

Rješenje zadatka 3.4 predmeta Strojno učenje

Siniša Biđin

5. siječnja 2013.

- (a) Svi su primjeri na skupu za učenje ispravno klasificirani, a pogreška unakrsne entropije iznosi 7.298×10^{-5} .

```
src $ python logreg.py train-set.txt tezinel.txt
```

- (b) Broj iteracija metode gradijentnog spusta ograničen je prisilno na 20. Postupak prekidamo ranije samo ukoliko je promjena pogreške unakrsne entropije između dva susjedna koraka “dovoljno malena”.

```
src $ python bestlambda.py train-set.txt validation-set.txt
```

λ	%	iteracija
0	3.75	2
0.01	3.75	20
0.02	3.75	20
0.05	3.75	20
0.1	3.75	20
0.2	3.75	20
0.5	3.75	20
1	3.75	20
2	3.75	20
5	3.75	20
10	3.75	20
20	3.75	20

- (c) Loše su klasificirana dva primjera (2.5%) iz skupa za testiranje. Dvadeset najreprezentativnijih riječi za kategoriju sport su: *utakmica, klub, igrač, reprezentacija, pobjeda, momčad, liga, olimpijski, turnir, trener, sezona, prvenstvo, kola, igrati, igra, dvoboj, prvak, svjetski, svladati, nogometan*.

```
src $ python final.py
```