



## **Задача «Прогнозирование оттока пользователей провайдера телекоммуникационных услуг»**

### **Введение**

Сфера телекоммуникаций — высококонкурентная динамичная среда, в которой уровень осведомленности пользователей позволяет легко получить информацию о других поставщиках услуг с аналогичным или наиболее высоким качеством по более выгодной цене. Отсюда и возникает проблема оттока пользователей.

Для любого провайдера важным аспектом является пользовательская база. В нее поступают новые пользователи, какое-то время активно потребляют услуги и через некоторый промежуток времени перестают, этот промежуток называется «жизненный цикл клиента». Отток пользователей выражается в сокращении пользовательской базы и снижении показателей выручки.

Анализ пользовательского поведения и удержание пользователей — ключ к улучшению продуктов и услуг. В современном мире темпы роста постоянно ускоряются, рынок заполнен огромным количеством конкурентов, что приводит к постоянному повышению требований потребителей к продукции и услугам; - привлечение новых пользователей дороже процедур удержания. На привлечение новых пользователей, как правило, нужно заложить рекламный бюджет, в то время как существующих можно активизировать специальным предложением с особыми условиями.

Анализ данных о пользовательском поведении по косвенным признакам — одна из самых востребованных на сегодняшний день технологий не только у бизнеса, но и у государственных структур. Она позволит определять востребованные пользователем услуги, а также поможет в реализации персонализированного уникального подхода к запросам граждан.

Причинами оттока могут служить технические проблемы (сбои/аварии), неудовлетворенность работой компании в решении возникающих проблем, изменение тарифных планов, в том числе повышение цен, а также более выгодные условия компании конкурента, демпинг. Нужно отметить, не всегда клиенты отказываются от услуг только потому что у

компании-конкурента более выгодные тарифы, чаще всего вышеперечисленные причины наслаиваются и в совокупности формируют желание отказаться от услуг.

На данный момент многие компании предотвращают отток путем обзвона клиентов, которые уже обратились с техническими проблемами или иными претензиями, мотивируют их не прекращать пользование услугами. Во-первых такой вариант ресурсозатратен, во-вторых не все клиенты сообщают о намерении отключиться. Именно поэтому одна из главных задач телекоммуникационных компаний — прогнозирование оттока, путем распознавания сигналов к уходу в поведении клиента или по другим собираемым данным о пользовании услугами компании.

### **Условие задачи**

Участникам чемпионата предстоит создать свое решение по прогнозированию оттока с использованием данных о клиентах. В вашем доступе находится информация о запросах пользователей к сайтам конкурентов и история обращений в компанию.

### **Описание входных значений**

train.csv — файл содержащий данные пользователей для тренировки. Где:

- 1 - клиент ушел,
- 0 - остался.

log.csv — содержит данные обращения пользователей.;

named.csv — лог днс-запросов к доменам конкурентов (rt.ru и sampo.ru).

type contract.csv - тип списания у пользователей, где:

- 1 - посуточная,
- 0 - месячная.

submission.csv — пример файла для отправки.

### **Метрика**

В качестве метрики выступает Recall по столбцу **blocked**.

Recall считается как:

$$recall = \frac{TP}{TP + FN}$$

где: TP (True Positive) - количество верно угаданных значений одного класса  
 FN (False Negative) - количество не правильно угаданных значений  
 класса

### **Правила чемпионата:**

1. С момента открытия датасета до момента завершения приема решений репозиторий участника, в котором он ведет разработку по задаче текущего чемпионата, должен оставаться закрытым.
2. Участник обязан открыть доступ к репозиторию на чтение по ссылке (которая была прикреплена в ЛК в поле «Ссылка на код (гитхаб)») не позднее чем в течение 12 часов с момента окончания дедлайна отправки решений на региональном чемпионате.
3. Согласно п. 5.8 Положения в процессе верификации решений организаторы и технические эксперты, проверяющие решения участников, вправе назначить интервью с участниками чемпионата. Участник получит приглашение и ссылку на интервью не позднее, чем за 12 часов до публикации итогового лидерборда. Пропуск интервью участником является поводом для дисквалификации.
4. Организаторы вправе исключить участника из призовых позиций лидерборда за непредоставление одного из артефактов решения задачи: тизера, скринкаста, презентации, ссылки на репозиторий.
5. Организаторы вправе дисквалифицировать участника в случае выявления плагиата кода или несоблюдения Положения конкурса.
6. Участник, получивший 2 дисквалификации за сезон проекта, попадает в чёрный список с дальнейшим отстранением от участия в чемпионатах до конца сезона.