



# Разбор задач

AI Journey 2019, экзамен по русскому языку

Татьяна Шаврина (Sberbank, HSE), Валентин Малых (Huawei)

- Антон Емельянов (Sberbank, MIPT)
- Алена Феногенова (Sberbank)
- Вадим Фомин (Sberbank, HSE)
- Владислав Михайлов (Sberbank, HSE)
- Владимир Ларин (Sberbank)
- Андрей Евлампиев (Sberbank)
- Александр Ватулин (ODS)
- Александр Натекин (ODS)
- Илья Демьянов (Sberbank)
- Петр Ромов (ODS)
- Андрей Черток (Sberbank)
- Амир Бакаров (Huawei)
- группа методологов из Института Образования ВШЭ
- две группы учителей и методологов госэкзаменов из Москвы и Новосибирска

и многие другие!

# Задача 2019

Цель: решить экзамен по русскому языку

26 заданий - с открытыми и закрытыми вариантами ответов (59% от оценки)

+ сочинение по тексту (41% от оценки)

Типы заданий:

- орфография
- логика
- семантика
- пунктуация
- орфоэпия (ударения)
- морфология и синтаксис
- составление / генерация текста



# Пример задачи

№ 9. Укажите варианты ответов, в которых во всех словах одного ряда пропущена безударная чередующаяся гласная корня. Запишите номера ответов.

- 1) зап..рать, р..стение, прил..гательное
- 2) сп..раль, заст..лить, к..мфорт
- 3) б..режок, ф..рмат, затв..рдеть
- 4) предв..рительный, прид..рожный, зам..чать
- 5) тв..рительный, з..рница, пл..вец

Ответ: 1, 5

# AI Journey в Числах

98 команд

Самые активные участники послали решение 206 раз!

Один из самых сложных бейслайнов. Авторы - 11 человек из Сбербанка и Huawei, и содержит 4000+ строк кода.

88 команд воспользовались бейслайном (30 баллов), при этом 60+ команд переработали его

Уровень решений участников очень разный - независимо от места. В топ-10 есть и команды компаний, и студенческие команды

# AI Journey в числах

Много пасхалок от участников:

- ❖ фото кота в итоговом докер-контейнере

Одна команда написала больше 200 условий if else для решения теста

Всего более 1000 сабмишенов







Eventually  
Russian AI ...  
got 69 out of 100  
for the exam!

# Лучшие решения

- сочинение - 66%
- тест - 61%

Перевод баллов в  
оценки:

2 - 0-23  
3 - 24-56  
4 - 57-71  
5 - 72-100

лучшее решение - 59 баллов из 100

лучшее сочинение + тест: 62-63 балла

Средний результат школьника - 68 баллов



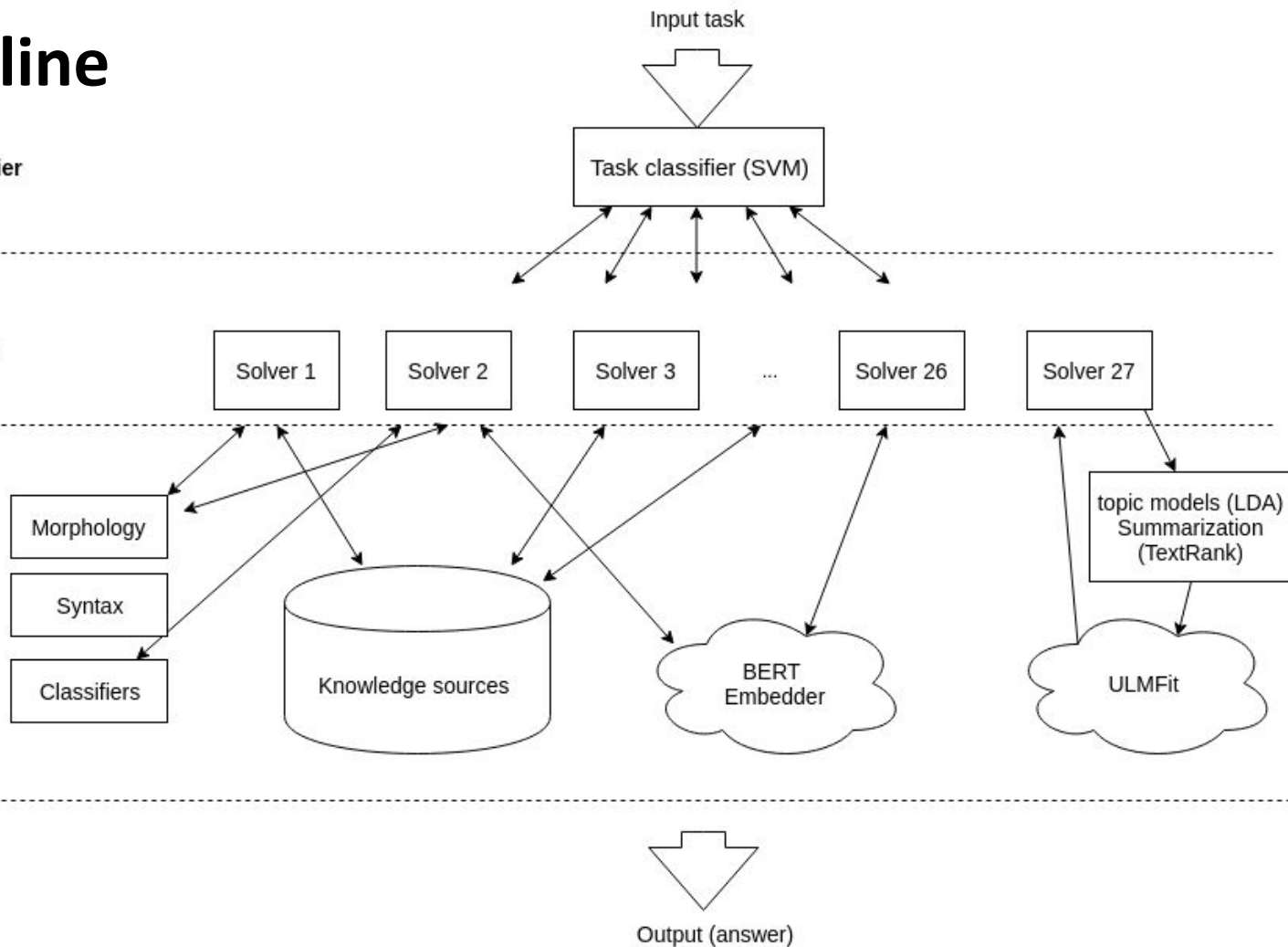
# Baseline

Task classifier

Task parsing  
and  
processing

Models and  
utils

Answer



if 'колбаса' in variant  
else

for w in slovarnie\_slova:  
...

# Решения участников

тест

pain = "

if "более" in x.lower():  
y\_pred =  
solver\_NAI\_BOLEE(x)

# Грамматические ошибки. Задание 8 (5 баллов)

## ГРАММАТИЧЕСКИЕ ОШИБКИ

- А. неправильное употребление падежной формы существительного с предлогом
- Б. нарушение в построении предложения с несогласованным приложением
- В. ошибка в построении предложения с однородными членами
- Г. неправильное построение предложения с деепричастным оборотом
- Д. нарушение построения предложения с причастным оборотом

## ПРЕДЛОЖЕНИЯ

1. Медвежонок плыл, зажмурившись, поэтому всё время натёкался то на Белку, то на Хомячка.
2. Так туча, закрывавшая небо целый день, внезапно уходит и обнажает огромный кусок пронзительной синевы.
3. Выйдя из переговорной, Радмиле многое показалось странным.
4. По прибытию на вокзал Анита немедленно бросилась звонить матери.
5. Дядя Сергей отличался от соседей только тем, что очень любил и гордился своей дочерью.
6. Алексею показалось, что не всё было в порядке с романом «Анной Карениной».
7. А вторая половина Земли, убегающая всё дальше и дальше трещиной, отъезжала с тяжким кряхтением, как, лопнув, разваливается перезрелый арбуз.
8. Никто даже не обратил внимания на заваленную тумбочку вещами.

# Грамматические ошибки. Задание 8 (5 баллов)

Какая из вышеперечисленных моделей лучше справляется с классификацией грамматических ошибок?

В ответ запишите номер модели.

- 1) классификатор на BERT embeddings предложения
- 2) куча if else
- 3) LightGBM + словари
- 4) stanfordnlp + правила синтаксических связей

Ответ:

# Грамматические ошибки. Задание 8 (5 баллов)

Какая из вышеперечисленных моделей лучше справляется с классификацией грамматических ошибок?

В ответ запишите номер модели.

- 1) классификатор на BERT embeddings предложения
- 2) куча if else
- 3) LightGBM + словари
- 4) stanfordnlp + правила синтаксических связей

# Пунктуация. Задания 16-21 (5 баллов)

## 2. Задание #T113



Расставьте все знаки препинания: укажите цифру(-ы), на месте которой(-ых) в предложении должна(-ы) стоять запятая(-ые).

Лето было поздним (1) и (2) хотя был уже конец июня (3) зелень парков была ещё новой, необтрепанной (4) и липа цвела запоздало...

Ответ



Проверить ответ

[Показать разбор и ответ](#)

# Пунктуация. Задания 16-21 (5 баллов)

1. Модель предсказания пунктуации на Bert
2. LightGBM-классификатор поверх мешка n-грам слов и мешка n-грам PoS тегов.
3. Stanfordnlp pipeline для предсказания мест запятых, для сложных случаев catboost
4. LightGBM-классификатор поверх векторов предложений на TF-IDF лемм (pymorphy)



# Пунктуация. Задания 16-21 (5 баллов)

1. Модель предсказания пунктуации на Bert
2. LightGBM-классификатор поверх мешка n-грам слов и мешка n-грам PoS тегов.
3. Stanfordnlp pipeline для предсказания мест запятых, для сложных случаев catboost
4. LightGBM-классификатор поверх векторов предложений на TF-IDF лемм (pymorphy)

# Литературные тропы. Задание 26 (4 балла)

*(1) На этот раз Гараське пришлось, видимо, преодолеть нелегкий путь. (2) Отрепья, делавшие вид, что они серьезно прикрывают его тощее тело, были все в грязи, еще не успевшей засохнуть. (3) Физиономия Гараськи, с большим отвислым красным носом, бесспорно служившим одной из причин его неустойчивости, покрытая жиденькой и неравномерно распределенной растительностью, хранила на себе вещественные знаки вещественных отношений к*

[Показать полностью](#)

«Рассказ Л.Н. Андреева пробуждает в читателе искренние чувства, демонстрируя удивительную способность литературы вызывать сопереживание. Используя такие тропы, как (А)\_\_\_\_ (предложения 21, 56) и (Б)\_\_\_\_ (2, 20), Л.Н. Андреев придаёт изобразительную глубину, часто ироническую, своему повествованию. Вероятно, с этой же целью автор применяет такое лексическое средство, как (В)\_\_\_\_ (в предложении 43). В синтаксисе Л.Н. Андреев передаёт экспрессивность повествования при помощи (Г)\_\_\_\_ (например, предложения 26, 39)».

## Список терминов

1. метафора
2. олицетворение
3. парцелляция
4. просторечие
5. ряды однородных членов предложения
6. восклицательные предложения
7. вопросно-ответная форма изложения

# Литературные тропы. Задание 26 (4 балла)

1. Словари + if else
2. BertMulticlassClassifier выбирает тип тропа
3. Bert embeddings предложений + PCA + логистическая регрессия

# Литературные тропы. Задание 26 (4 балла)

1. Словари + if else
2. BertMulticlassClassifier выбирает тип тропа
3. Bert embeddings предложений + PCA + логистическая регрессия

`my_position = [  
 'Откровенно говоря, я  
 не могу уверенно сказать ...']`

# Решения участников

## сочинение по тексту

`['Я считаю, что {}'.format(problem_thesis_dict[theme])]`

# Сочинение. Задание 27

11 критериев качества текста

Заданная структура сочинения:

вступление,

проблема

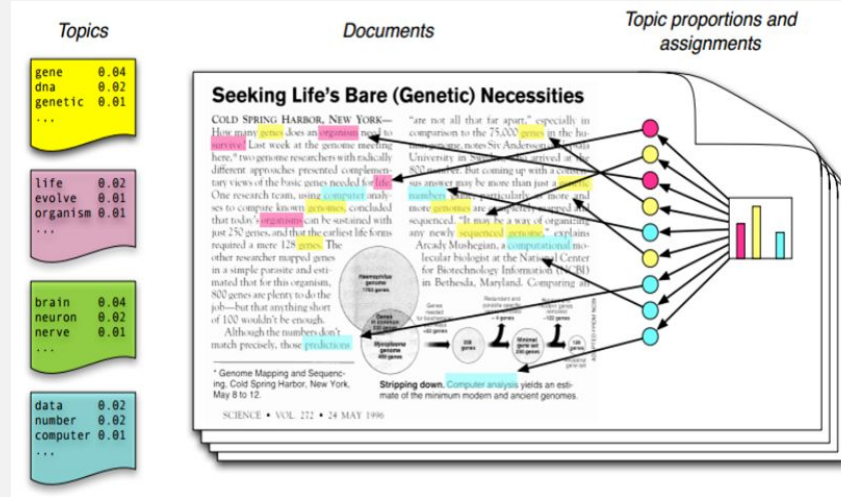
согласие/несогласие

аргументы

заключение

Baseline

сборник школьных сочинений



LDA, тематика текста +  
соотв. шаблон  
+TextRank

Доученный AWD LSTM  
(ULMFiT), продолжает  
сочинение, получив  
шаблон

# *Цитаты великих нейросетей*



“Кому на Руси жить хорошо? Кто такой Алексей Толстой? Я всегда верил, что Русь теперь уже обладает какой - нибудь другой теорией. “

“Драма «Горе от ума» — настоящее произведение Пушкина. Это «Евгений Онегин». Главный герой — Евгений Онегин. Этот факт не вызывает иронии, он не является автором, но и сам главным героем поэмы. Конец сюжета поэмы — это произведение в полном их развитии.”



# Лучшие решения:

- 1) Самописные шаблоны  
few-shot классификация проблем авторского текста  
(BERT embeddings)  
Поиск авторской позиции в тексте с помощью BERT
- 2) классификация проблем авторского текста с  
помощью извлечения ключевых слов (yake)  
Краткий пересказ текста в помощью summa  
База аргументов и шаблоны
- 3) эксперименты с GPT2  
TextRank  
NER от DeepPavlov  
нарезка готовых сочинений на предложения и составление из них текста  
перифразы кусков текста с помощью векторных моделей



# Шаблоны:

' Автор данного для анализа текста не случайно обратился к проблеме {}. Данная проблема занимала умы многих поколений людей и остается актуальной в современном мире.'.format(problem\_string\_dict[theme]))

'В глубине души я бы хотел отвергнуть авторскую позицию. Опровергнуть эти выводы из одного лишь духа противоречия! '


'Но нет... Нужно признать: автор прав. Не только лишь аргументы из данного текста - вся мировая литература против меня. '

'Как не вспомнить тут ещё одно произведение. {}\\nС такими авторитетами не поспоришь (да и не хочется).'

# Проблемы соревнования



# Взломать AI Journey

- 1) на первом этапе соревнования 2 раза взламывался evaluation script
  - 2) оценка сочинений:
    - a) плагиат
    - b) заготовленные тексты
    - c)  учителя
- Эффект зловещей долины у 50% учителей

Учителя:

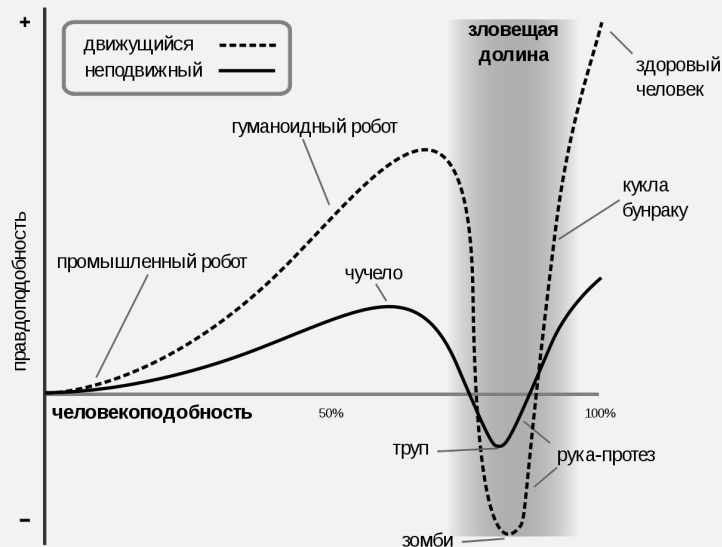
**Это не текст!**

# Проверить 300 сгенерированных сочинений

**Эффéкт «зловéщей доли́ны»** (англ. *uncanny valley*) — гипотеза, по которой **робот** или другой объект, выглядящий или действующий примерно как человек (но не точно так, как настоящий), вызывает неприязнь и отвращение у людей-наблюдателей.

Шаблонное начало и структура

Но: несуществующие персонажи, ошибки генерации, отсутствие общей мысли



# Вместо выводов:

- 1) Начинали с очень низких баллов, но закончили “четверкой” !
- 2) Среди команд в топ-10 очень разный уровень и подходы:
  - a) инженерные подходы и базы знаний
  - b) условия и правила
  - c) BERT everywhere!
- 3) Текущий уровень генерации текста уступает по качеству реальным текстам,  
эмпатичные шаблоны ее полностью победили

Критерий качества сгенерированного текста очень высок, ничто не должно выдавать его природу, чтобы разметчики оценили его.

Нужна

более управляемая генерация,  
генерация, зависящая от дискурсивных параметров.





**Спасибо!**  
**Это было здорово!**