

Документ подписан Министром науки и высшего образования Российской Федерации
Информация о владельце:
ФИО: Макаренко Елена Николаевна
Должность: Ректор
Дата подписания: 20.05.2024 11:04:49
Уникальный программный ключ:
c098bc0c1041cb2a4cf926cf171d6715d99a6ae00adc8e27b55cbe1e2dbd7c78

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

УТВЕРЖДАЮ
Начальник отдела лицензирования и
аккредитации
_____ Чаленко К.Н.
« ____ » _____ 20__ г.

**Рабочая программа дисциплины
Технологии машинного обучения**

основная профессиональная образовательная программа по направлению 02.03.02
Фундаментальная информатика и информационные технологии
02.03.02.01 "Теоретические основы информатики и компьютерные науки"

Для набора 2021 года

Квалификация
Бакалавр

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

Семестр (<Курс>.<Семестр на курсе>)	9 (5.1)		Итого	
Неделя	6			
Вид занятий	УП	РП	УП	РП
Лекции	8	8	8	8
Лабораторные	8	8	8	8
Итого ауд.	16	16	16	16
Контактная работа	16	16	16	16
Сам. работа	119	119	119	119
Часы на контроль	9	9	9	9
Итого	144	144	144	144

ОСНОВАНИЕ

Учебный план утвержден учёным советом вуза от 31.08.2021 протокол № 1.

Программу составил(и): д.э.н., проф., Долженко А.И. _____

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

Методическим советом направления: д.э.н., проф., Тищенко Е.Н. _____

1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

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

2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

ПК-1: Способность демонстрации общенаучных базовых знаний математических и естественных наук, фундаментальной информатики и информационных технологий

ПК-2: Способность понимать и применять в научно-исследовательской и прикладной деятельности современный математический аппарат, основные законы естествознания, современные языки программирования и программное обеспечение, операционные системы и сетевые технологии

ПК-3: Способность собирать, обрабатывать и интерпретировать экспериментальные данные, необходимые для проектной и производственно-технологической деятельности; способность к разработке новых алгоритмических, методических и технологических решений в конкретной сфере профессиональной деятельности

В результате освоения дисциплины обучающийся должен:

Знать:
методы машинного обучения основы аналитических работ в ИТ-проектах основы алгоритмизации и программирования
Уметь:
выбирать методы машинного обучения при решении профессиональных задач планировать аналитические работы в ИТ-проектах в профессиональной области разрабатывать алгоритмы и программы в конкретной сфере профессиональной деятельности
Владеть:
навыками применения методов машинного обучения при решении задач в профессиональной сфере навыками анализа работ в ИТ-проектах в профессиональной области навыками применения алгоритмических, программных и технологических решений в конкретной сфере профессиональной деятельности

3. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Компетенции	Литература
	Раздел 1. Методы машинного обучения				
1.1	Тема 1. Задачи и инструментарий машинного обучения. 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. В последние годы применение методов машинного обучения в повседневной жизни стало обыденным явлением. Многие современные веб-сайты и устройства используют алгоритмы машинного обучения, начиная с автоматических рекомендаций по просмотру фильмов, заказа еды или покупки продуктов, и заканчивая персонализированными онлайн радиотрансляциями и распознаванием друзей на фотографиях. Выйдя за пределы коммерческих приложений, машинное обучение уже оказало огромное влияние на научные исследования, управляемые данными. /Лек/	9	2	ПК-1 ПК-2 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>Цель регрессии состоит в том, чтобы спрогнозировать непрерывное число или число с плавающей точкой (floating - point number), если использовать термины программирования, или вещественное число (real number), если говорить языком математических терминов.</p> <p>Прогнозирование годового дохода человека в зависимости от его образования, возраста и места жительства является примером регрессионной задачи. /Лек/</p>	9	2	ПК-1 ПК-2 ПК-3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
1.3	<p>Тема 3.</p> <p>Методы машинного обучения без учителя.</p> <p>Машинное обучение без учителя включает в себя все виды машинного обучения, когда ответ неизвестен и отсутствует учитель, указывающий ответ алгоритму. В машинном обучении без учителя есть лишь входные данные и алгоритму необходимо извлечь знания из этих данных. /Лек/</p>	9	2	ПК-1 ПК-2 ПК-3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
1.4	<p>Тема 4.</p> <p>Глубокое машинное обучение.</p> <p>Семейство алгоритмов, известное как нейронные сети, недавно пережило свое возрождение под названием «глубокое обучение».</p> <p>Несмотря на то что глубокое обучение сулит большие перспективы в различных сферах применения машинного обучения, алгоритмы глубокого обучения, как правило, жестко привязаны к конкретным случаям использования. В данном разделе мы рассмотрим лишь некоторые относительно простые методы, а именно многослойные перцептроны для классификации и регрессии, которые могут служить отправной точкой в изучении более сложных методов машинного обучения. Многослойные перцептроны (MLP) также называют простыми (vanilla) нейронными сетями прямого распространения, а иногда и просто нейронными сетями. /Лек/</p>	9	2	ПК-1 ПК-2 ПК-3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
1.5	<p>Тема 1.</p> <p>Анализ данных климата с записными книжками.</p> <p>Записные книжки — это облачная платформа для создания и выполнения записных книжек. Jupyter — это среда на основе Python, облегчающая интерактивное программирование и анализ данных с использованием различных языков программирования, включая Python. Записные книжки Jupyter широко используются в научной среде для математического моделирования, машинного обучения, статистического анализа, а также для обучения написанию кода. Выполнение лабораторного задания с использованием LibreOffice.</p> <p>/Лаб/</p>	9	2	ПК-1 ПК-2 ПК-3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
1.6	<p>Тема 2.</p> <p>Прогнозирование задержки рейсов путем создания модели машинного обучения.</p> <p>Python является одним из наиболее популярных языков программирования в мире. Он широко используется в сообществе специалистов по обработке и анализу данных для машинного обучения и статистического анализа. Одна из причин его популярности — наличие тысяч библиотек с открытым кодом, таких как NumPy, Pandas, Matplotlib и scikit-learn, которые позволяют программистам и аналитикам изучать, преобразовывать, анализировать и визуализировать данные. Выполнение лабораторного задания с использованием LibreOffice.</p> <p>/Лаб/</p>	9	2	ПК-1 ПК-2 ПК-3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3

1.7	Тема 3. Анализ тональности отзывов с Keras. Keras — это API нейронных сетей высокого уровня, написанный на языке Python, который выполняется в дополнение к другим инструментам глубокого обучения, таким как TensorFlow. Keras используется для сборки нейронной сети, которая оценивает текст, например отзывы пользователей, для определения тональности. Выполнение лабораторного задания с использованием LibreOffice. /Лаб/	9	2	ПК-1 ПК-2 ПК-3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
1.8	Тема 4. Оценка неопределенности классификаторов. Использование интерфейса scikit-learn дает возможность вычислить оценки неопределенности прогнозов. Часто вас интересует не только класс, спрогнозированный моделью для определенной точки тестового набора, но и степень уверенности модели в правильности прогноза. В реальной практике различные виды ошибок приводят к очень разным результатам. Выполнение лабораторного задания с использованием LibreOffice. /Лаб/	9	2	ПК-1 ПК-2 ПК-3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
1.9	Тема: Математические основы машинного обучения Линейная алгебра для машинного обучения. Многомерное исчисление для машинного обучения. Теория вероятности для машинного обучения. /Ср/	9	119	ПК-1 ПК-2 ПК-3	Л1.1 Л1.2Л2.1 Л2.2 Л2.3
1.10	- /Экзамен/	9	9	ПК-1 ПК-2 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 Хенрик, Ричардс Джозеф, Феверолф Марк	Машинное обучение. — (Серия «Библиотека программиста»)	Санкт-Петербург: Питер, 2018	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	https://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=120298 неограниченный доступ для зарегистрированных пользователей
Л2.2	Кудинов, Ю. И.	Интеллектуальные системы: учебное пособие	Липецк: Липецкий государственный технический университет, ЭБС АСВ, 2014	https://www.iprbookshop.ru/55089.html неограниченный доступ для зарегистрированных пользователей

	Авторы, составители	Заглавие	Издательство, год	Колич-во
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=3666 35 неограниченный доступ для зарегистрированных пользователей

5.3 Профессиональные базы данных и информационные справочные системы

ИСС «КонсультантПлюс»

ИСС «Гарант» <http://www.internet.garant.ru/>

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

5.4. Перечень программного обеспечения

LibreOffice

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
ПК-1: Способность демонстрации общенаучных базовых знаний математических и естественных наук, фундаментальной информатики и информационных технологий			
З. методы машинного обучения	знает основные понятия и определения, методы, алгоритмы и технологии	полнота и содержательность ответа умение приводить примеры	Т – тест (1-5), Э – вопросы к экзамену (1-27)
У. выбирать методы машинного обучения при решении профессиональных задач	выполняет лабораторные задания, отвечает на вопросы, умеет применять полученные знания на практике	полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач	ЛЗ – лабораторные задания (1-4)
В. навыками применения методов машинного обучения при решении задач в профессиональной сфере	проводит обобщенный анализ информации и обработку данных	полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач	ЛЗ – лабораторные задания (1-4)
ПК-2: Способность понимать и применять в научно-исследовательской и прикладной деятельности современный математический аппарат, основные законы естествознания, современные языки программирования и программное обеспечение, операционные системы и сетевые технологии			
З. основы аналитических работ в ИТ-проектах	знает основные понятия и определения, методы, алгоритмы и технологии	полнота и содержательность ответа умение приводить примеры	Т – тест (6-10), Э – вопросы к экзамену (1-27)
У. планировать аналитические работы в ИТ-проектах в профессиональной области	выполняет лабораторные задания, отвечает на вопросы, умеет применять полученные знания на практике	полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач	ЛЗ – лабораторные задания (1-4)
В. навыками анализа работ в ИТ-проектах в профессиональной области	проводит обобщенный анализ информации и обработку данных	полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач	ЛЗ – лабораторные задания (1-4)
ПК-3: Способность собирать, обрабатывать и интерпретировать экспериментальные данные, необходимые для проектной и производственно-технологической деятельности; способность к разработке новых алгоритмических, методических и технологических решений в конкретной сфере профессиональной деятельности			
З. основы алгоритмизации и программирования	знает основные понятия и определения, методы, алгоритмы и технологии	полнота и содержательность ответа умение приводить примеры	Т – тест (11-15), Э – вопросы к экзамену (1-27)
У. разрабатывать алгоритмы и программы в конкретной сфере профессиональной деятельности	выполняет лабораторные задания, отвечает на вопросы, умеет применять полученные знания на практике	полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач	ЛЗ – лабораторные задания (1-4)
В. навыками применения алгоритмических, программных и технологических решений в конкретной сфере профессиональной деятельности	проводит обобщенный анализ информации и обработку данных	полнота и содержательность ответа умение приводить примеры умение самостоятельно находить решение поставленных задач	ЛЗ – лабораторные задания (1-4)

1.2 Шкалы оценивания:

Текущий контроль успеваемости и промежуточная аттестация осуществляется в рамках

накопительной балльно-рейтинговой системы в 100-балльной шкале:

- 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 экзамену

1. Основная идея, содержание алгоритма обучения без учителя.
2. Основные задачи, решаемые 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. Основные задачи, решаемые 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 «k ближайших соседей».
10. Основные параметры, определяющие точность прогнозов, метода классификации «k ближайших соседей».
11. Поясните требования, предъявляемые к наборам данным для обучения.
12. Как оценивается качество модели обучения?
13. Как сложность модели обучения с учителем зависит от таких характеристик как «недообучение» и «переобучение»?
14. Основная идея линейных методов машинного обучения.
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 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.

Тест

1 Какой метод машинного обучения целесообразно использовать при наличии данных в виде набора пар «объект - ответ» и необходимости получить прогноз ответа для нового объекта

- обучение с учителем
- обучение без учителя
- обучение с подкреплением
- стохастическое обучение

2 Для построения модели машинного обучения используется

- обучающий набор данных
- тестовый набор данных
- контрольный набор данных
- универсальный набор данных

3 При формировании обучающего и тестового наборов данных в пропорции 75 и 25 процентов из имеющегося набора данных что необходимо сделать?

- случайным образом перемешать данных примеры данных
- упорядочить наборы данных по возрастанию
- упорядочить наборы данных по убыванию
- ничего не делать с набором данных

4 Как в машинном обучении называют объект данных

- sample
- data point
- features
- label

5 Какая библиотека Python используется для обработки и анализа данных

- pandas
- NumPy
- SciPy
- IPython

6 Какой алгоритм машинного обучения без учителя позволяет выделить полезные характеристики

- алгоритм факторизации неотрицательных матриц
- алгоритм анализа главных компонент
- алгоритм стохастического вложения соседей
- алгоритм k-ближайших соседей

7 Какой метод агломеративной кластеризации выбирает и объединяет два кластера так, чтобы прирост дисперсии внутри кластеров был минимальным

метод Варда
метод средней связи
метод полной связи
метод максимальной связи

8 Какой метод агломеративной кластеризации объединяет два кластера, которые имеют наименьшее среднее значение всех расстояний, измеренных между точками двух кластеров

метод средней связи
метод Варда
метод полной связи
метод максимальной связи

9 Какой метод масштабирования пакета scikit-learn гарантирует, что для каждого признака среднее будет равно 0, а дисперсия будет равна 1

StandardScaler
RobustScaler
MinMaxScaler
Normalizer

10 Какой метод масштабирования пакета scikit-learn гарантирует, что вектор признаков имеет евклидову длину 1

Normalizer
RobustScaler
StandardScaler
MinMaxScaler

11 Укажите основное отличие между алгоритмами градиентного спуска RMSprop и Adadelta при изменении шага обучения использует корень из среднего от квадратов от градиентов при изменении шага обучения учитывается экспоненциальное среднее квадратов градиентов при изменении шага обучения учитывается сумма квадратов градиентов при изменении шага обучения учитывается логарифм суммы квадратов градиентов

12 Как называется функция, которая используется в нейронных сетях для введения нелинейности в преобразовании сигналов

функция активации
функция нормализации
функция регуляризации
функция стагнации

13 Если на вход операции субдискретизации подается матрица признаков размером (6 x 6), размер окна дискретизации 3 и шаг дискретизации 3, то размер матрицы после операции будет

(2 x 2)
(3 x 3)
(4 x 4)
(5 x 5)

14 Укажите НЕ правильное утверждение в отношении рекуррентных нейронных сетей LSTM свободный член забывающего фильтра инициализируются случайными числами близкими к нулю

архитектура сети позволяет решать проблему исчезающих градиентов
архитектура сети не защищает от неограниченного роста градиентов
веса сети инициализируются маленькими случайными числами

15 Какой фильтр добавлен в фазированной LSTM по сравнению в обычной LSTM

фильтр времени
забывающий фильтр

входной фильтр
выходной фильтр

Критерии оценивания:

Из имеющегося банка тестов формируется вариант, содержащий 10 вопросов для одного обучающегося.

17-20 б. – тест пройден на 85-100 %;

7-16 б. – тест пройден на 35-84 %;

0-6 б. – тест пройден на менее, чем 35 %.

Максимальное количество баллов за тест – 20.

Лабораторные задания

Лабораторное задание 1

Тема 1. Анализ данных климата с записными книжками.

Записные книжки Azure – это облачная платформа для создания и выполнения записных книжек Jupyter. Jupyter – это среда на основе IPython, облегчающая интерактивное программирование и анализ данных с использованием различных языков программирования, включая Python. Записные книжки Jupyter широко используются в научной среде для математического моделирования, машинного обучения, статистического анализа, а также для обучения написанию кода. Цель выполнения задания: создание записной книжки в записных книжках Azure; отправка, обработка и визуализация данных в записной книжке; предоставление совместного доступа к записным книжкам в Интернете.

Лабораторное задание 2

Тема 2. Прогнозирование задержки рейсов путем создания модели машинного обучения.

Python является одним из наиболее популярных языков программирования в мире. Он широко используется в сообществе специалистов по обработке и анализу данных для машинного обучения и статистического анализа. Одна из причин его популярности – наличие тысяч библиотек с открытым кодом, таких как NumPy, Pandas, Matplotlib и scikit-learn, которые позволяют программистам и аналитикам изучать, преобразовывать, анализировать и визуализировать данные. Цель выполнения задания: создание записной книжки Jupyter в записных книжках Azure, импорт данных и просмотр данных, загруженных в записную книжку; использование библиотеки Pandas для очистки и подготовки данных для модели машинного обучения; использование библиотеки scikit-learn для создания модели машинного обучения; использование библиотеки Matplotlib для визуализации эффективности модели.

Лабораторное задание 3

Тема 3. Анализ тональности отзывов с Keras.

Keras – это API нейронных сетей высокого уровня, написанный на языке Python, который выполняется в дополнение к другим инструментам глубокого обучения, таким как TensorFlow. Keras используется для сборки нейронной сети, которая оценивает текст, например, отзывы пользователей, для определения тональности. Цель выполнения задания: создание записной книжки Jupyter в записных книжках Azure, использование Keras для построения и обучения нейронной сети, которая будет выполнять анализ тональности, использование нейронных сетей для анализа текста для определения тональности.

Лабораторное задание 4

Тема 4. Оценка неопределенности классификаторов.

Использование интерфейса scikit-learn дает возможность вычислить оценки неопределенности прогнозов. Часто вас интересует не только класс, спрогнозированный моделью для определенной точки тестового набора, но и степень уверенности модели в правильности прогноза. В реальной практике различные виды ошибок приводят к очень разным результатам. Цель выполнения задания: провести медицинский тест для определения рака с использованием двух различных функции, с помощью которых можно оценить неопределенность прогнозов: `decision_function` и `predict_proba`.

Критерии оценивания (для каждого задания):

16-20 б. – задание выполнено верно;

10-15 б. – при выполнении задания были допущены неточности, не влияющие на результат;

6-9 б. – при выполнении задания были допущены ошибки;

0-5 б. – при выполнении задания были допущены существенные ошибки.

Максимальное количество баллов за лабораторные задания – 80 (4 задания по 20 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 экзамена.

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 и практическим занятиям.

В ходе лабораторных 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 студент могут воспользоваться консультациями преподавателя.

Вопросы, не рассмотренные на лекциях, лабораторных занятиях, должны быть изучены студентами в ходе самостоятельной работы. Контроль самостоятельной работы студентов над учебной программой курса осуществляется в ходе занятий методом теста 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.