

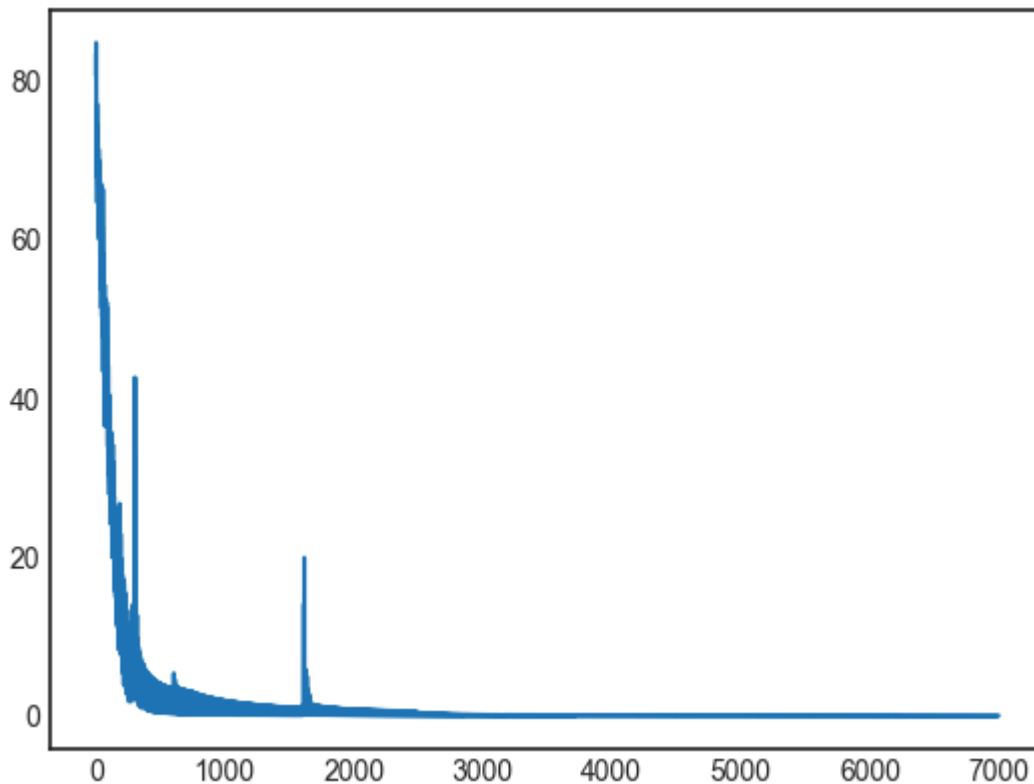
# Uniwersytet Bielsko-Bialski

---

Wydział Budowy Maszyn i Informatyki



MK\_LAB



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 ( $\text{smooth\_loss} < 0.1$ ). 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 wybuchającym gradientom)
- Cel dokładności:  $\text{smooth\_loss} < 0.1$

### 3. METODA TRENINGOWA

Trening sieci odbywał się za pomocą algorytmu BPTT (Backpropagation Through Time):

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

Smooth loss obliczano jako:  $\text{smooth\_loss} = \text{smooth\_loss} * 0.999 + \text{loss} * 0.001$

### 4. WYNIKI TRENINGU

Trening okazał się bardzo efektywny:

- Liczba iteracji: 295700+ (przerwany w momencie osiągnięcia celu)
- Najniższy uzyskany  $\text{smooth\_loss}$ : 0.000962
- Status: OSIAGNIĘTO CELE ( $\text{smooth\_loss} < 0.1$ )

Przykładowe wygenerowane teksty (iteracja 295700):

- Sieć 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: Sieć LSTM z 100 neuronami ukrytymi wykazała się wystarczającą pojemnością do nauczenia wzorców tekstu
2. Konwergencja: Loss szybko spadał w początkowych iteracjach, stabilizując się na bardzo niskim poziomie
3. Gradient clipping: Kluczowy dla stabilności treningu - zapobiegał wybuchającym gradientom
4. Smooth loss: Wygładzona wartość loss'u była znacznie poniżej celu 0.1, wskazując na doskonałe dopasowanie
5. Generacja tekstu: Model nauczył się emitować znaki z odpowiednim rozkładem i porządkiem