

## **Содержание**

<b>1</b>	<b>Краткое описание контрольной точки 1</b>	<b>2</b>
<b>2</b>	<b>Новая модель</b>	<b>3</b>
<b>3</b>	<b>Результаты на соревновании</b>	<b>4</b>
<b>4</b>	<b>Итоги</b>	<b>5</b>

## **1 Краткое описание контрольной точки 1**

Давайте немного вспомним какая наша цель.

Мы участвуем в соревновании по оценке семантической текстовой связи между предложениями. В датасете есть два предложения и их оценка. Метрикой качества является коэффициент ранговой корреляции Спирмена.

Для бейслайн решения мы использовали TF-IDF вместе с RandomForestClassifier. Получили корреляцию 0.435

Наша основная модель это Bert. Мы вычисляли для каждого предложения его Bert-эмбединги, затем оценивали схожесть между предложениями как косинусное сходство этих эмбедингов. Его выбрали из-за простоты реализации, а также потому, что он улучшает анализ слов относительно базового решения организаторов.

Для улучшение результатов мы решили использовать DistilBert, т.к. обучающая выборка была маленькой и более тяжелые модели на основе Bert (AlBert, RoBerta) могли переобучиться

## **2 Новая модель**

Мы использовали модель DistilBert из-за маленькой выборки входных данных. Она является более облегченной версией Bert, которая, практически, не теряет своей эффективности, но ощутимо увеличивает скорость. Из-за этого риск переобучиться будет ниже по сравнению с базовой и другими моделями.

Также мы попытались на практике посмотреть переобучатся ли тяжелые модели, такие как AlBert и RoBerta. Они показали результат хуже чем DistilBert, предположительно, как раз из-за переобучения.

### 3 Результаты на соревновании

Results						
#	User	Entries	Date of Last Entry	Team Name	Spearman Correlation ▲	Detailed Results
1	jgrmnmx	1	12/08/23		0.844 (1)	<a href="#">View</a>
2	aazhangmiaoran	5	11/18/23		0.826 (2)	<a href="#">View</a>
3	S_231110043	7	11/09/23		0.823 (3)	<a href="#">View</a>
4	Ayon	7	12/10/23	Unknown	0.815 (4)	<a href="#">View</a>
5	Syavaprd	5	11/11/23		0.814 (5)	<a href="#">View</a>
6	S_231110608	3	10/19/23		0.808 (6)	<a href="#">View</a>
7	Sls	15	12/05/23		0.806 (7)	<a href="#">View</a>
8	spoker	1	12/10/23		0.797 (8)	<a href="#">View</a>
9	janko0329	4	11/28/23	DDDDDDDD	0.781 (9)	<a href="#">View</a>
10	S_231110606	7	11/02/23		0.753 (10)	<a href="#">View</a>
11	ivan_belyaev	2	11/08/23		0.671 (11)	<a href="#">View</a>
12	Pollgorka	1	10/31/23		0.469 (12)	<a href="#">View</a>
13	Shilpa	1	10/17/23	SemanticLM	0.394 (13)	<a href="#">View</a>

#	Username	Score
1	S_231110608	0.675

Рис. 1: Leaderboard

Выделение синим - наш текущий результат 0.797 (User poker)

Наш прошлый результат - 0.671 (User ivan belyaev)

## **4 Итоги**

В конечном итоге у нас получилось достичь около 80% точности ответа. Как ни странно, облегчение модели помогло нам улучшить результат.

Большинство участников, которые нас превзошли, получили результаты буквально на пару процентов выше. От лучшего результата мы отстаем не более чем на 5%