

# strojno ocenjevanje kratkih odgovorov

Tomaž Tomažič (63100281)

17. december 2012

## 1 Uvod

Cilj naloge je oceniti kratke odgovore na neznana vprašanja tako, kot bi to storil pravi ocenjevalec.

## 2 Podatki

Podani so učni in testni podatki za učenje in napovedi ocen, v katerih so odgovori na 10 različnih vprašanj na katere je odgovorilo skoraj dva tisoč študentov. Odgovorov za učenje je 17043 in za vsak primerek je podana ocena dveh ocenjevalcev.

## 3 Metode

Nalogo sem rešil z uporabo logistične regresije. Najprej sem očistil besedilo tako da sem odstranil ločila, nepotrebne presledke, števila in določene znake. Potem sem za vsak odgovor preštel pojative štiriterk znakov. Ker jih je bilo zelo malo podobnih v različnih vprašanjih oziroma so zelo različne, sem podatke shranil v redko matriko, kar mi je omogočalo hitrejše računanje logistične regresije. Zatem sem zgradil 3 oziroma 4 modele za vsako vprašanje, odvisno od razpona ocene, katere sem uporabil za ocenjevanje testnih podatkov. Ker nisem implementiral k-prečnega preverjanja, sem napoved izračunal na le nekaj lambdah (4,6,7,7.5,8,10,12,16) in tako izbral najugodnejšo, glede na rezultat na učilnici. Zaradi istega razloga tudi nisem poiskal najboljših lambd za vsak tip odgovora posebej, ampak sem eno lambda uporabil za vse tipe.

## 4 Rezultati

Na učilnici kjer sem oddal napovedi, sem dosegel najvišji rezultat z točnostjo 0.65056

## 5 Izjava o izdelavi domače naloge

Domačo nalogo in pripadajoče programe sem izdelal sam.