Sprawozdanie SIiIW- Lista 5.

Jakub Cebula 266886

Zad 1. Przetworzenie danych i utworzenie zbioru uczącego i walidującego:

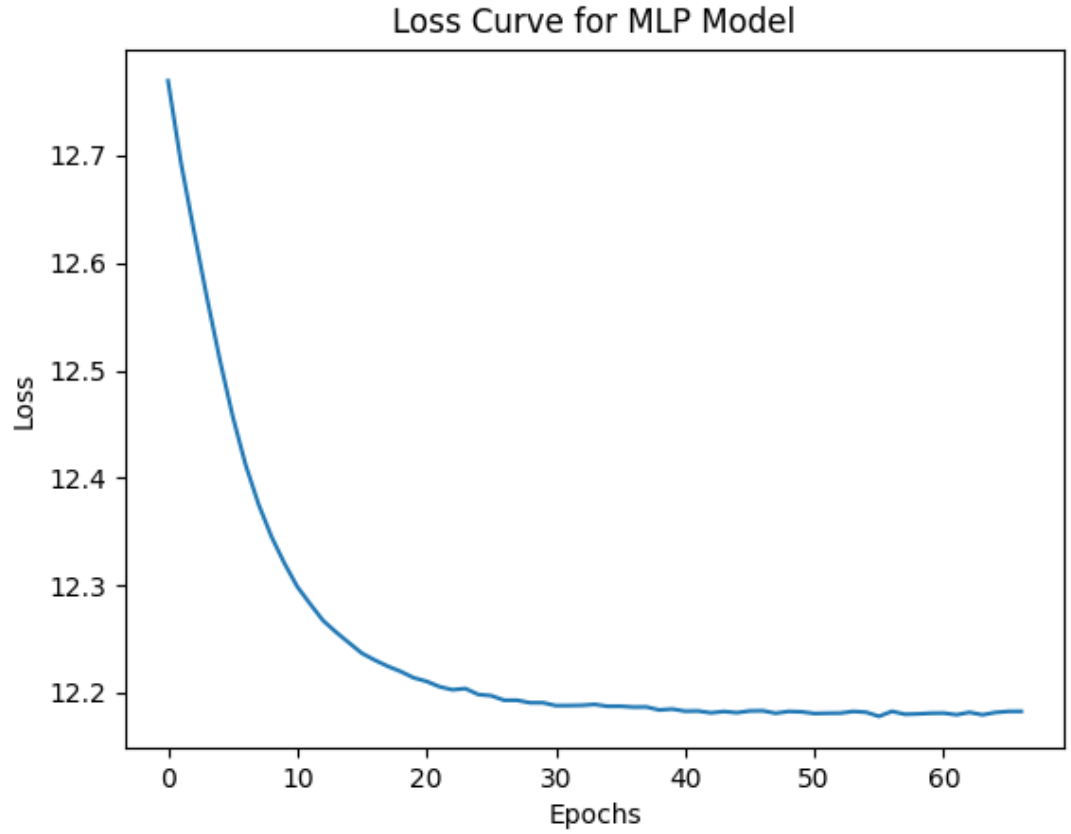
(135, 384)

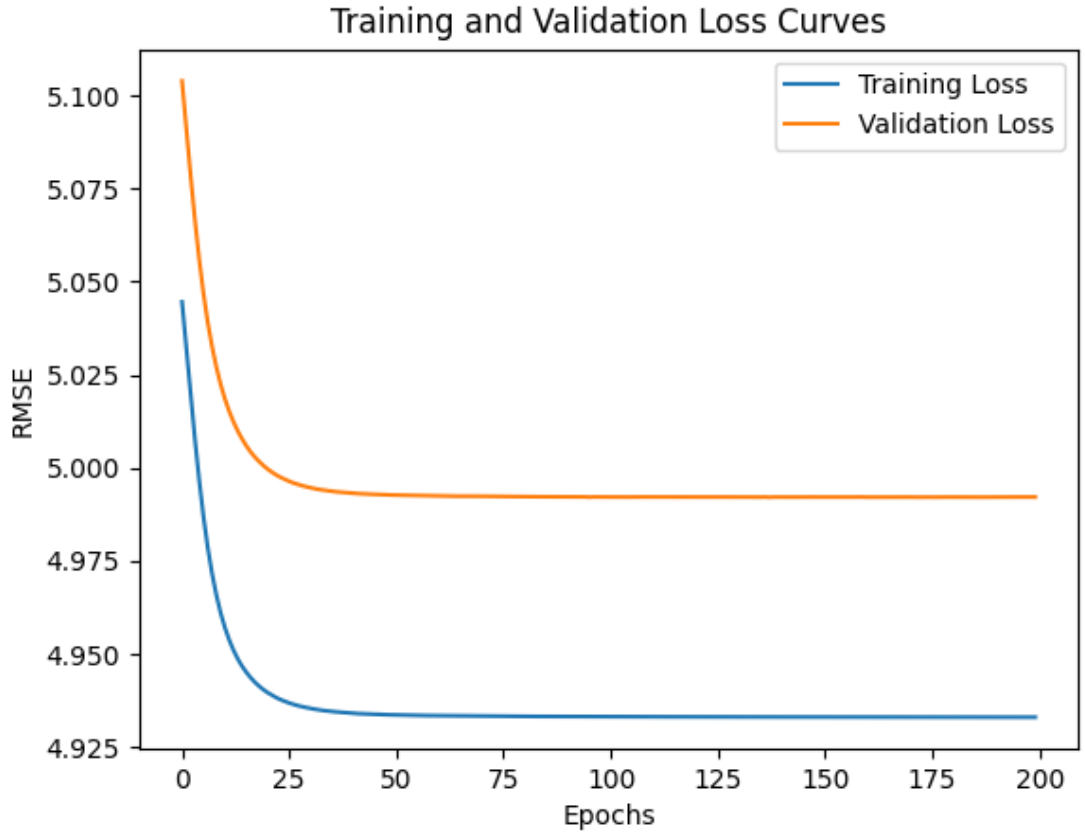
	1	99	99.1	99.2	99.3	...	99.152	99.153	99.154	99.155	99.156
0	1	99	99	99	99	...	99.0	99.00	99.00	99.00	99.00
1	4	99	99	99	99	...	99.0	99.00	99.00	99.00	99.00
2	47	99	99	99	99	...	0.0	99.00	0.00	5.93	4.19
3	13	99	99	99	99	...	0.0	99.00	99.00	99.00	0.00
4	33	99	99	99	99	...	99.0	-2.32	99.00	0.00	2.93
...	..	..	...	...	...	...	...	...	...	...	...
7693	27	99	99	99	99	...	99.0	0.00	0.00	0.00	99.00
7694	1	99	99	99	99	...	99.0	99.00	99.00	99.00	99.00
7695	26	99	99	99	99	...	3.5	6.26	99.00	-2.66	0.65
7696	64	99	99	99	99	...	0.0	2.97	0.57	0.00	0.00
7697	2	99	99	99	99	...	99.0	99.00	99.00	99.00	99.00

	rated_jokes	joke_id	rating
5	112	6	-4.45
38	75	6	-10.00
87	73	6	-5.76
91	67	6	9.04
109	121	6	0.00
...	...	...	...
1046906	33	157	0.00
1046911	67	157	2.73
1046919	50	157	0.41
1046925	26	157	0.65
1046926	64	157	0.00

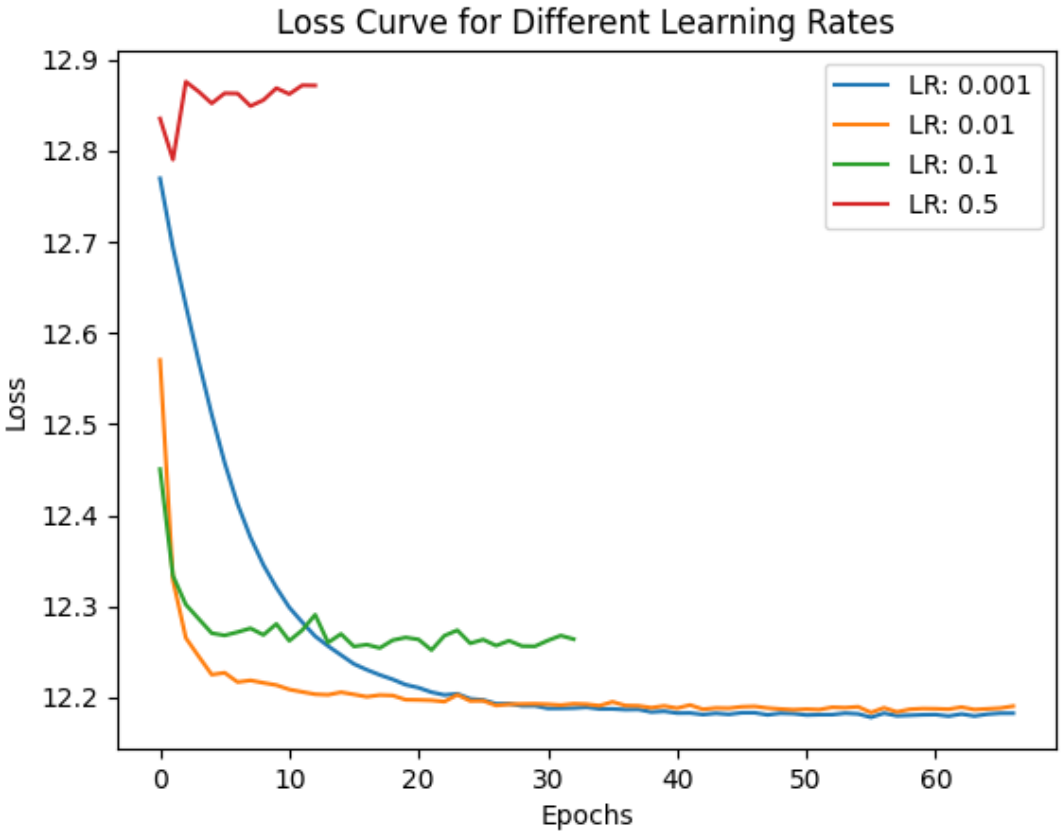
[106488 rows x 3 columns]

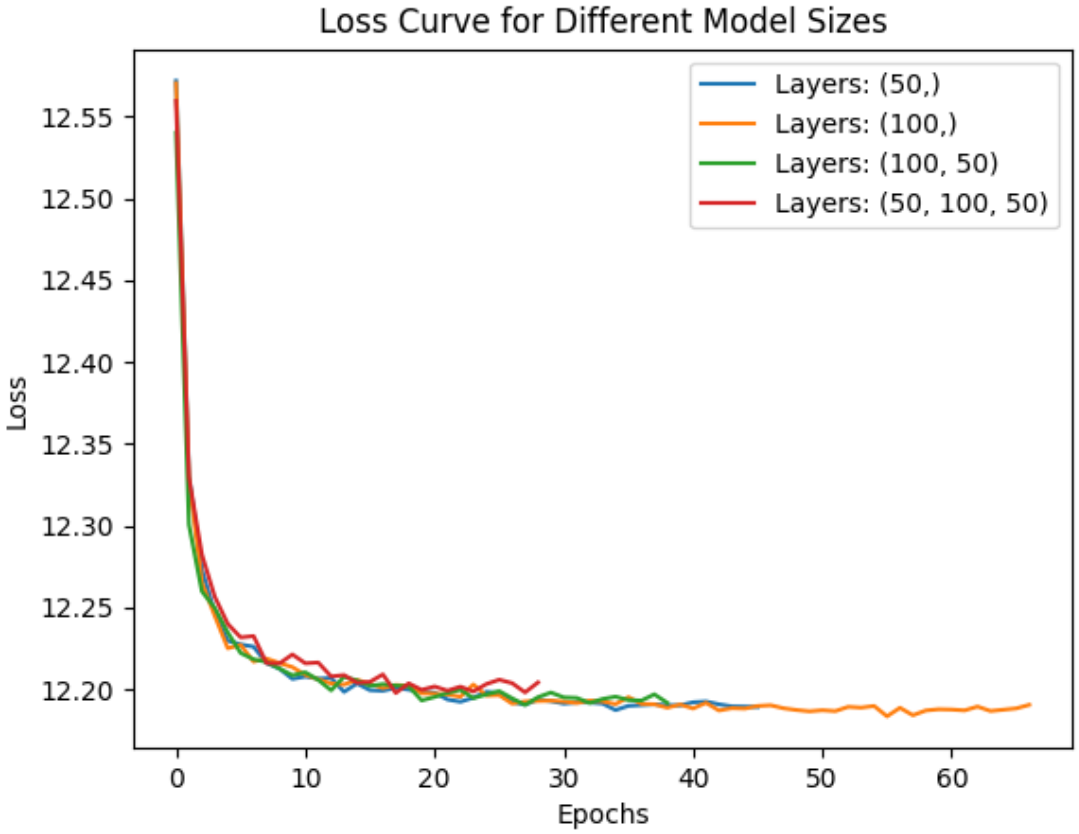
Początkowo miałem problem z pobraniem i uruchomieniem modelu Bert, stąd skorzystałem z jednego z domyślnych modeli biblioteki „all-MiniLM-L6-v2” wyniki jego zastosowania podaję poniżej:

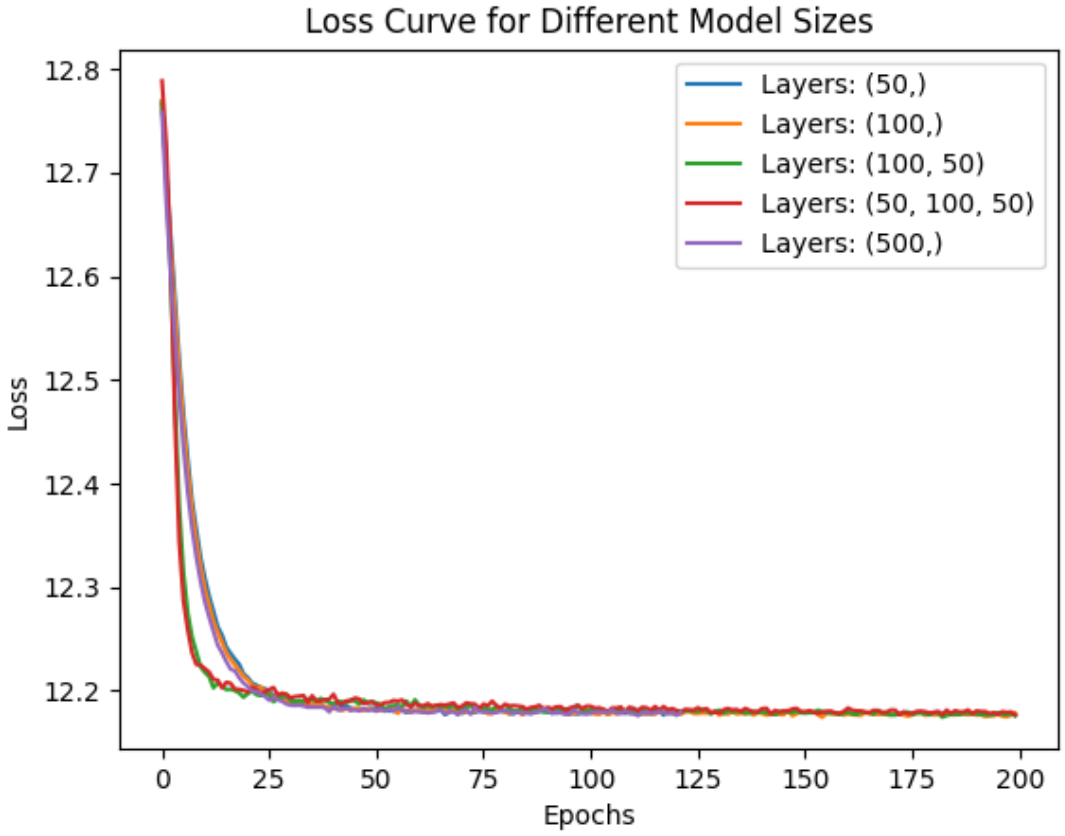

Loss Curve for MLP Model
Loss
12.7
12.6
12.5
12.4
12.3
12.2
0
10
20
30
40
50
60
Epochs


Training and Validation Loss Curves
Training Loss
Validation Loss
RMSE
5.100
5.075
5.050
5.025
5.000
4.975
4.950
4.925
0
25
50
75
100
125
150
175
200
Epochs

Zad 3.


Loss Curve for Different Learning Rates
LR: 0.001
LR: 0.01
LR: 0.1
LR: 0.5
Loss
Epochs
12.9
12.8
12.7
12.6
12.5
12.4
12.3
12.2
0
10
20
30
40
50
60


Loss Curve for Different Model Sizes
Layers: (50,)
Layers: (100,)
Layers: (100, 50)
Layers: (50, 100, 50)
Loss
12.55
12.50
12.45
12.40
12.35
12.30
12.25
12.20
0
10
20
30
40
50
60
Epochs

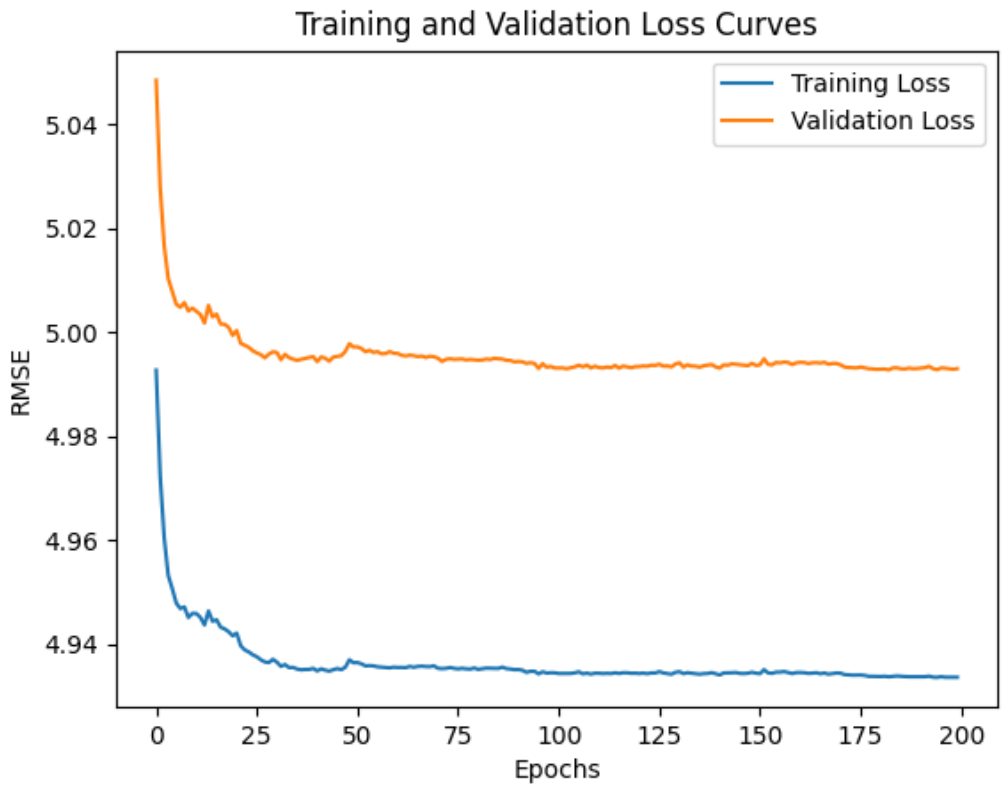

Loss Curve for Different Model Sizes
Layers: (50,)
Layers: (100,)
Layers: (100, 50)
Layers: (50, 100, 50)
Layers: (500,)
Loss
12.8
12.7
12.6
12.5
12.4
12.3
12.2
0
25
50
75
100
125
150
175
200
Epochs

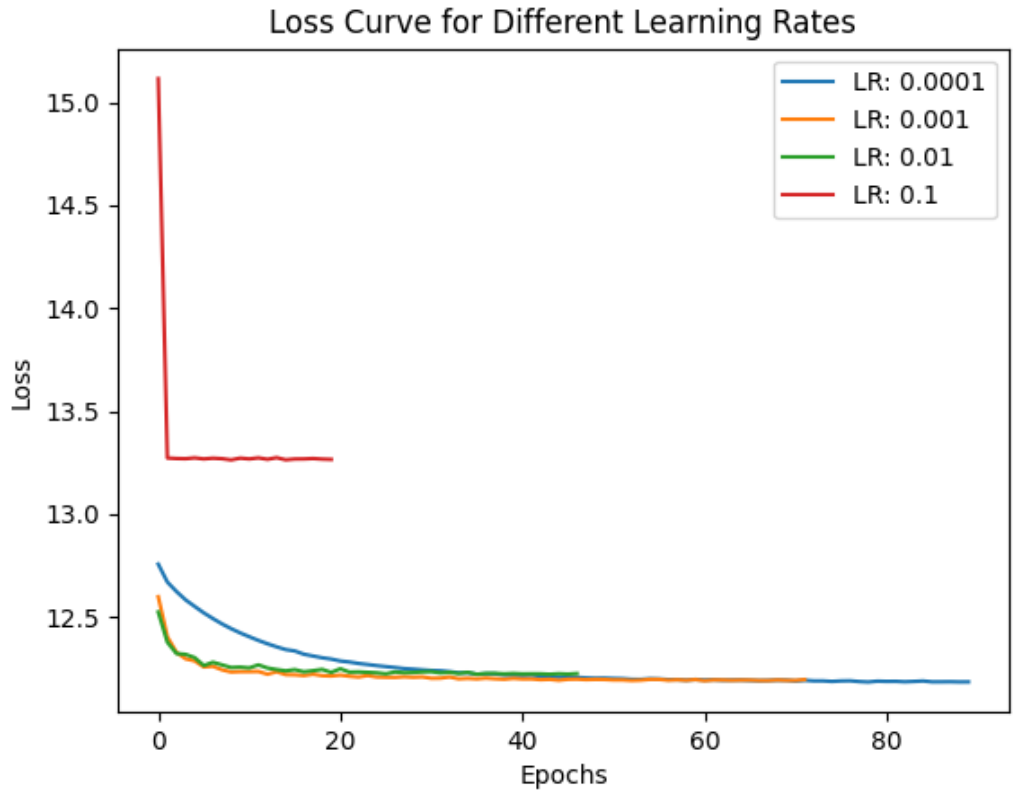
Zad 5.

Predicted rating for the new joke: 0.6436475515365601

Measured parameter	**RMSE**
RMSE on Training Set	4.933216025183661
RMSE on Validation Set	4.9922422282958445
Learning Rate: 0.001	4.992229179897027
Learning Rate: 0.01	4.992096284142123
Learning Rate: 0.1	5.002480956473221
Learning Rate: 0.5	5.121384117410883
Hidden Layers: (50,)	4.992546384528869
Hidden Layers: (100,)	4.993754085677515
Hidden Layers: (100, 50)	4.99206030062376
Hidden Layers: (50, 100, 50)	4.99205154563229
Hidden Layers: (500,)	4.992478277233869
Predicted rating for the new joke	0.6436475515365601

Ostatecznie udało mi się odpalić wersję (rithwik-db/bert-base-cased-10) modelu BERT. Wyniki są podobne, ale prezentuję je poniżej:





```
learning_rate_init=0.001
```