
Sprawozdanie laboratorium 3

Maciej Urbaniak

Abstract

Celem ćwiczenia jest zapoznanie się z budową i uczeniem sieci LSTM w bibliotece Tensorflow. W ćwiczeniu należy zbudować sieć LSTM, która jest w stanie na podstawie imienia określić płeć dziecka.

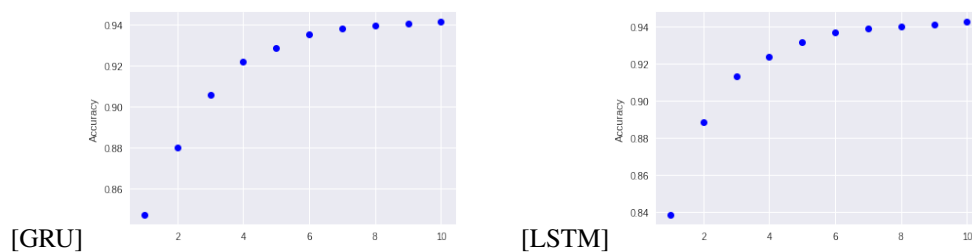
1 Model podstawowy

Parametry i struktura modelu podstawowego zostały dobrane zgodnie z wytycznymi zadania. Dodatkowo wykorzystano poniższe hiperparametry:

Batch size	200
AdamOptimizer	learning rate:0.001

2 LSTM GRU

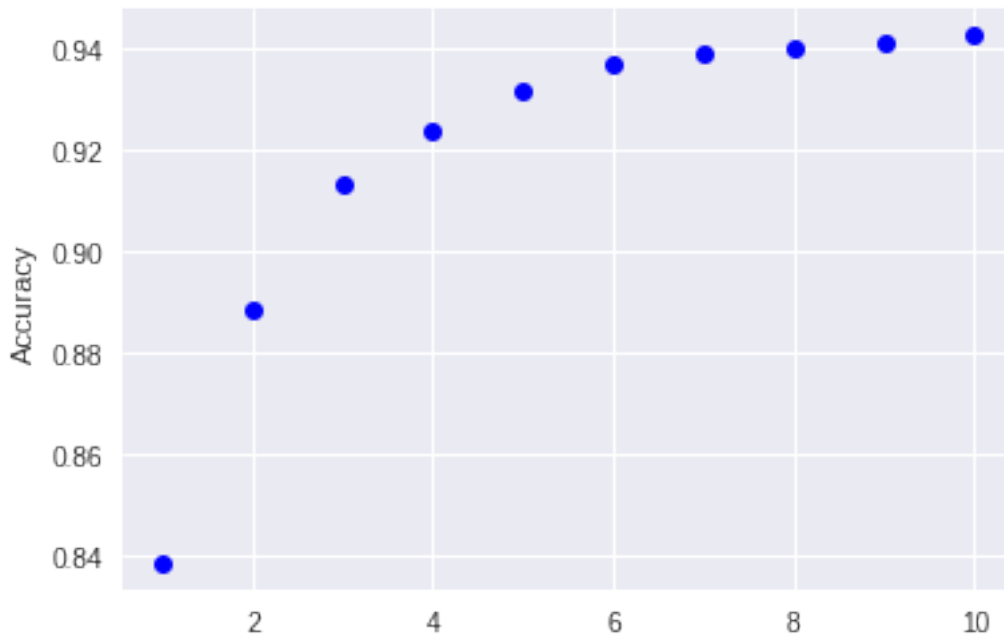
Porównanie LSTM z jedną warstwą i GRU. Jak widać na poniższych obrazkach różnice w celności są minimalne.



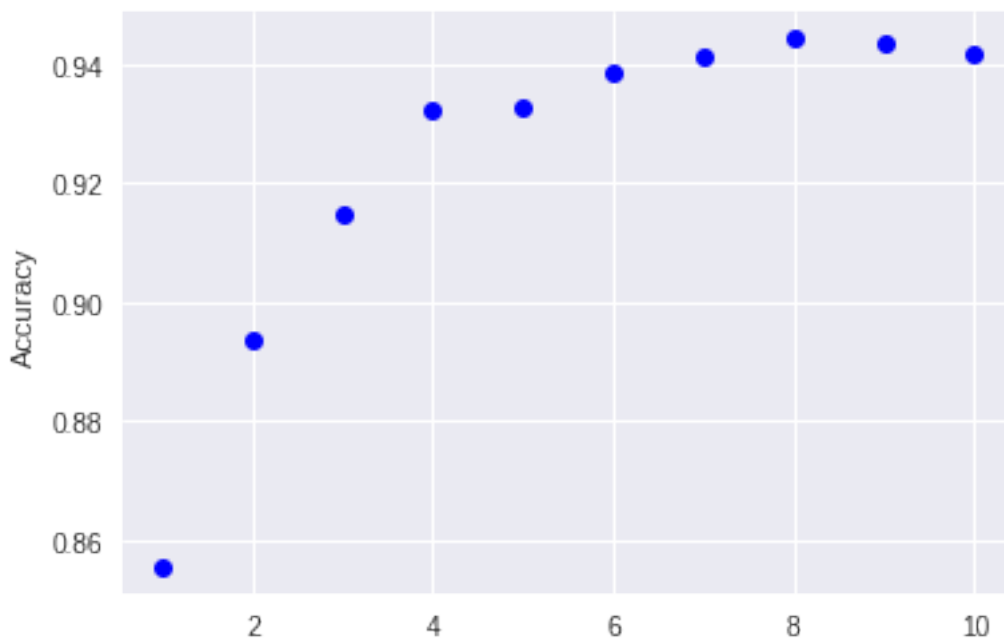
Rysunek 1: Porównanie celności modeli

3 Wpływ liczby warstw ukrytych na wyniki

Zwiększenie liczby warstw ukrytych z pewnością wpływa na czas obliczeń. Nastąpiła również poprawa celności na poziomie mniejszym od 1 punktu procentowego oraz wzrosła szybkość uczenia się modelu dla początkowych epok, co można zobaczyć na poniższych rysunkach.



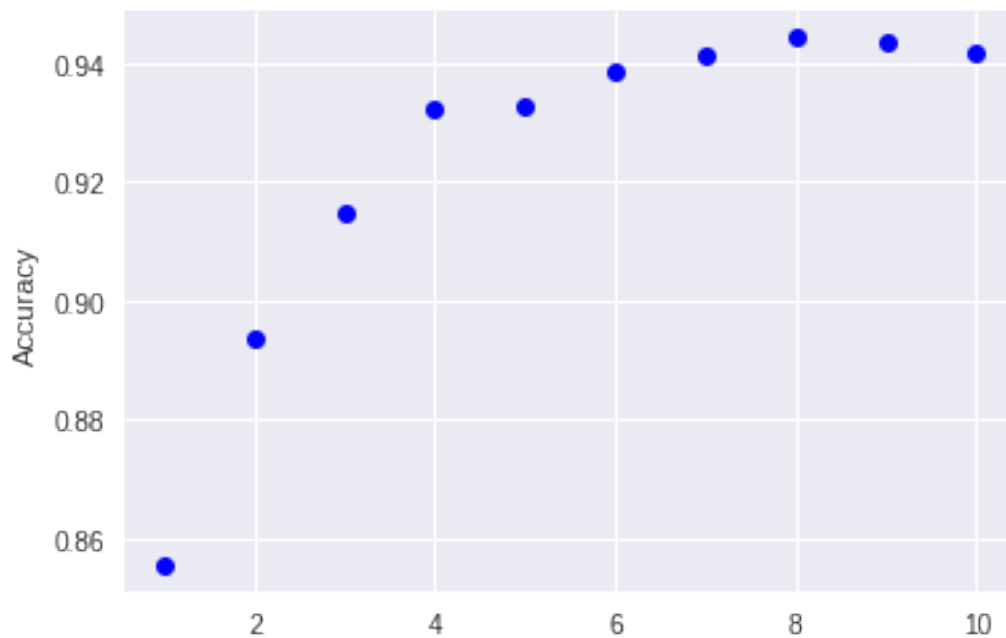
Rysunek 2: Celność dla modelu z 1 warstwą



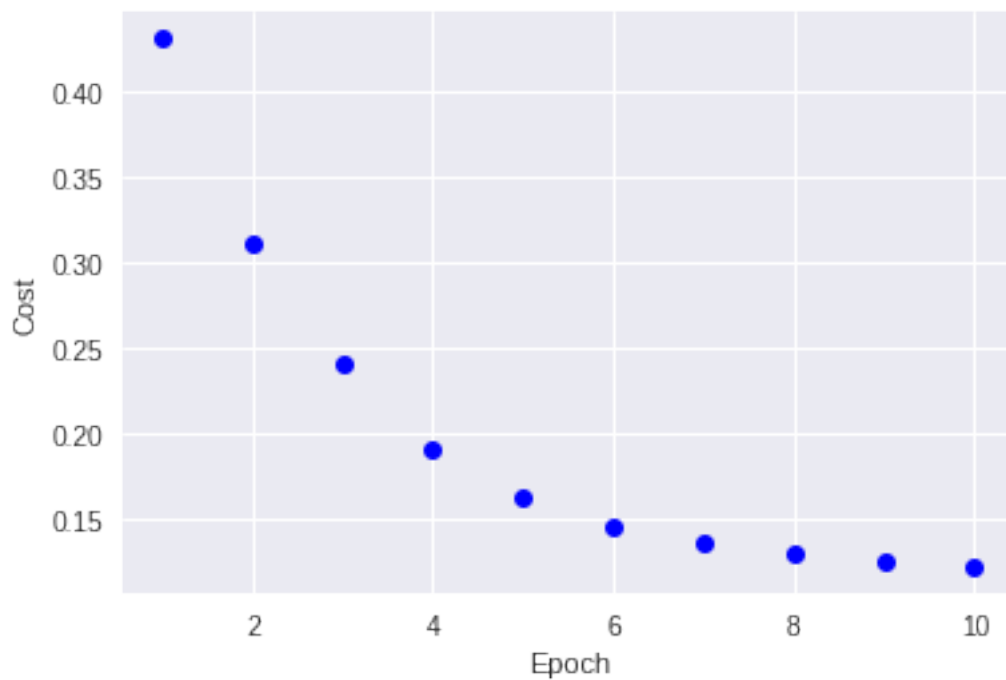
Rysunek 3: Celność dla modelu z 3 warstwami

4 Najlepszy model

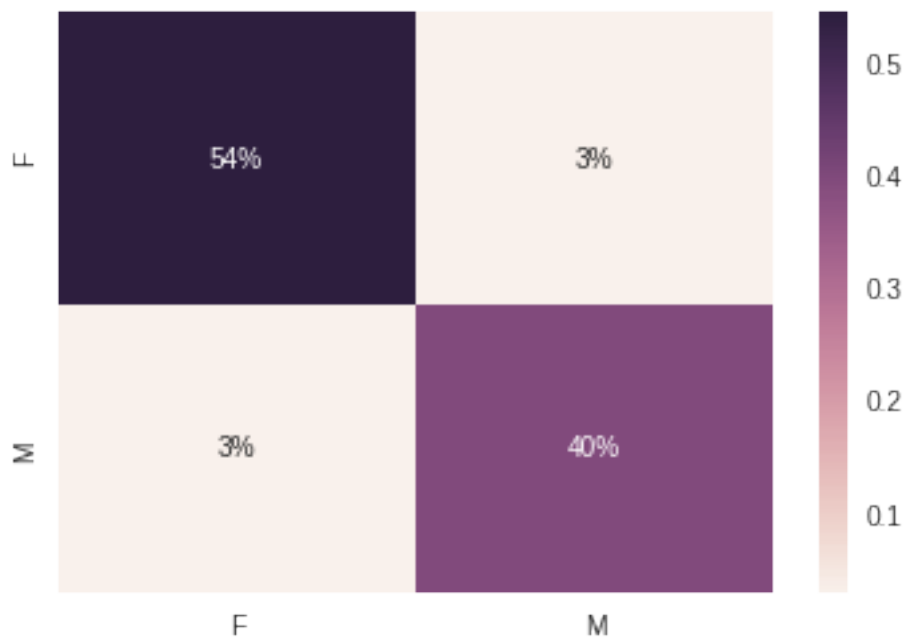
Najlepszym modelem został model oparty na LSTM z 3 warstwami osiągający celność na poziomie 94,43%. Na poniższych rysunkach zaprezentowane zostały wykresy celności, kosztu oraz macierz pomyłek.



Rysunek 4: Celność dla modelu końcowego.



Rysunek 5: Koszt dla modelu końcowego.



Rysunek 6: Macierz pomyłek dla modelu końcowego.

5 Wnioski

Mimo prostszej budowy komórki GRU nie ustępują tym na bazie LSTM. Na podstawie przeprowadzonych badań nie można jednak jednoznacznie określić, które rozwiązanie jest ewidentnie lepsze.