

Neuronske mreže

Predavanje 1: Osnove neuronskih mreža

Predmet: Inteligentni sistemi

Prof dr Zoran Ševarac

Univerzitet u Beogradu, Fakultet organizacionih nauka

Šta ćete saznati

- Osnovni pojmovi: arhitektura, slojevi i neuroni
- Trening neuronske mreže: učenje na greškama i back-propagation algoritam
- Primena neuronskih mreža za klasifikaciju i regresiju
- Matematičke osnove algoritma za učenje: optimizacija i gradijenti

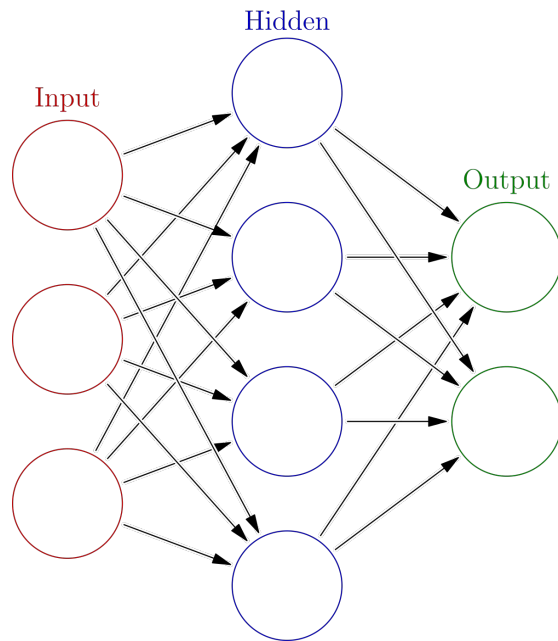
Neuronske mreže i arhitektura mreža

Neuronska mreža je algoritam mašinskog učenja koji za realizaciju modela koristi računarski graf.

Neuronska mreža ima sposobnost da nauči složene funkcije zavisnosti između skupa ulaza i izlaza datih u podacima za trening.

Čvorovi u grafu se nazivaju veštački neuroni i uređeni su hijerarhijski po **slojevima**.

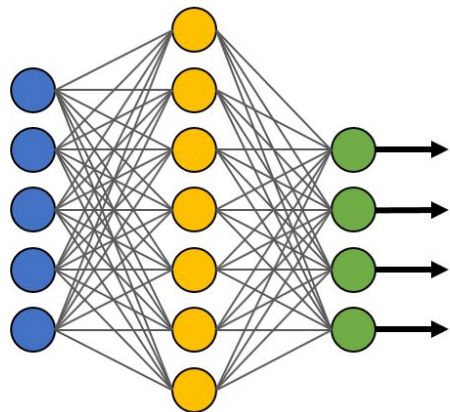
Arhitektura mreže definiše način povezanosti slojeva mreže odnosno čvorovima grafa. Najčešća je potpuna povezanost samo unapred ali moguća je rekurentna i lateralna.



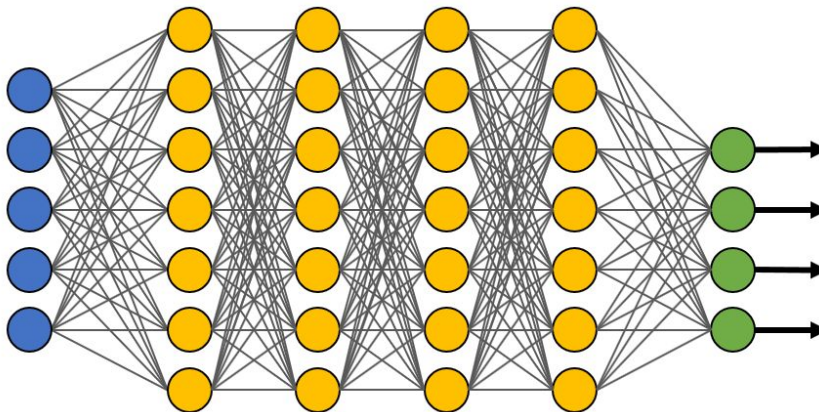
Duboke neuronske mreže / duboko učenje / *deep learning*

Neuronska mreža koja ima veći broj (preko 3) skrivenih slojeva

Simple Neural Network



Deep Learning Neural Network



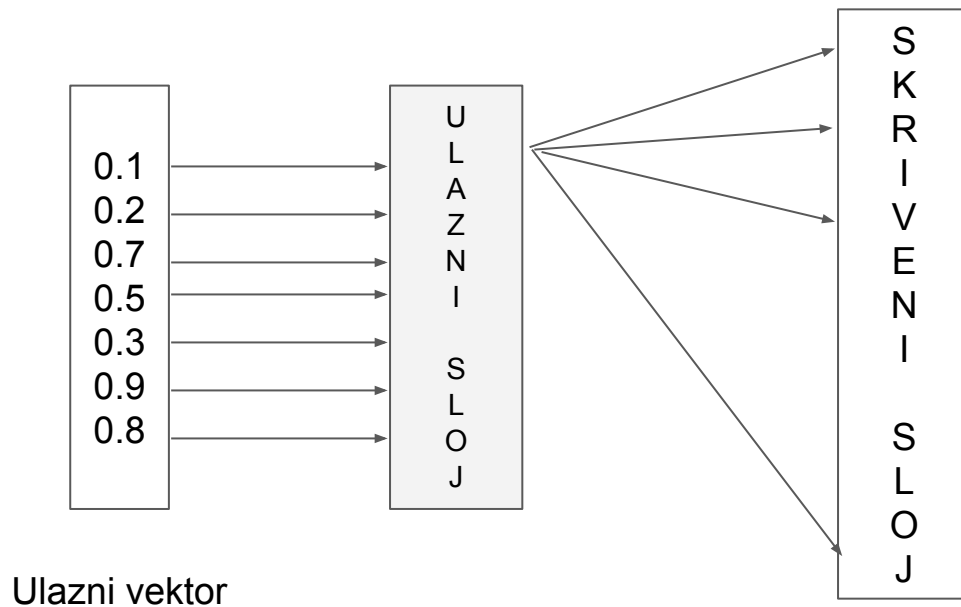
● Input Layer

● Hidden Layer

● Output Layer

Ulazni sloj neuronske mreže

Ulazni sloj ima ulogu da prihvati ulaz za neuronsku mrežu u obliku vektora i da ga prosledi na skrivene slojeve mreže na dalje procesiranje.



Skriveni slojevi neuronske mreže

Skriveni slojevi se nalaze između ulaznog i izlaznog sloja i one imaju glavnu ulogu za kreiranje modela i generisanje predikcije.

Osnovne komponente skrivenog sloja su:

- matrica težina (weights)
- funkcija aktivacije (tipično Relu)
- vektor izlaza

Proizvod matrice
težina i vektora ulaza

Matematički model slojeva: $Y = f(W * X + B)$

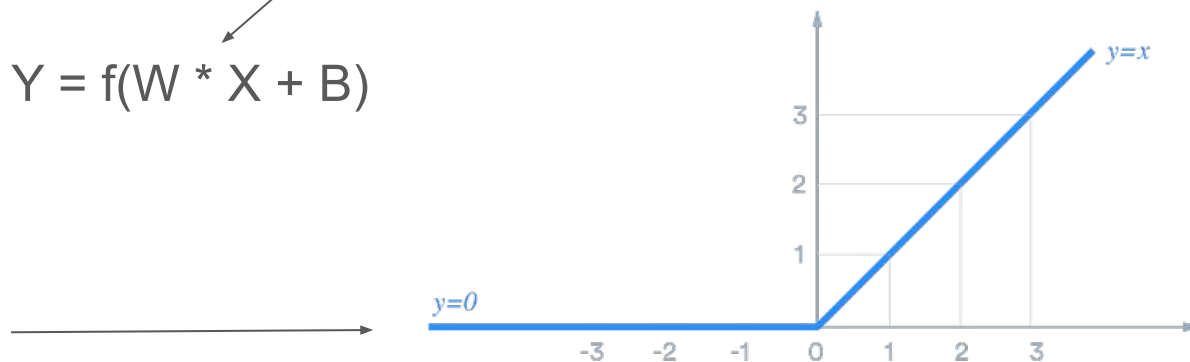
X - ulazni vektor (tenzor)

Y - izlazni vektor (tenzor)

W - matrica težina

B - bias vektor

f - funkcija aktivacije

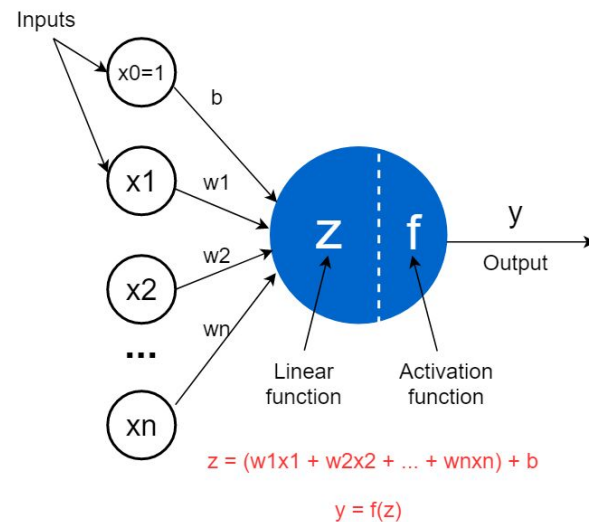


Veštački neuron

Svaki sloj sadrži niz veštačkih neurona koji su osnovni elementi procesiranja u neuronskoj mreži.

Glavni elementi neurona su: težine i aktivaciona funkcija.

Matematički model neurona: $y = f(x_1 * w_1 + \dots x_n * w_n + b)$



U suštini cela neuronska mreža se može posmatrati kao računarski graf u kome svaki čvor vrši linearnu ili logističku regresiju (u zavisnosti od funkcije aktivacije)

Izlazni sloj neuronske mreže

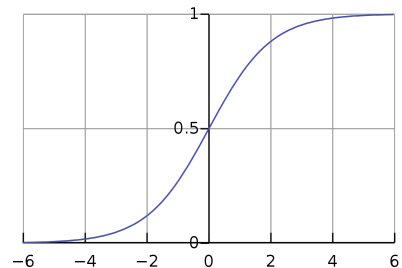
Izlazni sloj daje krajnji izlaz mreže odnosno konačnu predikciju modela.
Izlazni sloj takođe sadrži matricu težina.

U zavisnosti od vrste zadatka izlazni sloj ima odgovarajuću funkciju aktivacije:

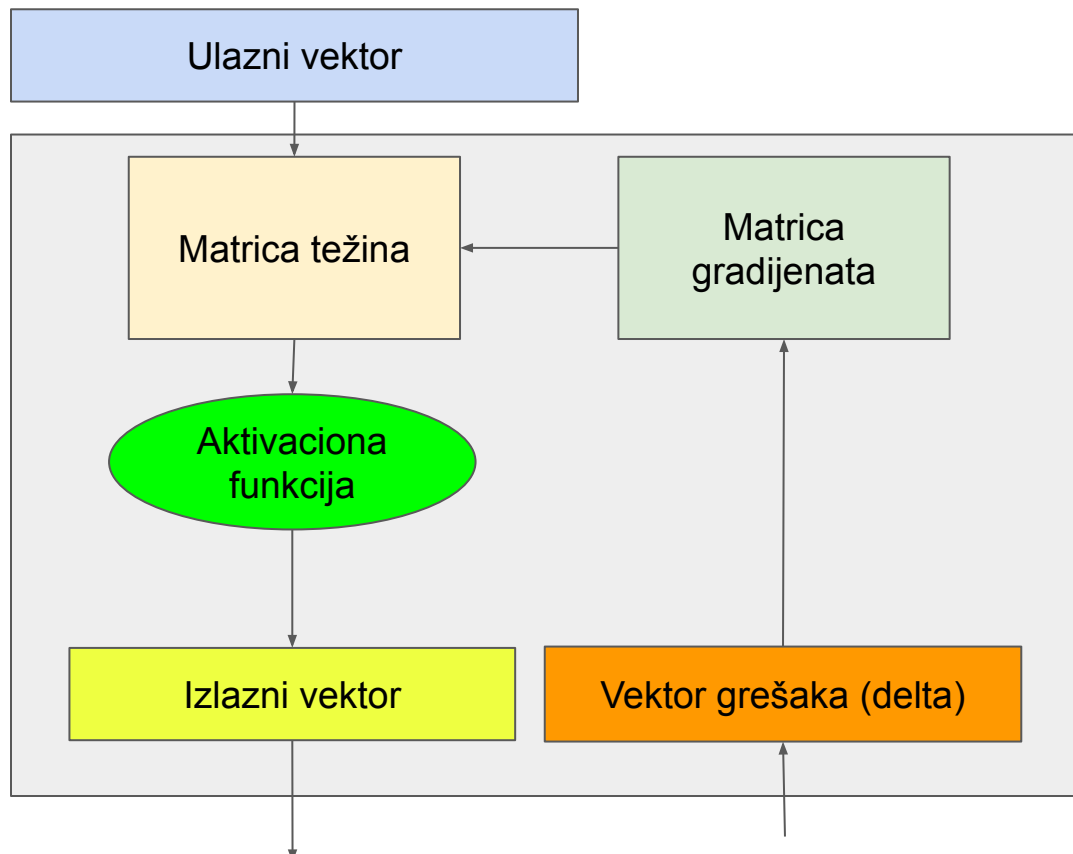
- Za **regresiju** - linearna funkcija $y=x$
- Za **binarnu klasifikaciju** sigmoidna \longrightarrow
- Za **više-klasni klasifikaciju** softmax

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

$$s(x_i) = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}}$$



Blok dijagram tipične implementacije sloja neuronske mreže



Trening neuronske mreže - učenje na greškama

Trening neuronske mreže je algoritam za automatsko podešavanje težinskih koeficijenata (matrica težina) u slojevima mreže, kako bi ukupna greška predikcije za zadati trening skup podataka bila što manja .

Algoritam se zasniva se na primeni matematičkih metoda za minimizaciju **funkcije greške**.

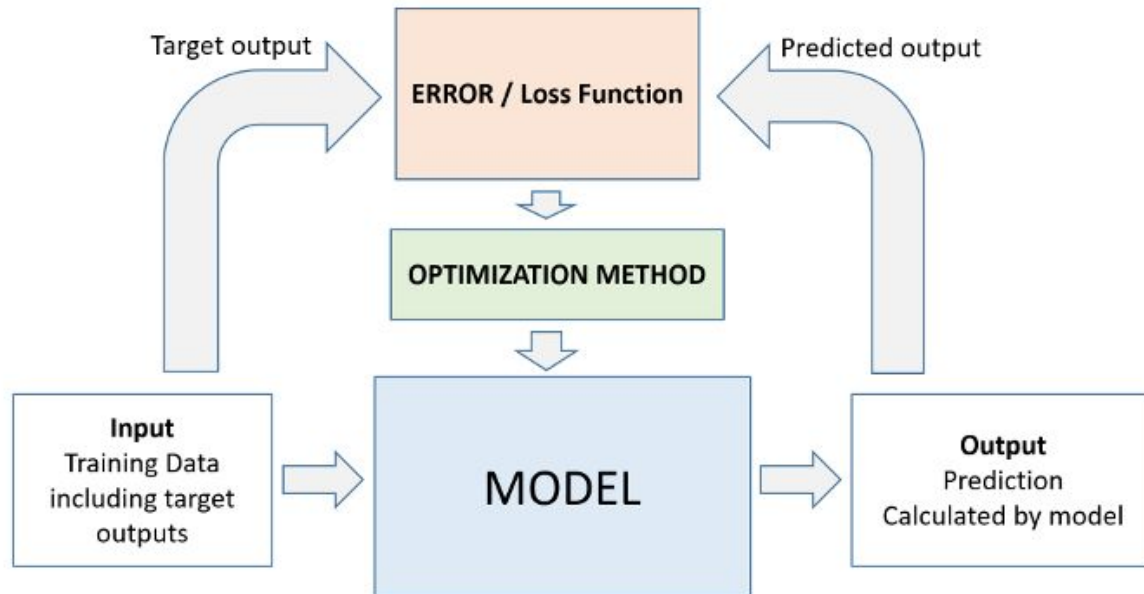
Primenom metoda **optimizacije** zasnovanih na **gradijentnom spustu** traži se minimum funkcije greške.

Algoritam se naziva Backpropagation - prostiranje **greške unazad**.

Ključni princip Backpropagation algoritma je mehanizam prenošenja greške unazad sa izlaznog na skrivene slojeve.

Blok dijagram backpropagation algoritma

Opšti princip: minimizacija ukupne greške kroz iterativnu proceduru



Funkcija greške

Funkcija greške izračunava ukupnu grešku predviđanja za ceo trening set.

Često korišćene funkcije su

Srednja kvadratna greška / Mean Squared Error

$$\text{MSE} = \overset{\text{Mean}}{\underbrace{\frac{1}{n} \sum_{i=1}^n}} \overset{\text{Error}}{\underbrace{(Y_i - \hat{Y}_i)}} \overset{\text{Squared}}{\underbrace{^2}}$$

Binarna Entropija

$$BCE(t, p) = -(t * \log(p) + (1 - t) * \log(1 - p))$$

Višeklasna entropija (Cross Entropy)

$$H = - \sum p(x) \log p(x)$$

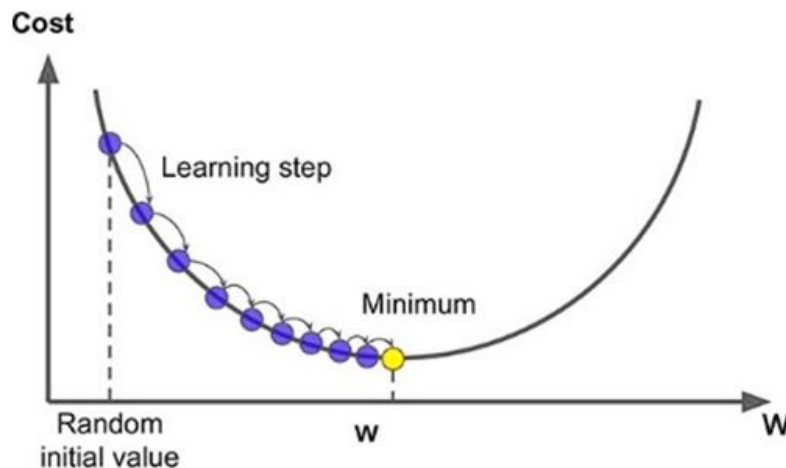
Osnovne formule Backpropagation algoritma

Korišćenje diferencijabilnih funkcija transfera (sigmoid), i primena pravila za izvod složenih funkcija omogućavaju prenošenje greške sa izlaza unazad, sa sloja na sloj.

$$w(k+1) = w(k) + u * dE/dw$$

$$dE/dw = dE/dy * dy/ds * ds / dw$$

Gradijentni spust (Gradient Descent)



$$w(k+1) = w(k) + u * dE/dw$$

U svakom koraku greška se smanjuje za izvesni procenat greške/odstupanja od ciljne vrednosti

Gradijentni spust u 2D sa ne-trivijalnom površinom

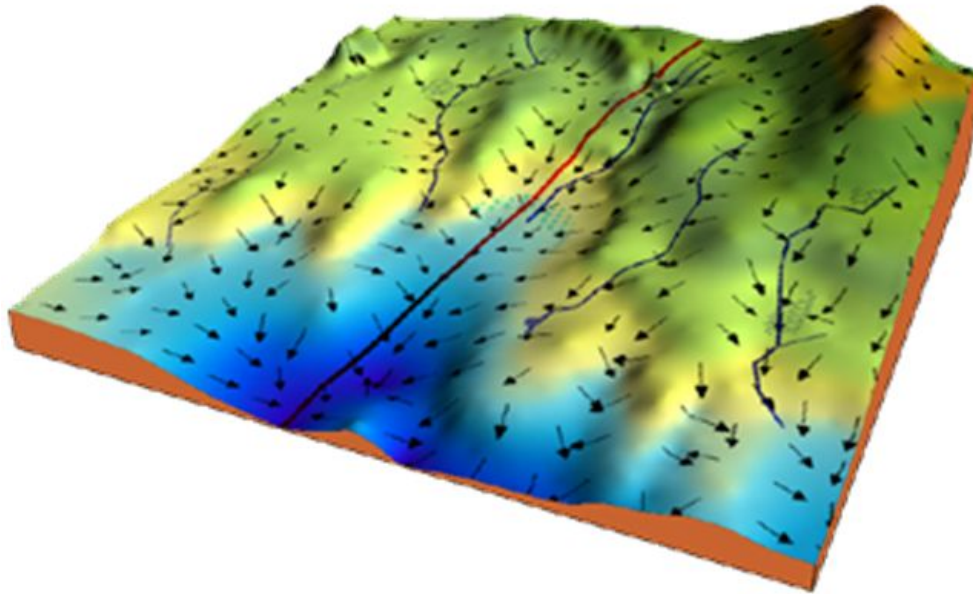
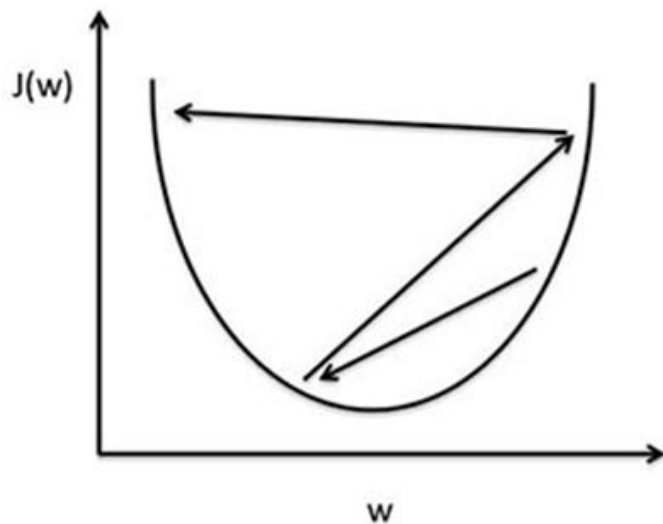


Image from: https://ml-cheatsheet.readthedocs.io/en/latest/gradient_descent.html

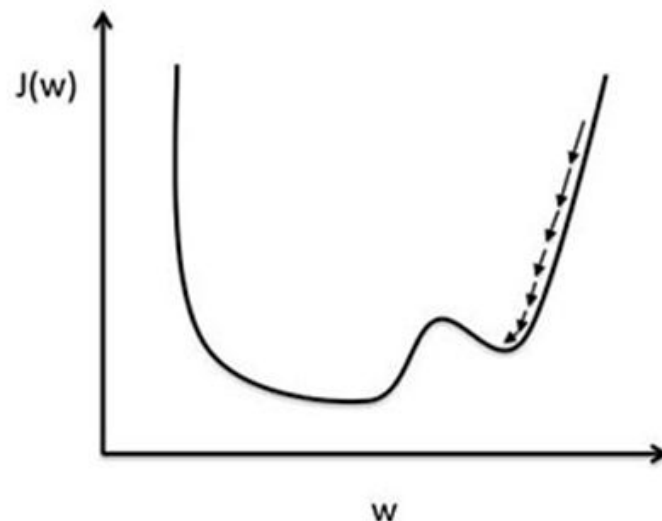
Osnovni parametri backpropagation algoritma

- **Epochs** – trening se zaustavlja posle zadatog broja iteracija/epoha učenja
- **Max error** – trening se zaustavlja kada ukupna greška padne ispod ove vrednosti
- **Learning rate** – koeficijent/korak učenja određuje u kojoj meri greška utiče na promenu težine

Uticaj koeficijenta/koraka učenja



Large Learning Rate



Small Learning Rate

Problemi Backpropagation algoritma

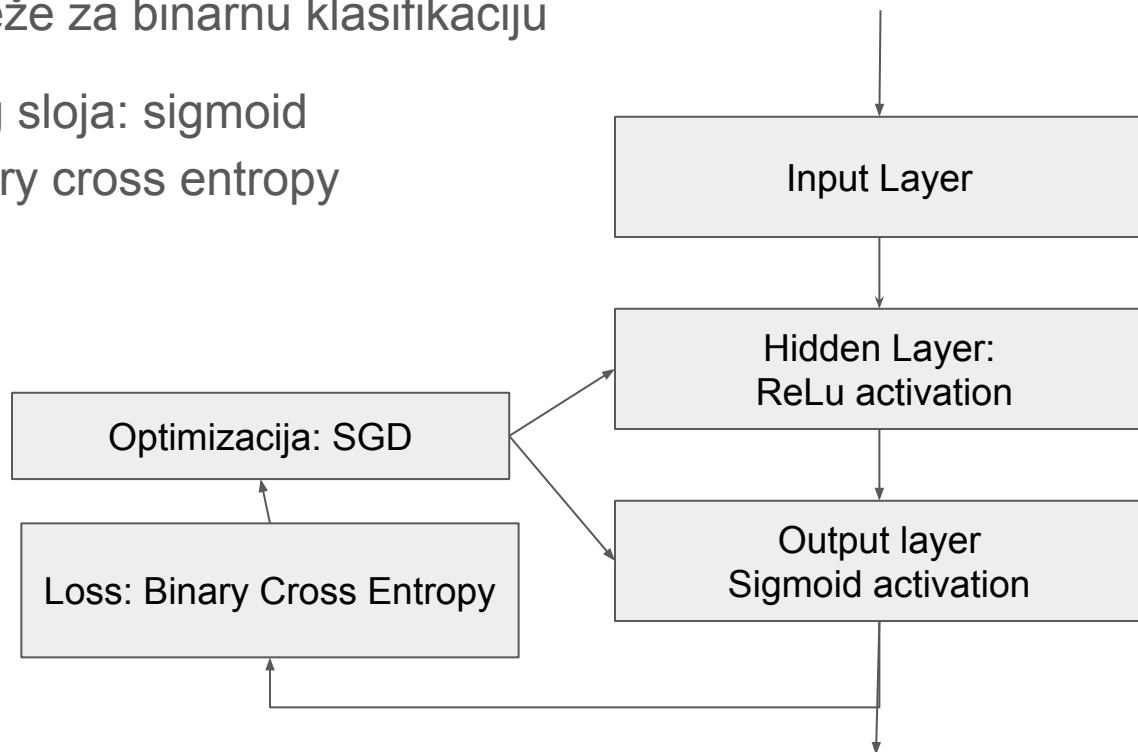
- Problem lokalnog minimuma
- Problem saturacije/nestajućih gradijenta
- Kada zaustaviti trening?
 - rast greške
 - suviše sporo učenje (smanjenje greške)
 - rano zaustavljanje - greška za validation set, izbegavanje overfittinga

Primena neuronske mreže za binarnu klasifikaciju

Konfiguracija neuronske mreže za binarnu klasifikaciju

Aktivaciona funkcija izlaznog sloja: sigmoid

Funkcija greške (Loss): Binary cross entropy

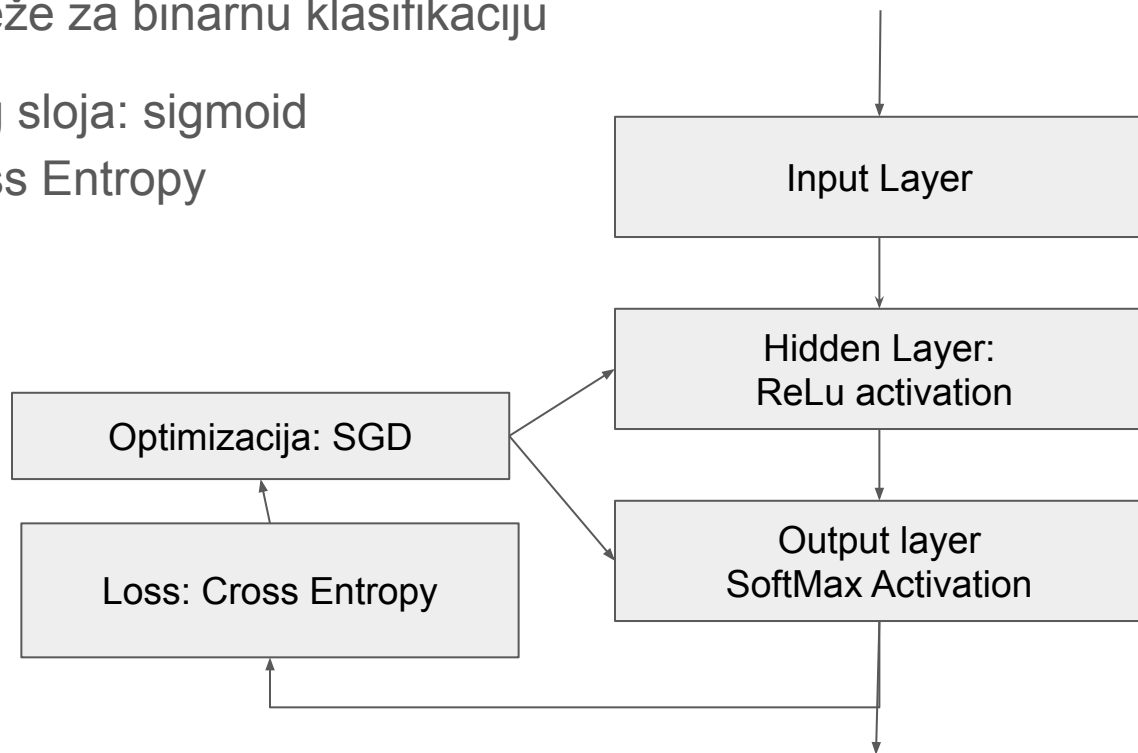


Primena neuronske mreže za višeklasnu klasifikaciju

Konfiguracija neuronske mreže za binarnu klasifikaciju

Aktivaciona funkcija izlaznog sloja: sigmoid

Funkcija greške (Loss): Cross Entropy

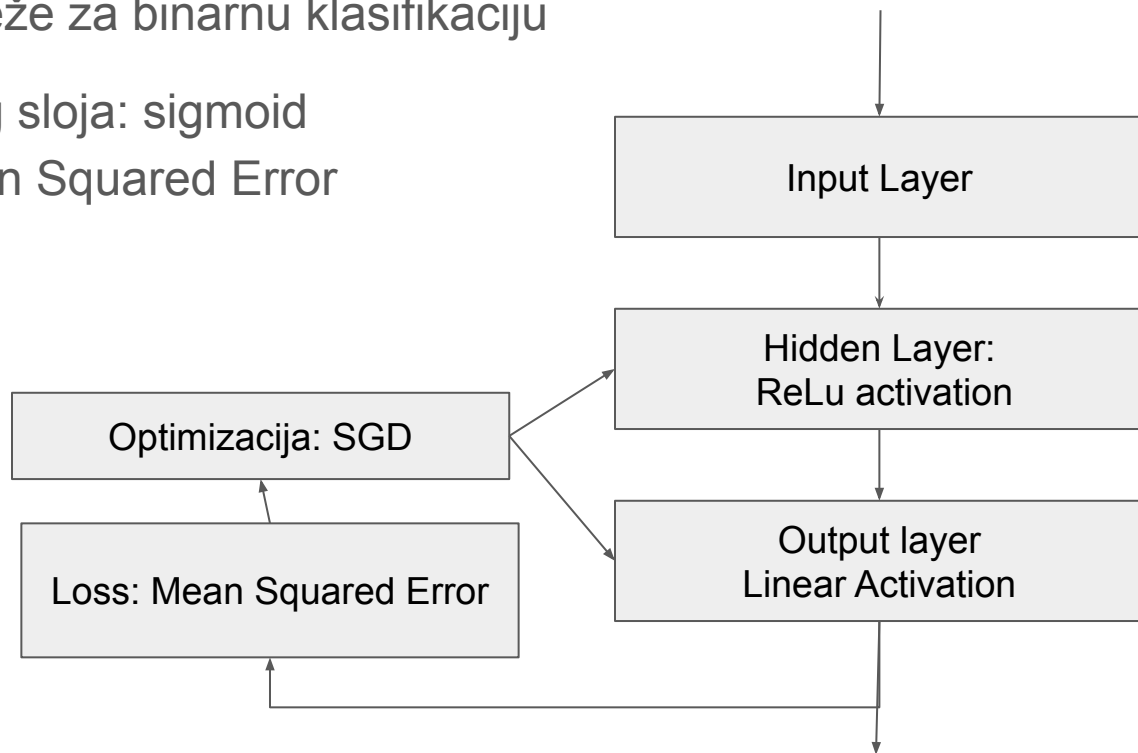


Primena neuronske mreže za regresiju

Konfiguracija neuronske mreže za binarnu klasifikaciju

Aktivaciona funkcija izlaznog sloja: sigmoid

Funkcija greške (Loss): Mean Squared Error



Prednosti i nedostaci neuronskih mreža

Prednosti:

- Jako dobre sa visokodimenzionalnim podacima
- Jako dobre sa velikom količinom podataka
- Široka primenljivost na tabelarne podatke, slike, zvuk, tekst

Nedostaci

- Nedostatak semantike u strukturi
- Teško objašnjiv
- Problemi sa određivanjem arhitekture, parametara algoritma za učenje i treningom za određenu primenu

Literatura

Neural Networks For Pattern Recognition, Christopher Bishop

Neural Networks and Deep Learning, free online book

<http://neuralnetworksanddeeplearning.com>

Deep learning, Ian Goodfellow, Yoshua Bengio, Aaron Courville

<https://www.deeplearningbook.org/>

[Aplikacioni okvir za razvoj neuronskih mreža](#), Zoran Ševarac, FON

Interaktivni online primeri

<https://playground.tensorflow.org>