Drugi domaci zadatak-Izvestaj

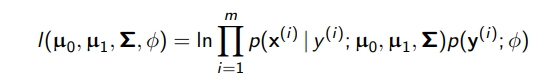
Softmax I GDA

Student:Tatjana Tabandzelic

Broj indeksa: 2021/3290

**Gausova diskriminaciona analiza-GDA**

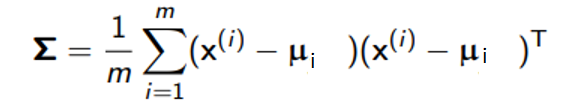
GDA je vrsta generativnog algoritma, cija je odlika ucenje raspodela klasa, tj. ne ucimo separacionu krivu vec ucimo parameter koji opisuju datu raspodelu. Predikore opisujemo Normalnom raspodelom, te su parametri raspodele: srednja vrednost I kovarijaciona matrica. Zelimo da nad trenirajucem skupu naucimo odgovarajuce parameter I da onda nad testnim skupom odredimo procenat uspesnosti nase procene. Parametre ucimo maksimizaciom verodostojnosti, pretpostavimo da su nam svi primeri nezavisni I onda odredimo maksimum sl. f-je:



GDA im analiticko resenje jer mozemo gradijent po svim parametrima f-je da izjednacimo sa 0 I resimo.

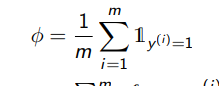
Srednja vrednost se odredjuje kao kolicnik sume svih primera te klase I broj primera te klase, tj. pronalzaenjem aritmeticke sredine za datu klasu .

Kovarijacionu matricu racunamo:



(gde je m broj pojavljivanja date klase a mi-y je srednja vrednost za datu klasu)

Takodje imamo I parematr φ koji predstavlja verovatnocu pojavljivanja date klase u celom trenig setu



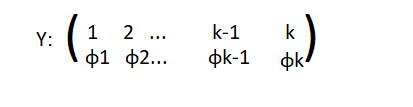
Nakon odredjivanja svih parametara kreira se gausova raspodela za svaku klasu.

Test set prolazi kroz svaku raspodelu i uzima onu sa najvecom verovatnocom.Na kraju dobijamo da je tacnost nad tesnom setu:

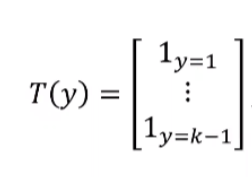


**Softmax**

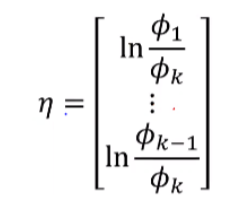
Softmax ili multinomijalna logisticka regresija je klasifikacioni metod koji generalizuje logisticku regresiju na problem vise klasa, odnosno izlanza promenljiva moze da ima vise od dve vrednosti. Multinomijalna logisticka regresija ima diskretnu raspodelu, sto znaci da Y moze da ima jednu od k diskretnih vednosti, parametri te raspodele su vreovatnoce da Y ima jedno od k-a vrednosti.



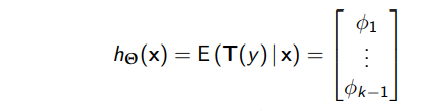
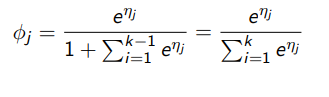
Ova raspodela takodje pripada eksponecijalnoj raspodeli. Dovoljna statistika je vektor koji preslikava skalarno Y u vektor cije komponente mogu da budu jedna od klasa



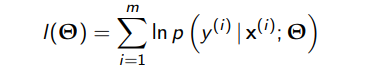
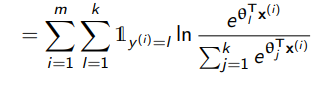
Prirodni parameter je vector, svaki element tog vektora je linearna fukcija prediktora:



Kao kod svakog drugog GLM hipoteza mora da ima sl oblik:

Softmaks mapira logite koji su realni brojevi u verovatnoce koje su izmedju 0 I 1 i ciji je zbir 1. Softmax nema paramatre sam po sebi jer je to preslikavanje realnih brojeva u verovatnoce, parameti naseg algoritma se nalaze u prirodnim parametrima I uce se maksimizacijom verodostojnosti.

Maksimizacija verodostojnosti se postize racunanjem stohastickog gradijetnog spusta sa mini-sarzama, sto znaci da ce mo ceo traing set podeliti na mini sarze, kao neke podskupove, I da se nakon svakog sarza racuna gradijent a potom azuriraju parametri.

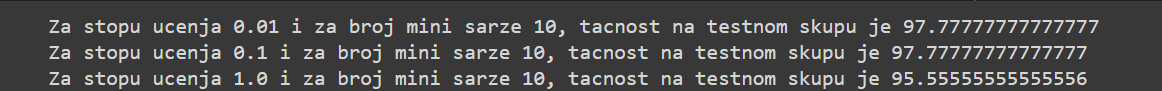
Chart

Description automatically generatedChart, histogram

Description automatically generatedChart

Description automatically generated

Na slikamo mozemo da vidimo uticaj learning rata-a na obucavajuci krivu. Za velike vrednosti learning rate-a dolazi do preskakanja maksimuma pa dolazi do divergecije, dok za malu velicinu jako sporo ucimo ali povecavamo sansu da dodjemo do najboljeg resenja.



Chart

Description automatically generatedChart, histogram

Description automatically generated

Chart, histogram

Description automatically generated

Na slikama mozemo da vidimo uticaj velicine mini sarze na obucavajucu krivu za learnig rate od 0.1.

Za manji broj mini sarzi tj kada traing dataset delimo na vece podsupove grafik lici na onaj koj bismo dobili da koristimo ceo treaing skup. Dok za veci broj mini sarzi tj kada dataset podelimo na vise delova, samim time traing se vrsi nad manjim brojem podataka dolazimo do nedostatka uzoraka i zbog toga dolazi do pojave pikova.

