3 Autenticarsi con le credenziali di ateneo
- 4 Scaricare il testo del compito
- 5 Realizzare il programma python come file `programma.py` e scrivere gli esercizi da linea di comando in un file di testo `linea_di_comando.txt`
- 6 Creare una cartella con *nome_cognome* e metterci i due file realizzati, comprimerla e caricarla su
`examina.icts.unitn.it`
- 7 Spegnere la macchina