

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Макаренко Елена Николаевна

Должность: Ректор

Дата подписания: 28.04.2021 16:37:52

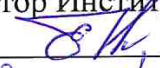
Уникальный программный ключ:

c098bc0c1041cb2a4cf926cf171d6715d99a6ae00adc8e27b55cbe1e2dbd7c78

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение высшего образования «Ростовский государственный экономический университет (РИНХ)»

УТВЕРЖДАЮ

Директор Института магистратуры

 Иванова Е.А.
« 30 » 05 2021 г.

**Рабочая программа дисциплины
Машинное обучение**

Направление 01.04.02 Прикладная математика и информатика
магистерская программа 01.04.02.03 "Искусственный интеллект в цифровой экономике"

Для набора 2021 года

Квалификация
Магистр

КАФЕДРА **Информационных систем и прикладной информатики****Распределение часов дисциплины по семестрам**

| Семестр (<Курс>. <Семестр на курсе>) | 2 (1.2) | | 3 (2.1) | | Итого | |
|--|---------|-----|---------|-----|-------|-----|
| | Неделя | | Неделя | | | |
| Неделя | 12 4/6 | | 16 4/6 | | | |
| Вид занятий | уп | рп | уп | рп | уп | рп |
| Лекции | 10 | 10 | 14 | 14 | 24 | 24 |
| Лабораторные | 10 | 10 | 14 | 14 | 24 | 24 |
| Итого ауд. | 20 | 20 | 28 | 28 | 48 | 48 |
| Контактная работа | 20 | 20 | 28 | 28 | 48 | 48 |
| Сам. работа | 84 | 84 | 107 | 107 | 191 | 191 |
| Часы на контроль | 4 | 4 | 9 | 9 | 13 | 13 |
| Итого | 108 | 108 | 144 | 144 | 252 | 252 |

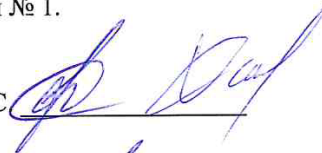
ОСНОВАНИЕ

Учебный план утвержден учёным советом вуза от 30.08.2021 протокол № 1.

Программу составил(и): д.э.н., проф., Долженко А.И.; доц., Хаймин Е.С.

Зав. кафедрой: д.э.н., доцент Щербаков С.М.

Методическим советом направления: д.ф.-м.н., проф., Стрюков М.Б.





1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

| | |
|-----|---|
| 1.1 | освоение обучающимися современных методов машинного обучения для систем интеллектуального анализа данных. |
|-----|---|

2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

УК-4:Способен применять современные коммуникативные технологии, в том числе на иностранном(ых) языке(ах), для академического и профессионального взаимодействия

ОПК-4:Способен комбинировать и адаптировать существующие информационно-коммуникационные технологии для решения задач в области профессиональной деятельности с учетом требований информационной безопасности

ПК-3:Способен разрабатывать и применять математические методы, системное и прикладное программное обеспечение для решения задач научной и проектно-технологической деятельности

В результате освоения дисциплины обучающийся должен:

| |
|--|
| Знать: |
| современные коммуникативные технологии (соотнесено с индикатором УК-4.1) требования информационной безопасности (соотнесено с индикатором ОПК-4.1) основные положения математических методов, системного и прикладного программного обеспечения (соотнесено с индикатором ПК-3.1) |
| Уметь: |
| применять современные коммуникативные технологии для решения задач в профессиональной области (соотнесено с индикатором УК-4.2) комбинировать и адаптировать существующие информационно-коммуникационные технологии для решения задач в области профессиональной деятельности (соотнесено с индикатором ОПК-4.2) применять математические методы, системное и прикладное программное обеспечение для решения задач в профессиональной области, в том числе в новой среде (соотнесено с индикатором ПК-3.2) |
| Владеть: |
| навыками применения современных коммуникативных технологий (соотнесено с индикатором УК-4.3) навыками адаптации технологий для решения задач в области профессиональной деятельности (соотнесено с индикатором ОПК-4.3) навыками использования математических методов, системного и прикладного программного обеспечения для решения тривиальных и нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте (соотнесено с индикатором ПК-3.3) |

3. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

| Код занятия | Наименование разделов и тем /вид занятия/ | Семестр / Курс | Часов | Компетенции | Литература |
|-------------|---|----------------|-------|--------------------|----------------------------|
| | Раздел 1. Методы машинного обучения. Часть 1 | | | | |
| 1.1 | Тема 1. Введение в машинное обучение. Задачи и инструментарий МО. Машинное обучение заключается в извлечении знаний из данных. Это научная область, находящаяся на пересечении статистики, искусственного интеллекта и компьютерных наук и также известная как прогнозная аналитика или статистическое обучение. Сферы применения машинного обучения: современные веб-сайты и устройства используют алгоритмы машинного обучения, начиная с автоматических рекомендаций по просмотру фильмов, заказа еды или покупки продуктов; МО оказывает огромное влияние на научные исследования, управляемые данными. Алгоритмы машинного обучения: обучение с учителем или контролируемое обучение, где пользователь предоставляет алгоритму пары объект- ответ, а алгоритм находит способ получения ответа по объекту; алгоритмы обучения без учителя или неконтролируемого обучения, где известны только объекты, а ответов нет. /Лек/ | 2 | 4 | УК-4 ОПК-4 ПК-3 | Л1.1 Л1.2Л2.1 Л2.2 Л2.3 |

| | | | | | |
|-----|--|---|----|--------------------|----------------------------|
| 1.2 | <p>Тема 2.</p> <p>Методы машинного обучения с учителем.</p> <p>Есть две основные задачи машинного обучения с учителем: классификация (classification) и регрессия (regression).</p> <p>Цель классификации состоит в том, чтобы спрогнозировать метку класса (class label), которая представляет собой выбор из заранее определенного списка возможных вариантов.</p> <p>Линейные методы классификации и регрессии. Класс задач для решения линейными методами.</p> <p>Модели обучения с учителем. Метод ближайших соседей.</p> <p>Оценка настройки параметров.</p> <p>Модель дерева и ансамбли деревьев (случайный лес и градиентный бустинг). Особенности переобучения деревьев.</p> <p>/Лек/</p> | 2 | 6 | УК-4 ОПК-4 ПК-3 | Л1.1 Л1.2Л2.1 Л2.2 Л2.3 |
| 1.3 | <p>Тема 1.</p> <p>Анализ данных с помощью блокнота Python.</p> <p>Блокнот Python — можно использовать среду выполнения Jupyter или Google Colab. Jupyter — это среда на основе IPython, облегчающая интерактивное программирование и анализ данных с использованием различных языков программирования, включая Python.</p> <p>Coogle Colab - облачный сервис для работы с интерактивным программированием и анализом данных на языке Python.</p> <p>Цель лабораторной работы:</p> <ul style="list-style-type: none"> • Создание блокнота в среде разработки Jupiter или Google Colab. • Отправка, обработка и визуализация данных в блокноте. • Предоставление совместного доступа к записным книжкам в Интернете./ <p>Выполнение заданий с Python, Anaconda 3, Tensor Flow, SciPy, NumPy, MatalibPlot. /Лаб/</p> | 2 | 4 | УК-4 ОПК-4 ПК-3 | Л1.1 Л1.2Л2.1 Л2.2 Л2.3 |
| 1.4 | <p>Тема 2.</p> <p>Реализация анализа данных с помощью базовых моделей машинного обучения с учителем.</p> <p>Метод ближайших соседей (k-neighbors); линейная классификация и регрессия; деревья и ансамбли деревьев (случайный лес и градиентный бустинг).</p> <p>Разработано тысяч библиотек на языке Python с открытым кодом, таких как NumPy, Pandas, Matplotlib и scikit-learn, которые позволяют программистам и аналитикам изучать, преобразовывать, анализировать и визуализировать данные.</p> <p>Цель лабораторной работы:</p> <ul style="list-style-type: none"> • Создание блокнота в среде разработки Jupiter или Google Colab. • Использование Pandas для очистки и подготовки данных для модели машинного обучения. • Использование scikit-learn для создания модели машинного обучения: метод ближайших соседей (k-neighbors); линейная классификация или регрессия; деревья или ансамбли деревьев (случайный лес или градиентный бустинг). • Использование Matplotlib для визуализации и сравнения эффективности моделей. <p>Выполнение заданий с Python, Anaconda 3, Tensor Flow, SciPy, NumPy, MatalibPlot. /Лаб/</p> | 2 | 6 | УК-4 ОПК-4 ПК-3 | Л1.1 Л1.2Л2.1 Л2.2 Л2.3 |
| 1.5 | Математические основы машинного обучения /Ср/ | 2 | 84 | УК-4 ОПК-4 ПК-3 | Л1.1 Л1.2Л2.1 Л2.2 Л2.3 |

| | | | | | |
|-----|--|---|-----|--------------------|----------------------------|
| 1.6 | - /Зачёт/ | 2 | 4 | УК-4 ОПК-4 ПК-3 | Л1.1 Л1.2Л2.1 Л2.2 Л2.3 |
| | Раздел 2. Методы машинного обучения. Часть 2 | | | | |
| 2.1 | Тема 3. Методы машинного обучения с учителем. Обучение нелинейных моделей методом опорных векторов SVM. Настройка модели и оценка параметров. Подготовка и масштабирование данных. Методы машинного обучения без учителя, которые включают в себя все виды машинного обучения, когда ответ неизвестен и отсутствует учитель, указывающий ответ алгоритму. Неконтролируемые преобразования – это алгоритмы, создающие новое представление данных, для облегчения анализа; Алгоритмы кластеризации (clustering algorithms) разбивают данные на отдельные группы схожих между собой элементов. /Лек/ | 3 | 8 | УК-4 ОПК-4 ПК-3 | Л1.1 Л1.2Л2.1 Л2.2 Л2.3 |
| 2.2 | Тема 4. Глубокое машинное обучение. Многослойные перцептроны (MLP) также называют простыми (vanilla) нейронными сетями прямого распространения, а иногда и просто нейронными сетями. Вводятся понятия – слой, узел, функции активации, построение сети. Настройка параметров слоев и их количество. Выбор функции активации и оценка модели. /Лек/ | 3 | 6 | УК-4 ОПК-4 ПК-3 | Л1.1 Л1.2Л2.1 Л2.2 Л2.3 |
| 2.3 | Тема 3. Особенности построения MLP модели. MLP можно рассматривать как обобщение линейных моделей, которое прежде, чем прийти к решению выполняет несколько этапов обработки данных. Настройка количество слоев, определение функции активации. Цели лабораторной работы: • Создание блокнота в среде разработки Jupiter или Google Colab. •Использование sklearn MLP для построения и обучения нейронной сети, которая будет выполнять анализ данных. •Использование нейронных сетей для анализа больших данных (данные о жертвах Титаника). Выполнение заданий с Python, Anaconda 3, Tensor Flow, SciPy, NumPy, MatalibPlot. /Лаб/ | 3 | 8 | УК-4 ОПК-4 ПК-3 | Л1.1 Л1.2Л2.1 Л2.2 Л2.3 |
| 2.4 | Тема 4. Оценка неопределенности классификаторов. Использование интерфейса scikit-learn дает возможность вычислить оценки неопределенности прогнозов. Предварительная обработка и преобразование данных. Выделение главных компонент. Кластеризация данных. Цель лабораторной работы: *Провести предварительную обработку обработку и масштабирование данных. Познакомиться с инструментом MinMaxScaler. Провести анализ главных компонент среди признаков. Разобрать пример распознавания лиц. Выполнение заданий с Python, Anaconda 3, Tensor Flow, SciPy, NumPy, MatalibPlot. /Лаб/ | 3 | 6 | УК-4 ОПК-4 ПК-3 | Л1.1 Л1.2Л2.1 Л2.2 Л2.3 |
| 2.5 | Методы глубокого обучения. /Ср/ | 3 | 107 | УК-4 ОПК-4 ПК-3 | Л1.1 Л1.2Л2.1 Л2.2 Л2.3 |

| | | | | | |
|-----|-----------|---|---|--------------------|----------------------------|
| 2.6 | /Экзамен/ | 3 | 9 | УК-4 ОПК-4 ПК-3 | Л1.1 Л1.2Л2.1 Л2.2 Л2.3 |
|-----|-----------|---|---|--------------------|----------------------------|

4. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

Структура и содержание фонда оценочных средств для проведения текущей и промежуточной аттестации представлены в Приложении 1 к рабочей программе дисциплины.

5. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

5.1. Основная литература

| | Авторы, составители | Заглавие | Издательство, год | Колич-во |
|------|---|---|--|--|
| Л1.1 | Бринк Х., Ричардс Д., Феверолф М. | Машинное обучение | Санкт-Петербург: Питер, 2017 | https://ibooks.ru/reading.php?short=1&productid=355472 неограниченный доступ для зарегистрированных пользователей |
| Л1.2 | Воронина, В. В., Михеев, А. В., Ярушкина, Н. Г., Святов, К. В. | Теория и практика машинного обучения: учебное пособие | Ульяновск: Ульяновский государственный технический университет, 2017 | http://www.iprbookshop.ru/106120.html неограниченный доступ для зарегистрированных пользователей |

5.2. Дополнительная литература

| | Авторы, составители | Заглавие | Издательство, год | Колич-во |
|------|--|---|---|--|
| Л2.1 | Емельянов А. А. | Прикладная информатика: журнал | Москва: Синергия ПРЕСС, 2006 | https://biblioclub.ru/index.php?page=book&id=120300 неограниченный доступ для зарегистрированных пользователей |
| Л2.2 | Чернышов, В. Н., Образцов, Д. В., Платёнкин, А. В. | Моделирование информационных процессов и исследование в ИТ: учебное пособие | Тамбов: Тамбовский государственный технический университет, ЭБС АСВ, 2017 | http://www.iprbookshop.ru/85960.html неограниченный доступ для зарегистрированных пользователей |
| Л2.3 | Элбон Крис | Машинное обучение с использованием Python. Сборник рецептов: Пер. с англ. | Санкт-Петербург: БХВ-Петербург, 2019 | https://ibooks.ru/reading.php?short=1&productid=366635 неограниченный доступ для зарегистрированных пользователей |

5.3 Профессиональные базы данных и информационные справочные системы

Национальная электронная библиотека (НЭБ) - <https://rusneb.ru/>

Консультант+

Гарант

5.4. Перечень программного обеспечения

язык Python

Anaconda 3

Tensor Flow

SciPy

NumPy

Matplotlib

5.5. Учебно-методические материалы для студентов с ограниченными возможностями здоровья

При необходимости по заявлению обучающегося с ограниченными возможностями здоровья учебно-методические материалы предоставляются в формах, адаптированных к ограничениям здоровья и восприятия информации. Для лиц с нарушениями зрения: в форме аудиофайла; в печатной форме увеличенным шрифтом. Для лиц с нарушениями слуха: в форме электронного документа; в печатной форме. Для лиц с нарушениями опорно-двигательного аппарата: в форме электронного документа; в печатной форме.

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Помещения для проведения всех видов работ, предусмотренных учебным планом, укомплектованы необходимой специализированной учебной мебелью и техническими средствами обучения. Для проведения лекционных занятий используется демонстрационное оборудование. Лабораторные занятия проводятся в компьютерных классах, рабочие места в которых оборудованы необходимыми лицензионными программными средствами и выходом в Интернет.

7. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)

Методические указания по освоению дисциплины представлены в Приложении 2 к рабочей программе дисциплины.

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

1. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания

1.1. Показатели и критерии оценивания компетенций:

| ЗУН, составляющие компетенцию | Показатели оценивания | Критерии оценивания | Средство оценивания |
|---|---|--|--|
| УК-4: Способен применять современные коммуникативные технологии, в том числе на иностранном(ых) языке(ах), для академического и профессионального взаимодействия | | | |
| З. современные коммуникативные технологии | знает содержание алгоритма обучения без учителя. основные задачи, решаемые с помощью алгоритма обучения с учителем, разновидности задач классификации при машинном обучении, содержание алгоритма обучения с учителем | полнота и содержательность ответа умение приводить примеры | О – задания для опроса (раздел 1: варианты 1-4), З – вопросы к зачету (1-5), Э – вопросы к экзамену (1-9) |
| У. применять современные коммуникативные технологии для решения задач в профессиональной области | применяет известные методы реализации алгоритмов обучения с учителем, известные методы реализации алгоритмов обучения без учителя | полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач | ЛЗ – лабораторные задания (раздел 1: задания 1-2) |
| В. навыками применения современных коммуникативных технологий | объясняет основную идею метода классификации «k ближайших соседей», поясняет требования, предъявляемые к наборам данным для обучения, оценивает качество модели обучения | полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач | ЛЗ – лабораторные задания (раздел 1: задания 1-2) |
| ОПК-4: Способен комбинировать и адаптировать существующие информационно-коммуникационные технологии для решения задач в области профессиональной деятельности с учетом требований информационной безопасности | | | |
| З. требования информационной безопасности | знает основную идею линейных методов машинного обучения, методы и алгоритмы предварительной обработки данных в машинном обучении, масштабирование обучающего и тестового набора | полнота и содержательность ответа умение приводить примеры | О – задания для опроса (раздел 2: варианты 5-9), З – вопросы к зачету (5-8), Э – вопросы к экзамену (10-18) |
| У. комбинировать и адаптировать существующие информационно-коммуникационные технологии для решения задач в области профессиональной деятельности | применяет кластеризацию данных при машинном обучении без учителя и алгоритм агломеративной кластеризации в машинном обучении без учителя | полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач | ЛЗ – лабораторные задания (раздел 1: задания 1-2) |
| В. навыками адаптации технологий для решения задач в области профессиональной деятельности | использует пакеты программ Tensor Flow, SciKit, NumPy, MatLibPlot для решения задач | полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач | ЛЗ – лабораторные задания (раздел 1: задания 1-2) |
| ПК-3: Способен разрабатывать и применять математические методы, системное и прикладное программное обеспечение для решения задач научной и проектно-технологической деятельности | | | |
| З. основные положения математических методов, системного и прикладного | | полнота и содержательность ответа | О – задания для опроса (раздел 1: варианты 1-4, раздел 2: варианты 5-9), З |

| | | | |
|---|--|--|--|
| программного обеспечения | | умение приводить примеры | – вопросы к зачету (10-14), Э – вопросы к экзамену (19-27) |
| У. применять математические методы, системное и прикладное программное обеспечение для решения задач в профессиональной области, в том числе в новой среде | | полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач | ЛЗ – лабораторные задания (раздел 2: задания 3-4) |
| В. навыками использования математических методов, системного и прикладного программного обеспечения для решения тривиальных и нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте | | полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач | ЛЗ – лабораторные задания (раздел 2: задания 3-4) |

1.2 Шкалы оценивания:

Текущий контроль успеваемости и промежуточная аттестация осуществляется в рамках накопительной балльно-рейтинговой системы в 100-балльной шкале:

50-100 баллов (зачет);

0-49 баллов (незачет).

84-100 баллов (оценка «отлично»);

67-83 баллов (оценка «хорошо»);

50-66 баллов (оценка «удовлетворительно»);

0-49 баллов (оценка «неудовлетворительно»).

2. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

Вопросы к зачету

1. Основная идея, содержание алгоритма обучения без учителя.
2. Класс задач, решаемых с помощью алгоритма обучения с учителем.
3. Разновидности задач классификации при машинном обучении.
4. Основная идея, содержание алгоритма обучения с учителем.
5. Назначение задач регрессии при машинном обучении.
6. Класс задач, решаемых с помощью алгоритма обучения без учителя.
7. Назовите известные вам методы реализации алгоритмов обучения с учителем.
8. Назовите известные вам методы реализации алгоритмов обучения без учителя.
9. Объясните основную идею метода классификации «к ближайших соседей».
10. Основные параметры, определяющие точность прогнозов, метода классификации «к ближайших соседей».
11. Поясните требования, предъявляемые к наборам данным для обучения.
12. Построение линейных моделей регрессии и классификации. Недостатки и преимущества.
13. Построение модели дерева. Построение графа дерева. Проблема переобучения на модели дерева.
14. Построение модели ансамбли деревьев. Особенности модели случайный лес. Принципы градиентного бустинга.

Зачетное задание включает два вопроса – один теоретический вопрос и одно практико-ориентированное задание из числа приведенных ниже практических заданий.

Критерии оценивания:

- 50-100 баллов («зачет») – изложенный материал фактически верен, наличие глубоких исчерпывающих знаний в объеме пройденной программы дисциплины в соответствии с поставленными программой курса целями и задачами обучения; правильные, уверенные действия по применению полученных знаний на практике, грамотное и логически стройное изложение материала при ответе, усвоение основной и знакомство с дополнительной литературой; наличие твердых и достаточно полных знаний в объеме пройденной программы дисциплины в соответствии с целями обучения, правильные действия по применению знаний на практике, четкое изложение материала, допускаются отдельные логические и стилистические погрешности, обучающийся усвоил основную литературу, рекомендованную в рабочей программе дисциплины; наличие твердых знаний в объеме пройденного курса в соответствии с целями обучения, изложение ответов с отдельными ошибками, уверенно исправленными после дополнительных вопросов; правильные в целом действия по применению знаний на практике;

- 0-49 баллов («незачет») – ответы не связаны с вопросами, наличие грубых ошибок в ответе, непонимание сущности излагаемого вопроса, неумение применять знания на практике, неуверенность и неточность ответов на дополнительные и наводящие вопросы.

Вопросы к экзамену

1. Основная идея, содержание алгоритма обучения без учителя.
2. Класс задач, решаемых с помощью алгоритма обучения с учителем.
3. Разновидности задач классификации при машинном обучении.
4. Основная идея, содержание алгоритма обучения с учителем.
5. Назначение задач регрессии при машинном обучении.
6. Класс задач, решаемых с помощью алгоритма обучения без учителя.
7. Назовите известные вам методы реализации алгоритмов обучения с учителем.
8. Назовите известные вам методы реализации алгоритмов обучения без учителя.
9. Объясните основную идею метода классификации «к ближайших соседей».
10. Основные параметры, определяющие точность прогнозов, метода классификации «к ближайших соседей».
11. Поясните требования, предъявляемые к наборам данным для обучения.
12. Построение линейных моделей регрессии и классификации. Недостатки и преимущества.
13. Построение модели дерева. Построение графа дерева. Проблема переобучения на модели дерева.
14. Построение модели ансамбли деревьев. Особенности модели случайный лес. Принципы градиентного бустинга.
15. Методы и алгоритмы предварительной обработки данных в машинном обучении.
16. Масштабирование обучающего и тестового набора.
17. Назовите методы снижения размерности наборов данных для машинного обучения.
18. Какие возможности предоставляет метод анализа главных компонент при машинном обучении без учителя?
19. Какие возможности предоставляет метод факторизации неотрицательных матриц при машинном обучении без учителя?
20. Поясните содержание задачи кластеризации данных при машинном обучении без учителя.
21. Поясните основную идею алгоритма кластеризации k-средних в машинном обучении без учителя.
22. Поясните основную идею алгоритма агломеративной кластеризации в машинном обучении без учителя.
23. Поясните основную идею плотностного алгоритма кластеризации пространственных данных с присутствием шума (DBSCAN) в машинном обучении без учителя.
24. Поясните содержание модели MLP многослойного персептрона.
25. Назовите известные вам функции активации нейронов в искусственной нейронной сети.
26. Назовите параметры нейронной сети, от которых может зависеть сложность модели и точность предсказаний.
27. Назовите основную идею настройки нейронной сети.

Экзаменационное задание включает три вопроса – два теоретических вопроса и одно практико-ориентированное задание из числа приведенных ниже лабораторных заданий.

Критерии оценивания:

- 84-100 баллов (оценка «отлично») – изложенный материал фактически верен, наличие глубоких исчерпывающих знаний в объеме пройденной программы дисциплины в соответствии с поставленными программой курса целями и задачами обучения; правильные, уверенные действия по применению полученных знаний на практике, грамотное и логически стройное изложение материала при ответе, усвоение основной и знакомство с дополнительной литературой;

- 67-83 баллов (оценка «хорошо») – наличие твердых и достаточно полных знаний в объеме пройденной программы дисциплины в соответствии с целями обучения, правильные действия по применению знаний на практике, четкое изложение материала, допускаются отдельные логические и стилистические погрешности, обучающийся усвоил основную литературу, рекомендованную в рабочей программе дисциплины;

- 50-66 баллов (оценка удовлетворительно) – наличие твердых знаний в объеме пройденного курса в соответствии с целями обучения, изложение ответов с отдельными ошибками, уверенно исправленными после дополнительных вопросов; правильные в целом действия по применению знаний на практике;

- 0-49 баллов (оценка неудовлетворительно) – ответы не связаны с вопросами, наличие грубых ошибок в ответе, непонимание сущности излагаемого вопроса, неумение применять знания на практике, неуверенность и неточность ответов на дополнительные и наводящие вопросы.

Задания для опроса

Раздел 1

Вариант 1

Основная идея, содержание алгоритма обучения без учителя.

Класс задач, решаемых с помощью алгоритма обучения с учителем.

Разновидности задач классификации при машинном обучении.

Вариант 2

Основная идея, содержание алгоритма обучения с учителем.

Назначение задач регрессии при машинном обучении.

Класс задач, решаемых с помощью алгоритма обучения без учителя.

Вариант 3

Назовите известные вам методы реализации алгоритмов обучения с учителем.

Назовите известные вам методы реализации алгоритмов обучения без учителя.

Объясните основную идею метода классификации «к ближайших соседей».

Вариант 4

Основные параметры, определяющие точность прогнозов, метода классификации «к ближайших соседей».

Поясните требования, предъявляемые к наборам данным для обучения.

Как оценивается качество модели обучения?

Построение линейных моделей регрессии и классификации. Недостатки и преимущества.

Критерии оценивания (для каждого варианта) (раздел 1):

18-20 б. – ответы на все вопросы даны верно;

13-17 б. – один из ответов с неточностями;

11-12 б. – 2 ответа с неточностями;

9-10 б. – 3 ответа с неточностями;

6-8 б. – нет ответа на один вопрос;

0-5 б. – нет ответа на 2 вопроса.

Максимальное количество баллов за опрос по разделу 1 – 20.

Раздел 2

Вариант 5

Основная идея линейных методов машинного обучения.

Методы и алгоритмы предварительной обработки данных в машинном обучении.

Масштабирование обучающего и тестового набора.

Вариант 6

Назовите методы снижения размерности наборов данных для машинного обучения.

Какие возможности предоставляет метод анализа главных компонент при машинном обучении без учителя?

Какие возможности предоставляет метод факторизации неотрицательных матриц при машинном обучении без учителя?

Вариант 7

Поясните содержание задачи кластеризации данных при машинном обучении без учителя.

Поясните основную идею алгоритма кластеризации k-средних в машинном обучении без учителя.

Поясните основную идею алгоритма агломеративной кластеризации в машинном обучении без учителя.

Вариант 8

Поясните основную идею плотностного алгоритма кластеризации пространственных данных с присутствием шума (DBSCAN) в машинном обучении без учителя.

Поясните содержание модели MLP многослойного персептрона.

Вариант 9

Назовите известные вам функции активации нейронов в искусственной нейронной сети.

Назовите параметры нейронной сети, от которых может зависеть сложность модели и точность предсказаний.

Назовите основную идею настройки нейронной сети.

Критерии оценивания (для каждого варианта) (раздел 2):

18-20 б. – ответы на все вопросы даны верно;

13-17 б. – один из ответов с неточностями;

11-12 б. – 2 ответа с неточностями;

9-10 б. – 3 ответа с неточностями;

6-8 б. – нет ответа на один вопрос;

0-5 б. – нет ответа на 2 вопроса.

Максимальное количество баллов за опрос по разделу 2 – 20.

Лабораторные задания

Раздел 1

Лабораторная работа №1

Подготовка данных для анализа классификации цветков Ириса

Лабораторная работа №2

Подготовка данных для прогнозирования вероятности болезни сердца у респондентов.

Критерии оценивания (для каждого задания 1-го раздела):

31-40 б. – задание выполнено верно;

21-30 б. – при выполнении задания были допущены неточности, не влияющие на результат;

11-20 б. – при выполнении задания были допущены ошибки;

0-10 б. – при выполнении задания были допущены существенные ошибки.

Максимальное количество баллов за лабораторные задания 1-го раздела – 80 (2 задания по 40 баллов).

Раздел 2

Лабораторная работа №3

Подготовка данных для анализа данных о пассажирах Титаника

Лабораторная работа №4

Подготовка данных для оценки неопределенности классификаторов

Критерии оценивания (для каждого задания 2-го раздела):

31-40 б. – задание выполнено верно;

21-30 б. – при выполнении задания были допущены неточности, не влияющие на результат;

11-20 б. – при выполнении задания были допущены ошибки;

0-10 б. – при выполнении задания были допущены существенные ошибки.

Максимальное количество баллов за лабораторные задания 2-го раздела – 80 (2 задания по 40 баллов).

3. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

Процедуры оценивания включают в себя текущий контроль и промежуточную аттестацию.

Текущий контроль успеваемости проводится с использованием оценочных средств, представленных в п. 2 данного приложения. Результаты текущего контроля доводятся до сведения студентов до промежуточной аттестации.

Промежуточная аттестация проводится в форме зачета, экзамена.

Зачет проводится по окончании теоретического обучения до начала экзаменационной сессии. Количество вопросов в зачетном задании – 2 (один теоретический вопрос и одно практико-ориентированное задание). Объявление результатов производится в день зачета. Результаты аттестации заносятся в зачетную ведомость и зачетную книжку студента. Студенты, не прошедшие промежуточную аттестацию по графику сессии, должны ликвидировать задолженность в установленном порядке.

Экзамен проводится по расписанию экзаменационной сессии в устном виде. Количество вопросов в экзаменационном задании – 3 (два теоретических вопроса и одно практико-ориентированное задание). Объявление результатов производится в день экзамена. Результаты аттестации заносятся в экзаменационную ведомость и зачетную книжку студента. Студенты, не прошедшие промежуточную аттестацию по графику сессии, должны ликвидировать задолженность в установленном порядке.

МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Учебным планом предусмотрены следующие виды занятий:

- лекции;
- лабораторные занятия.

В ходе лекционных занятий рассматриваются основные теоретические вопросы, даются рекомендации для самостоятельной работы и подготовке к лабораторным занятиям.

В ходе лабораторных занятий углубляются и закрепляются знания студентов по ряду рассмотренных на лекциях вопросов, развиваются навыки практической работы.

При подготовке к лабораторным занятиям каждый студент должен:

- изучить рекомендованную учебную литературу;
- изучить конспекты лекций;
- подготовить ответы на все вопросы по изучаемой теме.

В процессе подготовки к лабораторным занятиям студенты могут воспользоваться консультациями преподавателя.

Вопросы, не рассмотренные на лекциях и лабораторных занятиях, должны быть изучены студентами в ходе самостоятельной работы. Контроль самостоятельной работы студентов над учебной программой курса осуществляется в ходе занятий методом опроса и выполнения лабораторных заданий. В ходе самостоятельной работы каждый студент обязан прочитать основную и по возможности дополнительную литературу по изучаемой теме, дополнить конспекты лекций недостающим материалом, выписками из рекомендованных первоисточников, выделить непонятные термины, найти их значение в энциклопедических словарях.

Студент должен готовиться к предстоящему лабораторному занятию по всем обозначенным в рабочей программе дисциплины вопросам.

Для подготовки к занятиям, текущему контролю и промежуточной аттестации студенты могут воспользоваться электронно-библиотечными системами. Также обучающиеся могут взять на дом необходимую литературу на абонементе университетской библиотеки или воспользоваться читальными залами.