

Документ подписан простой электронной подписью  
 Информация о владельце:  
 ФИО: Косенок Сергей Михайлович  
 Должность: ректор  
 Дата подписания: 19.06.2024 06:18:08  
 Уникальный программный ключ:  
 e3a68f38aa1e62674b341499809903d6bfdcf836

**Тестовое задание для диагностического тестирования по дисциплине:**

**Машинное обучение, 3 семестр**

Код направления подготовки	09.04.02 ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ
Направленность (профиль)	УПРАВЛЕНИЕ ДАННЫМИ
Форма обучения	Очная
Кафедра-разработчик	Информатики и вычислительной техники
Выпускающая кафедра	Информатики и вычислительной техники

№ п-п	Проверяемая компетенция	Задание	Варианты ответов	Тип сложности вопроса
1.	ОПК-1 ОПК-2 ОПК-4	Метеоролог Вася хочет построить модель, прогнозирующую температуру воздуха на завтра. К какому типу относится данная задача?	Выберите один ответ: 1)Кластеризация 2)Регрессия 3)Ранжирование 3)Классификация	низкий
2.	ОПК-1 ОПК-2 ОПК-4	Доктор Петя хочет построить модель, прогнозирующую возможность развития осложнений у пациента после операции - все ли будет благополучно в течение нескольких следующих месяцев или нет. К какому типу относится данная задача?	Выберите один ответ: 1)Кластеризация 2)Регрессия 3)Ранжирование 4)Классификация	низкий
3.	ОПК-1 ОПК-2 ОПК-4	Астроном Витя хочет построить модель, которая сможет разбить известные науке звезды на группы по их характеристикам, чтобы лучше изучить их особенности. К какому типу относится данная задача?	Выберите один ответ: 1) Кластеризация 2) Регрессия 3) Ранжирование 4) Классификация	низкий
4.	ОПК-1 ОПК-2 ОПК-4	В задаче имеется 100 объектов. У каждого объекта 20 признаков. Укажите число настраиваемых параметров при обучении логистической регрессии, если <code>fit_intercept = True</code> .	Выберите один ответ: 1) 22 2) 20 3) 10 4) 21	низкий
5.	ОПК-1 ОПК-2 ОПК-4	Что происходит с ошибкой на тестовой части при увеличении гиперпараметра <code>n_estimators</code> (число деревьев)?	Выберите один ответ: 1) Ошибка все время увеличивается	низкий

			<p>2) Ошибка сначала увеличивается, потом начинает падать</p> <p>3) Ошибка сначала убывает, а затем выходит на один уровень</p>	
6.	ОПК-1 ОПК-2 ОПК-4	<p>Может ли случиться так, что все коэффициенты будут занулены в линейной модели, обучаемой с регуляризацией? Выберите все правильные ответы.</p>	<p>Выберите один или несколько ответов:</p> <p>1) Нет, не может, так как ненулевые коэффициенты всегда дают меньшую ошибку, чем нулевые</p> <p>2) Да, может, если целевая переменная не зависит от признаков и сосредоточена вокруг нуля</p> <p>3) Да, может, если выставить слишком большой вес у коэффициента регуляризации</p>	средний
7.	ОПК-1 ОПК-2 ОПК-4	<p>Может ли изменить среднеквадратичную ошибку масштабирование признаков в линейной регрессии, если мы пользуемся аналитическим решением?</p>	<p>Выберите один или несколько ответов:</p> <p>1) Да, может. Масштабированные признаки имеют лучшую предсказательную способность.</p> <p>2) Нет, не может. Масштабирование не добавляет в данные никак новых знаний, и поэтому не может увеличить качество предсказания.</p> <p>3) Да, может. Линейная регрессия лучше работает с признаками, значения которых близки к нулю.</p>	средний
8.	ОПК-1 ОПК-2 ОПК-4	<p>Какая из моделей приводит к отбору признаков?</p>	<p>Выберите один ответ:</p> <p>1) Lasso-регрессия</p> <p>2) Ridge-регрессия</p> <p>3) Линейная регрессия без регуляризации</p>	средний
9.	ОПК-1 ОПК-2 ОПК-4	<p>Выберите возможный способ получить оценку вероятности принадлежности объекта классу 1 при решении задачи бинарной классификации методом опорных векторов.</p>	<p>Выберите один ответ:</p> <p>1) Использовать калибровку Платта</p> <p>2) Применить сигмоидную функцию к отступу</p> <p>3) Использовать модуль отступа объекта</p> <p>4) Использовать модуль отступа объекта,</p>	средний

			поделенный на максимальное значение отступа всех объектов обучающей выборки	
10.	ОПК-1 ОПК-2 ОПК-4	<p>Предположим, что вы реализовали собственный стемминг для русского языка, который удаляет все гласные буквы на конце слова. Вы применили ваш стемминг к четырем документам.</p> <p>Он видел его семью своими глазами</p> <p>В семье нас было семь</p> <p>Семеро одного не ждут</p> <p>Семейный ужин был назначен на семь</p> <p>Укажите вектор-столбец в матрице объекты-признаки, который будет соответствовать слову "семь" после применения метода Bag Of Words.</p>	<p>Выберите один ответ:</p> <p>1) (1, 2, 1, 2)</p> <p>2) (0, 1, 1, 1)</p> <p>3) (1, 2, 0, 1)</p> <p>4) (1, 1, 0, 1)</p>	средний
11.	ОПК-1 ОПК-2 ОПК-4	<p>Почему бэггинг делают над глубокими решающими деревьями?</p>	<p>Выберите один ответ:</p> <p>1) Бэггинг уменьшает смещение алгоритмов, а глубокие решающие деревья имеют большое смещение</p> <p>2) Бэггинг уменьшает разброс и смещение алгоритмов, а глубокие решающие деревья имеют большой разброс и большое смещение</p> <p>3) Бэггинг уменьшает разброс алгоритмов, а глубокие решающие деревья имеют большой разброс</p>	средний
12.	ОПК-1 ОПК-2 ОПК-4	<p>Почему с помощью RandomForest удается достичь лучшего качества, чем при использовании бэггинга над решающими деревьями?</p>	<p>Выберите один ответ:</p> <p>1) В RandomForest деревья меньше скоррелированы, за счет того, что при построении каждого дерева используется случайное подмножество объектов.</p> <p>2) В RandomForest используют глубокие деревья.</p>	средний

			3) В RandomForest деревья меньше скоррелированы, за счет того, что оптимальное разбиение ищется среди случайных признаков.	
13.	ОПК-1 ОПК-2 ОПК-4	Выберите алгоритм у которого будет максимальный разброс (variance).	Выберите один ответ: 1) Решающее дерево с глубиной 2 2) Random Forest с глубиной деревьев 100 3) Решающее дерево с глубиной 100	средний
14.	ОПК-1 ОПК-2 ОПК-4	Почему в бэггинге не стоит использовать неглубокие деревья (глубины 1-2)?	Выберите один ответ: 1) Бэггинг уменьшает смещение алгоритма и не изменяет разброс. У неглубоких деревьев большой разброс. 2) Бэггинг уменьшает разброс алгоритма и не изменяет смещение. У неглубоких деревьев большое смещение и небольшой разброс. 3) Бэггинг уменьшает смещение и разброс алгоритма. У неглубоких деревьев большой разброс	средний
15.	ОПК-1 ОПК-2 ОПК-4	Сколько операций сравнения необходимо, чтобы вычислить ответ для композиции из 5000 Oblivious Decision Trees глубины 7?	30000	средний
16.	ОПК-1 ОПК-2 ОПК-4	Запишите сумму целевых переменных, на которые будет обучаться n-е дерево композиции в градиентном бустинге, если в выборке 5 объектов, сдвиги на них равны (1, 3, -1, 0, -9), а функционал ошибки MSE.	-6	высокий
17.	ОПК-1 ОПК-2 ОПК-4	В листе дерева оказываются 10 объектов, 8 из которых из одного класса, а 2 - из второго. Посчитайте (двоичную) энтропию получившейся выборки в листе. Ответ округлите до двух знаков после точки.	0.72	высокий
18.	ОПК-1 ОПК-2 ОПК-4	Перечислите недостатки решающих деревьев (выберите все подходящие варианты).	Выберите один или несколько ответов: 1) Можно применять только к задаче классификации 2) Высокая вероятность переобучения	высокий

			3) Не все объекты могут быть классифицированы 4) Высокая чувствительность к составу выборки (5) Сложно интерпретировать	
19.	ОПК-1 ОПК-2 ОПК-4	Решая задачу бинарной классификации, использовались данные с двумя признаками. На данных была обучена логистическая регрессия. Известно, что после обучения вектор весов $(w_0, w_1, w_2) = (0.2, 0.7, 0.1)$ . Для объекта $x$ с признаковым описанием $(3, 5)$ Посчитайте модуль отступа. Ответ округлите до десятых. Ответ вводите с точкой.	2.8	высокий
20.	ОПК-1 ОПК-2 ОПК-4	Решая задачу бинарной классификации, использовались данные с двумя признаками. На данных была обучена логистическая регрессия. Известно, что после обучения вектор весов $(w_0, w_1, w_2) = (0.2, 0.7, 0.1)$ . Для объекта $x$ с признаковым описанием $(3, 5)$ вычислите предсказание вероятности принадлежности классу 1. Ответ округлите до сотых. Ответ вводите с точкой.	0.94	высокий