



РАССТАНОВКА УДАРЕНИЙ В СЛОВАХ

Березин Даниил

АКТУАЛЬНОСТЬ

В наше время постановка ударений важная задача для следующих направлений:

- Автоматические аудио субтитры
- Прочие задачи озвучивания текста при помощи ИИ



ОПИСАНИЕ РЕАЛИЗОВАННОГО РЕШЕНИЯ НА ОСНОВЕ AI

В своем решении я рассмотрел два подхода

- Классический - ML (BoW + CatBoost)
- Deep Learning – RNN
с собственной архитектурой

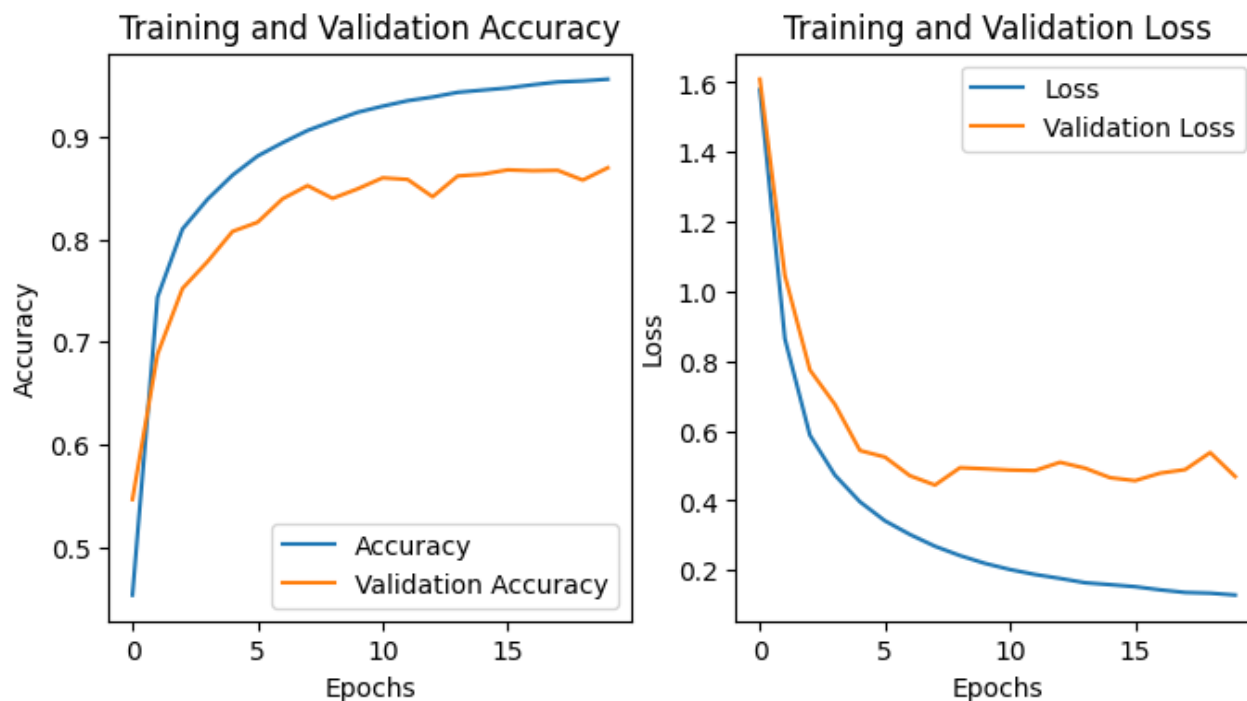
Все решения не содержит предобученных
моделей, все реализованно с 0



ТЕХНИЧЕСКИЕ ДЕТАЛИ

Ниже представлена структура, а так же график обучения RNN

В основе которой лежит двунаправленный LSTM



Layer (type)	Output Shape	Param #
text_vectorization_3 (Text Vectorization)	(None, None)	0
embedding_8 (Embedding)	(None, None, 128)	4224
bidirectional_32 (Bidirectional)	(None, None, 256)	263168
bidirectional_33 (Bidirectional)	(None, None, 128)	164352
dropout_24 (Dropout)	(None, None, 128)	0
bidirectional_34 (Bidirectional)	(None, None, 128)	98816
bidirectional_35 (Bidirectional)	(None, 128)	98816
dropout_25 (Dropout)	(None, 128)	0
dense_24 (Dense)	(None, 64)	8256
dense_25 (Dense)	(None, 32)	2080
dropout_26 (Dropout)	(None, 32)	0
dense_26 (Dense)	(None, 16)	528
Total params: 640240 (2.44 MB)		
Trainable params: 640240 (2.44 MB)		
Non-trainable params: 0 (0.00 Byte)		

Визуализация решения

- На данном этапе реализовано MVP в jupyter notebook



```
sample_text = 'социал'  
predictions = model.predict(np.array(list(sample_text)))  
pred(predictions, sample_text)
```



```
1/1 [=====] - 0s 26ms/step  
'социа^л'
```

РЕЗУЛЬТАТЫ

Так же важно отметить, что за счет разбиения слов на буквы а не слоги, размер словаря составляет всего 33 элемента и при не самой простой архитектуре число параметров RNN всего 640 240

MODEL	ACCURACY
CatBoost	0.69
RNN	0.90

Возможные улучшения

На данном этапе модель можно и нужно обучить на большем числе эпох для преодоления барьера в 0.95 accuracy

