

Implementacija povratne neuronske mreže i primjena na zadatku predviđanja sljedeće riječi

Ivan Martinović

Fakultet elektrotehnike i računarstva Sveučilišta u Zagrebu

5. srpnja 2021.



Fakultet
elektrotehnike i
računarstva

- ① Uvod i motivacija
- ② Implementacija
- ③ Skup podataka
- ④ Izgrađeni modeli i postignuti rezultati
- ⑤ Zaključak

- 1 Uvod i motivacija
- 2 Implementacija
- 3 Skup podataka
- 4 Izgrađeni modeli i postignuti rezultati
- 5 Zaključak

Uvod i motivacija

- Andrej Karpathy: The Unreasonable Effectiveness of Recurrent Neural Networks

Uvod i motivacija

- Andrej Karpathy: The Unreasonable Effectiveness of Recurrent Neural Networks
- Potpuno povezana unaprijedna slojevita neuronska mreža

Uvod i motivacija

- Andrej Karpathy: The Unreasonable Effectiveness of Recurrent Neural Networks
- Potpuno povezana unaprijedna slojevita neuronska mreža
- Algoritam propagacije pogreške unatrag

Uvod i motivacija

- Andrej Karpathy: The Unreasonable Effectiveness of Recurrent Neural Networks
- Potpuno povezana unaprijedna slojevita neuronska mreža
- Algoritam propagacije pogreške unatrag
- Povratna neuronska mreža

Uvod i motivacija

- Andrej Karpathy: The Unreasonable Effectiveness of Recurrent Neural Networks
- Potpuno povezana unaprijedna slojevita neuronska mreža
- Algoritam propagacije pogreške unatrag
- Povratna neuronska mreža
- Algoritam propagacije pogreške kroz vrijeme

① Uvod i motivacija

② Implementacija

Korištene tehnologije i alati

Razredi DenseLayer, RNNLayer i LSTMLayer

activations.py

optimizers.py

Razred CrossEntropyLoss

③ Skup podataka

④ Izgrađeni modeli i postignuti rezultati

⑤ Zaključak

① Uvod i motivacija

② Implementacija

Korištene tehnologije i alati

Razredi `DenseLayer`, `RNNLayer` i `LSTMLayer`

`activations.py`

`optimizers.py`

Razred `CrossEntropyLoss`

③ Skup podataka

④ Izgrađeni modeli i postignuti rezultati

⑤ Zaključak

Korištene tehnologije i alati



① Uvod i motivacija

② Implementacija

Korištene tehnologije i alati

Razredi `DenseLayer`, `RNNLayer` i `LSTMLayer`

`activations.py`

`optimizers.py`

Razred `CrossEntropyLoss`

③ Skup podataka

④ Izgrađeni modeli i postignuti rezultati

⑤ Zaključak

Razredi DenseLayer, RNNLayer i LSTMLayer

- DenseLayer – implementacija potpuno povezanog sloja
- RNNLayer – implementacija jednostavne povratne ćelije
- LSTMLayer – implementacija povratne ćelije s dugoročnim pamćenjem
- Unaprijedni prolaz – metoda forward
- Propagacija pogreške unatrag/kroz vrijeme - metoda backward

DenseLayer

```
class DenseLayer(object):

    def __init__(self, input_dim, output_dim, use_bias=True):
        sq = np.sqrt(1. / input_dim)
        self.use_bias = use_bias
        self.weights = np.random.uniform(-sq, sq, (output_dim, input_dim))
        if use_bias:
            self.bias = np.random.uniform(-sq, sq, output_dim)
        else:
            self.bias = np.zeros(output_dim)

        self.x_in = None

    def forward(self, x_in):
        self.x_in = x_in
        return np.einsum('b...i,ih->b...h', self.x_in, self.weights.T)
        + self.bias

    def backward(self, de_dy):

        axis = tuple(range(len(self.x_in.shape) - 1))
        de_dw = np.tensordot(de_dy, self.x_in, axes=(axis, axis))
        de_db = de_dy.sum(axis=axis)
        de_dx = np.einsum('b...h,hi->b...i', de_dy, self.weights)

        return de_dx, de_dw, de_db
```

① Uvod i motivacija

② Implementacija

Korištene tehnologije i alati

Razredi DenseLayer, RNNLayer i LSTMLayer

`activations.py`

`optimizers.py`

Razred CrossEntropyLoss

③ Skup podataka

④ Izgrađeni modeli i postignuti rezultati

⑤ Zaključak

activations.py

- Implementacija prijenosnih funkcija
- Razredi: TanH, ReLu, Sigmoid i Softmax
- Metode: forward, backward i backward_calculated

```
class Tanh(object):
```

```
    @staticmethod
    def forward(X_in):
        return np.tanh(X_in)

    @staticmethod
    def backward(X_in):
        return 1 - (np.tanh(X_in)) ** 2

    @staticmethod
    def backward_calculated(tanh_x_in):
        return 1 - tanh_x_in**2
```


① Uvod i motivacija

② Implementacija

Korištene tehnologije i alati

Razredi DenseLayer, RNNLayer i LSTMLayer

`activations.py`

`optimizers.py`

Razred CrossEntropyLoss

③ Skup podataka

④ Izgrađeni modeli i postignuti rezultati

⑤ Zaključak

optimizers.py

- Implementacija optimizatora
- Razredi: SGD, RMSProp i Adam
- Metoda `update_parameters`

```
class RMSProp(object):  
    def __init__(self, learning_rate=0.001, gamma=0.9, eta=1e-8):  
        self.learning_rate = learning_rate  
        self.gamma = gamma  
        self.last_momentum = []  
        self.eta = eta  
  
    def update_parameters(self, model_parameters, gradients):  
        for i in range(len(model_parameters)):  
            if len(self.last_momentum) < i + 1:  
                self.last_momentum.append(  
                    np.zeros_like(model_parameters[i]))  
  
            new_momentum = self.gamma * self.last_momentum[i]  
                + (1 - self.gamma) * (gradients[i] ** 2)  
  
            model_parameters[i] -= self.learning_rate  
                * (gradients[i] / (np.sqrt(new_momentum)  
                    + np.full_like(new_momentum, self.eta, dtype=np.double)))  
  
            self.last_momentum[i] = new_momentum
```

① Uvod i motivacija

② Implementacija

Korištene tehnologije i alati

Razredi DenseLayer, RNNLayer i LSTMLayer

activations.py

optimizers.py

Razred CrossEntropyLoss

③ Skup podataka

④ Izgrađeni modeli i postignuti rezultati

⑤ Zaključak

CrossEntropyLoss

- Funkcija gubitka – *unakrsna entropija*
- Metode: forward i backward

```
import numpy as np
from activations import Softmax

class CrossEntropyLoss(object):
    def __init__(self):
        self.y_pred = None

    def forward(self, y, o):
        self.y_pred = Softmax.forward(o)
        arr = np.array(y.shape[0:len(y.shape)-1])
        return np.sum(-y * np.log(self.y_pred + 1e-15))/
            (np.prod(arr))

    def backward(self, y):
        return (self.y_pred - y)/
            (np.prod(np.array(y.shape[0:len(y.shape)-1])))
```

① Uvod i motivacija

② Implementacija

③ Skup podataka

④ Izgrađeni modeli i postignuti rezultati

⑤ Zaključak

Skup podataka – DailyDialog

- DailyDialog dataset
- 13118 dijaloga
- Prosječan broj riječi po dijalogu: 114.7
- 10 tema
- Niska razina šuma među podacima

A: I'm worried about something.

B: What's that?

A: Well, I have to drive to school for a meeting this morning, and I'm going to end up getting stuck in rush hour traffic.

B: That's annoying, but nothing to worry about.

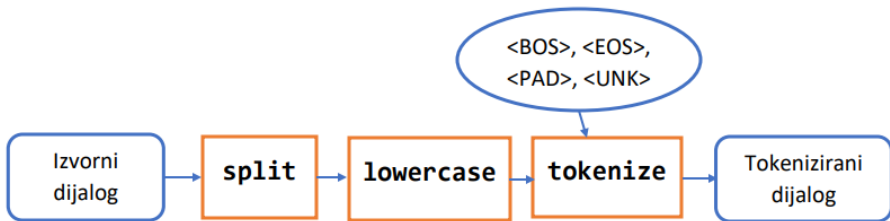
Just breathe deeply when you feel yourself getting upset.

A: Ok, I'll try that.

B: Is there anything else bothering you?

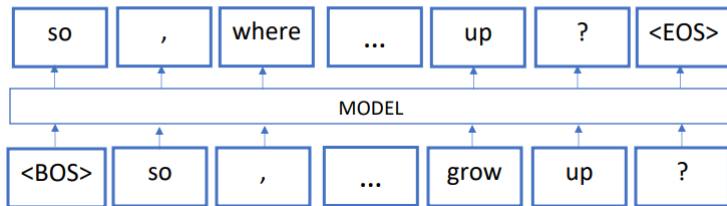
A: ...

Predobrada skupa podataka



Primjer zapisa iz skupa podataka nakon provedene predobrade

Rečenica	['<BOS>', 'so', ',', 'where', 'did', 'you', 'grow', 'up', '?', '<EOS>']
Ulaz	['<BOS>', 'so', ',', 'where', 'did', 'you', 'grow', 'up', '?']
Izlaz	['so', ',', 'where', 'did', 'you', 'grow', 'up', '?', '<EOS>']



① Uvod i motivacija

② Implementacija

③ Skup podataka

④ Izgrađeni modeli i postignuti rezultati

Klasifikatori `RNNClassifier` i `LSTMClassifier`

Mjere uspješnosti

Optimizacija hiperparametara

Učenje modela

Iskorištavanje modela

Generirajući model

⑤ Zaključak

① Uvod i motivacija

② Implementacija

③ Skup podataka

④ Izgrađeni modeli i postignuti rezultati

Klasifikatori `RNNClassifier` i `LSTMClassifier`

Mjere uspješnosti

Optimizacija hiperparametara

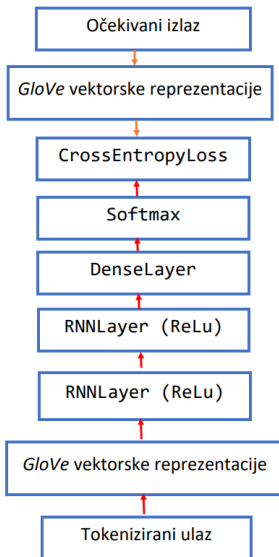
Učenje modela

Iskorištavanje modela

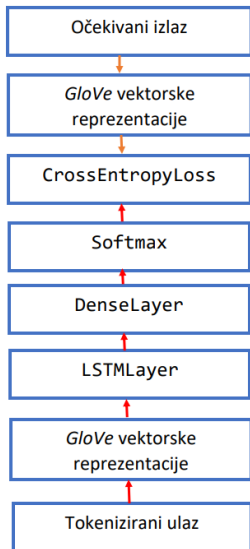
Generirajući model

⑤ Zaključak

Razred RNNClassifier



Razred LSTMClassifier



① Uvod i motivacija

② Implementacija

③ Skup podataka

④ Izgrađeni modeli i postignuti rezultati

Klasifikatori `RNNClassifier` i `LSTMClassifier`

Mjere uspješnosti

Optimizacija hiperparametara

Učenje modela

Iskorištavanje modela

Generirajući model

⑤ Zaključak

Mjere uspješnosti

Točnost k najboljih (engl. *accuracy at k*)

$$\text{acc@}k = \frac{N_k}{N}, \quad N - \text{broj primjera}$$

Mjere uspješnosti

Točnost k najboljih (engl. accuracy at k)

$$\text{acc@}k = \frac{N_k}{N}, \quad N - \text{broj primjera}$$

Mjera zbunjenosti (engl. perplexity)

$$\text{perplexity} = e^{\text{loss}} = e^{L(t_i, y_i)}$$

① Uvod i motivacija

② Implementacija

③ Skup podataka

④ Izgrađeni modeli i postignuti rezultati

Klasifikatori `RNNClassifier` i `LSTMClassifier`

Mjere uspješnosti

Optimizacija hiperparametara

Učenje modela

Iskorištavanje modela

Generirajući model

⑤ Zaključak

Optimizacija hiperparametara

- Optimizirani hiperparametri:
 - lr – stopa učenja
 - hd – dimenzionalnost skrivenog stanja
 - $batch\ size$ – veličina grupe
- Optimizacija u dva koraka
- Optimizator *Adam*

Optimizacija hiperparametara

- Optimizirani hiperparametri:
 - lr – stopa učenja
 - hd – dimenzionalnost skrivenog stanja
 - $batch\ size$ – veličina grupe
- Optimizacija u dva koraka
- Optimizator *Adam*

Model	lr	hd	$batch\ size$	$acc@10$	$perplexity$
RNNClassifier	0.003	100	128	0.629	45.271
LSTMClassifier	0.006	200	128	0.634	41.525

① Uvod i motivacija

② Implementacija

③ Skup podataka

④ Izgrađeni modeli i postignuti rezultati

Klasifikatori `RNNClassifier` i `LSTMClassifier`

Mjere uspješnosti

Optimizacija hiperparametara

Učenje modela

Iskorištavanje modela

Generirajući model

⑤ Zaključak

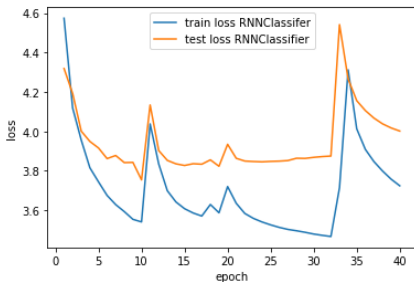
Učenje modela

- Podjela skupa podataka
- Hiperparametri iz prethodnog koraka
- 40 epoha
- Veličina vokabulara: 10000
- Dimenzija vektorskih reprezentacija: 100

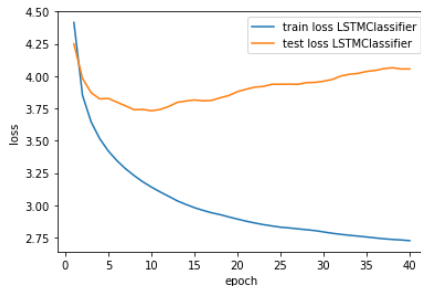
Model	Broj epohe	Točnost 10 najboljih	Mjera zbunjenosti
LSTMClassifier	10	0.637	44.401
RNNClassifier	10	0.624	46.661

Prosječni iznos funkcije gubitka na skupu za učenje i skupu za testiranje kroz epohe

- Model RNNClassifier

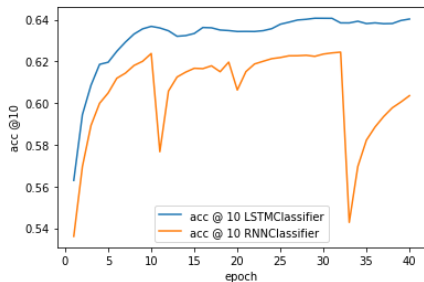


- Model LSTMClassifier

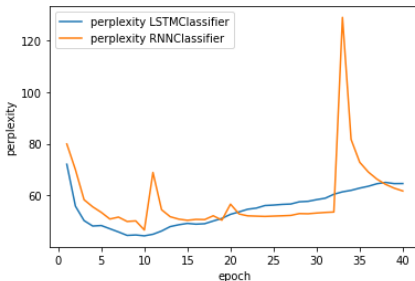


Usporedba modela LSTMClassifier i RNNClassifier kroz epohe prema *točnosti 10 najboljih i mjeri zbunjenosti*

- *točnost 10 najboljih*



- *mjera zbunjenosti*



① Uvod i motivacija

② Implementacija

③ Skup podataka

④ Izgrađeni modeli i postignuti rezultati

Klasifikatori `RNNClassifier` i `LSTMClassifier`

Mjere uspješnosti

Optimizacija hiperparametara

Učenje modela

Iskorištavanje modela

Generirajući model

⑤ Zaključak

Primjeri predviđanja sljedeće riječi – ulazna rečenica i tri najvjerojatnije riječi na izlazu

Model	Ulazna rečenica	Predviđeni izlaz
LSTMClassifier	['are', 'you', 'having', 'a', 'good']	['job', 'impression', 'day']
RNNClassifier	['are', 'you', 'having', 'a', 'good']	['time', 'idea', 'job']
LSTMClassifier	['how', 'is', 'your', 'day']	[',', '?', 'after']
RNNClassifier	['how', 'is', 'your', 'day']	['?', 'going', ';']
LSTMClassifier	['could', 'you', 'say', 'that', 'again', ',', '']	['we', 'please', 'and']
RNNClassifier	['could', 'you', 'say', 'that', 'again', ',', '']	['i', 'and', 'you']
LSTMClassifier	['thank', 'you', 'so', 'much', 'for', 'your']	['job', 'company', 'family']
RNNClassifier	['thank', 'you', 'so', 'much', 'for', 'your']	['parents', 'wife', 'attitude']
LSTMClassifier	['i', 'am', 'well', ',', 'how', 'are']	['you', 'they', 'we']
RNNClassifier	['i', 'am', 'well', ',', 'how', 'are']	['you', 'the', 'we']
LSTMClassifier	['you', 'have', 'a', 'great']	['time', 'progress', 'idea']
RNNClassifier	['you', 'have', 'a', 'great']	['day', 'favor', 'time']

① Uvod i motivacija

② Implementacija

③ Skup podataka

④ Izgrađeni modeli i postignuti rezultati

Klasifikatori `RNNClassifier` i `LSTMClassifier`

Mjere uspješnosti

Optimizacija hiperparametara

Učenje modela

Iskorištavanje modela

Generirajući model

⑤ Zaključak

Generirajući model izgrađen pomoću modela LSTMClassifier

- Arhitektura *sljed u sljed*
- Jednostavna strategija generiranja

Generirajući model izgrađen pomoću modela LSTMClassifier

- Arhitektura *slijed u slijed*
- Jednostavna strategija generiranja

i ' ll just bring me up .

do we see , sir . what can you suggest that you are in touch ? we ' ve never had to make a new - exchange options in the company for certain each other .

we ' ve lost it on a couple - free - up call you , and i ' re to the company for your first time , especially in your country is problem , but it will help you ? you should be very familiar with them ! do you think you can do anything ?

yes ! what ' s the problem , mr ? i ' ve always had the best friends to be honest .

well , it was excellent to be the most important . we should try on you .

i have been feeling about a month , but it is the most suitable.

① Uvod i motivacija

② Implementacija

③ Skup podataka

④ Izgrađeni modeli i postignuti rezultati

⑤ Zaključak

Zaključak

- Izazovi pri izradi završnog rada

Zaključak

- Izazovi pri izradi završnog rada
- Moguća poboljšanja

Zaključak

- Izazovi pri izradi završnog rada
- Moguća poboljšanja
- Budući rad

Zahvala

- izv. prof. dr. sc Jan Šnajder
- mag. ing. Josip Jukić

Reference

- <http://karpathy.github.io/2015/05/21/rnn-effectiveness/>
- <https://aclanthology.org/I17-1099/>
- <https://www.python.org/>
- <https://pytorch.org/>
- <https://takelab.fer.hr/podium/>
- <https://numpy.org/>