Вопросы к экзамену по модулю «Проблема конфиденциальности данных при использовании ML»

|  |  |
| --- | --- |
| № | Вопрос |
| 1 | Выберите наиболее правильное определение федеративного обучения:   1. Метод машинного обучения на основе обучения среды, которая взаимодействует с агентом 2. **Метод машинного обучения, который функционирует на нескольких децентрализованных периферийных устройствах, хранящих локальные данные, без обмена ими** 3. Метод машинного обучения на основе обучения интеллектуального агента, который действует во внешней среде 4. Метод машинного обучения с частичным привлечением учителя |
| 2 | Что означает термин «дифференциальная приватность»?   1. **Математическое определение понятия «наличия приватности»** 2. Подход, при котором конфиденциальные данные разделяются на части и хранятся в разных хранилищах 3. **Совокупность методов, направленных на минимизацию возможности идентификации отдельных записей в базе данных** 4. Свойство процесса, алгоритма, модели позволяющее надежно шифровать данные |
| 3 | С какой целью используют модели учителей в алгоритме PATE?   1. Для увеличения скорости обучения и сокращения объема передаваемых по сети данных 2. Для повышения общей точности предсказания алгоритма 3. **Для последующей разметки общедоступных данных для модели студента** 4. Для усреднения этих моделей в модель студента |
| 4 | Назовите проблемы, характерные для федеративного обучения?   1. **Скорость обучения** 2. Передача тренировочных данных на сервер 3. **«Отравление» модели** 4. Высокие требования к пропускной способности канала связи |
| 5 | На какие две группы разделяется информация согласно № 149-ФЗ?   1. Конфиденциальная информация 2. **Ограниченного доступа** 3. Государственная тайна 4. **Общедоступная** |
| 6 | Выберете то, что относится к персональным данным:   1. **Фамилия, имя, отчество** 2. **Размер одежды** 3. **Образование** 4. **Модель телефона** |
| 7 | К какой категории персональных данных относится информация о состоянии здоровья человека?   1. **Специальные категории персональных данных** 2. Биометрические персональные данные 3. Иные персональные данные 4. Общедоступные персональные данные |
| 8 | Что означает термин «распространение персональных данных»?   1. **Действия, направленные на раскрытие персональных данных неопределенному кругу лиц** 2. Действия, направленные на раскрытие персональных данных определенному кругу лиц 3. Действия, направленные на получение доступа к персональным данным 4. Действия, связанные с обработкой запроса на предоставление персональных данных |
| 9 | Какой шум чаще всего применяют в алгоритмах дифференциальной приватности?   1. Равномерный шум 2. Белый шум 3. **Шум Лапласа** 4. Гауссовский шум |
| 10 | Выберите верные отверждения:   1. За нарушение законодательства в области персональных данных предусмотрена только административная ответственность 2. **Клавиатура Gboard является одним из первых масштабных применений федеративного обучения** 3. **В основе протокола безопасной агрегации Google лежит схема разделения секрета Шамира** 4. Обезличенные персональные данные можно по-другому назвать анонимизированными |