

1. Opis wyboru optymalnej architektury:

Model był trenowany iteracyjnie to znaczy kryterium stopu była zadana liczba epok. Na wejściu model otrzymywał sekwencję znaków w postaci ich liczb-identyfikatorów ze słownika, następnie tworzył wielowymiarowe osadzenia (nn.Embedding), które przekazywane były do sieci LSTM, wyniki tej sieci następnie były przekazywane do warstwy liniowej, która rzutowała je na identyfikatory znaków z ze słownika, tzn. Sieć zwracała wektor o rozmiarze naszego słownika i indeks, na którym była największa wartość to identyfikator przewidywanego znaku. Do treningu użyto optymalizatora Adam ze współczynnikiem learning_rate=0.0001 oraz funkcji celu CrossEntropy.

Do wyboru najlepszych hiperparametrów zastosowano autorską metodę na wzór GridSearch, która tworzy wszystkie możliwe modele na podstawie zadanych list parametrów i testuje je na zbiorze testowym za pomocą miary accuracy. Trening modelu kończył się po zadanej jako parameter liczbie epok. Rozważane hiperparametry:

EMBEDDING_DIM = (rozmiar osadzenia w module nn.Embedding); HIDDEN_DIM = (rozmiar osadzenia w module nn.LSTM); LSTM_LAYERS = (ilość warstw w module nn.LSTM); EPOCHS = (liczba epok); SEQUENCE_LENGTH = (długość przetwarzanej sekwencji)

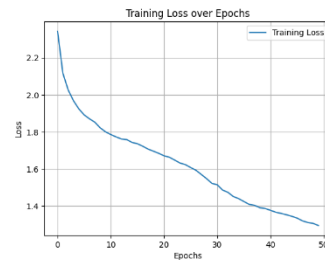
Wyniki accuracy zbioru testowego dla badanych parametrów:

EMBEDDING_DIM	SEQUENCE_LENGTH	HIDDEN_DIM	LSTM_LAYERS	EPOCHS	ACCURACY
50	50	128	1	10	40.90
50	50	128	1	50	40.72
50	50	128	2	10	41.59
50	50	128	2	50	48.23
50	50	256	1	10	43.55
50	50	256	1	50	47.90
50	50	256	2	10	46.16
50	50	256	2	50	45.33
100	50	128	1	10	42.45
100	50	128	1	50	47.59
100	50	128	2	10	43.41
100	50	128	2	50	48.50
100	50	256	1	10	45.54
100	50	256	1	50	47.56
100	50	256	2	10	46.70
100	50	256	2	50	45.43
50	100	128	1	10	40.81
50	100	128	1	50	47.75
50	100	128	2	10	46.75
50	100	128	2	50	48.47
50	100	256	1	10	44.39
50	100	256	1	50	48.02
50	100	256	2	10	46.05
50	100	256	2	50	45.93
100	100	128	1	10	41.73
100	100	128	1	50	47.81
100	100	128	2	10	43.17
100	100	128	2	50	48.68
100	100	256	1	10	45.58
100	100	256	1	50	48.17
100	100	256	2	10	46.98
100	100	256	2	50	46.24

Optymalny model: EMBEDDING_DIM = 100; HIDDEN_DIM = 128;

LSTM_LAYERS = 2; EPOCHS = 50; SEQUENCE_LENGTH = 100

Wykres funkcji celu dla zbioru treningowego dla optymalnego modelu:



2. Wyniki dla optymalnego modelu i porównanie wygenerowanych

tekstów

Wyniki:

```
Accuracy: 0.4867890649548862
Precision(weighted): 0.48908306407187346
Precision(macro): 0.3820605212679732
Recall(weighted): 0.4867890649548862
Recall(macro): 0.26843218704684724
F1(weighted): 0.4743807123556834
F1(macro): 0.2862424996848985
```

Porównanie wygenerowanego tekstu pomiędzy EMBEDDING_DIM równym 50 a 100:

Input : "zostaniesz przez resztę żywota
bardzo samotny, stary, wdowiec i sierota!
powiedz, czym ci ten drogi"

Poprawny output(100): "dar mam wynagrodzić

I czym twoje sieroctwo i wdowstwo osłodzić?"

"Czy ja Cybulski? – rzecze na to Kl"

Output(100): "z pole starego szabał,
i w polskie stary podkomorzy w stole podkonu,
i z polak wielkie stary pod st"

Output(100): "serwis potem –
rzekł sędzia – ja przeproszę wojskiego soplicy!
koto już minni! jesteś to mi serca n"

Porównanie wygenerowanego tekstu pomiędzy HIDDEN_DIM równym 128 a 256:

Input: "esz dla mnie nic oprócz dobroci,
możem te zaręczyny czas jakiś odwlekać;
wzięć twej woli nie chcę;"

Poprawny output(100): "będziem, Zosiu, czekać.

Nic nas nie nagli, zwłaszcza że wczora wieczorem

Dano mi rozkaz zostać w Li"

Output(100): jako ona biała
wywodzi na koniec – rzekł na to rzekł do stoliku.
daremnie koniec, na stolnika wypowi

Output(100): "jako ona biała
wywodzi na koniec – rzekł na to rzekł do stolicy
dwa ręką ten wojski w głowę od pani"

Porównanie wygenerowanego tekstu pomiędzy LSTM_LAYERS równym 1 a 2:

Input: "hta, dobrzyńscy! ja nie będę radził
nie a nic; powiem tylko, po com was zgromadził:
a co robić, jak"

Poprawny output(100): "robić, decydujcie sami.

Wście, słuch dawno chodzi między zaściankami,

Że się na wielkie rzeczy zan"

Output(100): "wielki zawsze szlachta w stronie
i z polak wielkie star wielki zawsze szlachta w stronie
i z polak wielkie stary pod strony obrazy.

więc się z tyłu za"

Output(100): "po obiedzi wojenne!
jeśli po wierzchu dobry do stolicem wiele.

już tam na stolniku i tak milej"

QuoVadis:

Input: "rozdział pierwszy

petroniusz obudził się zaledwie koło południa i jak zwykle, znęcony bardzo. popr"

Poprawny output(100): "zedniego dnia był na uczcie u Nerona, która przeciągnęła się do późna w noc. Od pewnego czasu zdrowi"

Output(100): "awiał w stroniczneml

i w polskie stary pod stolek pod nlebo strony,

i z polak wielkie stary pod sto"

Input: "e słynny współnik sewerusa, celer, rozszerzył mu, przebudował i urządził z tak nadzwyczajnym smakiem"

Poprawny output(100): "iż sam Nero przyznawał im wyższość nad cezariańskimi, chociaż cezariańskie były obszerniejsze i ur"

Output(100): "w strony z polewania – z podarne strony w polowanie – to pod stolekiem w polowanie – pod stolekiej"

Input: "gemny nie przyszły. petroniusz znów przynął oczy i wydał rozkaz, by przeniesiono go do tepidarium,"

Poprawny output(100): żądzone z nierównie większym przepychem. Po owej więc uczcie, na której znudziwszy się blaznowaniem"

Output(100): "

nie wierz mi stary podkomorzy w stolnik na pole,

i z polak wielkie stary pod stole pod strony

i z p"

Tekst angielski podawany na wejście:

Input: "minister and deputy interior minister inside the presidential palace in warsaw in a day of unprecede"

Poprawny output(100): "nted political theatre. Mariusz Kaminski and Maciej Wasik were sentenced to two years jail last mont"

Output(100): "nowanie – pod stolekiem wyrzekł pod stolek pod stolek – pod stolekiem wyrzekł pod soplicowie – pod s"

Input: "terior minister inside the presidential palace in warsaw in a day of unprecedented political theatre"

Poprawny output(100): ". Mariusz Kaminski and Maciej Wasik were sentenced to two years jail last month for abuse of power w"

Output(100): "m o to zabawia – z podniesione pod stolekiem w polowanie – rzekł się z polego w polskiej pod stolek"

Input: "stice (pis) party and the new pro-eu coalition.

the men, who were elected pis mps in october, refus"

Poprawny output(100): "ed to recognise last months court decision because President Andrzej Duda, a PIS ally, pardoned them"

Output(100): "ze,

zabawiała się w polskiej pod stolek pod strony

i z polaki starego zawsze strony oderzem,

na konti"

Obserwując wyniki wszystkich rozważanych modeli możemy wyciągnąć następujące wnioski:

- Długość sekwencji 50, a 100 nie miała większego znaczenia na wyniki
- Lepiej sprawdzały się modele z 2 warstwami LSTM
- Lepiej sprawdzał się EMBEDDING_DIM = 100 niż 50
- Lepiej sprawdzał się HIDDEN_DIM = 128 niż 256