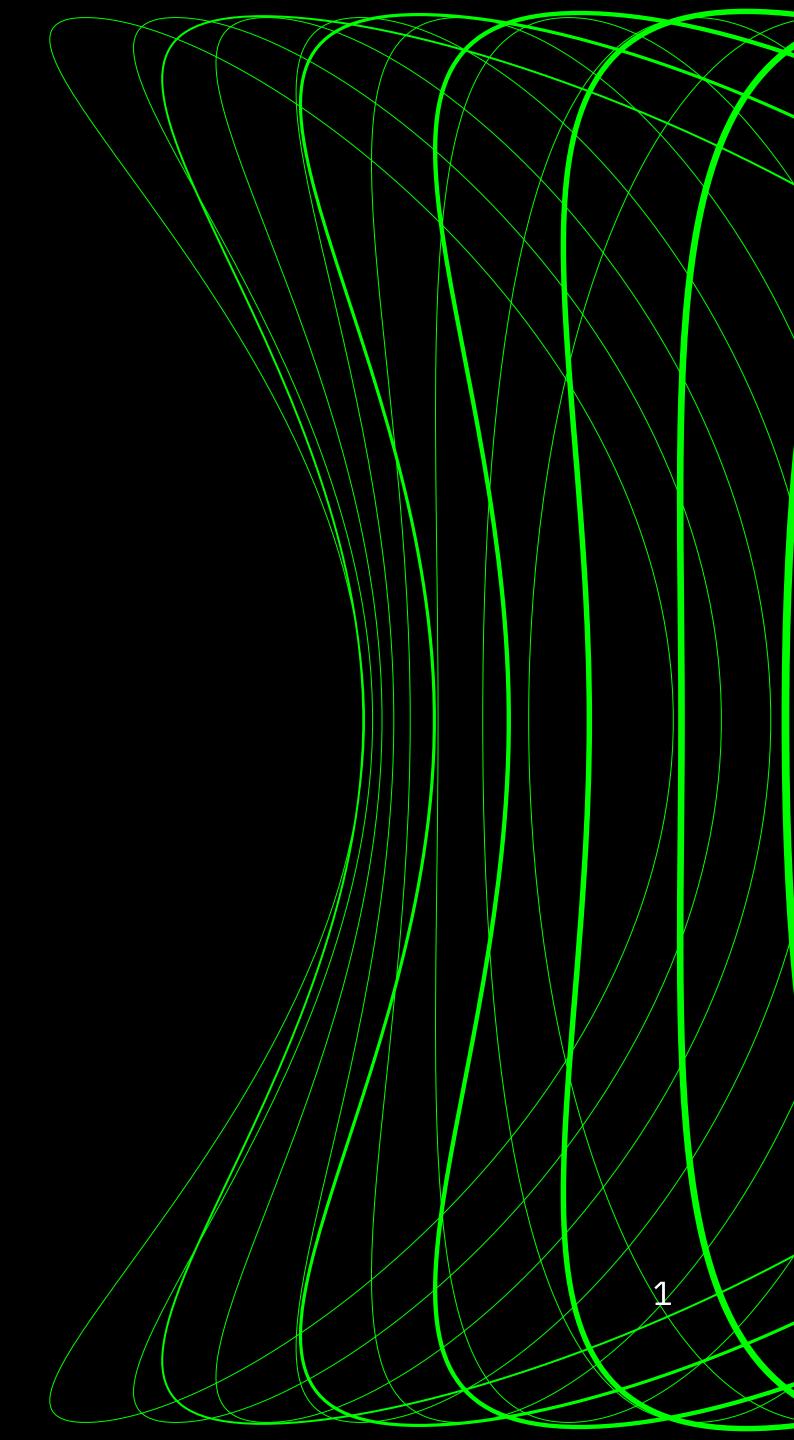
цифровой сезон: ии

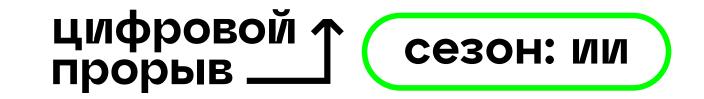
ПАО «РОСТЕЛЕКОМ»

Переводчик с естественного языка на SQL









Кейсодержатель

Название организации

ПАО «Ростелеком»

Сфера деятельности

Финансовая, выдача кредитных продуктов, открытие банковских счетов, услуги

Краткое описание кейса

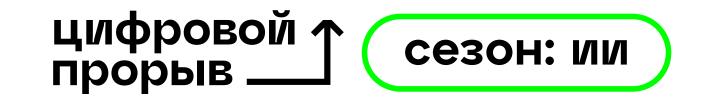
Кластеризация индивидуальных предпринимателей РФ с целью оценки доходности на основе открытых источников данных.

Сайт организации

https://rtkit.ru







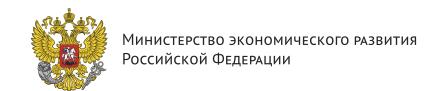
Постановка задачи

Современные технологии позволяют легко решать задачи перевода естественных человеческих языков. Доступно большое количество технологий и уже обученных нейросетей для решения подобных задач, но стабильно работающего переводчика с человеческой речи на язык программирования — нет.

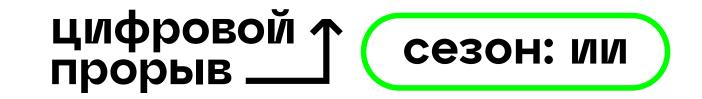
Решения в datadriven-компаниях принимаются на основании проверки гипотез на данных. При этом большая часть таких гипотез не является сложными — такие

проверки могут осуществлять сами участники команд со стороны бизнеса, но они не всегда обладают знаниями языков программирования.

Участникам предлагается реализовать переводчикс языкабизнес-запросов на язык SQL.







Проблематика

Разработанное участниками решение позволит сократить время при принятии решений, ускорит бизнес-процессы в компании ПАО «Ростелеком», а также позволит командам быстрее обучаться работе с данными и сократит нагрузку на дата-аналитиков.

Решение

Решение кейса представляет собой модель которая, используя текст запроса, сможет его перевести на язык запросов SQL.





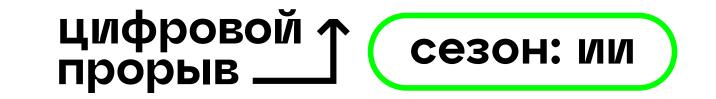


Стек технологий, рекомендуемых/обязательных к использованию

Программы и библиотеки с открытым исходным кодом (open source) и общедоступные публичные данные, язык программирования - Python







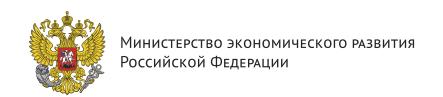
Методы и критерии оценивания

Для оценки решений применяется метод экспертных оценок и автоматизированные средства оценивания.

Жюри состоит из отраслевых экспертов и/или представителей Кейсодержателя.

На основании описанных ниже характеристик, жюри выставляет оценки 0-3 балла.

Итоговая оценка определяется, как сумма баллов всех экспертов, отраслевых и/или представителей Кейсодержателя.



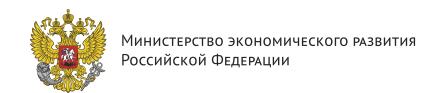




Отраслевой эксперт и/или представитель Кейсодержателя оценивает решение по следующим критериям:

- 1. Релевантность поставленной задаче (команда погрузилась в отрасль, проблематику; предложенное решение соответствует поставленной задаче; проблема и решение структурированы)
- 2. Уровень реализации (концепция/ прототип и тд)
- 3. Проработка пользовательских историй (UX/ UI)
- 4. Реализация в решении требований Заказчика

5. Выступление команды, (умение презентовать результаты своей работы, строить логичный, понятный и интересный рассказ для презентации результатов своей работы)





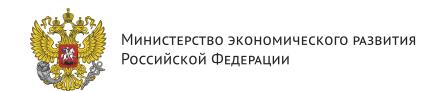


Технический эксперт оценивает решение по следующим критериям:

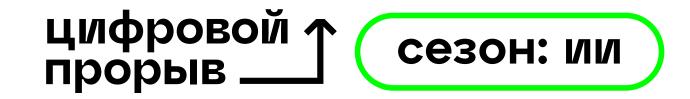
- 1. Запускаемость кода
- 2. Обоснованность выбранного метода (описание подходов к решению, их обоснование и релевантность задаче)
- 3. Точность работы алгоритма (возможность оценить формальной метрикой с обоснованием выбора)
- 4. Адаптивность/ Масштабируемость

- 5. Отсутствие в решении импортного ПО и библиотек, кроме свободно распространяемого с обоснованием выбора
- 6. Наличие интеграционных интерфейсов, в первую очередь интерфейсов загрузки данных

Итоговая оценка определяется как итоговый балл жюри, умноженный на оценку автоматизированной системы.







Ссылки на датасет



https://ai-data.obs.ru-moscow-1.hc.sbercloud.ru/Переводчик_с_естественного_языка_на_SQL.zip

cesoh: MM

