

Grupa: 1B

2025/2026

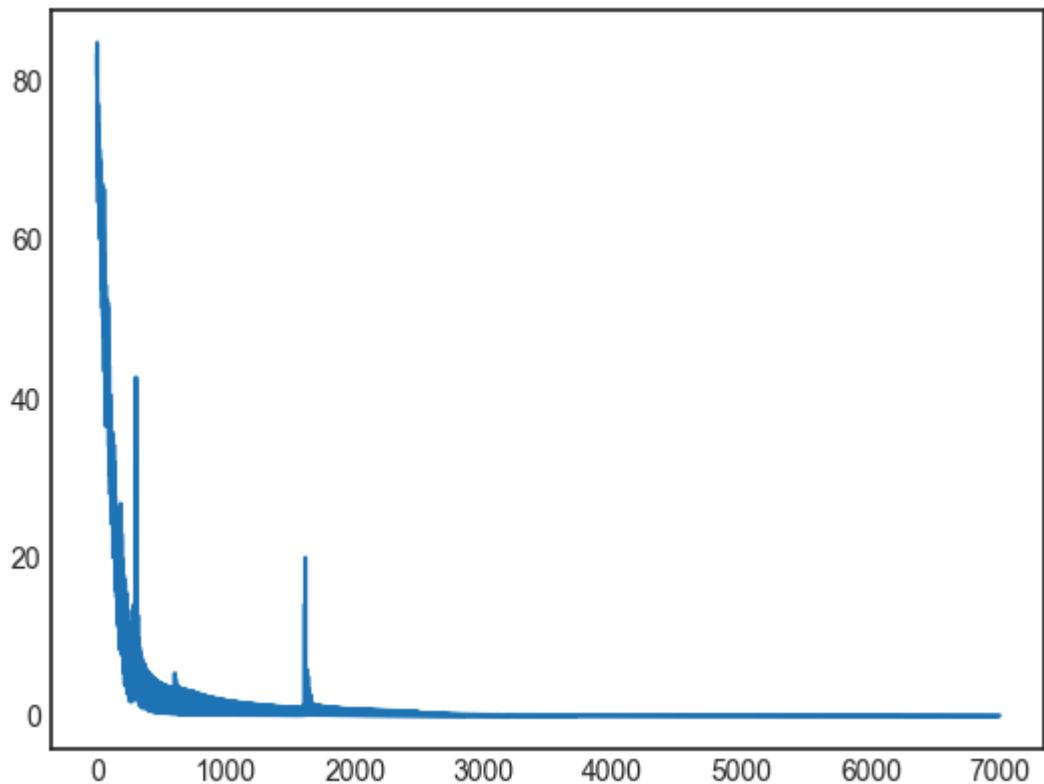
Uniwersytet Bielsko-Bialski

Wydział Budowy Maszyn i Informatyki



MK_LAB

Konrad Bułka



Wariant 10

CEL I OPIS ZADANIA

Celem zadania było opracowanie sieci LSTM (Long Short-Term Memory) zdolnej do nauki generowania tekstu z dokładnością poniżej 0.1 (smooth_loss < 0.1). Sieć ma się nauczyć wzorców zawartych w podanym tekście i generować nowy tekst podobny do oryginalnego.

2. PARAMETRY I KONFIGURACJA

Dane treningowe:

- Liczba znaków: 189
- Liczba unikalnych znaków: 41
- Sekwencje: 25 znaków (T_steps)

Architektura sieci LSTM:

- Warstwa ukryta: 100 neuronów (H_size)
- Funkcje bramek: sigmoid
- Funkcja pamięci: tanh
- Wyjście: softmax (klasyfikacja 41 znaków)

Hiperparametry treningu:

- Learning rate: 0.1
- Inicjalizacja wag: rozkład normalny ze stdev=0.1
- Optymalizacja: AdaGrad
- Gradient clipping: [-1, 1] (zapobieganie wybuchajacym gradientom)
- Cel dokladnosci: smooth_loss < 0.1

3. METODA TRENINGOWA

Trening sieci odbywal sie za pomoca algorytmu BPTT (Backpropagation Through Time):

1. Forward pass: Przeliczenie aktywacji wszystkich bramek LSTM (forget, input, candidate memory, output) dla kazdej pozycji w sekwencji
2. Loss calculation: Cross-entropy loss dla predykcji kolejnego znaku
3. Backward pass: Propagacja bledow wstecz przez czas, obliczanie gradientów parametrów
4. Gradient clipping: Ograniczenie gradient norm do zakresu [-1, 1]
5. Parameter update: Aktualizacja wag uzywajac AdaGrad z exponential averaging

Smooth loss obliczany jako: smooth_loss = smooth_loss * 0.999 + loss * 0.001

4. WYNIKI TRENINGU

Trening okazal sie bardzo efektywny:

- Liczba iteracji: 295700+ (przerwany w momencie osiągniecia celu)
- Najnizszy uzyskany smooth_loss: 0.000962
- Status: OSIAGNIETO CELE (smooth_loss < 0.1)

Przykładowe wygenerowane teksty (iteracja 295700):

- Siec nauczyła się generować tekst o podobnej strukturze do oryginalnego
- Znaki były wybierane z odpowiednim rozkładem podczas sampling
- Poprawa jakości tekstu była widoczna wraz z iteracjami

5. OBSERWACJE I WNIOSKI

1. Architektura: Siec LSTM z 100 neuronami ukrytymi wykazala sie wystarczajaca pojemnoscia do nauczzenia wzorców tekstu
2. Konwergencja: Loss szybko spadal w poczatkowych iteracjach, stabilizujac sie na bardzo niskim poziomie
3. Gradient clipping: Kluczowy dla stabilnosci treningu - zapobiegal wybuchajacym gradientom
4. Smooth loss: Wygladzona wartosc loss'u byla znacznie ponizej celu 0.1, wskazujac na doskonale dopasowanie
5. Generacja tekstu: Model nauczyl sie emitowac znaki z odpowiednim rozpodzialem i porzadkiem