## Softmax – zadatak 2

**Priprema**

1. Učitavanje podataka iz data-class.csv
2. Razdvajanje podataka: Skup podataka je podeljen na karakteristike (X) i oznake (y).
3. Standardizacija: Karakteristike su standardizovane
4. Dodavanje biasa: dodata je kolona jedinica na početak matrice karakteristika X.
5. Podela podataka na trening, validacione i test skupove: podaci su podeljeni tako da je 60% podataka za trening, 20% za validaciju i 20% za test

* Početna log-verodostojnost na testu: -37.525302255200735
* Početna tačnost: 33,84%

Cilj optimizacije hiper-parametara u ovom projektu bio je dvostruk:

1. **Identifikacija Optimalnih Parametara**:
   * Pronaći optimalne vrednosti za dva ključna hiper-parametra: veličinu mini-šarže i stopu učenja (alpha). Optimalne vrednosti ovih parametara omogućavaju modelu da efikasno konvergira ka najvišoj mogućoj tačnosti i najvećoj log-verodostojnosti, minimizujući broj uzoraka potrebnih za treniranje. Algoritam se izvodi sve dok se ne dostigne maksimalna log-verodostojnost. Napravljene su optimizacije kako bi se sprečilo da model zaglavi u lokalnim maksimumima i kako bi se postigla stabilnija konvergencija.
2. **Analiza Uticaja Promene Parametara**:
   * Ispitati kako varijacije u veličini mini-šarže i stopi učenja utiču na brzinu konvergencije, tačnost, i log-verodostojnost modela.

**Rezultati**

* **Optimalna stopa učenja: 0.3384857145670226**
* **Optimalna mini šarža: 16**
* **Konvergencija: Log-verodostojnost dostiže vrednost -7.16 nakon 106 uzoraka.**
* **Log-verodostojnost na test skupu: -2.19**
* **Tačnost na trening skupu: 96.2%**
* **Tačnost na test skupu: 97.3%**

Model brzo postiže log-verodostojnost od -7.16 već nakon 106 uzoraka, uz visoku tačnost i stabilnu konvergenciju. Ovi rezultati ukazuju na dobru generalizaciju i efikasnost modela u učenju bez prekomernog treniranja.

A graph with a line

Description automatically generated

*Prikaz log-verodostojnosti u zavisnosti od broja uzoraka za optimalne parametre*

* **Prevelika stopa učenja: 1**
* **Optimalna mini šarža: 16**
* **Konvergencija: Log-verodostojnost dostiže vrednost -12.24 nakon 106 uzoraka.**
* **Log-verodostojnost na test skupu: -5.42**
* **Tačnost na trening skupu: 97.17%**
* **Tačnost na test skupu: 97.30%**

Ovi rezultati pokazuju da, uprkos visokim vrednostima tačnosti na trening i test skupu, prevelika stopa učenja uzrokuje oscilacije u log-verodostojnosti, što može dovesti do nestabilne konvergencije.

A graph with a line

Description automatically generated

*Prikaz log-verodostojnosti u zavisnosti od broja uzoraka za preveliku stopu učenja*

* **Premala stopa učenja: 0.001**
* **Optimalna mini šarža: 16**
* **Konvergencija: Log-verodostojnost dostiže vrednost -6.48 nakon 4802 uzorka.**
* **Log-verodostojnost na test skupu: -2.31**
* **Tačnost na trening skupu: 98.11%**
* **Tačnost na test skupu: 100%**

Iako je stopa učenja niska, model je uspeo da postigne visoku tačnost (100% na testu) i relativno visoku log-verodostojnost uz stabilnu konvergenciju. Potrebno je mnogo više uzoraka (oko 4800) da bi se postigla stabilnost u poređenju sa drugim vrednostima parametara. Ovi rezultati pokazuju da niža stopa učenja može dovesti do stabilnije konvergencije, ali po cenu sporijeg učenja.

A graph with a blue line

Description automatically generated

*Prikaz log-verodostojnosti u zavisnosti od broja uzoraka za premalu stopu učenja*

* **Prevelika mini šarža: 100**
* **Optimalna stopa učenja: 0.3384857145670226**
* **Konvergencija: Log-verodostojnost dostiže vrednost -12.25 nakon 206 uzorka.**
* **Log-verodostojnost na test skupu: -4.13**
* **Tačnost na trening skupu: 96.23%**
* **Tačnost na test skupu: 97.30%**

Veća mini-šarža (100) uzrokuje sporiju konvergenciju i nešto lošije rezultate na test skupu, uprkos tome što daje stabilnije vrednosti log-verodostojnosti.

.

A graph with a red dot

Description automatically generated

*Prikaz log-verodostojnosti u zavisnosti od broja uzoraka za preveliku mini šaržu*

* **Premala mini šarža: 1**
* **Optimalna stopa učenja: 0.3384857145670226**
* **Konvergencija: Log-verodostojnost dostiže vrednost -8.59 nakon 40 uzorka.**
* **Log-verodostojnost na test skupu: -3.75**
* **Tačnost na trening skupu: 94.34%**
* **Tačnost na test skupu: 100%**

Iako premala mini-šarža daje brzu konvergenciju i visoku tačnost, postoji rizik od prenaučenosti što može uticati na performanse modela na novim, nepoznatim podacima.

A graph with a line

Description automatically generated

*Prikaz log-verodostojnosti u zavisnosti od broja uzoraka za premalu mini šaržu*