

# Stock Forecasting

Michał Górecki  
Jakub Karczewski



# Założenia projektu

Sprawdzić w jaki sposób możemy przewidzieć wahania notowań spółek giełdowych przy użyciu deep learningu.



# Status prac

Aktualnie pracujemy nad badaniem zależności pomiędzy wizerunkiem spółek publicznych na Twitterze a ceną ich akcji.

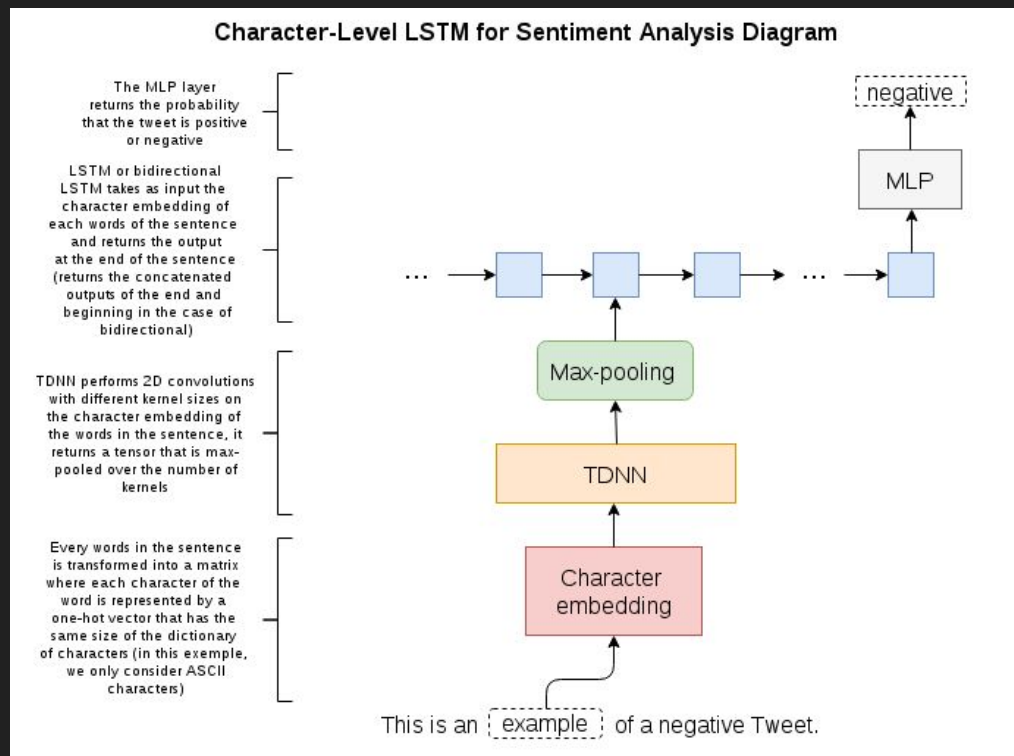
Do klasyfikacji używamy głębokiej sieci neuronowej posiadającej warstwy konwolucyjne i rekurencyjne komórki LSTM.

# Źródła danych

- Dataset Stanford Sentiment140 ~1M otagowanych tweetów do treningu klasyfikatora sentymentu tj. zabarwienia emocjonalnego tweeta.
- Twitter API
- Yahoo Finance



# Architektura sieci



# Przykładowe wyniki z LSTM

- Wyjście sieci to 2 wartości, które interpretujemy jako prawdopodobieństwo tego, że dany tekst jest pozytywny lub negatywny.
- Sieć “rozumie” sens leżący w kolejności słów:
  - I like used cars. [ 0.06939355 0.93060648]
  - I used to like cars. [ 0.73696208 0.26303795]
- Sieć jest względnie odporna na literówki:
  - Bad cake [ 0.98962688 0.0103731]
  - Badd cake [ 0.98835272 0.01164733]
  - Baad cake [ 0.38543931 0.61456066]

OUTPUT:  
[negatywny, pozytywny]

# Plany na przyszłość

- Wykorzystanie wyjścia z klasyfikatora sentymentu do poprawienia predykcji innych modeli próbujących przewidzieć notowania giełdowe na podstawie historycznych danych giełdowych.
- Eksploracja innych portali społecznościowych niż Twitter.

Dziękuję za uwagę