

# Deep Learning in NLP

## Russian Aspect-Based Sentiment Analysis



Бунин Д.А.

Весна 2020

## 1 Описание задачи

- ABSA
- ATSC

## 2 Применяемые подходы

## 3 Наш подход

- Английский язык
- Русский язык

## 4 Результаты

- Английский язык
- Русский язык

## 5 Выводы



Пример отзыва на ресторан:

*Прекрасный ресторан! Уютная обстановка и атмосфера, красивый вид на залив. Обслуживание хорошее, вежливый персонал, все пожелания выполнялись. Вкусная еда и порции нормальные, готовили и приносили довольно быстро. В целом остались очень хорошие и приятные впечатления от посещения, обязательно придем еще и не один раз. Спасибо!*

- Еда – positive
- Интерьер – positive
- Цены – absense



Пример отзыва на ресторан:

*Прекрасный ресторан! **Уютная обстановка и атмосфера**, красивый вид на залив. Обслуживание хорошее, вежливый персонал, все пожелания выполнялись. **Вкусная еда и порции** нормальные, готовили и приносили довольно быстро. В целом остались очень хорошие и приятные впечатления от посещения, обязательно придем еще и не один раз. Спасибо!*

- Уютная – интерьер – positive
- обстановка – интерьер – positive
- атмосфера – интерьер – positive
- Вкусная – еда – positive
- еда – еда – positive
- порции – еда – positive



- 1 ATAЕ-LSTM
- 2 TD-LSTM
- 3 IAN
- 4 RAM
- 5 Deep Memory Network (MemNet)
- 6 AEN(Attention Encoder Network)
- 7 BERT



Применяемые подходы:

- ELMoClassifierByToken – усреднение эмбеддингов целевых токенов, линейный слой
- ELMoClassifierByContext – конкатенация average pooling, max pooling от эмбеддингов контекста, линейный слой
- ELMoCNN – CNN на эмбеддингах контекста



## Модели ELMo:

- DeepPavlov (Russian WMT news) – эксперимент прошел неудачно
- RusVectores (НКРЯ и Википедия, токены) – основная модель



- Отзывы о ресторанах: соревнование
- Соотношение классов:

Datasets	Positive	Negative	Neutral
Train	2164	637	807
Test	728	196	196





■ Результаты:

Models	Accuracy	Macro-F1
TD-LSTM	75.6	-
ATAE-LSTM	77.2	-
IAN	78.6	-
RAM	80.2	70.8
MemNet	78.2	65.8
AEN-BERT	83.1	73.8
BERT-SPC	84.5	77.0
BERT-ADA Rest	<b>87.9</b>	<b>81.1</b>
ELMoCNN	80.7	70.7



- Отзывы о ресторанах: соревнование
- Соотношение классов:

Datasets	Positive	Negative	Both
Train	2719	746	55
Test	3600	939	87

- Результаты:

Models	Micro-F1	Macro-F1
Baseline	71.0	26.7
Best Competition	82.5	<b>55.5</b>
ELMoCNN	<b>85.5</b>	52.0



## Английский язык

- До SOTA далеко

## Русский язык

- Уровень лучшего решения соревнования
- Проблемы с несбалансированностью классов