Вопросы к экзамену по модулю «Методы классификации из ML в ИБ»

|  |  |
| --- | --- |
| № | Вопрос |
| 1 | Выберите классы методов машинного обучения:   1. **«Без учителя» и «с учителем»** 2. «С явным программированием» и «без явного программирования» 3. «Графические» и «графовые» 4. «Обучения» и «предсказания» |
| 2 | Выберите наиболее правильное определение «классификация»:   1. **Предсказание для объекта одного из конечного списка классов** 2. Предсказание для объекта одного из бесконечного списка классов 3. Предсказание для класса его объектов 4. Предсказание для объекта вероятности попадания в интервал |
| 3 | Выберите наиболее правильное определение «регрессия»:   1. Предсказание для объекта одного из конечного списка классов 2. **Предсказание для объекта одного из бесконечного списка классов** 3. Предсказание для класса его объектов 4. Предсказание для объекта вероятности попадания в интервал |
| 4 | Выберите качества, присущие деревьям решений:   1. Нечувствительность к шумам 2. **Возможность интерполировать знания** 3. Возможность экстраполировать знания 4. **Интерпретируемость** |
| 5 | Выберите качества, присущие методу K-ближайших соседей   1. **Устойчивость к выбросам** 2. Теоретическая обоснованность выбора всех параметров 3. **Зависимость от способа измерения расстояния между объектами** 4. **Интерпретируемость** |
| 6 | Выберите качества, присущие методу опорных векторов:   1. **Хорошие результаты работы в условиях большого количества признаков** 2. Нечувствительность к шуму 3. **Возможность замены ядра** 4. Наличие общепринятого метода подбора ядра |
| 7 | Выберите качества, присущие наивному байесовскому классификатору:   1. Малое количество гиперпараметров 2. **Отсутствие гиперпараметров** 3. Умение обнаруживать категории, отсутствовавшие при обучении 4. Работа в предположении о том, что признаки зависимы |
| 8 | Выберите качества, присущие линейной регрессии:   1. Возможность работы с конечным количеством классов 2. Эффективная работа при любой функции зависимости целевой переменной от входной 3. **Метод регрессии, а не классификации** 4. Метод классификации, а не регрессии |
| 9 | Выберите качества, присущие логистической регрессии:   1. **Возможность работы с конечным количеством классов** 2. Эффективная работа при любой функции зависимости целевой переменной от входной 3. Метод регрессии, а не классификации 4. **Метод классификации, а не регрессии** |
| 10 | Выберите характерные для метода ансамблирования «Бэггинг» особенности:   1. Каждая слабая модель принимает решение о классе объекта на основе ограниченного набора признаков, затем добавляет недостающие, принимает решение ещё раз и передаёт неизменённый исходный набор данных следующей модели. Решение последней модели является общим решением. 2. **Слабые модели обучаются на ограниченном наборе признаков, общее решение принимается по усреднению набора решений.** 3. Каждая слабая модель добавляет к исходным данным своё решение и передаёт следующей. Общее решение – это решение последней модели. 4. Каждая слабая модель принимает решение о классе объекта, а ещё одна модель, принимая в качестве входных данных ответы всех предыдущих моделей, принимает общее решение. |
| 11 | Выберите характерные для метода ансамблирования «Бустинг» особенности:   1. Каждая слабая модель принимает решение о классе объекта на основе ограниченного набора признаков, затем добавляет недостающие, принимает решение ещё раз и передаёт неизменённый исходный набор данных следующей модели. Решение последней модели является общим решением. 2. Слабые модели обучаются на ограниченном наборе признаков, общее решение принимается по усреднению набора решений. 3. **Каждая слабая модель добавляет к исходным данным своё решение и передаёт следующей. Общее решение – это решение последней модели.** 4. Каждая слабая модель принимает решение о классе объекта, а ещё одна модель, принимая в качестве входных данных ответы всех предыдущих моделей, принимает общее решение |
| 12 | Выберите характерные для метода ансамблирования «Стэкинг» особенности:   1. Каждая слабая модель принимает решение о классе объекта на основе ограниченного набора признаков, затем добавляет недостающие, принимает решение ещё раз и передаёт неизменённый исходный набор данных следующей модели. Решение последней модели является общим решением. 2. Слабые модели обучаются на ограниченном наборе признаков, общее решение принимается по усреднению набора решений. 3. Каждая слабая модель добавляет к исходным данным своё решение и передаёт следующей. Общее решение – это решение последней модели. 4. **Каждая слабая модель принимает решение о классе объекта, а ещё одна модель, принимая в качестве входных данных ответы всех предыдущих моделей, принимает общее решение**. |
| 13 | Выберите типовые случаи применения машинного обучения в информационной безопасности:   1. **Сложно выявить детерминированные правила для работы системы** 2. Требуется абсолютная точность работы системы 3. **Повышенные требования к скорости создания правил для работы системы** 4. Любой другой случай – методы машинного обучения имеет смысл использовать в любой задаче |
| 14 | Что нужно сделать в задаче выявления спама при создании модели после получения размеченных данных от пользователей?   1. Передать полученные данные в качестве входных данных для обучения модели 2. **Очистить полученные данные** 3. Передать полученные данные в качестве входных данных для предсказания модели 4. Определить верный ответ для каждого сэмпла в данных |
| 15 | Как имеет смысл обработать полученные от пользователей данные в задаче выявления спама?   1. Оставить как есть – модель сама разберётся 2. Выделить текстовые признаки 3. Выделить экспертные признаки 4. **Выделить текстовые и экспертные признаки** |
| 16 | Выберите наиболее правильное завершение определения. «Эвристический анализ – это метод поиска вирусов, заключающийся в …»:   1. генерации вирусов и обучении модели поиску сгенерированных сэмплов 2. выявлении последовательностей команд, встречавшихся ранее в обнаруженных вирусах 3. **выявлении последовательностей команд, предположительно присущих вредоносной программе** 4. выявлении сигнатур, соответствующих исходным кодам вирусов |
| 17 | Выберите, на что опирались в работе первые DLP-системы:   1. **Характеристики файла** 2. Контекст 3. Представление текстового содержимого файла, приведённое в двумерное пространство 4. Особенности операционной системы, в которой установлена DLP-система |
| 18 | Выберите наиболее правильное утверждение:   1. Решение модели машинного обучения важнее, чем решение сотрудника отдела ИБ 2. В случае расхождения решений модели машинного обучения и сотрудника отдела ИБ не обходимо обратиться к руководителю отдела ИБ 3. **Решение модели машинного обучения менее важно, чем решение сотрудника отдела ИБ** 4. Решение модели машинного обучения не может отличаться от решения сотрудника отдела ИБ |
| 19 | Выберите одно или несколько правильных утверждений:   1. Межсетевой экран, UTM, NGFW – это синонимы 2. UTM – это развитие NGFW 3. **NGFW – это развитие UTM** 4. **UTM и NGFW – это развитие межсетевых экранов** |
| 20 | Что реализация NGFW может предложить в случае появления неопознанного узла в защищаемой сети?   1. Классификация по первому пакету 2. **Классификация по слепку сетевого трафика** 3. Классификация с применением активного сканирования 4. Классификация на основе экспертного знания об адресном пространстве |