

Министерство науки и высшего образования РФ
Федеральное государственное автономное образовательное учреждение
высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»



УТВЕРЖДАЮ

Проректор по учебной работе

/Д.С. Гуц/

«28» марта 2022 года

ПРОГРАММА
вступительного испытания для поступающих в аспирантуру

1.2 Компьютерные науки и информатика

шифр и наименование группы научных специальностей

1.2.1 Искусственный интеллект и машинное обучение

шифр и наименование научной специальности

Перечень вопросов вступительного испытания

1 Задача обучения по прецедентам. Объекты и признаки. Ответы и типы задач. Модель алгоритмов и метод обучения. Этап обучения и этап применения. Функционал качества. Примеры. Эмпирический риск. Переобучение, его возможные причины и способы его минимизации.

2 Сведение задачи обучения к задаче оптимизации. Примеры.

3 Обобщающая способность. Формализации понятия «обобщающая способность». Проблема обобщающей способности.

4 Типы задач, причисляемые к машинному обучению. Понятие обучения с учителем и без. Примеры алгоритмов. Особенности. Постановка задачи классификации. Постановка задачи регрессии. Примеры.

5 Вероятностный подход в машинном обучении. Наивный Байесовский классификатор. Оценка его оптимальности и возможности применения на практике.

6 Основные алгоритмы, применяемые в задачах классификации. Основные алгоритмы, применяемые в задачах регрессии. Основные алгоритмы, применяемые в задачах кластеризации.

7 Основные проблемы исходных данных и способы их решения (пропуски, выбросы, несовместимые с алгоритмом типы данных). Способы визуализации данных.

8 Методы минимизации числа признаков. Метод главных компонент (РСА).

9 Понятие интерпретируемости алгоритма и принимаемого им решения по отдельному прецеденту. Поясняющие примеры задач.

10 Нейронные сети. Структура сети.

11 Подходы к реализации многоклассового классификатора на основе бинарных классификаторов.

12 Метрические методы классификации. Гипотеза компактности. Обобщенный метрический классификатор. Метод ближайшего соседа и его обобщения. Особенности, преимущества и недостатки. Метрика. Примеры.

13 Логические алгоритмы классификации. Понятия закономерности, информативности и интерпретируемости. Основные виды закономерностей. Примеры критериев информативности. Возможные их недостатки. Пример алгоритма поиска информативных закономерностей.

14 Бинарное решающее дерево. Пример алгоритма построения бинарного решающего дерева (ID3). Его достоинства и недостатки. Обработка пропусков в данных на стадии обучения и на стадии применения алгоритма. Понятие редукции решающего дерева. Цель. Пример алгоритма.

15 Бинаризация вещественного признака. Цель. Способы разбиения области значений признака на зоны. Пример алгоритма. Линейные методы

16 Линейный метод опорных векторов (linearSVM). Нелинейное обобщение SVM (kernelSVM). Ядра. Особенности метода проверки, является функция ядром или нет. Конструктивные методы синтеза ядер. Примеры ядер.

17 Определение ROC-кривой и AUC. Характеристики ROC-кривой. Эффективный алгоритм вычисления AUC и его асимптотическая сложность.

18 Композиция классификаторов. Корректирующие операции: простое голосование, взвешенное голосование, смесь алгоритмов. Примеры.

19 Бустинг. Основная идея. Бустинг для бинарной задачи классификации. Основная теорема бустинга (для AdaBoost). Переобучение при применении бустинга. Особенности в случаях построения комбинации простых и сложных алгоритмов. Алгоритм AdaBoost. Эвристики и рекомендации.

20 Постановка задачи кластеризации. Цели. Некорректность задачи кластеризации. Основные типы кластерных структур. Функции качества кластеризации.

21 Агломеративная иерархическая кластеризация. Формула Ланса-Уильямса. Визуализация кластерных структур: диаграмма вложения, дендрограмма.

22 EM-алгоритм. Метод k-средних (k-means).

Список рекомендованных источников

1. Ян Лекун. Как учится машина. Революция в области нейронных сетей и глубокого обучения. — М.: Альпина Диджитал, 2021.
2. Гладилин, П. Е. Технологии машинного обучения : учебно-методическое пособие / П. Е. Гладилин, К. О. Боченина. — Санкт-Петербург : НИУ ИТМО, 2020. — 75 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/190885>
3. Флах, П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / П. Флах. — Москва : ДМК Пресс, 2015. — 400 с. — ISBN 978-5-97060-273-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/69955>
4. Воронина, В. В. Теория и практика машинного обучения : учебное пособие / В. В. Воронина. — Ульяновск : УлГТУ, 2017. — 290 с. — ISBN 978-5-9795-1712-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/165053>
5. Шалев-Шварц, Ш. Идеи машинного обучения : учебное пособие / Ш. Шалев-Шварц, Бен-Давид Ш. ; перевод с английского А. А. Слинкина. — Москва : ДМК Пресс, 2019. — 436 с. — ISBN 978-5-97060-673-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/131686>
6. Боровская, Е. В. Основы искусственного интеллекта : учебное пособие / Е. В. Боровская, Н. А. Давыдова. — 4-е изд. — Москва : Лаборатория знаний, 2020. — 130 с. — ISBN 978-5-00101-908-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/151502>
7. Гаврилова, И. В. Основы искусственного интеллекта : учебное пособие / И. В. Гаврилова, О. Е. Масленникова. — 3-е изд., стер. — Москва : ФЛИНТА, 2019. — 283 с. — ISBN 978-5-9765-1602-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/115839>

Составители программы:

д.т.н., профессор кафедры информационных систем Масич И.С.

