

Минобрнауки России

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
(ФГБОУ ВО «ВГУ»)**

УТВЕРЖДАЮ



Заведующий кафедрой
Сирота Александр Анатольевич

Кафедра технологий обработки и защиты информации

29.06.2021

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

Б1.В.ДВ.02.01 Разработка приложений для систем машинного обучения

1. Код и наименование направления подготовки/специальности:

09.04.02 Информационные системы и технологии

2. Профиль подготовки/специализация:

Системы прикладного искусственного интеллекта

3. Квалификация (степень) выпускника:

Магистратура

4. Форма обучения:

Очная

5. Кафедра, отвечающая за реализацию дисциплины:

Кафедра технологий обработки и защиты информации

6. Составители программы:

Дрюченко Михаил Анатольевич, к.т.н., доцент

7. Рекомендована:

протокол № 5 от 10.03.21

8. Учебный год:

2022-2023

9. Цели и задачи учебной дисциплины:

изучение теоретических основ машинного обучения, вопросов практической реализации и применения алгоритмов машинного обучения при решении реальных задач, связанных и с распознаванием, классификацией и обработкой данных; получение профессиональных компетенций в области современных технологий машинного обучения.

Основные задачи дисциплины:

- знакомство студентов с теоретическими аспектами машинного обучения;
- изучение основных алгоритмов машинного обучения и особенностей их применения при разработке автоматизированных алгоритмов обработки данных;

овладение практическими навыками применения алгоритмов машинного обучения при разработке приложений для распознавания образов, классификации данных и прогнозирования трендов.

10. Место учебной дисциплины в структуре ООП:

учебная дисциплина относится к блоку обязательных дисциплин вариативной части.

Входные знания в области информатики, математического анализа, теории вероятности, математической статистики, технологий обработки информации, обработки изображений.

11. Планируемые результаты обучения по дисциплине/модулю (знания, умения, навыки), соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями выпускников) и индикаторами их достижения:

Код и название компетенции	Код и название индикатора компетенции	Знания, умения, навыки
ПКВ-8 Способен разрабатывать новые технологии проектирования информационных систем, прогнозировать развитие информационных систем и технологий	ПКВ-8.1 Знает инструменты и методы моделирования бизнес-процессов, современные подходы и стандарты автоматизации организации, отраслевую документацию, основы реинжиниринга бизнес-процессов организации	<p>Знать: основные модели и алгоритмы машинного обучения, специфику их использования.</p> <p>Уметь: применять на практике методы машинного обучения для задач моделирования, анализа данных, прогнозирования, в том числе при моделировании бизнес-процессов.</p> <p>Владеть: практическими навыками применения методов машинного обучения в задачах моделирования, анализа, прогнозирования данных и бизнес-процессов.</p>

Код и название компетенции	Код и название индикатора компетенции	Знания, умения, навыки
ПКВ-10 Способен определять варианты структур программного обеспечения информационных систем (программного средства), необходимые информационные потоки и исследовать варианты структур с использованием моделей различного уровня	ПКВ-10.1 Умеет проводить анализ внешнесистемных требований, возможностей их реализации, определяет концептуальный и функциональный облик системы (программного средства), выявление и анализ известных аналогов	<p>Знать: современные методы и средства разработки приложений на языках высокого уровня, известные библиотеки машинного обучения.</p> <p>Уметь: проводить анализ методологических и технологических требований к проектам, использующим методы машинного обучения, оценивать возможности реализации и выявлять возможные риски, формировать функциональный облик программного обеспечения.</p> <p>Владеть: практическими навыками анализа требований с учетом специфики проектов, навыками формирования функционального облика разрабатываемого программного обеспечения.</p>

12. Объем дисциплины в зачетных единицах/час:

3/108

Форма промежуточной аттестации:

Зачет

13. Трудоемкость по видам учебной работы

Вид учебной работы	Семестр 4	Всего
Аудиторные занятия	56	56
Лекционные занятия	28	28
Практические занятия	14	0
Лабораторные занятия	28	28
Самостоятельная работа	52	52
Курсовая работа		0
Промежуточная аттестация	0	0
Часы на контроль		0
Всего	108	108

13.1. Содержание дисциплины

п/п	Наименование раздела дисциплины	Содержание раздела дисциплины	Реализация раздела дисциплины с помощью онлайн-курса, ЭУМК
1. Лекции			
1.1	Нет		
2. Практические занятия			
2.1	Введение в машинное обучение	Обзор предметной области, основных понятий машинного обучения, моделей и алгоритмов. Примеры решаемых прикладных задач (классификации, регрессии, прогнозирования).	Создан онлан электронный курс, размещены материалы к лабораторным и практическим занятиям.
2.2	Обзор известных библиотек и фреймворков для машинного обучения	Пакеты для машинного обучения в Matlab. Фреймворки и библиотеки машинного обучения, использующие Python. Фреймворки и библиотеки машинного обучения, использующие C/C++/C#. Фреймворки и библиотеки машинного обучения, использующие Java. Вопросы взаимодействия библиотек машинного обучения с разрабатываемыми программными средствами. Импорт/экспорт обученных моделей. Вопросы производительности обучения, использования GPU и SIMD расширений CPU.	Создан онлан электронный курс, размещены материалы к лабораторным и практическим занятиям.

п/п	Наименование раздела дисциплины	Содержание раздела дисциплины	Реализация раздела дисциплины с помощью онлайн-курса, ЭУМК
2.3	Разработка приложений для классификации, распознавания и обработки данных с использованием нейронных сетей, решающих деревьев, случайного леса и метода опорных векторов	<p>Принципы работы решающих деревьев, случайного леса, метода опорных векторов, однослойных и многослойных нейронных сетей прямого распространения.</p> <p>Обзор механизмов регуляризации, критериев качества классификации, скользящего контроля, ROC, AUC, точности и полноты.</p> <p>Программная реализация многослойных нейронных сетей, алгоритмов обучения RProp, QProp, Левенберга-Марквардта.</p> <p>Компьютерное зрение. Алгоритмы обработки цифровых изображений в задачах распознавания и классификации мультимедиа информации. Обзор возможностей библиотеки OpenCV.</p> <p>Создание автоматизированных алгоритмов обработки данных (на основе методов машинного обучения) для решения прикладных задач.</p>	Создан онлан электронный курс, размещены материалы к лабораторным и практическим занятиям.
2.4	Разработка приложений для решения задач обработки данных с использованием методов глубокого обучения	<p>Обзор технологии глубокого обучения, сверточных и рекуррентных нейронных сетей.</p> <p>Проблема разметки исходных данных. Аугментация данных.</p> <p>Создание автоматизированных алгоритмов обработки данных (на основе глубоких нейронных сетей) для решения прикладных задач.</p>	Создан онлан электронный курс, размещены материалы к лабораторным и практическим занятиям.
3. Лабораторные работы			

п/п	Наименование раздела дисциплины	Содержание раздела дисциплины	Реализация раздела дисциплины с помощью онлайн-курса, ЭУМК
3.1	Изучение возможностей известных библиотек и фреймворков для машинного обучения	1. Изучение пакетов для машинного обучения в Matlab, фреймворков и библиотек машинного обучения, использующих Python, C/C++/C#, Java. Изучение вопросов подключения и использования библиотек машинного обучения в пользовательских проектах.	Создан онлан электронный курс, размещены материалы к лабораторным и практическим занятиям.
3.2	Нейронные сети	2. Изучение принципов работы искусственных нейронных сетей. Реализация нейронной сети прямого распространения и ее использование для решения задач классификации. 3. Использование различных библиотек машинного обучения и компьютерного зрения для решения задачи обработки и распознавания мультимедиа информации. Изучение вопросов выделения, анализа и отбора признаковых описателей.	Создан онлан электронный курс, размещены материалы к лабораторным и практическим занятиям.
3.2	Глубокое обучение (Deep learning)	4. Практическое изучение технологии глубокого обучения, использование сверточных нейронных сетей из библиотеки TensorFlow (Keras) для задачи анализа объектов на изображениях.	Создан онлан электронный курс, размещены материалы к лабораторным и практическим занятиям.

13.2. Темы (разделы) дисциплины и виды занятий

№ п/п	Наименование темы (раздела)	Лекционные занятия	Практические занятия	Лабораторные занятия	Самостоятельная работа	Всего
1	Введение в машинное обучение	6	0	0	8	14

№ п/п	Наименование темы (раздела)	Лекционные занятия	Практические занятия	Лабораторные занятия	Самостоятельная работа	Всего
2	Обзор известных библиотек и фреймворков для машинного обучения	6	0	4	12	22
3	Разработка приложений для классификации, распознавания и обработки данных с использованием нейронных сетей, решающих деревьев, случайного леса и метода опорных векторов	10	0	14	16	40
4	Разработка приложений для решения задач обработки данных с использованием методов глубокого обучения	6	0	10	16	32
		28	0	28	52	108

14. Методические указания для обучающихся по освоению дисциплины

1) При изучении дисциплины рекомендуется использовать следующие средства:

- рекомендуемую основную и дополнительную литературу;
- методические указания и пособия;
- контрольные задания для закрепления теоретического материала;
- электронные версии учебников и методических указаний для выполнения лабораторно - практических работ (при необходимости материалы рассылаются по электронной почте).

2) Для максимального усвоения дисциплины рекомендуется проведение письменного опроса

(тестирование, решение задач) студентов по материалам лекций и практических работ. Подборка вопросов для тестирования осуществляется на основе изученного теоретического материала. Такой подход позволяет повысить мотивацию студентов при конспектировании лекционного материала.

3) При проведении практических работ обеспечивается максимальная степень соответствия с материалом лекционных занятий и осуществляется экспериментальная проверка методов, алгоритмов и технологий обработки информации, излагаемых в рамках лекций.

4) При использовании дистанционных образовательных технологий и электронного обучения выполнять все указания преподавателей, вовремя подключаться к online занятиям, ответственно подходить к заданиям для самостоятельной работы.

15. Перечень основной и дополнительной литературы, ресурсов интернет, необходимых для освоения дисциплины

№ п/п	Источник
1	<i>Сирота, Александр Анатольевич. Методы и алгоритмы анализа данных и их моделирование в MATLAB : [учебное пособие] / А.А. Сирота .— Санкт-Петербург : БХВ-Петербург, 2016 .— 381 с. : ил. — Библиогр.: с. 371-374 .— Предм. указ.: с. 377-381 .— ISBN 978-5-9775-3778-0.</i>
2	Практикум по курсу " Нейросетевые технологии обработки информации" [Электронный ресурс] : учебно-методическое пособие для вузов : [для магистров фак. компьютер. наук днев. формы обучения; для направления 09.04.02 - Информ. системы и технологии. / Воронеж. гос. ун-т ; сост.: А.А. Сирота, Е.Ю. Митрофанова , М.А. Дрюченко .— Электрон. текстовые дан. — Воронеж : Издательско-полиграфический центр Воронежского государственного университета, 2013

б) дополнительная литература:

№ п/п	Источник
1	Осовский С. Нейронные сети для обработки информации / С. Осовский. - М. : Горячая Линия - Телеком, 2017. - 448.
2	Анализ данных и процессов / А. Барсегян [и др.]. - СПб. : БХВ-ПИТЕР, 2009. - 512 с.
3	Алгазинов, Эдуарт Константинович. Анализ и компьютерное моделирование информационных процессов и систем : [учебное пособие для студ. вузов, обуч. по специальности 080801 "Приклад. информатика" и др. междисциплинар. специальностям] / Э.К. Алгазинов, А.А. Сирота ; под общ. ред. А.А. Сироты .— М. : Диалог-МИФИ, 2009 .— 416 с. : ил .— Библиогр. в конце разд. — ISBN 978-5-86404-233-5
4	Осинга Д. Глубокое обучение. Готовые решения / Д. Осинга. - М. : Диалектика, 2019. - 288 с.
5	Элбон К. Машинное обучение с использованием Python. Сборник рецептов / К. Элбон. - БХВ-Петербург, 2019. - 384.

в) информационные электронно-образовательные ресурсы:

№ п/п	Источник
1	Электронный каталог Научной библиотеки Воронежского государственного университета. – (http // www.lib.vsu.ru/).
2	Образовательный портал «Электронный университет ВГУ».- (https://edu.vsu.ru/)
3	ЭБС «Издательства «Лань», Договор №3010-06/71-14 от 25.11.2014, ЭБС «Университетская библиотека online», Договор №3010-06/70-14 от 25.11.14, Национальный цифровой ресурс «РУКОНТ», Договор №ДС-208 от 01.02.2012

16. Перечень учебно-методического обеспечения для самостоятельной работы

№ п/п	Источник
1	Практикум по курсу " Нейросетевые технологии обработки информации" [Электронный ресурс] : учебно-методическое пособие для вузов : [для магистров фак. компьютер. наук днев. формы обучения; для направления 09.04.02 - Информ. системы и технологии. / Воронеж. гос. ун-т ; сост.: А.А. Сирота, Е.Ю. Митрофанова , М.А. Дрюченко .— Электрон. текстовые дан. — Воронеж : Издательско-полиграфический центр Воронежского государственного университета, 2013

17. Образовательные технологии, используемые при реализации учебной дисциплины, включая дистанционные образовательные технологии (ДОТ), электронное обучение (ЭО), смешанное обучение):

Для реализации учебного процесса используются:

1) ПО Microsoft в рамках подписок «Imagine», ежегодные сублицензионные договоры № 56035/ВРН3739 и № 56036/ВРН3739 от 07.10.2016.

2) ПО MATLAB Classroom ver. 7.0, 10 конкурентных бессрочных лицензий на каждый, компоненты: Matlab, Simulink, Stateflow, 1 тулбокс, N 21127/VRN3 от 30.09.2011 (за счет проекта ЕК TEMPUS/ERAMIS).

3) ПО Матлаб в рамках подписки "Университетская лицензия на программный комплекс для ЭВМ - MathWorks, Headcount – 25 ": лицензия до 31.01.2022, сублицензионный контракт 3010-07/01-19 от 09.01.19.

4) При проведении занятий в дистанционном режиме обучения используются технические и информационные ресурсы Образовательного портала "Электронный университет ВГУ (<https://edu.vsu.ru/>), базирующегося на системе дистанционного обучения Moodle, развернутой в университете.

18. Материально-техническое обеспечение дисциплины:

1) Мультимедийная лекционная аудитория (корп.1а, ауд. № 380), ПК-Intel-G3420, рабочее место преподавателя: проектор, видеокоммутатор, специализированная мебель: доска меловая 1 шт., столы 31 шт., стулья 64 шт.; выход в Интернет, доступ к фондам учебно-методической документации и электронным изданиям.

2) Компьютерный класс (один из корп. 1а, ауд. № 291, 293, 295, 387, 381), ПК-Intel-Core2/i3 14 шт., специализированная мебель: доска маркерная 1 шт., столы 14 шт., стулья 28 шт.; доступ к фондам

учебно-методической документации и электронным изданиям, доступ к электронным библиотечным системам, выход в Интернет.

19. Оценочные средства для проведения текущей и промежуточной аттестаций

Порядок оценки освоения обучающимися учебного материала определяется содержанием следующих разделов дисциплины:

№ п/п	Разделы дисциплины (модули)	Код компетенции	Код индикатора	Оценочные средства для текущей аттестации
1	Разделы 1-4 Введение в машинное обучение. Обзор известных библиотек и фреймворков для машинного обучения. Разработка приложений для классификации, распознавания и обработки данных с использованием нейронных сетей, решающих деревьев, случайного леса и метода опорных векторов. Разработка приложений для решения задач обработки данных с использованием методов глубокого обучения.	ПКВ-8	ПКВ-8.1	Устный опрос. Лабораторные работы 1-4
2	Разделы 1-4 Введение в машинное обучение. Обзор известных библиотек и фреймворков для машинного обучения. Разработка приложений для классификации, распознавания и обработки данных с использованием нейронных сетей, решающих деревьев, случайного леса и метода опорных векторов. Разработка приложений для решения задач обработки данных с использованием методов глубокого обучения.	ПКВ-10	ПКВ-10.1	Устный опрос. Лабораторные работы 1-4

Промежуточная аттестация

Форма контроля - Зачет

Оценочные средства для промежуточной аттестации

Перечень вопросов, практическое задание

20 Типовые оценочные средства и методические материалы, определяющие процедуры оценивания

20.1 Текущий контроль успеваемости

Текущий контроль успеваемости по дисциплине осуществляется с помощью следующих оценочных средств:

Устный опрос на практических занятиях

Примерный перечень применяемых оценочных средств

№ п/п	Наименование оценочного средства	Представление оценочного средства в фонде	Критерии оценки
1	Устный опрос на практических занятиях	Вопросы по темам/разделам дисциплины	Правильный ответ – зачтено, неправильный или принципиально неточный ответ – не зачтено
2	Контрольная работа по разделам дисциплины	Теоретические вопросы по темам/разделам дисциплины	Шкала оценивания соответствует таблице, приведенной ниже
3	Лабораторная работа	Содержит 4 лабораторных задания, предусматривающих разработку и тестирование (на модельных и реальных данных) приложений на языках C++ и Python с использованием известных фреймворков и библиотеки машинного обучения	При успешном выполнении работ в течение семестра фиксируется возможность оценивания только теоретической части дисциплины в ходе промежуточной аттестации (экзамена), в противном случае проверка задания по лабораторным работам выносится на экзамен.

Пример задания для выполнения лабораторной работы

Лабораторная работа № 1

«Изучение фреймворков и библиотек для машинного обучения»

Цель работы

Изучить принципы работы и использования в пользовательских приложениях известных библиотек для машинного обучения.

Форма контроля

Опрос в устной форме по исходному коду и результатам работы реализованной программы.

Количество отведённых аудиторных часов - 4

Содержание работы

Получить у преподавателя вариант задания, написать код, реализующий соответствующий алгоритм обработки информации. Провести тестирование реализованного алгоритма. Проанализировать полученные результаты и сформулировать выводы по проделанной работе.

Пример варианта задания:

С использованием библиотеки FANN написать приложение для распознавания рукописных букв и цифр. Использовать нейронную сеть класса многослойный перцептрон.

Примеры контрольных вопросов:

1. На что влияет количество и размер скрытых слоев сети?
2. Какие активационные функции можно применять при обучении сети методом обратного распространения ошибки.

20.2 Промежуточная аттестация

Промежуточная аттестация может включать в себя проверку теоретических вопросов, а также, при необходимости (в случае невыполнения в течение семестра), проверку выполнения установленного перечня лабораторных заданий, позволяющих оценить уровень полученных знаний и/или

практическое (ие) задание(я), позволяющее (ие) оценить степень сформированности умений и навыков.

Для оценки теоретических знаний используется перечень контрольно-измерительных материалов. Каждый контрольно-измерительный материал для проведения промежуточной аттестации включает два задания - вопросов для контроля знаний, умений и владений в рамках оценки уровня сформированности компетенции. При оценивании используется количественная шкала. Критерии оценивания приведены в таблице, приведенной ниже.

Для оценивания результатов обучения на зачете используются следующие содержательные показатели (формулируется с учетом конкретных требований дисциплины):

1. знание теоретических основ учебного материала, основных определений, понятий и используемой терминологии;
2. умение связывать теорию с практикой, умение выявлять и анализировать основные закономерности, полученные, в том числе, в ходе выполнения практических заданий;
3. умение обосновывать свои суждения и профессиональную позицию по излагаемому вопросу;
4. владение навыками программирования в рамках выполняемых практических заданий;
5. владение навыками проведения компьютерного эксперимента, тестирования алгоритмов обработки информации.

Различные комбинации перечисленных показателей определяют критерии оценивания результатов обучения (сформированности компетенций) на государственном экзамене:

- высокий (углубленный) уровень сформированности компетенций;
- повышенный (продвинутый) уровень сформированности компетенций;
- пороговый (базовый) уровень сформированности компетенций.

Для оценивания результатов обучения на зачете используется - зачтено, не зачтено по результатам тестирования.

Соотношение показателей, критериев и шкалы оценивания результатов обучения на государственном экзамене представлено в следующей таблице.

Критерии оценивания компетенций и шкала оценок

Критерии оценивания компетенций	Уровень сформированности компетенций	Шкала оценок
Обучающийся демонстрирует полное соответствие знаний, умений, навыков по приведенным критериям свободно оперирует понятийным аппаратом и приобретенными знаниями, умениями, применяет их при решении практических задач. Успешно выполнены лабораторные работы в соответствии с установленным перечнем.	Повышенный уровень	Зачтено
Ответ на контрольно-измерительный материал не полностью соответствует одному из перечисленных выше показателей, но обучающийся дает правильные ответы на дополнительные вопросы. При этом обучающийся демонстрирует соответствие знаний, умений, навыков приведенным в таблицах показателям, но допускает незначительные ошибки, неточности, испытывает затруднения при решении практических задач. Успешно выполнены лабораторные работы в соответствии с установленным перечнем.	Базовый уровень	Зачтено

Обучающийся демонстрирует неполное соответствие знаний, умений, навыков приведенным в таблицах показателям, допускает значительные ошибки при решении практических задач. При этом ответ на контрольно-измерительный материал не соответствует любым двум из перечисленных показателей, обучающийся дает неполные ответы на дополнительные вопросы. Успешно выполнены лабораторные работы в соответствии с установленным перечнем.	Пороговый уровень	Зачтено
Ответ на контрольно-измерительный материал не соответствует любым трем из перечисленных показателей. Обучающийся демонстрирует отрывочные, фрагментарные знания, допускает грубые ошибки. Не выполнены лабораторные работы в соответствии с установленным перечнем.	-	Не зачтено

Примерный перечень вопросов к экзамену

№	Содержание
1	Задачи, решаемые алгоритмами машинного обучения. Методика анализа данных.
2	Классификация алгоритмов машинного обучения.
3	Постановка и методы решения задачи классификации.
4	Постановка и методы решения задачи кластеризации.
5	Задача анализа временных рядов.
6	Принципы работы метода опорных векторов.
7	Структура деревьев решений.
8	Алгоритм Random Forest.
9	Нейронные сети прямого распространения. Структура сети, обучение по алгоритму обратного распространения ошибки.
10	Извлечение и генерация признаков на примере задачи анализа цифровых изображений.
11	Извлечение и генерация признаков на примере задачи анализа текстов.
12	Алгоритмы отбора и преобразования признаков.
13	Функционалы качества (log loss, ROC, AUC, accuracy, precision, recall и т.д.).
14	Глубокое обучение.
15	Сверточные и рекуррентные нейронные сети.
16	Подготовка обучающих наборов данных. Аугментация.
17	Библиотеки и фреймворки для машинного обучения
18	Взаимодействие библиотек машинного обучения с разрабатываемыми программными средствами. Импорт/экспорт обученных моделей.

Пример контрольно-измерительного материала

УТВЕРЖДАЮ

Заведующий кафедрой технологий обработки и защиты информации

_____ А.А. Сирота

__._.2021

Направление подготовки / специальность 09.04.02 Информационные системы и технологии

Дисциплина Б1.В.ДВ.02.01 Разработка приложений для систем машинного обучения

Форма обучения Очное

Вид контроля Зачет

Вид аттестации Промежуточная

Контрольно-измерительный материал № 1

1. Классификация алгоритмов машинного обучения.
2. Алгоритмы отбора и преобразования признаков.

Преподаватель _____ М.А. Дрюченко