

Министерство науки и высшего образования Российской Федерации
Муромский институт (филиал)
федерального государственного бюджетного образовательного учреждения высшего образования
**«Владимирский государственный университет
имени Александра Григорьевича и Николая Григорьевича Столетовых»**
(МИ ВлГУ)

Кафедра *ФПМ*

«УТВЕРЖДАЮ»
Заместитель директора по УР
_____ Д.Е. Андрианов
_____ 23.05.2023

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Технологии машинного обучения

Направление подготовки

*01.03.02 Прикладная математика и
информатика*

Профиль подготовки

Интеллектуальный анализ данных

Семестр	Трудоем- кость, час./зач. ед.	Лек- ции, час.	Практи- ческие занятия, час.	Лабора- торные работы, час.	Консультация, час.	Конт- роль, час.	Всего (контакт- ная работа), час.	СРС, час.	Форма промежу- точного контроля (экз., зач., зач. с оц.)
6	180 / 5	16	16		1,6	0,25	33,85	146,15	Зач. с оц.
7	108 / 3	16	16		3,6	0,35	35,95	36,4	Экз.(35,65)
Итого	288 / 8	32	32		5,2	0,6	69,8	182,55	35,65

Муром, 2023 г.

1. Цель освоения дисциплины

Цель дисциплины: приобретение студентами знаний в области применения машинного обучения для интеллектуального анализа данных, а также формирование навыков разработки программных средств на основе машинного обучения в различных предметных областях с учётом актуальных прикладных задач.

Задачами дисциплины являются:

- ознакомление студентов с понятием, целями и задачами машинного обучения;
- изучение базовых алгоритмов, применяемых в машинном обучении;
- овладение методами решения практических задач с использованием технологий машинного обучения.

2. Место дисциплины в структуре ОПОП ВО

Курс базируется на знаниях компьютерных технологий, основ программирования и знаниях, приобретенных при изучении дисциплин «Информатика», «Математика», «Специальные главы математики». На данном курсе базируется написание выпускной квалификационной работы.

3. Планируемые результаты обучения по дисциплине

Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения ОПОП (компетенциями и индикаторами достижения компетенций)

Формируемые компетенции (код, содержание компетенции)	Планируемые результаты обучения по дисциплине, в соответствии с индикатором достижения компетенции		Наименование оценочного средства
	Индикатор достижения компетенции	Результаты обучения по дисциплине	
ПК-3 Способен выполнять разработку программных модулей интеллектуального анализа данных	ПК-3.1 Программирует модули интеллектуального анализа и обработки данных	Владеет методами программирования модулей интеллектуального анализа и обработки данных (ПК-3.1)	вопросы к устному опросу, тест

4. Структура и содержание дисциплины

Общая трудоемкость дисциплины составляет 8 зачетных единиц, 288 часов.

4.1. Форма обучения: очная

Уровень базового образования: среднее общее.

Срок обучения 4г.

4.1.1. Структура дисциплины

№ п\п	Раздел (тема) дисциплины	Семестр	Контактная работа обучающихся с педагогическим работником							Самостоятельная работа	Форма текущего контроля успеваемости (по неделям семестра), форма промежуточной аттестации(по семестрам)
			Лекции	Практические занятия	Лабораторные работы	Контрольные работы	КП / КР	Консультация	Контроль		
1	Технологии машинного обучения	6	16	16						146,15	устный опрос, тестирование
Всего за семестр		180	16	16				1,6	0,25	146,15	Зач. с оц.
2	Технологии машинного обучения	7	16	16						36,4	устный опрос, тестирование
Всего за семестр		108	16	16				3,6	0,35	36,4	Экз.(35,65)
Итого		288	32	32				5,2	0,6	182,55	35,65

4.1.2. Содержание дисциплины

4.1.2.1. Перечень лекций

Семестр 6

Раздел 1. Технологии машинного обучения

Лекция 1.

Введение в машинное обучение (2 часа).

Лекция 2.

Основные алгоритмы машинного обучения (2 часа).

Лекция 3.

Линейная регрессия (2 часа).

Лекция 4.

Метрики регрессии (2 часа).

Лекция 5.

Базовые задачи регрессии в машинном обучении (2 часа).

Лекция 6.

Признаки в машинном обучении (2 часа).

Лекция 7.

Классификация в машинном обучении (2 часа).

Лекция 8.

Базовые задачи классификации в машинном обучении (2 часа).

Семестр 7

Раздел 2. Технологии машинного обучения

Лекция 9.

Метод kNN (2 часа).

Лекция 10.

Метрики классификации (2 часа).

Лекция 11.

Метод опорных векторов (2 часа).

Лекция 12.

Вероятностные модели (2 часа).

Лекция 13.

Метод Байеса (2 часа).

Лекция 14.

Решающие деревья (2 часа).

Лекция 15.

Кластеризация (2 часа).

Лекция 16.

Базовые задачи классификации в машинном обучении (2 часа).

4.1.2.2. Перечень практических занятий

Семестр 6

Раздел 1. Технологии машинного обучения

Практическое занятие 1

Анализ данных с использованием Python (2 часа).

Практическое занятие 2

Использование библиотеки Pandas для анализа данных (2 часа).

Практическое занятие 3

Метод линейной регрессии (2 часа).

Практическое занятие 4

Реализация методов линейной регрессии на языке Python (2 часа).

Практическое занятие 5

Реализация методов нелинейной регрессии на языке Python (2 часа).

Практическое занятие 6

Реализация метода классификации kNN на языке Python (2 часа).

Практическое занятие 7

Метод опорных векторов (2 часа).

Практическое занятие 8

Реализация метода опорных векторов на языке Python (2 часа).

Семестр 7

Раздел 2. Технологии машинного обучения

Практическое занятие 9

Задачи кластеризации данных (2 часа).

Практическое занятие 10

Решение задач кластеризации с использованием Python (2 часа).

Практическое занятие 11

Наивный классификатор Байеса (2 часа).

Практическое занятие 12

Реализация наивного классификатора Байеса на языке Python (2 часа).

Практическое занятие 13

Метод решающих деревьев (2 часа).

Практическое занятие 14

Реализация метода решающих деревьев на языке Python (2 часа).

Практическое занятие 15

Дискриминантный анализ (2 часа).

Практическое занятие 16

Реализация дискриминантного анализа на языке Python (2 часа).

4.1.2.3. Перечень лабораторных работ

Не планируется.

4.1.2.4. Перечень тем и учебно-методическое обеспечение самостоятельной работы

Перечень тем, вынесенных на самостоятельное изучение:

1. Применение машинного обучения в промышленности.
2. Языковые модели в машинном обучении.
3. Распознавание образов.
4. Кластеризация в практическом применении.
5. Дискриминантные методы классификации.
6. Графовые нейронные сети.
7. Сверточные нейронные сети.
8. Обучение с подкреплением.

Для самостоятельной работы используются методические указания по освоению дисциплины и издания из списка приведенной ниже основной и дополнительной литературы.

4.1.2.5. Перечень тем контрольных работ, рефератов, ТР, РГР, РПР

Не планируется.

4.1.2.6. Примерный перечень тем курсовых работ (проектов)

Не планируется.

5. Образовательные технологии

В процессе изучения дисциплины применяется контактная технология преподавания (за исключением самостоятельной работы студентов). При проведении практических работ применяется имитационный подход с совместным с преподавателем разбором проблемных ситуаций на конкретных примерах, типовые примеры решения задач демонстрируются при помощи мультимедийной техники. Затем студенты самостоятельно решают аналогичные задания.

6. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины.

Фонды оценочных материалов (средств) приведены в приложении.

7. Учебно-методическое и информационное обеспечение дисциплины.

7.1. Основная учебно-методическая литература по дисциплине

1. Ракитский А.А. Методы машинного обучения : учебно-методическое пособие / Ракитский А.А.. — Новосибирск : Сибирский государственный университет телекоммуникаций и информатики, 2018. — 32 с. - <https://www.iprbookshop.ru/90591.html>
2. Теория и практика машинного обучения : учебное пособие / В.В. Воронина [и др.].. — Ульяновск : Ульяновский государственный технический университет, 2017. — 291 с. — ISBN 978-5-9795-1712-4. - <https://www.iprbookshop.ru/106120.html>

3. Неделько В.М. Основы статистических методов машинного обучения : учебное пособие / Неделько В.М.. — Новосибирск : Новосибирский государственный технический университет, 2010. — 72 с. — ISBN 978-5-7782-1385-2. - <https://www.iprbookshop.ru/45418.html>

7.2. Дополнительная учебно-методическая литература по дисциплине

1. Барский, А. Б. Введение в нейронные сети : учебное пособие / А. Б. Барский. — 3-е изд. — Москва, Саратов : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. — 357 с. — ISBN 978-5-4497-0309-5. - <https://www.iprbookshop.ru/89426.html>

2. Теория и практика машинного обучения : учебное пособие / В. В. Воронина, А. В. Михеев, Н. Г. Ярушкина, К. В. Святлов. — Ульяновск : Ульяновский государственный технический университет, 2017. — 291 с. — ISBN 978-5-9795-1712-4. - <https://www.iprbookshop.ru/106120.html>

7.3. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень программного обеспечения и информационных справочных систем

В образовательном процессе используются информационные технологии, реализованные на основе информационно-образовательного портала института (www.mivlgu.ru/iop), и инфокоммуникационной сети института:

- предоставление учебно-методических материалов в электронном виде;
- взаимодействие участников образовательного процесса через локальную сеть института и Интернет;
- предоставление сведений о результатах учебной деятельности в электронном личном кабинете обучающегося.

Информационные справочные системы:

Информационно-образовательный портал МИ ВлГУ <https://www.mivlgu.ru/iop/>

Образовательный сайт - Численные методы http://service.rintd.ru/chislennyye_metody

Портал знаний <http://statistica.ru/branches-maths/obzor-chislennykh-metodov/>

Образовательный математический сайт - <http://www.exponenta.ru/soft/Mathcad/UsersGuide/0.asp>.

Математический форум Math Help Planet <http://mathhelpplanet.com/viewforum.php?f=22>.

Образовательный портал "Единое окно доступа к образовательным ресурсам" - <http://window.edu.ru/>

Программное обеспечение:

LibreOffice (Mozilla Public License v2.0)

7-Zip (GNU LGPL)

Mathcad Education – University Edition (100 pack) v.15 (Государственный контракт №1 от 10.01.2012 года)

Google Chrome (Лицензионное соглашение Google)

Mozilla Firefox (MPL)

Free Commander XE (Лицензионное соглашение FreeCommander)

Pot Player (Daum PotPlayer EULA)

Microsoft Windows 10 Professional (Программа Microsoft Azure Dev Tools for Teaching (Order Number: IM126433))

Oracle VirtualBox (GNU GPL)

Microsoft Visual Studio (Программа Microsoft Azure Dev Tools for Teaching (Order Number: IM126433))

GIMP (GNU GPL 3.0)

INKSCAPE (GNU GPL)

FireBird (Initial Developer's Public License и InterBase Public Licence)

Microsoft SQL Server (Программа Microsoft Azure Dev Tools for Teaching (Order Number: IM126433))

K-Lite Mega Codec Pack (Freeware)
Adobe Acrobat Reader DC (Общие условия использования продуктов Adobe)
Unity (свободная (ограниченная версия))
Open Office (Apache License 2.0)
Yandex (EULA)
Python 3.9.4 (Python Software Foundation License)

7.4. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

iprbookshop.ru
service.rintd.ru
statistica.ru
exponenta.ru
mathhelpplanet.com
mivlgu.ru/iop

8. Материально-техническое обеспечение дисциплины

Кабинет информатики, технологий и методов программирования
Персональный компьютер - 12 шт.; коммутатор TRENDnet TEG-S24G; видеопроектор SANYO PLC-XU355; экран Lumien Master Picture LMP-100109. Доступ к сети Интернет

Компьютерный класс
Персональный компьютер - 12 шт.; коммутатор TRENDnet TEG-S24G; видеопроектор SANYO PLC-XU355; экран Lumien Master Picture LMP-100109. Доступ к сети Интернет

Помещение для самостоятельно работы обучающихся
Персональный компьютер - 12 шт.; коммутатор TRENDnet TEG-S24G; видеопроектор SANYO PLC-XU355; экран Lumien Master Picture LMP-100109. Доступ к сети Интернет

9. Методические указания по освоению дисциплины

Для успешного освоения теоретического материала обучающийся: знакомится со списком рекомендуемой основной и дополнительной литературы; уточняет у преподавателя, каким дополнительным пособиям следует отдать предпочтение; ведет конспект лекций и прорабатывает лекционный материал, пользуясь как конспектом, так и учебными пособиями.

На практических занятиях пройденный теоретический материал подкрепляется решением задач по основным темам дисциплины. Занятия проводятся в компьютерном классе, используя специальное программное обеспечение. Каждой подгруппе обучающихся преподаватель выдает задачу, связанную с разработкой и программной реализацией алгоритмов обработки информации. В конце занятия обучающие демонстрируют полученные результаты преподавателю и при необходимости делают работу над ошибками.

Самостоятельная работа оказывает важное влияние на формирование личности будущего специалиста, она планируется обучающимся самостоятельно. Каждый обучающийся самостоятельно определяет режим своей работы и меру труда, затрачиваемого на овладение учебным содержанием дисциплины. Он выполняет внеаудиторную работу и изучение разделов, выносимых на самостоятельную работу, по личному индивидуальному плану, в зависимости от его подготовки, времени и других условий.

Форма заключительного контроля при промежуточной аттестации – экзамен. Для проведения промежуточной аттестации по дисциплине разработаны фонд оценочных средств и балльно-рейтинговая система оценки учебной деятельности студентов. Оценка по дисциплине выставляется в информационной системе и носит интегрированный характер, учитывающий результаты оценивания участия студентов в аудиторных занятиях, качества и своевременности выполнения заданий в ходе изучения дисциплины и промежуточной аттестации.

Программа составлена в соответствии с требованиями ФГОС ВО по направлению *01.03.02 Прикладная математика и информатика* и профилю подготовки *Интеллектуальный анализ данных*

Рабочую программу составил *к.т.н. Макаров М.В.*_____

Программа рассмотрена и одобрена на заседании кафедры *ФПМ*

протокол № 19 от 26.04.2023 года.

Заведующий кафедрой *ФПМ* _____ *Орлов А.А.*

(Подпись)

Рабочая программа рассмотрена и одобрена на заседании учебно-методической комиссии факультета

протокол № 9 от 19.05.2023 года.

Председатель комиссии *ФИТР* _____ *Рыжкова М.Н.*

(Подпись)

(Ф.И.О.)

Фонд оценочных материалов (средств) по дисциплине
Технологии машинного обучения

1. Оценочные материалы для проведения текущего контроля успеваемости по дисциплине

Темы для устного опроса:

1. Основные понятия машинного обучения. Основные постановки задач. Примеры прикладных задач.
2. Линейные пространства. Векторы и матрицы. Линейная независимость. Обратная матрица.
3. Производная и градиент функции. Градиентный спуск. Выпуклые функции.
4. Случайные величины. Дискретные и непрерывные распределения. Примеры.
5. Оценивание параметров распределений, метод максимального правдоподобия. Бутстрэппинг.
6. Линейные методы классификации и регрессии: функционалы качества, методы настройки, особенности применения.
7. Метрики качества алгоритм регрессии и классификации.
8. Оценивание качества алгоритмов. Отложенная выборка, ее недостатки. Оценка полного скользящего контроля. Кросс-валидация. Leave-one-out.
9. Деревья решений. Методы построения деревьев. Их регуляризация.
10. Композиции алгоритмов. Разложение ошибки на смещение и разброс.
11. Случайный лес, его особенности.
12. Методы поиска выбросов в данных. Методы восстановления пропусков в данных.
- Работа с несбалансированными выборками.
13. Задача анализа потребительской корзины. Поддержка и достоверность. Частые, замкнутые и максимальные частые множества.
14. Задача кластеризации. Алгоритм K-Means. Оценки качества кластеризации.
15. Нейронные сети для анализа изображений.

Общее распределение баллов текущего контроля по видам учебных работ для студентов

Рейтинг-контроль 1	устный опрос из 3 вопросов	до 15 баллов
Рейтинг-контроль 2	устный опрос из 3 вопросов	до 15 баллов
Рейтинг-контроль 3	устный опрос из 3 вопросов	до 15 баллов
Посещение занятий студентом		0
Дополнительные баллы (бонусы)		0
Выполнение семестрового плана самостоятельной работы	устный опрос из 3 вопросов	до 15 баллов

2. Промежуточная аттестация по дисциплине

Перечень вопросов к экзамену / зачету / зачету с оценкой.

Перечень практических задач / заданий к экзамену / зачету / зачету с оценкой (при наличии)

Фонд оценочных средств расположен на ИОП МИ ВлГУ по адресу
<https://www.mivlgu.ru/iop/question/edit.php?courseid=3115&cat=36696%2C105911>

Методические материалы, характеризующие процедуры оценивания

На основе перечня тестовых вопросов программным комплексом информационно-образовательного портала МИ ВлГУ формируются в автоматическом режиме тестовые задания для студентов из 20 вопросов. Программный комплекс формирует индивидуальные задания для каждого зарегистрированного в системе студента и устанавливает время прохождения тестирования. Каждый ответ оценивается в 2 балла. Результатом тестирования является сумма баллов, которая складывается с индивидуальным семестровым рейтингом студента и определяет оценку за зачет или экзамен.

51 - 65 балла – «удовлетворительно»;

66 – 81 баллов – «хорошо»;

81 – 100 баллов – «отлично».

Максимальная сумма баллов, набираемая студентом по дисциплине равна 100.

Оценка в баллах	Оценка по шкале	Обоснование	<i>Уровень сформированности компетенций</i>
Более 80	«Отлично»	Содержание курса освоено полностью, без пробелов, необходимые практические навыки работы с освоенным материалом сформированы, все предусмотренные программой обучения учебные задания выполнены, качество их выполнения оценено числом баллов, близким к максимальному	<i>Высокий уровень</i>
66-80	«Хорошо»	Содержание курса освоено полностью, без пробелов, некоторые практические навыки работы с освоенным материалом сформированы недостаточно, все предусмотренные программой обучения учебные задания выполнены, качество выполнения ни одного из них не оценено минимальным числом баллов, некоторые виды заданий выполнены с ошибками	<i>Продвинутый уровень</i>

50-65	«Удовлетворительно»	Содержание курса освоено частично, но пробелы не носят существенного характера, необходимые практические навыки работы с освоенным материалом в основном сформированы, большинство предусмотренных программой обучения учебных заданий выполнено, некоторые из выполненных заданий, возможно, содержат ошибки	<i>Пороговый уровень</i>
Менее 50	«Неудовлетворительно»	Содержание курса не освоено, необходимые практические навыки работы не сформированы, выполненные учебные задания содержат грубые ошибки	<i>Компетенции не сформированы</i>

3. Задания в тестовой форме по дисциплине

Примеры заданий:

1. Что такое градиентный спуск?
 - Алгоритм постепенного обновления весов и пороговых смещений нейронной сети во время ее обучения
 - Алгоритм случайного выбора весов и пороговых смещений нейронной сети
 - Алгоритм выполнения предсказания с помощью нейронной сети
 - Алгоритм выбора объектов для обучения нейронной сети
2. Что происходит во время обучения нейронной сети?
 - Настраивается число слоев и число нейронов в каждом слое
 - Настраивается длина шага градиентного спуска
 - Настраивается входной объект
 - Настраиваются параметры слоев: веса и пороговые смещения
3. Вычислите размер после применения операции свёртки к картинке (100, 50, 1), размер фильтра (1,1), шаг 2, нет пэдинга. Ответ запишите без пробелов и скобок, разделяя числа запятыми, например 100,100,100.
4. Сколько весовых коэффициентов содержится в нейронной сети, структура слоев которой задана следующим кодом?


```
self.layer1 = nn.Linear(3, 24)
self.layer2 = nn.Linear(24, 24)
self.layer3 = nn.Linear(24, 5)
```

Полный перечень тестовых заданий с указанием правильных ответов, размещен в банке вопросов на информационно-образовательном портале института по ссылке <https://www.mivlgu.ru/iop/question/edit.php?courseid=3115&cat=36696%2C105911>

Оценка рассчитывается как процент правильно выполненных тестовых заданий из их общего числа.