# Sprawozdanie 7

Krzysztof Maciejewski 260449

## Wstęp

Zadanie polegało na porównaniu predykcji sentymentu recenzji za pomocą modelu sieci rekurencyjnej z wybranym typem warstwy rekurencyjnej (RNN vs. LSTM). Zbiór danych składał się z 25000 recenzji filmów opisanych czy są one negatywne albo pozytywne.

Implementacja modelu

model = Sequential()  
model.add(Embedding(input\_dim=max\_features, output\_dim=32, input\_length=max\_sentence\_length))  
model.add(SimpleRNN(100))  
*# model.add(LSTM(units=100))*model.add(Dense(1, activation='sigmoid'))  
  
model.compile(loss='binary\_crossentropy', optimizer='adam', metrics=['accuracy'])  
print(model.summary())  
  
model.fit(X\_train, y\_train, validation\_data=(X\_test, y\_test), epochs=3, batch\_size=64)  
  
scores = model.evaluate(X\_test, y\_test, verbose=0)  
print("Accuracy: %.2f%%" % (scores[1]\*100))

Posiada on warstwę Embedding, która konwertuje wartości ze skończonego zbioru recenzji na wektory.

## Doświadczenia

## • Typ wybranej warstwy rekurencyjnej (RNN vs. LSTM)

## • Wymiar warstwy rekurencyjnej

## • Maksymalna długość sekwencji

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Długość sekwencji | Wymiar warstwy | Accuracy |
| RNN | pełna | 100 | 76.20% |
| LSTM | pełna | 100 | 88.00% |
| RNN | pełna | 20 | 67.15% |
| LSTM | pełna | 20 | 88.01% |
| RNN | pełna | 300 | 62.77% |
| LSTM | pełna | 300 | 85.98% |
| RNN | 10 | 100 | 50.38% |
| LSTM | 10 | 100 | 53.59% |
| RNN | 20 | 100 | 51.37% |
| LSTM | 20 | 100 | 55.93% |

## Wnioski

Podczas doświadczeń o wiele lepsze wyniki uzyskał LSTM, ale wymagał również więcej czasu na naukę. LSTM ma zdolność przechowywania informacji dłużej niż RNN i nie ma problemu wybuchającego gradientu, co jest istotne w analizie większych sekwencji, takich jak recenzje filmów. W momencie gdy zmniejszyłem długość sekwencji, rozbieżność między LSTM, a RNN nie była już tak znacząca.

Najbardziej znaczący wpływ na wyniki sieci miał parametr długości sekwencji. Mniejsza długość sekwencji znacząco obniżała zdolności predykcyjne modeli. Dlatego kluczowe okazało się dostarczenie jak najdłuższego ciągu słów.

Zwiększanie rozmiaru warstwy rekurencyjnej nieznacznie poprawiało dokładność modelu, jednakże bardzo mocno spowalniało uczenie.