В 2012 году в средней школе Минервы, город Питтсбург, штат Пенсильвания, в которой учились почти 3000 учеников, существенно вырос процент отчислений – 9% учеников, которые поступили на первый курс, ушли, так и не закончив его. То есть ушёл каждый 10-ый. Совет школы отреагировал на мрачную статистику требованием принять срочные меры. В противном случае г-ну Вулкани – директору, пригрозили перестать финансировать школу или вовсе её закрыть.

Г-н Вулкани встретился с членами школьного совета. Они предложили ему собрать данные об учениках: дисциплинарные взыскания, посещаемость, оценки, покупки закусок или обедов и т. п. Для сбора данных могла использоваться в том числе школьная сеть Wi-Fi: она отслеживала использование учащимися интернета и их цифровой след.

Члены школьного совета предположили, что с помощью технологий они смогут пролить свет на причины высокого уровня отчислений. А уже потом можно думать, какие меры предпринимать.

Г-н Вулкани и школьный совет согласились предоставить компании по обработке данных Hephaestats информацию за всю историю существования школы, а доступ к более свежим данным открывать по мере сбора.

Но учащихся и их родителей об этом не уведомили. Соответственно, отказаться от сбора данных они не могли. Г-н Вулкани сознательно не стал выносить этот вопрос на обсуждение: практика показывала, что родители и администрация с большим трудом приходят к консенсусу по любой инициативе, а решить вопрос с отчислениями нужно было в кратчайшие сроки.

Что нужно сделать?

Ответьте на вопросы и напишите небольшое эссе.

1. Как в этой ситуации сделать решение о внедрении технологий ИИ этичным?
2. Повлияет ли на результаты тот факт, что заказчик – совет школы и директор, являются заинтересованными сторонами?
3. Как решение не включать опцию отказа о предоставлении данных влияет на легитимность школьной программы?
4. Нарушила ли школа конфиденциальность данных учеников, поделившись ими с Hephaestats? Поясните свой ответ.
5. Как вы думаете, если бы родители учащихся подали в суд, кто бы вероятнее всего выиграл дело?

В каком формате присылать результат

1. Создайте гугл-документ и напишите небольшое эссе – 1-2 страницы, 12 кеглем.
2. Опишите ход ваших мыслей, как бы вы рассуждали, решая такую задачу в реальности. А потом – как бы объясняли ваше решение заказчику.
3. Отправьте ссылку на гугл-документ в форму для сдачи работы.
4. Не забудьте настроить доступ к документу, чтобы ментор смог ознакомиться с его содержанием.

Говоря об этической составляющей внедрения технологий искусственного интеллекта в практику анализа данных, и, в частности, персональных данных, на наш взгляд, необходимо выделить несколько основных моментов. Во-первых, важно понимать, какие данные используются, и насколько эти данные чувствительны и значимы для человека, а также, каким образом они могут отразиться на конкретном человеке в случае их «утечки». Во-вторых, насколько непредвзятым будет предсказание модели, которая использует эти данные.

Если первый вопрос решается их анонимизацией (обезличиванием), путем удаления или преобразования информации таким образом, чтобы по полученным данным нельзя было определить источник информации и, к какому именно лицу они относятся, то вопрос предвзятости модели не имеет однозначного решения.

В контексте исследуемой проблемы, отметим, что, по нашему мнению, такие сведения как дисциплинарные взыскания, посещаемость, оценки учеников, и даже, покупки в школьной столовой, скорее всего могли быть использованы школой без согласия учеников и их родителей. Очевидно, что любая компания, учреждение или предприятие использует внутреннюю информацию для формирования отчетов, оценки своей деятельности. Как минимум, сбор, хранение и обработка вышеперечисленных сведений является необходимостью и обязанностью школы.

Другое дело личная информация, например, посещаемые учениками интернет ресурсы, сбор и обработка которой которая могла отслеживаться школой негласно при использовании школьной сети Wi-Fi. Естественно, что анализ этих сведений мог бы помочь в изучении интересов учеников, их пристрастий, поведения вне школы и так далее, что позволило бы выделить причины, которые могут влиять на вероятность отчисления учеников, относящихся к определенной группе, например, по материальному статусу или с похожими интересами, или одинаковой моделью поведения. Для примера, оценка сведений о посещаемых интернет ресурсах прямо или косвенно может указать на то, что отчисления учеников связаны с недостатком часов по тем или иным образовательным предметам. Однако, подобная информация, безусловно, является конфиденциальной и не может использоваться школой без согласия учеников.

Переходя к вопросу о заинтересованности заказчика и предвзятости предсказаний модели скажем, что, по нашему мнению, заинтересованность как совета школы, так и директор существенно не влияют на результаты, поскольку у них имеется общая цель.

Другое дело, что в результате анализа данных может оказаться, что причина роста числа отчислений учеников кроется не в деятельности школы, а в самих учениках. Например, можно предположить, что в школе большое число детей из семей иммигрантов с недостаточным знанием английского языка, что в определенный момент вынуждает их искать другую школу с более низкими требованиями. С точки зрения алгоритма он выдает статистически обоснованный результат, который, однако, может расцениваться некоторыми людьми как предвзятый по расовому, национальному или иному признаку.

Последние три вопроса, на наш взгляд, больше связаны с юридической стороной проблемы. Как уже отмечалось выше, сбор, хранение и анализ сведений, связанных с деятельностью школы, включая сведения об учениках является не только необходимостью, но и обязанностью школы.

Из описания не до конца понятно, какие данные собирались школой. В частности, остается вопросом какие сведения подразумеваются под фразой «и т. п.». Если основываться на российском законодательстве, то полный перечень сведений указывается в согласии на обработку персональных данных. Имеется ли подобный аналог в американском законодательстве нам не известно. При этом не исключено, что в американском законодательстве право на обработку таких данных включено в договор между школой и родителями по умолчанию.

Если все же предположить, что школа нарушила конфиденциальность данных учеников, не включать опцию отказа о предоставлении данных или передав их данные компании Hephaestats, то, здесь важно определить, в каком виде были переданы эти данные компании и как они обрабатывались в Hephaestats. При этом, по нашему мнению, решение не включать опцию отказа о предоставлении данных не влияет на легитимность школьной программы. В большей степени это влияет на репутацию школы.

На наш взгляд, поскольку ситуация с отчислениями учеников требовала срочного решения и при этом процесс сбора данных мог привести к длительному нахождению консенсусу между школой и родителями или недостаточному объему данных, количественному дисбалансу между разными группами учеников в силу большого числа отказов и так далее, допустимо, что данные были собраны и проанализированы без оповещения об этом  учеников и их родителей.

Другое дело, что, при этом, передаваемая информация и/или результаты ее анализа должна быть обезличены, а также гарантировать, что информация не будет передана кому либо, кроме школы и компании Hephaestats. В идеальном случае Hephaestats после завершения работы с ней, должна была бы «уничтожить» первичные данные для избегания возможной утечки данных.

Исходя из этого, скажем, что результат обращения в суд родителей по поводу нарушения конфиденциальность данных учеников зависит от того, считается ли информация, о которой говорилось выше, персональной или нет. Безусловно, если среди собранных и переданных данных были персональные, то закон будет на стороне родителей и учеников.