Оценочные материалы по дисциплине (модулю)

по направлению: Радиотехника

профиль подготовки: Аппаратно-программная инженерия радиолокационных и навигационных систем

Сетевое обучение

Кафедра «Технологий проектирования сложных технических систем»

курс: 4

квалификация: бакалавр

Семестр, форма промежуточной аттестации: 7 (осенний) – зачет

Разработчик: К.В. Воронцов, канд. физ.-мат. наук, профессор

1. Компетенции, формируемые в процессе изучения дисциплины (модуля)

Код и наименование компетенции	Индикаторы достижения компетенций
ОПК-4 Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности	ОПК-4.1 Использует при решении задач профессиональной деятельности современные информационные технологии и программные средства, в том числе отечественного производства
ОПК-5 Способен разрабатывать алгоритмы и компьютерные программы, пригодные для практического применения	ОПК-5.1 Понимает требования к алгоритмам, суть процесса алгоритмизации задач
	ОПК-5.2 Выполняет разработку алгоритмического и программного обеспечения для решения прикладных задач

2. Показатели оценивания компетенция

В результате освоения дисциплины Учебная практика (Программирование на языке Python), обучающиеся должны:

- знать:
 - основные принципы и проблематику теории обучения машин;
 - основные методы и алгоритмы решения задач обучения по прецедентам;
 - основные области применения этих методов и алгоритмов;
 - классификации, кластеризации и регрессии;
- уметь:
 - формализовать постановки прикладных задач анализа данных;
 - использовать методы обучения по прецедентам для решения практических задач;
 - оценивать точность и эффективность полученных решений;
- владеть:

- основными понятиями теории машинного обучения;
- навыками самостоятельной работы при решении типовых задач;
- культурой постановки и моделирования практически значимых задач;
- навыками теоретического анализа реальных задач, решаемых с помощью алгоритмов обучения по прецедентам.

3. Перечень типовых (примерных) вопросов, заданий, тем для подготовки к текущему контролю

1. Записать общую формулу байесовского классификатора (надо помнить формулу).
2. Какие вы знаете три подхода к восстановлению плотности распределения по выборке?
3. Что такое наивный байесовский классификатор?
4. Что такое оценка плотности Парзена-Розенблатта (надо помнить формулу). Выписать формулу алгоритма классификации в методе парзеновского окна.
5. На что влияет ширина окна, а на что вид ядра в методе парзеновского окна?
6. Многомерное нормальное распределение (надо помнить формулу). Вывести формулу квадратичного дискриминанта. При каком условии он становится линейным?
7. На каких предположениях основан линейный дискриминант Фишера?
8. Что такое «проблема мультиколлинеарности», в каких задачах и при использовании каких алгоритмов она возникает? Какие есть подходы к её решению?
9. Что такое «смесь распределений» (надо помнить формулу)?
10. Что такое EM-алгоритм, какова его основная идея? Какая задача решается на E-шаге, на M-шаге? Каков вероятностный смысл скрытых переменных?
11. Последовательное добавление компонент в EM-алгоритме, основная идея алгоритма.
12. Что такое стохастический EM-алгоритм, какова основная идея? В чём его преимущество (какой недостаток стандартного EM-алгоритма он устраняет)?
13. Что такое сеть радиальных базисных функций?
14. Что такое «выбросы»? Как осуществляется фильтрация выбросов?
15. Что такое обобщённый алгоритм классификации (надо помнить формулу)? Какие вы знаете частные случаи?
16. Как определяется понятие отступа в метрических алгоритмах классификации?
17. Что такое окно переменной ширины, в каких случаях его стоит использовать?
18. Что такое метод потенциальных функций? Идея алгоритма настройки. Сравните с методом радиальных базисных функций.
19. Зачем нужен отбор опорных объектов в метрических алгоритмах классификации?
20. Основная идея алгоритма СТОЛП.
21. Что такое функция конкурентного сходства? Основная идея алгоритма FRiS-СТОЛП.
22. Приведите пример метрического алгоритма классификации, который одновременно является байесовским классификатором.
23. Приведите пример метрического алгоритма классификации, который одновременно является линейным классификатором.
24. Что такое модель МакКаллока-Питтса (надо помнить формулу)?
25. Метод стохастического градиента. Расписать градиентный шаг для квадратичной функции потерь и сигмоидной функции активации.
26. Недостатки метода SG и как с ними бороться?
27. Что такое линейный адаптивный элемент ADALINE?
28. Что такое правило Хэбба?
29. Что такое «сокращение весов»?
30. Обоснование логистической регрессии (основная теорема), основные посылки (3) и следствия (2). Как выражается апостериорная вероятность классов (надо помнить формулу).

31. Как выражается функция потерь в логистической регрессии (надо помнить формулу).
32. Две мотивации и постановка задачи метода опорных векторов. Уметь вывести постановку задачи SVM (рекомендуется помнить формулу постановки задачи).
33. Какая функция потерь используется в SVM? В логистической регрессии? Какие ещё функции потерь Вы знаете?
34. Что такое ядро в SVM? Зачем вводятся ядра? Любая ли функция может быть ядром?
35. Какое ядро порождает полимиальные разделяющие поверхности?
36. Что такое ROC-кривая, как она определяется? Как она эффективно вычисляется?
37. В каких алгоритмах классификации можно узнать не только классовую принадлежность классифицируемого объекта, но и вероятность того, что данный объект принадлежит каждому из классов?
38. Каков вероятностный смысл регуляризации? Какие типы регуляризаторов Вы знаете?
39. Что такое принцип максимума совместного правдоподобия данных и модели (надо помнить формулу)?
40. Что такое ядерное сглаживание?
41. Что есть общего между ядром в непараметрической регрессии и ядром SVM?
42. На что влияет ширина окна, а на что вид ядра в непараметрической регрессии?
43. Что такое окна переменной ширины, и зачем они нужны?
44. Что такое «выбросы»? Как осуществляется фильтрация выбросов в непараметрической регрессии?
45. Постановка задачи многомерной линейной регрессии. Матричная запись.
46. Что такое сингулярное разложение? Как оно используется для решения задачи наименьших квадратов?
47. Что такое «проблема мультиколлинеарности» в задачах многомерной линейной регрессии? Какие есть три подхода к её устранению?
48. Сравнить гребневую регрессию и лассо. В каких задачах предпочтительнее использовать лассо?
49. Какую проблему решает метод главных компонент в многомерной линейной регрессии? Записать матричную постановку задачи для метода главных компонент.
50. Как свести задачу многомерной нелинейной регрессии к последовательности линейных задач?

Примеры задач

1. Задана цена отказа от классификации. Выписать модифицированную формулу байесовского классификатора.
2. Вывести формулу линейного дискриминанта для случая независимых признаков.
3. Вывести формулу наивного байесовского классификатора для случая бинарных признаков (доказать, что он линеен).
4. Вывести формулу градиентного шага в методе логистической регрессии для задачи классификации с двумя классами. Сравнить с правилом Хэбба.
5. Вывести формулу непараметрической регрессии Надарая-Ватсона.
6. Вывести формулу регуляризованного решения задачи многомерной линейной регрессии через сингулярное разложение.
7. Вывести градиентный метод обучения в логистической регрессии.

4. Перечень типовых (примерных) вопросов и тем для проведения промежуточной аттестации обучающихся

1. В чем отличия внутренних и внешних критериев?
2. Разновидности внешних критериев.
3. Разновидности критерия скользящего контроля.

4. Что такое критерий непротиворечивости? В чём его недостатки?
5. Что такое многоступенчатый выбор модели по совокупности критериев?
6. Основная идея отбора признаков методом полного перебора. Действительно ли это полный перебор?
7. Основная идея отбора признаков методом добавлений и исключений.
8. Что такое шаговая регрессия? Можно ли её использовать для классификации, в каком методе?
9. Основная идея отбора признаков методом поиска в глубину.
10. Основная идея отбора признаков методом поиска в ширину.
11. Что такое МГУА?
12. Основная идея отбора признаков с помощью генетического алгоритма.
13. Основная идея отбора признаков с помощью случайного поиска.
14. В чём отличия случайного поиска от случайного поиска с адаптацией?
15. Приведите пример выборки, которую невозможно классифицировать без ошибок с помощью линейного алгоритма классификации. Какова минимальная длина выборки, обладающая данным свойством? Какие существуют способы модифицировать линейный алгоритм так, чтобы данная выборка стала линейно разделимой?
16. Почему любая булева функция представима в виде нейронной сети? Сколько в ней слоёв?
17. Метод обратного распространения ошибок. Основная идея. Основные недостатки и способы их устранения.
18. Как можно выбирать начальное приближение в градиентных методах настройки нейронных сетей?
19. Как можно ускорить сходимость в градиентных методах настройки нейронных сетей?
20. Что такое диагональный метод Левенберга-Марквардта?
21. Что такое «паралич» сети, и как его избежать?
22. Как выбирать число слоёв в градиентных методах настройки нейронных сетей?
23. Как выбирать число нейронов скрытого слоя в градиентных методах настройки нейронных сетей?
24. В чём заключается метод оптимального прореживания нейронной сети? Какие недостатки стандартного алгоритма обратного распространения ошибок позволяет устранить метод ODB?
25. Дать определение алгоритмической композиции (помнить формулу). Какие типы корректирующих операций вы знаете?
26. Какие типы голосования вы знаете? Какой из них наиболее общий? (помнить формулу)
27. Как обнаружить объекты-выбросы при построении композиции классификаторов для голосования по большинству?
28. Как обеспечивается различность базовых алгоритмов при голосовании по большинству?
29. Как обеспечивается различность базовых алгоритмов при голосовании по старшинству?
30. Какие возможны стратегии выбора классов базовых алгоритмов при голосовании по старшинству?
31. Какие две эвристики лежат в основе алгоритма AdaBoost?
32. Как обнаружить объекты-выбросы в алгоритме AdaBoost?
33. Достоинства и недостатки алгоритма AdaBoost.
34. Основная идея алгоритма AnyBoost.
35. Основная идея метода bagging.
36. Основная идея метода случайных подпространств.
37. Что такое смесь экспертов (помнить формулу)?
38. Приведите примеры выпуклых функций потерь. Почему свойство выпуклости помогает строить смеси экспертов?
39. Что такое логическая закономерность? Приведите примеры закономерностей в задаче распознавания спама.

40. Часто используемые типы логических закономерностей.
41. Дайте определение эпсилон-дельта-логической закономерности (помнить формулы).
42. Дайте определение статистической закономерности (помнить формулы).
43. Сравните области статистических и логических закономерностей в (p, n) -плоскости.
44. С какой целью делается бинаризация?
45. В чём заключается процедура бинаризации признака?
46. Как происходит перебор в жадном алгоритме синтеза информативных конъюнкций?
47. Какие критерии информативности используются в жадном алгоритме синтеза информативных конъюнкций и почему?
48. Как приспособить жадный алгоритм синтеза конъюнкций для синтеза информативных шаров?
49. Что такое стохастический локальный поиск?
50. В чём отличия редукции и стабилизации? В чём их достоинства и недостатки?
51. Что такое решающий список?
52. Какие критерии информативности используются при синтезе решающего списка и почему?
53. Достоинства и недостатки решающих списков.
54. Что такое решающее дерево?
55. Какие критерии информативности используются при синтезе решающего дерева и почему?
56. Достоинства и недостатки решающих деревьев.
57. Зачем делается редукция решающих деревьев?
58. Какие есть два основных типа редукции решающих деревьев?
59. Как преобразовать решающее дерево в решающий список, и зачем это делается?
60. Что такое ADT (alternating decision tree)? Как происходит построение ADT?
61. Основная идея алгоритма КОРА.
62. Почему возникает проблема предпочтения признаков с меньшими номерами в алгоритме КОРА? Как она решается?
63. Основная идея алгоритма ТЭМП.
64. Какие критерии информативности используются в алгоритме ТЭМП и почему?
65. Почему возникает проблема дублирования закономерностей в алгоритме ТЭМП? Как она решается?
66. Достоинства и недостатки алгоритма ТЭМП.
67. Как использовать алгоритм AdaBoost для построения взвешенного голосования закономерностей?
68. Какой критерий информативности используется в алгоритме AdaBoost?
69. Структура алгоритма вычисления оценок (АВО).
70. Что такое ассоциативное правило? Приведите пример ассоциативного правила в задаче анализа потребительских корзин.
71. Основная идея алгоритма поиска ассоциативных правил APriory.
72. Каковы основные цели кластеризации?
73. Основные типы кластерных структур. Приведите для каждой из этих структур пример алгоритма кластеризации, который для неё НЕ подходит.
74. В чём заключается алгоритм кратчайшего незамкнутого пути? Как его использовать для кластеризации? Как с его помощью определить число кластеров? Всегда ли это возможно?
75. Основная идея алгоритма ФорЭл.
76. Как вычисляются центры кластеров в алгоритме ФорЭл, если объекты — элементы метрического (не обязательно линейного векторного) пространства?
77. Какие существуют функционалы качества кластеризации и для чего они применяются?
78. Основные отличия алгоритма k -средних и EM-алгоритма. Кто из них лучше и почему?
79. Основная идея иерархического алгоритма Ланса-Вильямса.

80. Какие основные типы расстояний между кластерами применяются в алгоритме Ланса-Вильямса?
81. Какие расстояния между кластерами, применяемые в алгоритме Ланса-Вильямса, лучше и почему?
82. Что такое дендрограмма? Всегда ли её можно построить?

Билет 1

1. Как приспособить EM-алгоритм для решения задачи с частичным обучением?
2. Какие способы решения задачи с частичным обучением Вы знаете?

Билет 2

1. Почему задачи с частичным обучением выделены в отдельный класс? Приведите примеры, когда методы классификации и кластеризации дают неадекватное решение задачи с частичным обучением.
2. Как приспособить графовые алгоритмы кластеризации для решения задачи с частичным обучением?

Билет 3

1. Как устроена самоорганизующаяся карта Кохонена?
2. Как интерпретируются карты Кохонена?

Билет 4

1. Какой функционал качества оптимизируется сетью Кохонена? (помнить формулу)
2. В чем отличия правил мягкой и жесткой конкуренции? В чём преимущества мягкой конкуренции?

Критерии оценивания

«зачтено» – выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но недостаточно грамотно обосновывает полученные результаты;

«не зачтено» – выставляется студенту, который не знает основного содержания учебной программы дисциплины, допускает грубейшие ошибки в формулировках базовых понятий дисциплины и вообще не имеет навыков решения типовых практических задач.

5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности

Во время проведения зачета обучающиеся могут пользоваться программой дисциплины, а также справочной литературой, вычислительной техникой, конспектами лекций. Зачет может проводиться по итогам текущей успеваемости и сдачи заданий или путем организации специального опроса, проводимого в устной форме.