

Gendered Pronoun Resolution

Pair pronouns to their correct entities

Кирилл Бродт

Задача

Lucy went to the gym. She was happy.



Lucy went to the gym. Her dog stayed at home.



In May, Fujisawa joined *Mari Motohashi's* rink as the team's skip, moving back from Karuizawa to Kitami where **she** had spent her junior days.

Метрика и данные

$$\text{logloss} = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^M y_{ij} \log(p_{ij}).$$

	ID	Text	Pronoun	Pronoun- offset	A	A-offset	A-coref	B	B-offset	B-coref	URL
0	test-1	Upon their acceptance into the Kontinental Hoc...	His	383	Bob Suter	352	False	Dehner	366	True	http://en.wikipedia.org /wiki/Jeremy_Dehner
1	test-2	Between the years 1979-1981, River won four lo...	him	430	Alonso	353	True	Alfredo Di St'fano	390	False	http://en.wikipedia.org /wiki/Norberto_Alonso
2	test-3	Though his emigration from the country has aff...	He	312	Ali Aladhadh	256	True	Saddam	295	False	http://en.wikipedia.org /wiki/Aladhadh

Этапы

- Первый этап: разработать модель и зафиксировать для второго этапа
- Второй этап: сделать предсказание для новых данных
 - важно, что здесь нельзя было менять параметры модели, дообучать и использовать любые другие трюки.












Этап 1. Бейзлайн

Публичные данные: train (2k) + dev (~400) + test (2k)

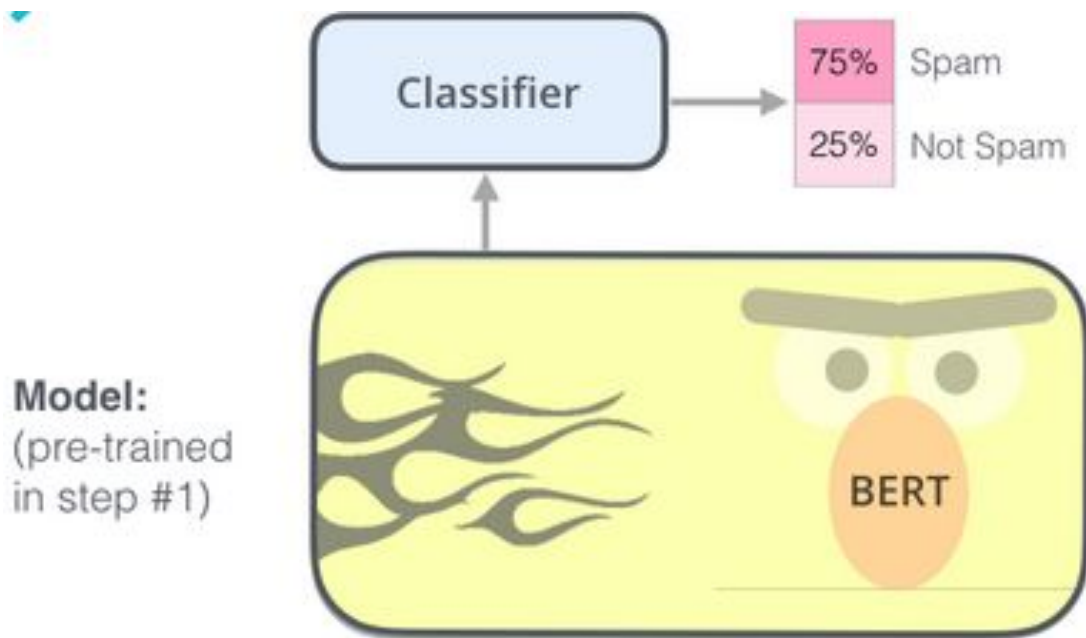
Модель учим используя эти данные и фиксируем на конец первого этапа.

Но зачем учиться, если можно залить тест?

Этап 1. Бейзлайн или игра вслепую (~850)

#	Team Name	Kernel	Team Members	Score ?	Entries	Last
1	Michael Vine	</> Pronoun Subject Dist...		0.00000	1	11d
2	Shadab Hussain			0.00000	2	11d
3	DL			0.00000	1	11d
4	Your Name			0.00000	3	11d
5	TeamCoeus		 	0.00000	2	7d
6	Adel Valiullin			0.00000	3	5d
7	Pranav Pandya			0.00000	2	7d
8	shishirranjan			0.00000	2	7d
9	Aman.cyberia			0.00000	1	11d
10	Leaner			0.00000	3	7d

Этап 1. Бейзлайн 2



Отступление

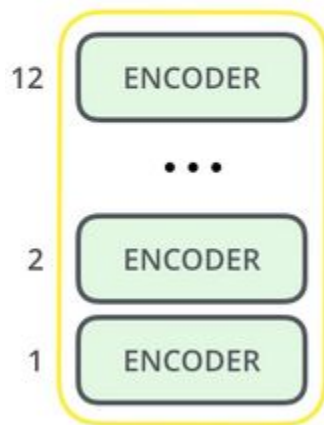


BERT_{BASE}

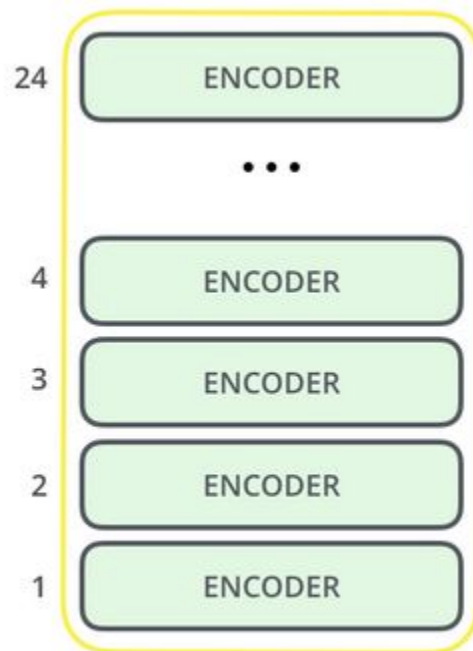


BERT_{LARGE}

Отступление

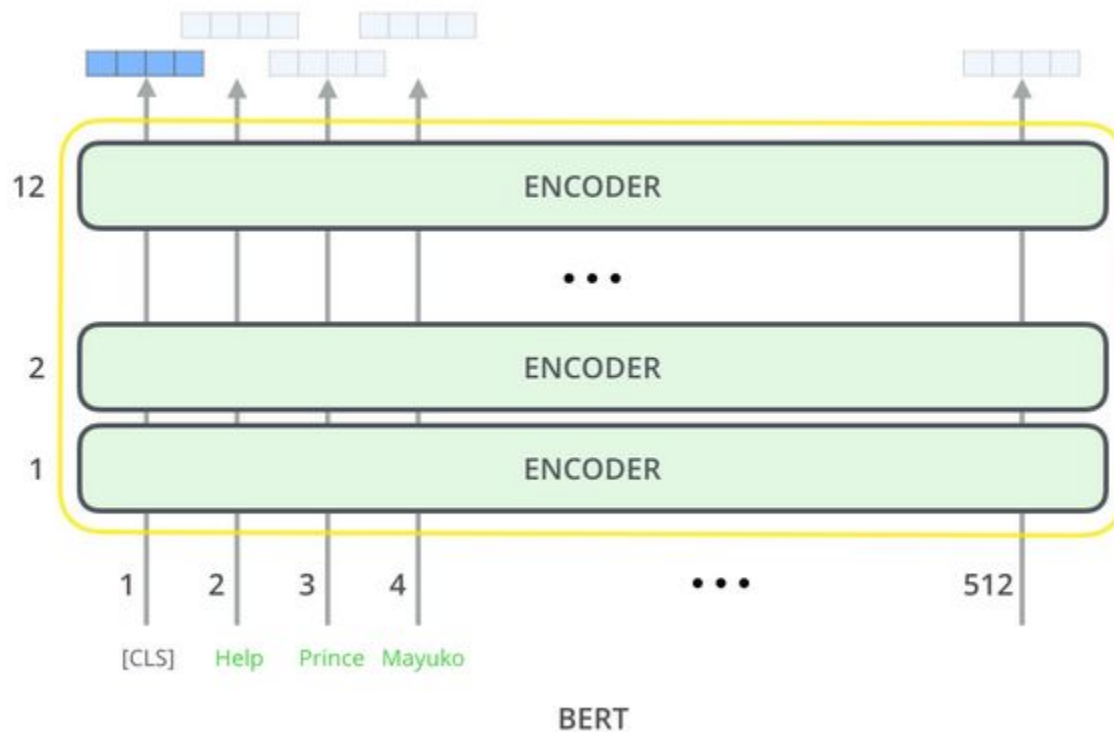


BERT_{BASE}



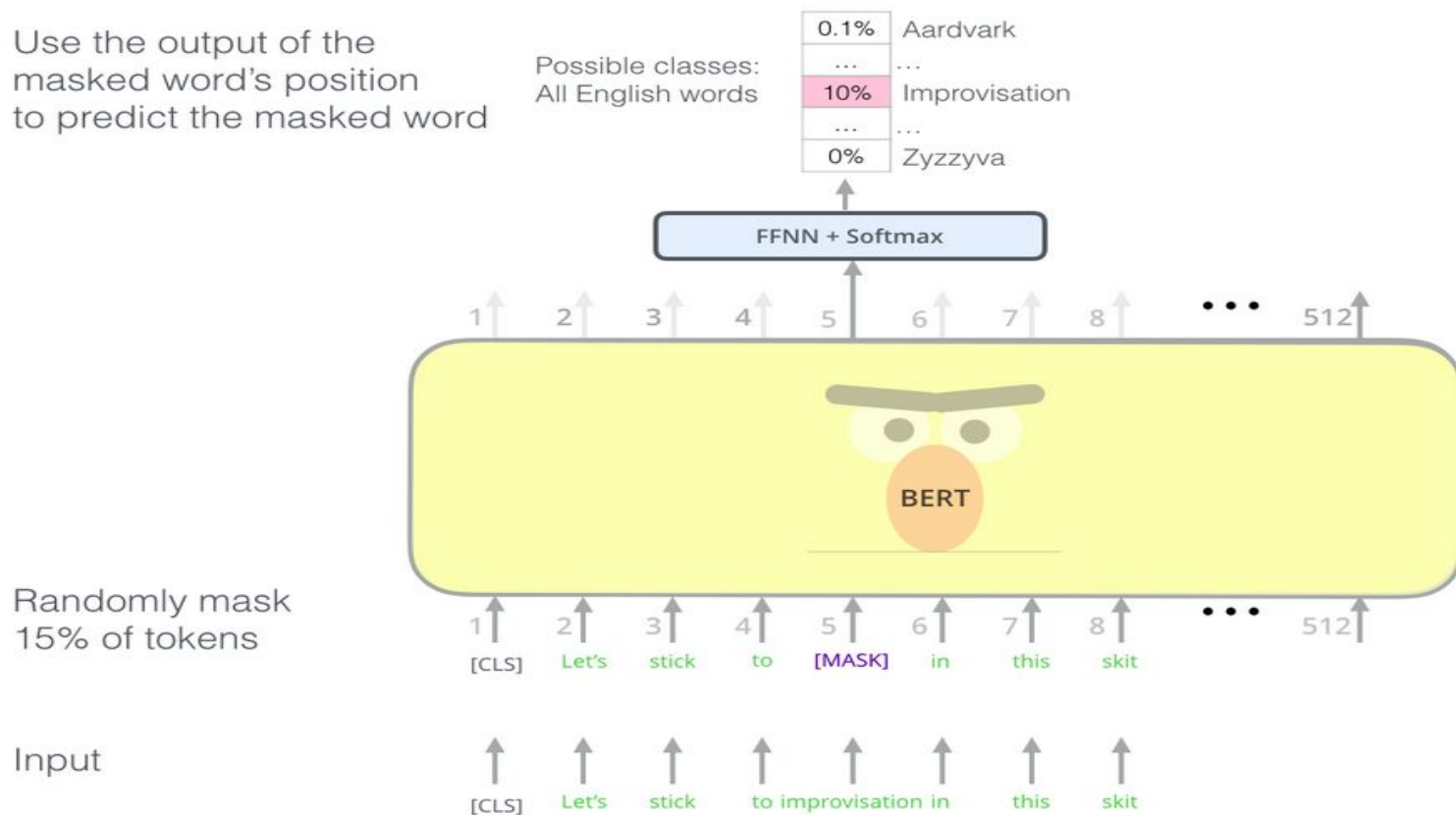
BERT_{LARGE}

Отступление



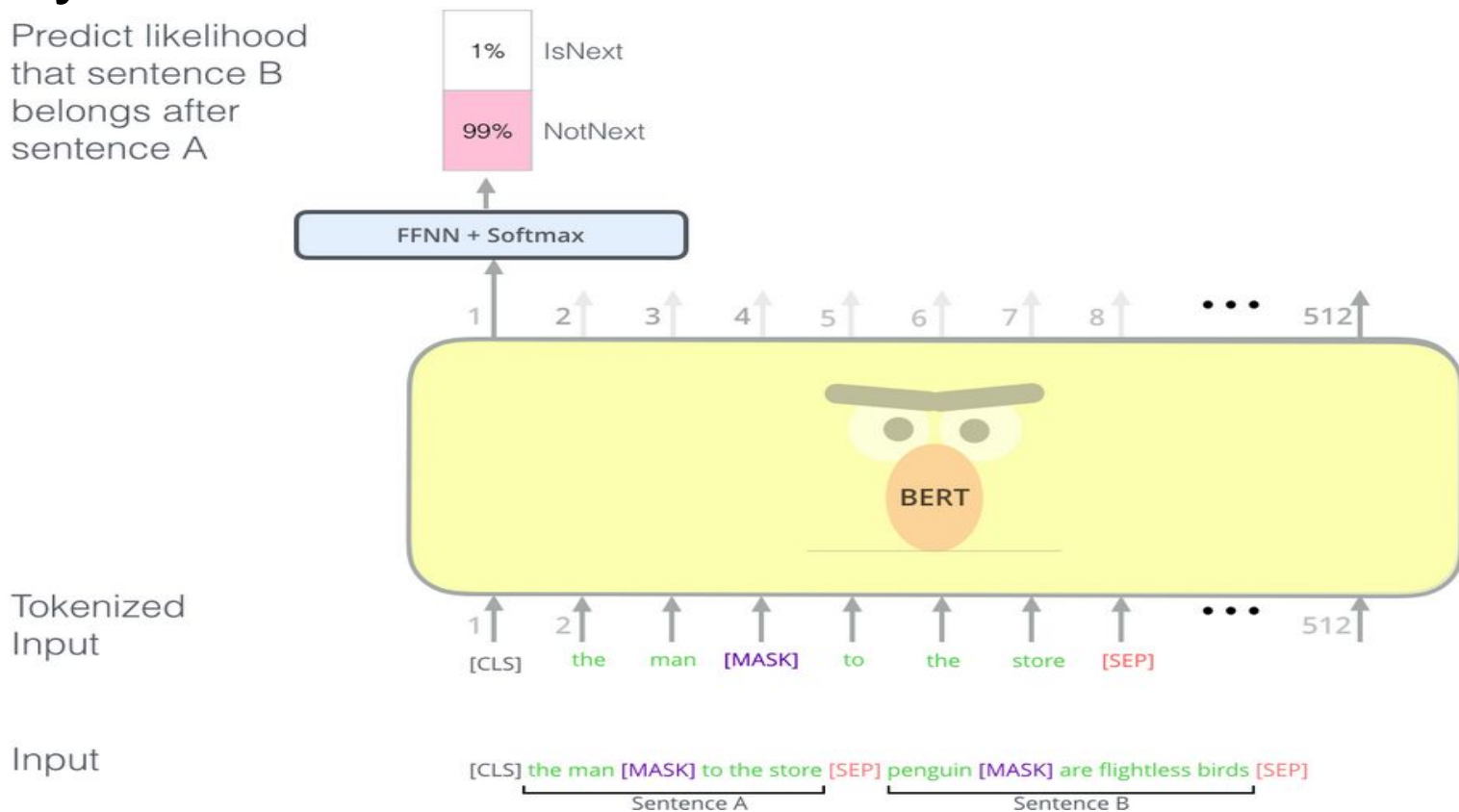
Отступление

Use the output of the masked word's position to predict the masked word

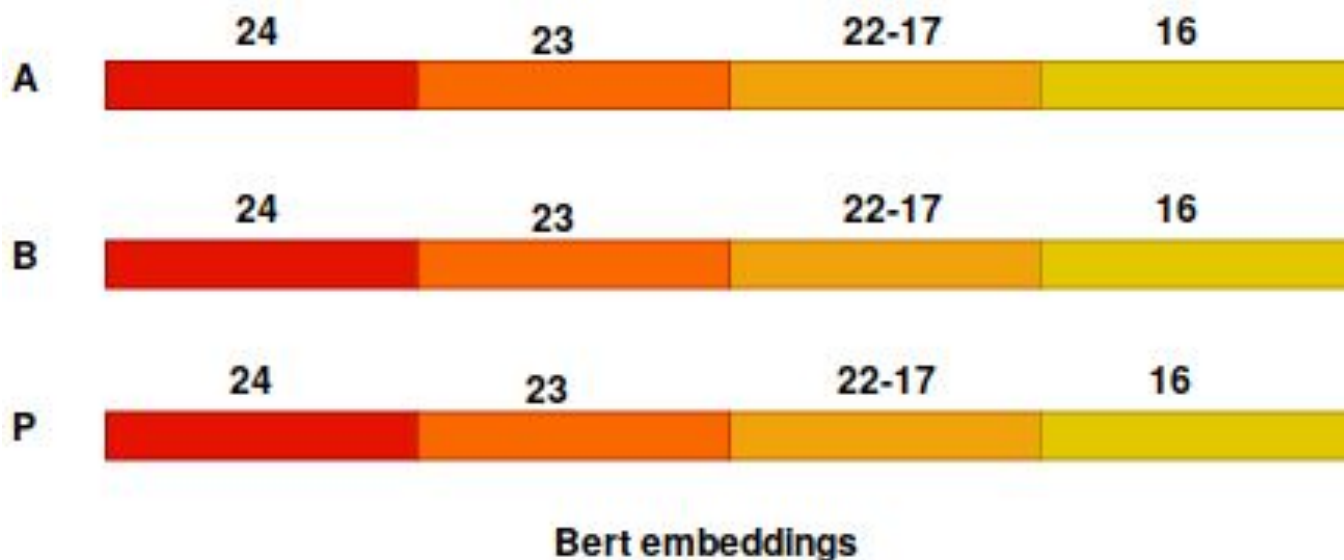


Отступление

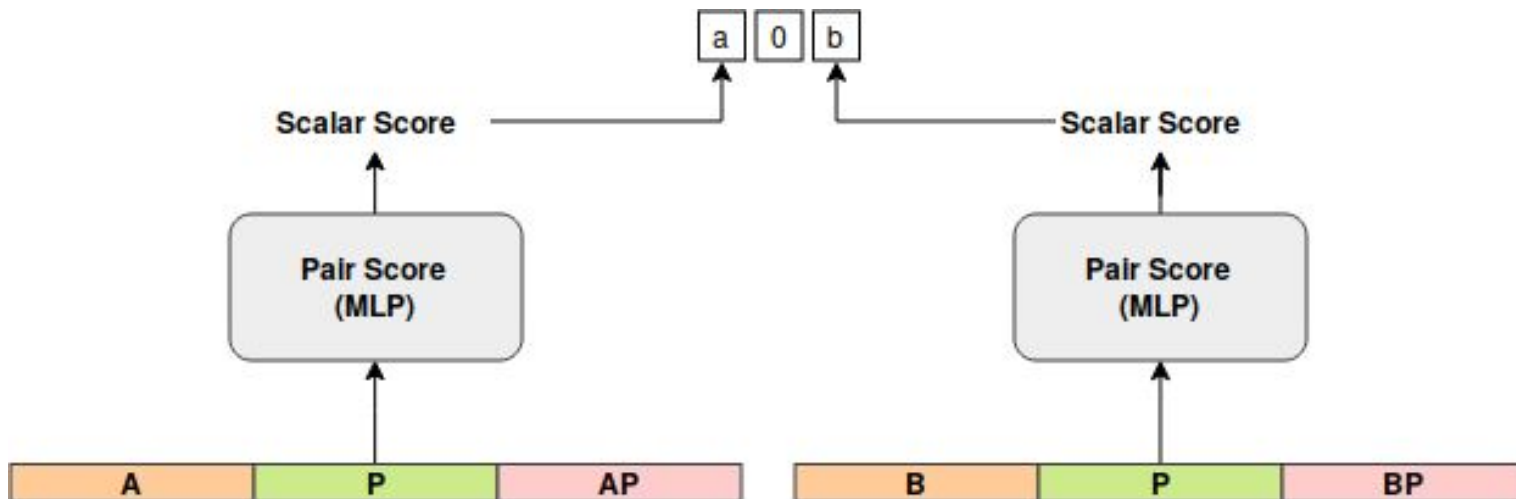
Predict likelihood
that sentence B
belongs after
sentence A



Этап 1. Бейзлайн 2



Этап 1. Бейзлайн 2



Этап 1. Бейзлайн 2

- Эмбединги берта для трёх слов (кандидат А, В и местоимение Р)
- Двухслойная нейронная сеть (MLP)
 - hidden dim: 2048
 - dropout: 0.8
- Батч: 16
- Adam $2e-4$
- 20 эпох
- 10 фолдов
- усреднение по лучшим моделям
- логлосс на публичном тесте: ~ 0.35 (лучший паблик кёрнел ~ 0.475)

Этап 1. Попытки улучшить модель

- Ручные лингвистические признаки
 - расстояние от местоимения до кандидатов
 - присутствие кандидатов в URL
 - из кёрнелов
- Аугментация данных
 - spacy -> NER -> другие персонажи как кандидаты
- Дообучение берта
 - один раз получилось, но не смог потом воспроизвести (ошибка оказалась в том, что дообучал свою текущую модель, а не более простую с последним эмбедингом)\
- Исправление датасета












Идеи, которые не успел попробовать

- Урезание длины контекста














Этап 2

Сделать предсказание на тесте (12.4k) моделью, разработанной на первом этапе.




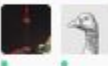





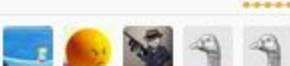
Игра вслепую 2

#	Team Name	Kernel	Team Members	Score ?	Entries	Last
1	Michael Vine	</> Pronoun Subject Dist...		0.00000	1	11d
2	Shadab Hussain			0.00000	2	11d
3	DL			0.00000	1	11d
4	Your Name			0.00000	3	11d
5	TeamCoeus		 	0.00000	2	7d
6	Adel Valiullin			0.00000	3	5d
7	Pranav Pandya			0.00000	2	7d
8	shishirranjan			0.00000	2	7d
9	Aman.cyberia			0.00000	1	11d
10	Leaner			0.00000	3	7d

Итог (263)

#	△pub	Team Name	Kernel	Team Members	Score ?	Entries	Last
1	▲ 213	Sandeep Attree			0.13667	2	6d
2	▲ 42	Zili Wang			0.17289	7	7d
3	▲ 149	[ods.ai] abzaliev			0.18397	2	8d
4	▲ 157	Justin Yang			0.18498	1	9d
5	▲ 208	Ken Krige			0.19189	5	6d
6	▲ 56	Cookpad		   	0.19458	2	11d
7	▲ 117	Bo			0.19473	2	9d
8	▲ 21	LouisClouatre			0.20138	3	7d
9	▲ 241	Rakesh Chada			0.20238	2	5d
10	▲ 25	tko			0.20454	4	10d

Итог (263)

11	▲196	Kobe is Better		0.20472	3	6d
12	▲221	[ods.ai] five ones		0.20574	3	5d
13	▲34	VietAI		0.20650	3	6d
14	▲173	We used bert!		0.20758	2	8d
15	▲87	Ilia		0.21074	4	9d
16	▲164	Some creative name		0.21394	5	8d
17	▲95	NSK & fugusuki		0.21426	2	10d
18	▲48	Kirill Brodt (shad nsk)		0.21618	2	11d
19	▼1	Guanshuo Xu		0.22099	2	11d
20	▲181	A or B or P		0.22183	3	6d

Итог (263)

- 1. 0.13667
 - использовал блэндинг всех известных моделей для этой задачи
- 2. 0.17289
- 3. 0.18397
 - -5 и -6 эмбеддинги, тренировка головы, потом дообучение -5 и -6 слоя, блендинг
- Золото. [0.18498, 0.20472]
 - (4) input dropout + lower projection 0.18498
 - (5) урезание длины контекста 0.19189
 - (8) блэндинг дообученных моделей на других задачах 0.20138
 - (10) дообучение берта 0.20454
- Серебро. [0.20574, 0.32606] (0.21618)
 - лингвистические признаки
- Бронза. [0.32673, 0.47059]

Работа над ошибками

- Всегда думайте, куда текут градиенты
- Нужно иметь больше опыта в алгоритмах и разработке
- Не заходить в конкурс за 4 дня до окончания, лучше недели за 2, чтобы побрейнштормить идеи и их проверить
- Изучить мультигпу