

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ПО ДИСЦИПЛИНЕ

МАШИННОЕ ОБУЧЕНИЕ В АГРОИНЖЕНЕРИИ

Направление подготовки: 35.03.06 Агроинженерия

Профиль подготовки: Искусственный интеллект

Квалификация (степень)
выпускника: Бакалавр

Череповец, Вологда – Молочное
2023

1 Паспорт фонда оценочных средств по дисциплине (модулю)

1.1 Текущий контроль

| № п/п | Раздел (тема) дисциплины | Результаты обучения (компетенции) | Наименование оценочного средства / Форма текущего контроля | Метод контроля |
|-------|--|-----------------------------------|--|----------------|
| 1. | Введение в машинное обучение | ПК-14, ПК-15 | Лабораторное занятие | Устный опрос |
| 2. | Исследование данных, их визуализация и интерпретация | ПК-14, ПК-15 | Лабораторное занятие | Устный опрос |
| 3. | Методы классификации | ПК-14, ПК-15 | Лабораторное занятие | Устный опрос |
| 4. | Методы числового прогнозирования | ПК-14, ПК-15 | Лабораторное занятие | Устный опрос |
| 5. | Обнаружение закономерностей на основе ассоциативных правил | ПК-14, ПК-15 | Лабораторное занятие | Устный опрос |
| 6. | Методы кластеризации | ПК-14, ПК-15 | Лабораторное занятие | Устный опрос |
| 7. | Методы понижения размерности данных | ПК-14, ПК-15 | Лабораторное занятие | Устный опрос |

1.2 Промежуточная аттестация

В соответствии с учебным планом промежуточная аттестация по учебной дисциплине «Машинное обучение в агроинженерии» предусматривает проведение экзамена. Для оценки результатов обучения используется метод тестирования.

**2 Комплект оценочных материалов для проведения текущего контроля
оценки знаний, умений и уровня сформированности компетенций**

Министерство сельского хозяйства Российской Федерации
федеральное государственное бюджетное образовательное учреждение
высшего образования «Вологодская государственная
молочнохозяйственная академия имени Н.В. Верещагина»

Инженерный факультет

Кафедра «Энергетические средств и технический сервис»

Вопросы для контроля освоения компетенции ПК-14

«Способен использовать системы искусственного интеллекта в решении задач анализа, прогнозирования, планирования, синтеза и принятия решений»

1. Как Вы считаете, чем машинное обучение отличается от интеллектуального анализа данных (если эти понятия отличаются друг от друга)?
2. Приведите описание этапов решения задач с использованием машинного обучения: сбор данных: исследование и подготовка данных: обучение модели: оценка модели: улучшение модели.
3. Приведите примеры использования методов машинного обучения.
4. Приведите описание критериев эффективности и качества функционирования системы искусственного интеллекта: точность, релевантность, достоверность, целостность, быстрота решения задач, надежность, защищенность функционирования систем искусственного интеллекта.
5. Приведите описание методов постановки задач, проведения и анализа тестовых и экспериментальных испытаний работоспособности систем искусственного интеллекта, в том числе систем машинного обучения.
6. Перечислите методы и критерии оценки качества моделей машинного обучения.
7. Приведите содержание унифицированных и обновляемых методологии описания, сбора и разметки данных, а также механизмов контроля за соблюдением указанных методологий.
8. Приведите примеры задач, решаемых с использованием метода k-NN.
9. Приведите пример использования метода k-NN для классификации данных.
10. Приведите примеры задач, решаемых с использованием наивного байесовского классификатора.
11. Приведите пример использования наивного байесовского алгоритма для классификации данных.
12. Приведите примеры задач, решаемых с использованием деревьев.
13. Приведите пример использования дерева решений для классификации данных.

14. Приведите примеры задач, решаемых с использованием регрессии.
15. Приведите описание множественной линейной регрессии. В чем заключаются преимущества и недостатки данного метода?
16. Приведите пример использования линейной регрессии для числового прогнозирования.
17. В каких библиотеках Python реализован метод ассоциативных правил?
18. Какие задачи решаются с использованием методов кластеризации?
19. Приведите пример решения задачи с использованием метода k-средних.
20. Какие задачи решаются с использованием алгоритма понижения размерности данных?
21. В каких библиотеках Python реализован данный алгоритм?

Министерство сельского хозяйства Российской Федерации
федеральное государственное бюджетное образовательное учреждение
высшего образования «Вологодская государственная
молочнохозяйственная академия имени Н.В. Верещагина»
Инженерный факультет
Кафедра «Энергетические средств и технический сервис»

МАШИННОЕ ОБУЧЕНИЕ В АГРОИНЖЕНЕРИИ

Вопросы для контроля освоения компетенции ПК-15

«Способен разрабатывать и применять методы машинного обучения для решения задач»

1. Дайте описание типов входных данных, используемых при решении задач с помощью методов машинного обучения.
2. Перечислите типы алгоритмов машинного обучения.
3. Как подбирается метод машинного обучения для решения конкретной прикладной задачи? Что влияет на выбор метода?
4. Каково назначение и возможности библиотек библиотеки Python для машинного обучения (дайте заключение на основе анализа документации разработчиков библиотек).
5. Для каких целей выполняется интерпретация данных?
6. Что такое структура данных?
7. Какие базовые наборы изменений обычно применяются в числовом данным?
8. Почему в ходе исследования данных запрашивают как средние, так и медианные значения числовых переменных?
9. Что такое «пятичисловая сводка»? Для каких целей она используется?
10. Что отображает диаграмма размаха?
11. Что отображает гистограмма?
12. Как выглядит гистограмма равномерного распределения?
13. Как выглядит кривая нормального распределения?
14. Что измеряется стандартным отклонением?
15. Что гласит правило «68-95-99.7»?
16. Что отображает таблица частотности?
17. Для каких целей строится диаграмма разброса?
18. Что показывают перекрестные таблицы (кросс-таблицы, таблицы сопряженности)?
19. В чем заключается суть метода k-NN?
20. Каковы преимущества метода k-NN?
21. Каковы недостатки метода k-NN?
22. Как измеряется степень сходства между экземплярами набора данных?
23. Каким образом выбирается подходящее k?
24. Что такое «минимаксная» нормализация?

25. Что такое «фиктивное» кодирование?
26. Почему алгоритм k-NN называют ленивым?
27. Изучите документацию разработчиков библиотеки Scikit-learn (в части реализации метода k-NN).
28. Что такое «вероятностное обучение»?
29. В чем заключается суть работы наивного байесовского классификатора?
30. Каковы преимущества наивного байесовского классификатора?
31. Каковы недостатки наивного байесовского классификатора?
32. Почему алгоритм называют наивным?
33. Для каких целей используются методы деревьев?
34. Почему группа методов получила такое название?
35. Что такое «рекурсивное сегментирование»?
36. Каким образом работает алгоритм дерева решений C5.0?
37. Каким образом выбирается лучшее разделение?
38. С какой целью выполняется «сокращение» дерева решений?
39. В чем заключается суть работы со случайными лесами?
40. Как работает градиентный бустинг?
41. Для каких целей используются методы регрессии?
42. Приведите понятие регрессии.
43. Как определяется простая линейная регрессия?
44. Приведите описание оценки методом наименьших квадратов.
45. Как рассчитывается коэффициент корреляции Пирсона?
46. В чем заключается суть метода ассоциативных правил?
47. Какие задачи решаются с использованием данного метода?
48. К какому типу методов машинного обучения относится метод ассоциативных правил?
49. В чем заключается суть метода Apriori?
50. Каким образом производится обучение модели?
51. Какие параметры необходимо указать для запуска обучения?
52. Как проверить эффективность модели?
53. Что необходимо сделать, чтобы повысить эффективность модели?
54. Как сохранить ассоциативные правила в файл или фрейм данных?
55. Что такое «кластеризация»?
56. Чем кластеризация отличается от классификации?
57. Перечислите известные Вам методы кластеризации.
58. В чем заключаются суть метода k-средних?
59. Перечислите достоинства и недостатки метода k-средних
60. В каких библиотеках Python реализован метод k-средних?
61. В чем заключается принцип работы алгоритма понижения размерности данных?

3 Комплект оценочных материалов для проведения промежуточной аттестации по итогам изучения учебной дисциплины (модуля)

Министерство сельского хозяйства Российской Федерации
федеральное государственное бюджетное образовательное учреждение
высшего образования «Вологодская государственная
молокохозяйственная академия имени Н.В. Верещагина»
Инженерный факультет
Кафедра «Энергетические средств и технический сервис»

МАШИННОЕ ОБУЧЕНИЕ В АГРОИНЖЕНЕРИИ

Тесты для контроля освоения компетенции ПК-14

«Способен использовать системы искусственного интеллекта в решении задач анализа, прогнозирования, планирования, синтеза и принятия решений»

Что находится в нелистовых вершинах решающего дерева?

объекты
разделяющие вопросы
признаки

В каком случае построенный алгоритм решающего дерева легче проанализировать человеку?

в случае маленького числа вершин дерева
в случае большого числа вершин дерева
в случае большого числа ребер дерева
в случае маленького числа ребер дерева

Автоматическая оценка стоимости квартиры является актуальной задачей в банковской сфере. Определите к какому типу она относится

задача классификации
задача регрессии
задача кластеризации
задача прогнозирования

Что называется переобучением алгоритма машинного обучения?

характеристика метода машинного обучения, означающая, что алгоритм максимально точно описывает данные
проблема алгоритмов машинного обучения, с ней необходимо бороться
характеристика метода машинного обучения, означающая, что алгоритм обучается слишком долго

Предположим, что имелась некоторая автоматизированная система, которая работала с использованием большого количества эвристик. Что можно сделать с эвристиками, если было принято решение заменить автоматизированную систему на систему машинного обучения?

использовать эвристические правила для извлечения признаков, которые затем будут подаваться в алгоритм машинного обучения

удалить весь код эвристических правил в целях освобождения дополнительного места для хранения данных
необходимо отменить принятое решение, потому что эвристических правил

достаточно для решения любых задач

Предположим, что необходимо узнать, какие изменения в мобильном приложении банка хочет увидеть среднестатистический россиянин. Для этого решено провести социальный опрос среди большого числа студентов нескольких вологодских вузов: ВГМА, ВГУ и ЧГУ. Будут ли при этом достоверные результаты или стоит скорректировать аудиторию, на которой проводится опрос?

не нужно ничего корректировать, результаты будут достоверные для достоверности результатов необходимо подключить к опросу студентов вузов из других областей России

для достоверности результатов необходимо опрашивать все слои населения России, использующие мобильное приложение банка

Чем отличается стохастический градиентный спуск от обычного градиентного спуска?

используется случайный выбор шага градиентного спуска

просматривается выборка в случайном порядке, чаще обновляется вектор параметров

выбирается каждая n-я точка (настраиваемый параметр) случайным образом, чтобы не застревать в локальных минимумах

Предположим, необходимо решать задачу определения числа людей на фотографии. Задачи подобного рода возникают, например, с целью определения количества посетителей помещения в разное время. Какой класс методов является наиболее эффективным для этой задачи?

свёрточные нейронные сети

решающие деревья

рекуррентные нейронные сети

Выберите определение термина Deep Learning

раздел машинного обучения, который изучает многослойные нейронные сети

раздел машинного обучения, к которому относятся алгоритмы, основанные на

глубоком понимании предметной области

раздел машинного обучения, который изучает методы классификации живых существ по фотографиям

Министерство сельского хозяйства Российской Федерации
федеральное государственное бюджетное образовательное учреждение
высшего образования «Вологодская государственная
молочнохозяйственная академия имени Н.В. Верещагина»
Инженерный факультет
Кафедра «Энергетические средств и технический сервис»

МАШИННОЕ ОБУЧЕНИЕ В АГРОИНЖЕНЕРИИ

Тесты для контроля освоения компетенции ПК-15

«Способен разрабатывать и применять методы машинного обучения для
решения задач»

**Класс задач, относящихся к обучению без учителя — это задачи, для
которых нам известны...**

только данные, а ответы неизвестны или вообще не существуют
данные и ответы
только ответы, а данные неизвестны.
обучение происходит под контролем человек

**Класс задач, относящихся к обучению с учителем – это задачи, для
которых нам известны...**

данные и ответы
только данные, а ответы неизвестны или вообще не существуют
обучение происходит под контролем человека
только признаки

Алгоритм, избыточно подстроившийся под данные, называют...

переобученным
недообученным
самообученным
сверхообученным

Выберите верное утверждение

обычно выбросы лучше удалять, но в некоторых случаях выбросами могут
быть важные объекты, которые следует обрабатывать отдельно
выбросы всегда следует удалять.
выбросы всегда следует обрабатывать отдельно.
выбросы являются основой алгоритмов машинного обучения без учителя

**Модели, в которых произошла ### это модели, которые на самом деле кое-
что поняли о том, как устроены данные, и могут выдавать полезные
предсказания для объектов, которых не видели**

генерализация
реструктуризация
регрессия
оптимизация

Задачи классификации и регрессии можно сформулировать как поиск отображения из множества объектов X в множество возможных чисел (целевых переменных). Эти числа называют ###

- таргетами
- векторами признаков
- фичами
- маркерами

Если вы построили линейную модель, вы неплохо можете объяснить заказчику те или иные её результаты. Это качество моделей называют:

- интерпретируемостью
- надежностью
- эффективностью
- точностью

Достаточно легко судить, как влияют на результат те или иные признаки у линейной модели. Скажем, если вес ω_i положителен, то с ростом i -го признака таргет в случае регрессии будет ###, а в случае классификации наш выбор будет сдвигаться в пользу одного из классов

- увеличиваться
- уменьшаться
- оставаться неизменным
- зависеть от других факторов

Регрессия, сочетающая регуляриность L1 и регуляриность L2 называется:

- Elastic Net
- логистическая
- бинарная
- лассо
- Ридж

Что из перечисленного не является метрикой качества модели регрессии?

- F1-Мера
- MSE
- MAE
- R2

Какой метод машинного обучения использует метрические методы для классификации объектов?

- Метод ближайших соседей
- Деревья решений
- Метод k -средних
- Метод опорных векторов

Что такое метрика в метрических методах?

- функция расстояния между объектами
- алгоритм, который находит ближайших соседей
- метод, который используется для классификации объектов
- функция, которая находит оптимальный порог для классификации

Что такое метод k -ближайших соседей?

метод, который использует k-ближайших соседей для классификации объектов

метод, который использует k-ближайших соседей для регрессии

метод, который находит k-соседей с наименьшим расстоянием до объекта

метод, который находит k-соседей с наибольшим расстоянием до объекта

Что такое взвешенный метод k-ближайших соседей?

метод, который учитывает веса каждого соседа при классификации объекта

метод, который находит k-ближайших соседей, которые имеют наибольший вес

метод, который находит k-ближайших соседей, которые имеют наименьший вес

метод, который находит соседей только из определенного класса

Какая из метрик существует в методе k-ближайших соседей?

Манхэттенская метрика

Бруклинская метрика

Куинсанская метрика

Бронксонская метрика

Какую метрику расстояния следует использовать для задачи классификации текстовых документов?

косинусное расстояние

Манхэттенская метрика

Метрика Минковского

Расстояние Жаккара

Какой тип метрического метода используется для задачи регрессии?

k-Nearest Neighbors

k-Means

линейная регрессия

решающие деревья

Какие из следующих форм алгоритм k-Means может не агрегировать?

спиральное распределение

круговое распределение

ленточное распространение

выпуклое распределение полигонов

В чём заключается задача кластеризации?

разбить обучающую выборку на непересекающиеся подмножества, так, чтобы каждое множество состояло из объектов, близких по метрике $math$, а объекты разных множеств существенно отличались

распределить все обучающее множество на несколько классов

упростить обработку данных, т.е. воссоединить множества в один класс

Что, из нижеперечисленного, служит целями кластеризации?

упрощение дальнейшей обработки данных

сокращение объема хранимых данных;

построение иерархии множества объектов;

распределение обучающего множества на несколько классов;
упрощение входных параметров

Дробление крупных кластеров на более мелкие называется задачей:

наименьших квадратов
сопряженного градиента
аппроксимации
таксономии

Фонд оценочных средств составлен в соответствии с требованиями ФГОС ВО и с учетом рекомендаций ОПОП ВО по направлению подготовки 35.03.06 Агроинженерия.

Разработчик: канд. техн. наук, доцент Шабалов В.А

Фонд оценочных средств одобрен на заседании кафедры математического и программного обеспечения ЭВМ 25 сентября 2023 года, протокол № 2.

Зав. кафедрой: Зав. кафедрой: доктор техн. наук, профессор Ершов Е.В