

Generiranje naselja pomoću LSTM mreže

Čogelja, Granić, Lubina, Jurković, Juvančić, Logarušić

21. listopada 2024.

prijedlog

(I.) Ishod projekta

Ishod projekta je LSTM rekurzivna neuronska mreža na razini znakova koja generira realistična imena hrvatskih naselja.

Mreža radi sa vektorima koji predstavljaju slova hrvatske abecede proširene specijalnim znakovima $\Sigma = \{\text{hrv. abeceda}\} \cup \{\langle start \rangle, " \setminus 0" \}$. Ulaz mreže je one-hot vektor $\mathbf{x}^{(t)}$ dimezije $|\Sigma| = 32$.

$$\mathbf{x}_i^{(t)} = \begin{cases} 1, & \text{ako } i = j \\ 0, & \text{inače} \end{cases} \quad (1)$$

Izlaz dobiven na kraju pojedinog vremenskog koraka t je vektor vjerojatnosti pojave pojedinog znaka abecende.

$$\hat{\mathbf{y}}^{(t)} = \begin{bmatrix} p(c_0) \\ p(c_1|c_0) \\ \vdots \\ p(c_{|\Sigma|-1} | \bigcap_{i=0}^{|\Sigma|-2} c_i) \end{bmatrix} \quad \text{Gdje } c \in \Sigma \quad (2)$$

Vjerojatnosti su dobivene softmax funkcijom parametriziranom hiperparametrom temperature τ .

Na temelju tih vjerojatnosti se uzorkuje konačni izlazni vektor $\mathbf{y}^{(t)}$, odnosno t-ti znak u imenu naselja.

$$\mathbf{y}^{(t)} \sim \sigma_{\tau}(\hat{\mathbf{y}}^{(t)}) \quad (3)$$

Temperaturno uzorkovanje je izabrano, jer omogućava eksperimentiranje i generiranje zanimljivih toponima.

Izlaz mreže je niz znakova $\{\mathbf{y}^{(t)}\}_{t=0}^{T-1}$, odnosno ime naselja.

(II.) Tema i kratki opis

Tema projekta je generiranje realističnih imena hrvatskih naselja.

U tu svrhu će se izraditi LSTM mreža parametrizirana sljedećim hiperparametrima:

1. Dimenzije skrivenog stanja: $|\mathbf{a}|$
2. Stopa učenja: μ
3. Temperatura: τ
4. Broj LSTM jedinica

(III.) Zadatci na projektu i raspodjela posla

Ostvarenje projekta podrazumijeva slijedeće zadatke:

	ime	opis
pisani rad	kratki uvod	N/A
	opis problema	N/A
	opis eksperimentalnih rezultata	N/A
	diskusija i usporedba rezultata	N/A
	lektoriranje	N/A
	zaključak	N/A
administrativni poslovi		Lubina
izrada prezentacije		N/A
implementacija	priprema skupa podataka	N/A
	implementacija mreže	N/A
	implementacija uzorkovanja	N/A
treniranje	implementacija algoritma učenja	N/A
	treniranje	N/A
validacija	ručna validacija modela	N/A
	podešavanje hiperparametara	N/A

Tablica 1: zadatci i raspored

Mreža će biti implementirana pomoću TensorFlow radnog okvira.
Podešavanje hiperparametara se odvija paralelno sa implementacijom mreže u radnom okviru Keras.