An isometric illustration on a dark blue background featuring several laptops and a central server unit. The server has a glowing green and red grid on its top surface. Thin, glowing blue and purple lines resembling circuitry or data paths connect the devices and extend across the background.

Придумай инструмент,
который сможет находить
связи между параграфами

SEARCH IT! **МТС**

● Быть лучше каждый день

Количество существующих нормативных документов огромно, и иногда сложно выбрать из них наиболее актуальные для использования при разработке корпоративной политики компании. В кейсе вам предстоит помочь МТС с управлением регуляторными изменениями и разработать инструмент поиска противоречий между документацией. Вам необходимо создать модель, способную находить связи между параграфами так же точно, как лучшие комплаенс-эксперты делают это вручную.



ВВЕДЕНИЕ

3



О КОМПАНИИ

6



**КАРЬЕРНЫЕ ВОЗМОЖНОСТИ
В МТС**

10



ВВЕДЕНИЕ



Морозным февральским вечером Максим¹, руководитель отдела в Центре искусственного интеллекта МТС, спешил покинуть офис компании. Он задержался, так как хотел дожидаться результатов эксперимента с новой моделью. Проходя по коридору, Максим увидел, что дверь в кабинет Ирины, руководителя комплаенс-отдела², приоткрыта, и там горит свет.

— К тебе можно? — постучал в дверь Максим. — Домой не собираешься?

— Да, заходи, — откликнулась Ирина. — Вот закончу проверку нашего внутреннего документа на соответствие изменениям в законодательстве и уйду.

— Это своего рода регуляторная гильотина³, да? — спросил Максим.

— Можно и так сказать, — улыбнулась Ирина. — Понимаешь, нам очень важно вовремя вносить соответствующие поправки во внутреннюю нормативную базу МТС. От этого зависит наша инвестиционная привлекательность и репутация. Мы не можем позволить, чтобы комплаенс-риски повлияли на операционную деятельность и финансовое состояние компании, — Ирина снова погрузилась в свои документы.

— Только не говори, что ты сверяешь документацию вручную, — нахмурился Максим. — Немудрено, что ты часто задерживаешься в офисе допоздна.

— Да, я всегда проверяю документы вручную, чтобы ни одно изменение в регуляциях не осталось незамеченным, — серьезно ответила Ирина. — И не я одна. Ты не представляешь, сколько у нас документов, в которые с каждым изменением в законодательстве нужно вносить правки. Мы регулярно привлекаем к этим задачам тысячи юристов.

«Должен быть способ делать это быстрее», — подумал Максим, а вслух произнес:

— Ты знаешь, у нас в Центре искусственного интеллекта есть целый отдел, который занимается автоматической обработкой естественного языка, и сравнение параграфов является одной из типовых задач. Так что, думаю, мы сможем тебе помочь.

Моя идея — для каждого параграфа в новой регуляции найти логически связанные с ним параграфы в наших внутренних документах. Эта задача давно прорабатывается многими учеными по всему миру, в статьях ее обычно называют *Natural Language Inference*. Наиболее распространенным подходом к решению является представление каждого параграфа в документе в виде точки в многомерном пространстве и последующий расчет расстояния между этими точками. Чем ближе расположены точки, тем выше шанс, что параграфы логически связаны. Ты бы смогла отказать от привычной рутины в пользу современных технологий?

ФАКТ 1



МТС — первый российский сотовый оператор, установивший станцию сотовой связи в Антарктиде и развернувший на станции «Прогресс» сеть *NB-IoT* для передачи данных. Трекеры, мультисенсоры и датчики, работающие на сети, позволят исследовать движение ледников и снежного покрова, контролировать целостность снеговых взлетно-посадочных полос и микроклимат в жилых и рабочих помещениях комплекса.

Максим замолчал, чтобы дать коллеге подумать. Он отошел к окну — с третьего этажа открывался вид на проспект Андропова. Максим надеялся, что Ирина по достоинству оценит идею автоматизированного поиска логических связей между параграфами в регуляциях и внутренних документах, и они смогут воплотить проект в жизнь.

— Звучит многообещающе, — нарушила паузу Ирина. — Я понимаю, что за этими технологиями будущее. Мы в МТС инвестируем в такие перспективные направления, как ИИ и большие данные.

¹ Все имена и названия вымышленные, любые совпадения случайны. Данные в кейсе могут быть изменены в целях конфиденциальности.

² Комплаенс — это комплекс инициатив, направленных на предупреждение противоречащих закону действий сотрудников компании и внедрение корпоративной бизнес-этики, основанной на соблюдении законодательства.

³ Термин означает инвентаризацию всех действующих и обязательных для бизнеса требований с целью понять, соответствуют ли они современным реалиям. Если соответствуют, то нормы остаются, если нет, то они отменяются или изменяются (<https://knd.ac.gov.ru/about/>).

Так что внедрение технологий обработки естественного языка — хорошая возможность сделать наш повседневный труд более технологичным. Но за свою долгую карьеру я так привыкла к стандартной схеме работы с документацией. Боюсь, мне будет непросто перестроиться. Как я смогу быть уверена, что автоматизированный поиск действительно нашел все параграфы, которые должны быть изменены с учетом изменений в законах, и мне не придется потом все исправлять? И ведь нам нужно не просто найти логически связанные параграфы, а еще понять тип этих связей. На самом деле нас интересуют только те параграфы в наших внутренних документах, которые противоречат обновлениям в законодательстве.

— Ладно, давай сделаем вот что, — произнес Максим. — Ты даешь мне одну неделю, за которую я предложу готовый инструмент поиска противоречий между документами. Я создам прототип модели на основе открытых данных, но позже мы сможем адаптировать ее и для нашей внутренней документации. А пока мне предстоит выбрать наиболее подходящий алгоритм из множества доступных вариантов и проверить его работу на открытых данных.

— Хорошо, я согласна попробовать. Спасибо, Максим! — улыбнулась в ответ Ирина. — Если в апреле я смогу использовать твою модель, а не сверять документы вручную, значит, ты настоящий волшебник.

— Не волшебник, а дата-сайентист, — отозвался Максим. — Завтра пришлю тебе на почту описание того, что планирую сделать в ближайшие семь дней по нашей задаче.

Тема: Разработка инструмента, который сможет находить связи между параграфами

— □ ×

От кого: Максим Обухов

Кому: Ирина Михайлова



Ирина, доброе утро!

На начальном этапе проекта хочу сконцентрироваться на задаче классификации связей между параграфами. В качестве исходных данных собираюсь использовать открытый и хорошо известный датасет SNLI (Stanford Natural Language Inference Dataset), который был подготовлен группой ученых из Стэнфордского университета. По выходным я посещаю онлайн-курс доцента Сэма Боумана, и он посоветовал начать наш проект с апробации модели на открытых данных.

Описание датасета и ссылка для скачивания доступны на странице. Датасет представляет собой таблицу (столбцы разделены символом табуляции '\t', а строки — символом переноса строки '\n'), в каждой строке которой находится пара параграфов и тип связи между ними:

1. entailment (из параграфа 1 следует параграф 2)
2. contradiction (параграф 1 противоречит параграфу 2)
3. neutral (в параграфе 1 и параграфе 2 содержится схожая по смыслу информация)

Датасет был размечен вручную, каждая пара параграфов оценивалась пятью независимыми экспертами. Столбцы *label1...label5* соответствуют оценке каждого человека, а столбец *gold_label* содержит обобщенную оценку (majority voting). Помимо этого, в датасете можно найти лексический разбор каждого параграфа, что может оказаться очень полезным при подготовке фичей.



← Ответить

↗ Переслать

Тема: Разработка инструмента, который сможет находить связи между параграфами

— □ ×

От кого: Максим Обухов

Кому: Ирина Михайлова



Датасет разбит на три части:

1. train — для обучения
2. dev — для валидации
3. test — для финальной оценки.

Нам необходимо обучить модель, которая будет автоматически определять один из трех вышеперечисленных типов связей между параграфами. Можно пробовать различные подходы — как классические (word2vec, Word Mover's Distance, Jaccard), так и нейросетевые (LSTM, Transformers).

При этом нужно:

1. Убедиться, что модель одинаково хорошо работает как на обучающей, так и на тестовой выборках и оценить степень уверенности модели в выдаваемых предсказаниях.
2. Произвести анализ данных, посчитать основные статистики.
3. Подобрать метрики для оценки качества, которые будут достаточно точно отображать качество работы модели для всех классов.
4. Произвести тюнинг модели, то есть оценить, как влияет изменение параметров модели на качество их работы.

Буду держать тебя в курсе.
Хорошего дня!

С уважением,
Максим



← Ответить

↗ Переслать

ФАКТ 2



В 2020 году 19 руководителей из разных направлений бизнеса МТС вошли в ежегодный рейтинг «1000 лучших топ-менеджеров России».



О КОМПАНИИ



МТС — ведущая цифровая компания, создающая на базе основного телеком-бизнеса, где компания является безусловным лидером российского рынка более 20 лет, продуктовую экосистему цифровых сервисов. В активе компании большое количество инновационных решений в сфере интернета вещей, больших данных, искусственного интеллекта, облачного хранения, медиа, развлечений, онлайн-образования, финансовых сервисов и телемедицины.

ФАКТ 3



Осенью 2020 года МТС и Сколтех запустили в «Сколково» пилотную зону 5G для создания отечественных решений в IT-сфере. Совместный проект ускорит выход на рынок готовых промышленных 5G-решений и сервисов для применения в различных отраслях экономики.

МТС ежегодно укрепляет лидирующие позиции на рынке по таким ключевым показателям, как количество абонентов и выручка. В 2019 году Группа МТС продемонстрировала рекордные финансовые результаты. Выручка компании выросла на 5,5 % и составила 476,1 млрд руб., а количество абонентов превысило 85 млн в России, Беларуси и Армении.

В 2019 году МТС приняла новую стратегию развития CLV 2.0, которая подразумевает предоставление клиенту множества цифровых сервисов за пределами телекоммуникаций и ориентирована на создание на базе основного телеком-бизнеса продуктовой экосистемы с бесшовным переходом между сервисами. МТС развивает клиентоцентричный подход, который повысит продолжительность нахождения клиента в экосистеме, и сфокусирована на долгосрочной прибыли от клиента за счет концентрации на его потребностях, повышении лояльности и предоставлении выгодных условий при использовании комплексом продуктов. Благодаря этому МТС сможет вывести на рынок больше конвергентных решений, способствуя повышению уровня жизни и развитию цифровой экономики в России.

Сегодня экосистема МТС состоит из нескольких ключевых компонентов. Базовый телекоммуникационный бизнес компании — услуги мобильной и фиксированной связи. Финтех, где по сравнению с другими телеком-игроками МТС имеет су-

ЭКОСИСТЕМА МТС:

- 78 миллионов мобильных абонентов
- 20 млн пользователей Мой МТС
- Более 5 млн участников программы МТС Cashback
- Более 1 млн пользователей музыкальных сервисов
- 2,5 млн проданных билетов
- Более 4,5 млн пользователей ТВ
- Около 2 млн пользователей интерактивного ТВ
- Более 1 млн пользователей OTT ТВ
- Более 2 млн клиентов МТС Банка
- 1 млн пользователей мобильного приложения банка
- Более 1,5 млн пользователей Daily Banking

щественное преимущество благодаря МТС Банку. Развлечения — медийные сервисы и киберспорт. Цифровые сервисы для бизнеса, которые в будущем будут наиболее тесно связаны с проникновением 5G, интернета вещей и облачными решениями.

Развитие этих направлений и работа почти над каждым продуктом МТС основывается на анализе больших данных, искусственном интеллекте, клиентском опыте, едином ID, IT-разработке и программах лояльности.

О КОМПАНИИ



DATA (ДАННЫЕ)

Данные остаются важным направлением для развития долгосрочной конкурентоспособности в базовом бизнесе. МТС продолжает усиленно инвестировать в развитие телекоммуникационной инфраструктуры с фокусом на улучшение покрытия и качества связи. В конце 2019 года охват населения России сетями LTE МТС достиг 74 %. Проникновение смартфонов в сети МТС достигло 74,5 %, а проникновение пользователей мобильного интернета увеличилось до 62,9 %. Объем трафика составил более 9 Гб в среднем на одного пользователя в месяц.

В 2020 году компания развернула первую в России промышленную 5G-зону частной сети LTE на КамАЗе в Набережных Челнах. Задача системы — оптимизировать производство грузовых автомобилей и повысить его конкурентоспособность. Частная (private) беспроводная сеть объединила различные элементы производственной экосистемы предприятия в закрытый контур, что должно обеспечить повышенную безопасность и надежность. На первом этапе в сети развернуты системы видеонаблюдения и групповой связи, защищенный доступ к локальным информационным ресурсам и виртуальное решение для удаленного обучения персонала.

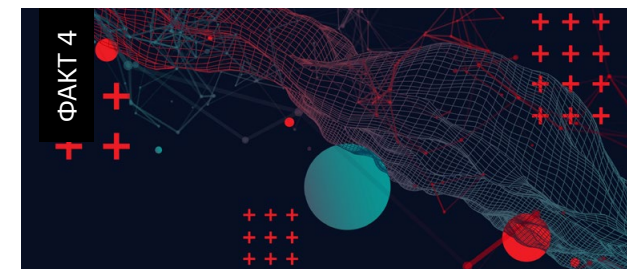
В 2019 году МТС заключила соглашение с правительством Москвы о развитии 5G. МТС намерена развернуть постоянно действующие сети 5G в столице до 2022 года, инвестировав в создание сетей пятого поколения более 20 млрд рублей.

DIGITAL (ЦИФРОВИЗАЦИЯ)

Компания постоянно внедряет цифровые и конвергентные сервисы. Это позволяет обеспечивать клиентов самыми актуальными сервисами и поддерживать эффективность деятельности компании как цифровой экосистемы.

На конец 2019 года количество пользователей конвергентных продуктов (фиксированные услуги + мобильные услуги + спутниковое ТВ) достигло более 950 тыс., а количество подписчиков «МТС Музыка» в партнерстве с Apple Music превысило 1 млн. При этом отток подобных мультипродуктовых абонентов снизился более чем в 3 раза, а средний чек вырос более чем в 2 раза. Доля абонентов, пользующихся нетелеком-продуктами МТС, составила свыше 20 %.

Рост экосистемы МТС реализуется как через развитие цифровых направлений бизнеса, так и через партнерства, собственные разработки инновационных продуктов и экспансию на новые рынки. Так, компания имеет контрольный пакет в LiteBox — разработчике облачных касс для корпоративных клиентов, владеет двумя ведущими российскими билетными операторами на рынке культурно-развлекательных мероприятий — «Пономиналу.ру» и Ticketland.ru, инвестировала в один из крупнейших российских онлайн-сервисов Youdo.com по поиску исполнителей для решения бытовых и бизнес-задач и в один из лидирующих киберспортивных клубов мира Gambit Esports. А в 2019 году запустила собственную интерактивную медиаплатформу для кибер-



МТС и Московская школа управления «Сколково» совместно реализуют программу дополнительного бизнес-образования **MOOVE**, направленную на развитие компетенций управления проектами в сфере цифровых технологий. В рамках второго набора МТС предоставит студентам гранты на обучение в размере 18 млн рублей.

спортсменов и геймеров WASD.TV. Помимо этого, МТС купила киностудию «Кинопалис» и планирует создать собственного производителя видеоконтента «МТС Studio».

DIVIDENDS (ДИВИДЕНДЫ)

Высокие дивиденды для акционеров являются еще одним приоритетом компании. Общая сумма выплаченных дивидендов за календарный 2019 год составила 28,66 рубля на одну обыкновенную акцию. Для повышения доходности по акциям компания отслеживает возможности неорганического роста: МТС обладает достаточным опытом приобретения и интеграции компаний, что позволяет максимально быстро получать положительный экономический эффект.



РИСУНОК 1. ЭКОСИСТЕМА ЦИФРОВЫХ ПРОДУКТОВ МТС

О КОМПАНИИ

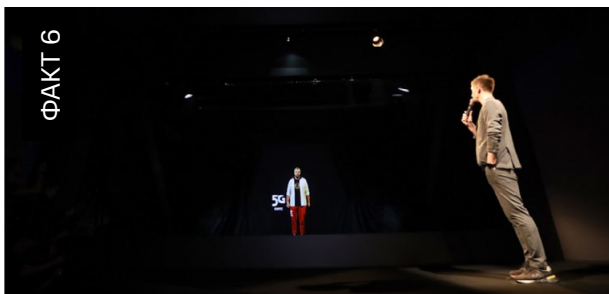


ФАКТ 5



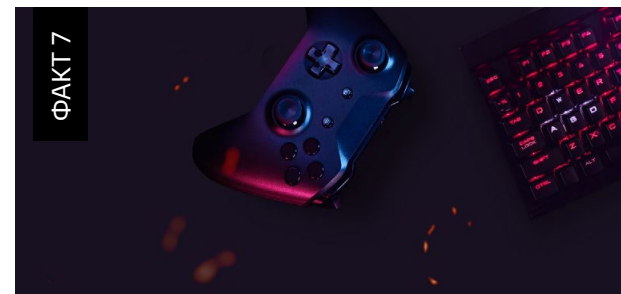
МТС реализует всероссийскую программу стажировок. Ежегодно более 200 студентов стажировались в IT-подразделениях компании. Среди них есть участники кейс-чемпионатов Changellenge >>.

ФАКТ 6



В 2019 году в ходе конференции «Неделя российского интернета» МТС провела международный двусторонний голографический телемост на сети 5G. Участники форума в Москве и их собеседники в Кронштадте наблюдали голографическое 3D-изображение друг друга в прямом эфире.

ФАКТ 7



В июле 2020 МТС запустила МТС Cloud gaming — маркетплейс, в котором собраны сервисы для облачного гейминга. Технология позволяет играть почти на любом компьютере удаленно. Главное — иметь стабильный и быстрый интернет.

ФАКТ 8



МТС запускает образовательную программу по разработке устройств для сети интернета вещей (IoT) вместе с ведущими техническими вузами России. Студенты получают оборудование для разработки энергоэффективных устройств, а преподаватели — экспертную и консультационную поддержку от специалистов компании.

ФАКТ 9



Для людей, больных диабетом, МТС разработала специальное приложение «ДиаБаланс». Решение позволяет самостоятельно контролировать заболевание и вносить в смартфон данные об уровне глюкозы. Записи в дневники питания и приема инсулина можно производить с помощью технологии распознавания голоса.

ФАКТ 10



МТС входит в топ-7 российских компаний по раскрытию информации об изменении климата в рейтинге Carbon Disclosure Project (CDP). В 2020 году компания повысила свою оценку до уровня «менеджмент», разделяя позиции с АФК «Система», ПАО «Газпром нефть» и «Русал».

КАРЬЕРНЫЕ ВОЗМОЖНОСТИ В МТС



МТС — один из крупнейших работодателей в России.

Наша команда — это более 60 тыс. профессионалов, которые ежедневно обеспечивают миллионы клиентов качественными услугами, развивают новые digital-направления, создают экосистему цифровых продуктов МТС.

Ежегодно более 400 студентов и выпускников вузов проходят стажировки в МТС, и более 30 % из них продолжают развиваться вместе с нами.

В этом году мы запускаем МТС INTRO — программу стажировок, свободную от стереотипов.



Почему МТС INTRO



Мы не вписываем стажеров в шаблоны, а хотим, чтобы каждый сам смог определить, что ему по душе. Перед тобой множество актуальных направлений и команд с самыми разными людьми — выбор за тобой.



Найди единомышленников, которые помогут тебе раскрыться и воплощать в жизнь самые смелые проекты.



Если тебе понравится работать с нами, оставайся и стань частью одной из команд бизнесов экосистемы МТС.

Условия



365. Начни карьеру зимой/весной/летом/осенью.



Совмещай. Выбирай график, который удобен тебе. Меняй его на протяжении стажировки по мере твоей загруженности в учебе.



Без границ. Построй свою карьеру прямо из дома в любом регионе.



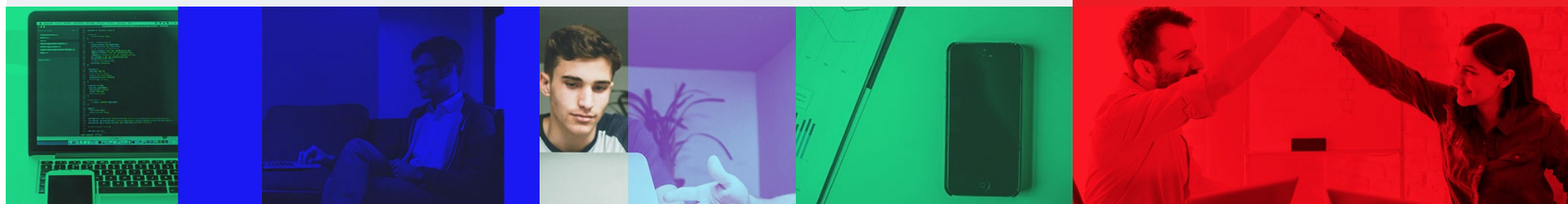
Занимайся тем, что нравится. Более 16 направлений стажировки — выбирай!



Стоит ли говорить про сложные **задачи, реальные проекты, оплату и крутую команду?**

НАЙДИ СЕБЯ В МТС INTRO

1. Разработка.
2. Тестирование.
3. Искусственный интеллект.
4. Big Data.
5. Аналитика.
6. Системное администрирование.
7. Информационная безопасность.
8. Дизайн.
9. Продуктовый менеджмент.
10. Маркетинг.
11. PR.
12. Работа с корпоративными клиентами.
13. Технический блок.
14. HR.
15. Финансы и инвестиции.
16. Юриспруденция.



CHANGELLENGE >>

Кейс написан и опубликован
Changellenge >> —
ведущей организацией
по кейсам в России.

www.changellenge.com
info@changellenge.com
vk.com/changellengeglobal
facebook.com/changellenge



Кейс создан по заказу
ПАО «Мобильные ТелеСистемы»

www.mts.ru