

LAB 03 - WPROWADZENIE DO GŁĘBOKIEGO UCZENIA - CNN + MLP

Karol Działowski

nr albumu: 39259
przedmiot: Uczenie Maszynowe 2

Szczecin, 21 stycznia 2021

Spis treści

1 Cel laboratorium	1
2 Wyniki eksperymentu wpływu liczby epok i współczynnika uczenia	2
2.1 Wnioski	2
3 Porównanie modeli	3
3.1 Wnioski	3
Bibliografia	3

1 Cel laboratorium

Celem laboratorium było zapoznanie się z funkcjonalnością pakietów *TensorFlow* i *Keras*. W tym celu zaimplementowano prosty model sieci splotowej o wielu wejściach różnych typów.

Zadaniem głównym było zbudowanie prostego modelu sieci splotowej o zadanej architekturze, który integruje dane typu alfanumerycznego i dane obrazowe. Dla tego numeru zbadać wpływ epok uczenia na jakość klasyfikacji i wpływ kroku uczenia.

Badania przeprowadzono na zbiorze danych *Houses dataset* [1] dostępnym pod adresem <https://github.com/emanhamed/Houses-dataset>. Jest to testowy zbiór danych dla szacunkowych cen nieruchomości który łączy dane wizualne oraz informacje tekstowe. Każdy dom reprezentowany jest przez 4 zdjęcia dla sypialni, łazienki, kuchni i widok frontalny domu. Zbiór zawiera 2140 obrazów, 4 obrazy na każdą z nieruchomości. Dodatkowo, oprócz obrazków, dostępny jest dokument tekstowy, który zawiera tekstowe metadane zbioru. Każ-

dy wiersz w pliku reprezentuje jeden z domów. Wartości w wierszy to kolejno: liczba sypialni, liczba łazienek, powierzchnia domu, kod pocztowy i cena.

2 Wyniki eksperymentu wpływu liczby epok i współczynnika uczenia

Stworzono model łączący wyjścia z sieci MLP oraz CNN. Nauczono go na zbiorze testowym dla liczby epok $epochs \in \{10, 100, 1000\}$ oraz dla korku uczenia $lr \in \{0.0001, 0.001, 0.01\}$. Testowany model jako wejścia miał obrazek złożony z czterech zdjęć nieruchomości, który był wejściem do sieci CNN oraz dane numeryczne, które były wejściem do MLP. Wyjścia z tych sieci zostały połączone w jeden model.

Kod źródłowy 1: Złączenie modelu MLP oraz CNN

Źródło: Opracowanie własne

```
1 def create_combined(mlp_shape):
2     # Create the MLP and CNN models
3     mlp = create_mlp(mlp_shape)
4     cnn = create_cnn(64, 64, 3)
5
6     combinedInput = concatenate([mlp.output, cnn.output])
7
8     x = Dense(4, activation="relu")(combinedInput)
9     x = Dense(1, activation="linear")(x)
10
11    model = Model(inputs=[mlp.input, cnn.input], outputs=x)
12    return model
```

Liczba epok	Krok uczenia	Błąd
10	0.0001	157.837543
100	0.0001	52.595277
1000	0.0001	25.830520
10	0.0010	24.312008
100	0.0010	23.146181
1000	0.0010	22.573178
10	0.0100	24.773907
100	0.0100	19.307650
1000	0.0100	20.932601

Tabela 1: Przetestowane modele dla 100 epok

2.1 Wnioski

Najlepszym modelem okazał się model nauczony na 100 epokach dla kroku uczenia 0.01. W tym przypadku zaobserwowano ciekawe zjawisko, gdzie większa liczba epok ma negatywny

wpływ na jakość modelu.

Generalnie, krok uczenia ma kluczowy wpływ na jakość modelu. Jednak po pewnej liczbie epok wpływ ten maleje i dalszy uczenie nie przynosi zauważalnego zysku.

Dla badanego modelu, krok uczenia miał wpływ na jakość modelu. Dla badanego zestawu kroków uczenia najlepszy okazał się 0.01.

3 Porównanie modeli

Oprócz wcześniej badanego modelu, zaimplementowano model w którym wejściem do sieci splotowej był jeden obrazek przedstawiający front nieruchomości. Porównano wyniki tej sieci oraz sieci z wcześniejszego eksperymentu oraz sieci składowych, czyli pojedynczej sieci CNN uczonej tylko na obrazach oraz pojedynczej sieci MLP uczonej tylko na danych numerycznych.

Wszystkie modele zostały nauczone na 100 iteracjach dla kroku uczenia 0.001.

Model	Błąd
MLP	24.52%
CNN	45.78%
Combined 4 obrazki	36.37 %
Combined 1 obrazek	23.34 %

Tabela 2: Przetestowane modele dla 100 epok

3.1 Wnioski

Błędy składowych sieci MLP oraz CNN wynoszą odpowiednio 24.52% oraz 45.78%. Predykcja ceny jest lepsza korzystając z samych danych tekstowych niż korzystając z samych obrazków.

Złączenie sieci z dwóch powyższych rezultatów nie polepsza wyniku najlepszej z nich. Wynik 36.37% jest gorszy od wyniku sieci MLP wynoszącego 24.52%.

Dopiero zbudowanie sieci działającej na jednym obrazku, reprezentującym frontalny widok budynku, zwiększa nieznacznie skuteczność sieci zmniejszając błąd do 23.34%.

Otrzymane wyniki jednak nie są dostatecznie miarodajne, bo porównanie zostało przeprowadzone na modelach nauczonych tylko na 100 epokach. Uznałem, że dłuższe uczenie jednak nie będzie potrzebne na potrzeby tego porównania.

Bibliografia

- [1] Ahmed E., Moustafa M.: House price estimation from visual and textual features. *arXiv preprint arXiv:1609.08399*, 2016.