

Generisanje Teksta LSTM od nule

Ivan Pop-Jovanov, 85/2018

Septembar 2022

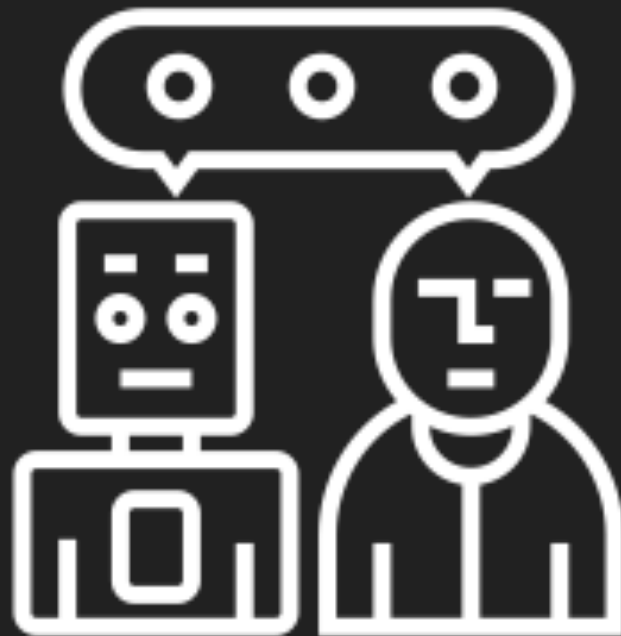
Matematički Fakultet, Univerzitet u Beogradu

Jezičko modelovanje

NLP - Presek lingvistike i veštačke inteligencije

Modeli jezika i slučajevi upotrebe

- Provera pravopisa
- Detekcija neželjene elektronske pošte
- Automatsko odgovaranje na pitanja
- Klasifikacija emocija teksta
- I mnoge druge...



Jezičko modelovanje

S = Where are we going



Previous words
(Context)



Word being
predicted

$$P(S) = P(\text{Where}) \times P(\text{are} \mid \text{Where}) \times P(\text{we} \mid \text{Where are}) \times P(\text{going} \mid \text{Where are we})$$

Sekvencijalna priroda

Pristup karakter po karakter

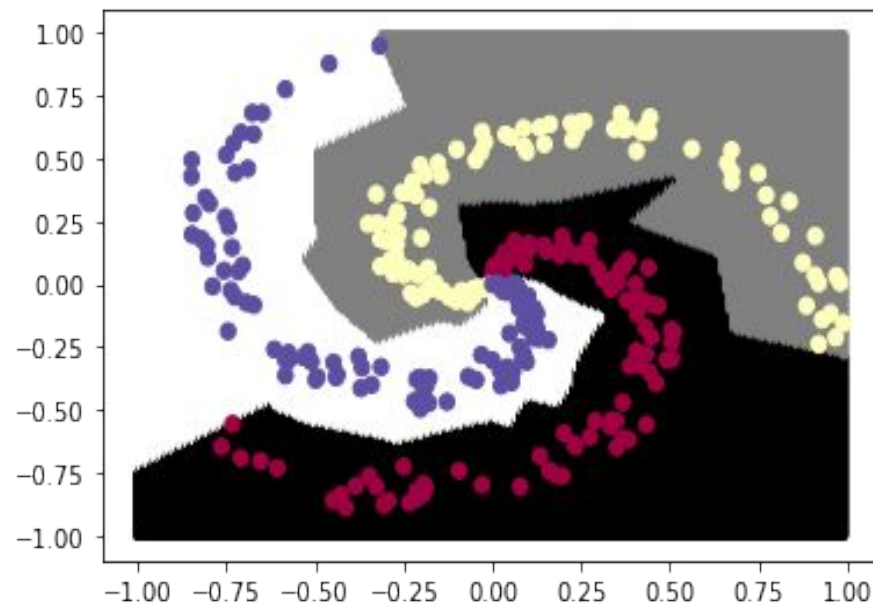
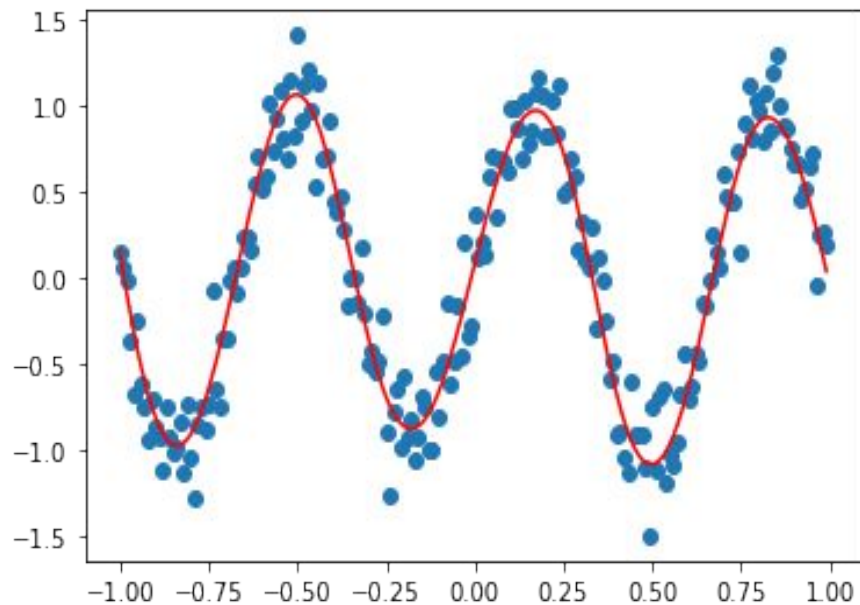
Pristup reč po reč

Bag of words

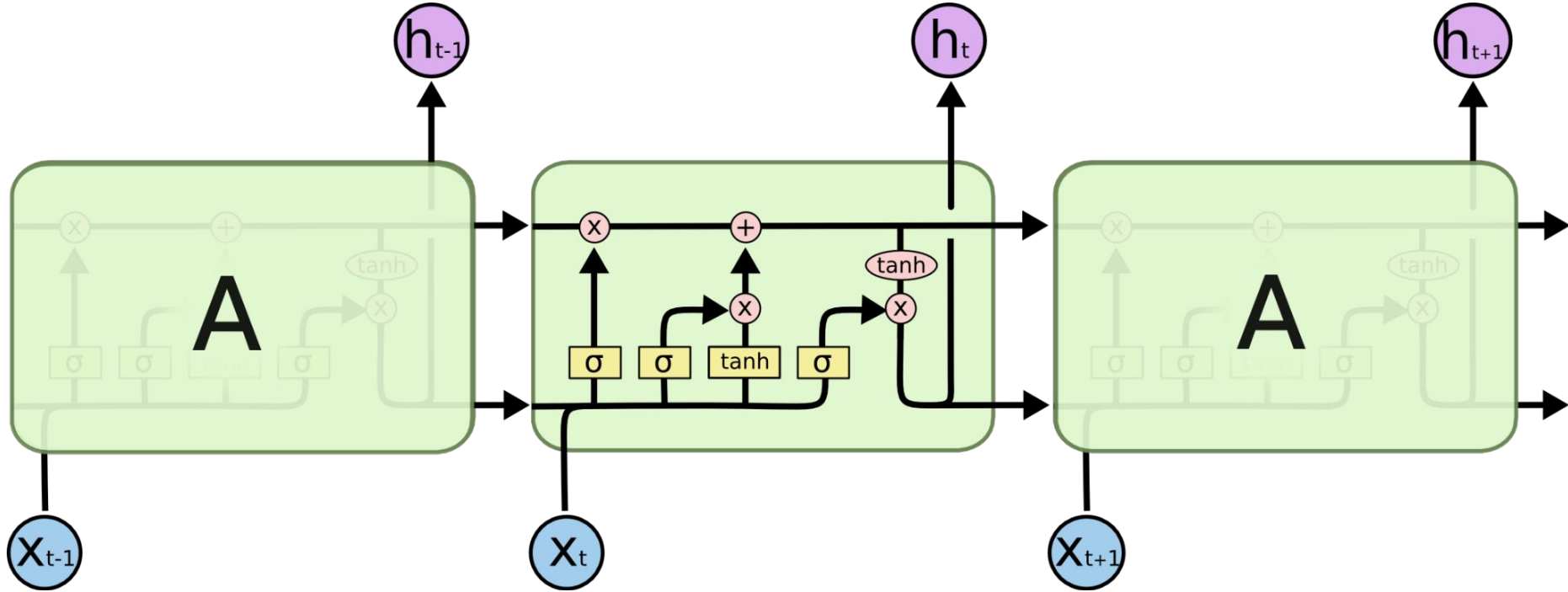
Vektor reči

word2vec

Moja Biblioteka



Long Short-Term Memory Unit



Model

```
optimizer = OptimizerAdam(learning_rate=0.001, batch_size = 128, loss=LossCrossEntropy())
```

```
hidden_size = 256
```

```
model = Sequential(optimizer)
```

```
model.add(LSTM(vocab_size, hidden_size, activation = ActivationLinear()))
```

```
model.add(Dense(hidden_size, hidden_size, activation = ActivationTanh()))
```

```
model.add(Dense(hidden_size, vocab_size, activation = ActivationSoftmax()))
```

Rezultati

"the moong has somewhat oo hir woing ho wounh saint in tn hised tide aiseso homedhinealy. He opent ahat hishing hosks but tir reseegas . Wuths he wonhing whan he wished yome had rome ent conessine and Monersnt entiess and Monk sing work, but the donn, hunlell, Cunseell an tranlenglangs an to frig, he wrdion to Mre hn oone and yry neng dook accinginaling in the gornt in whise saine iistigest ano monhid woniing and yo f Frie, hun equenteainglon hise and honement on tia were saditanle and made sigatlecc"

Hvala na pažnji!