

Документ подписан простой электронной подписью Информация о владельце: ФИО: Таскаев Сергей Валерьевич Должность: Ректор Дата подписания: 16.06.2025 16:07:04 Уникальный программный ключ: 04c19ed8bfb98f3b6cb77a486b9a8788b8322323	МИНОВЕРНАУКИ РОССИИ Федеральное государственное бюджетное образовательное учреждение высшего образования «Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)	Рабочая программа дисциплины "Машинное обучение и анализ данных" по направлению подготовки (специальности) 09.04.03 "Прикладная информатика" направленности (профилю) Управление решениями на основе данных ФГБОУ ВО «ЧелГУ»	стр. 1
--	--	--	--------

Рабочая программа дисциплины (модуля)*

Машинное обучение и анализ данных

Направление подготовки (специальность)

09.04.03 Прикладная информатика

Направленность (профиль)

Управление решениями на основе данных

Присваиваемая квалификация (степень)

магистр

Форма обучения

очная

Год(ы) набора

*Рабочая программа дисциплины (модуля) адаптирована для инклюзивного обучения инвалидов и лиц с ограниченными возможностями здоровья

Челябинск 2024 г.

09.04.03 Прикладная информатика, Управление решениями на основе данных, магистр, *Машинное обучение и анализ данных*, 2025, очная

Проректор по учебной работе утверждено 24.02.2025 А.А. Саламатов

Ученым советом института информационных технологий

Протокол заседания № 6 от 20.02.2025

Председатель Ученого совета
института информационных
технологий

согласовано

Ю. В. Петриченко

Заседанием кафедры информационных технологий и экономической информатики

Протокол заседания № 6 от 20.02.2025

И. о. заведующего кафедрой

согласовано

С.А. Скрипов

Автор (составитель)

А.В. Вохминцев

Структура рабочей программы соответствует приказу ректора ФГБОУ ВО «ЧелГУ» от «13» апреля 2021 г. № 247-1



Содержание

1. Цели освоения дисциплины
2. Место дисциплины в структуре ОПОП
3. Компетенции обучающегося, формируемые в результате освоения дисциплины (модуля)
4. Объем дисциплины (модуля)
5. Структура и содержание дисциплины (модуля)
6. Фонд оценочных средств
 - 6.1. Перечень видов оценочных средств
 - 6.2. Типовые контрольные задания и иные материалы для текущей аттестации
 - 6.3. Типовые контрольные вопросы и задания для промежуточной аттестации
 - 6.4. Критерии оценивания
7. Учебно-методическое и информационное обеспечение дисциплины (модуля)
 - 7.1. Рекомендуемая литература
 - 7.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"
 - 7.3. Перечень информационных технологий
8. Материально-техническое обеспечение дисциплины (модуля)
9. Методические указания для обучающихся по освоению дисциплины (модуля)
10. Специальные условия освоения дисциплины обучающимися с инвалидностью и ограниченными возможностями здоровья



1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Цель – сформировать у студентов навыки работы с данными и решения прикладных задач, дать представление об основных методах машинного обучения и видах задач, решаемых ими.

Задачи: 1. Ознакомить студентов с основными задачами машинного обучения. 2. Дать представление об основных методах машинного обучения, выборе модели для конкретной задачи, оценке качества модели и ее настройке. 3. Сформировать практические навыки решения задач машинного обучения, показать готовые реализации методов машинного обучения в современных библиотеках.

Результаты обучения по дисциплине направлены на достижение индикаторов:

ПК-3.1. Обладает знаниями методов и алгоритмов машинного обучения и интеллектуального анализа данных, средств и технологий сбора, обработки и представления данных

ПК-3.2. Демонстрирует умение выбора методов машинного обучения и технологий анализа данных, технологий бизнес-аналитики исходя из требований к решению прикладных задач

ПК-3.3. Имеет практический опыт разработки решений в прикладных задачах сбора, обработки и интеллектуального анализа данных

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

Цикл (раздел) ОПОП: Б1.В.01

2.1 Требования к предварительной подготовке обучающегося:

Для изучения дисциплины обучающийся должен владеть знаниями из дисциплин бакалавриата, связанными с информационными технологиями.

2.2 Дисциплины и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:

Знания, полученные в данной дисциплине, могут быть использованы для написания магистерской диссертации.

Выполнение и защита выпускной квалификационной работы

3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

ПК-3: Способен собирать, обрабатывать и интерпретировать данные, используя методы машинного обучения и интеллектуального анализа данных, технологии бизнес-аналитики

Знать:

ключевые термины в области искусственного интеллекта и интеллектуальных систем (ИС);
основные методики формализации и представления данных
методики и алгоритмы обработки, представления, анализа данных, представленных различными способами и с разной степенью формализации

Уметь:

извлекать знания из экспертов, текстов, а также других различных источников информации;
выбирать соответственно поставленной задаче и использовать изученные ранее программные способы обработки и хранения информации
собирать и анализировать статистические данные, являющиеся основой для построения интеллектуальной системы;

Владеть:

навыками использования специального программного обеспечения для решения практических задач
навыками разработки компонентов информационных систем

В результате освоения дисциплины обучающийся должен

3.1 Знать:

3.2 Уметь:

3.3 Владеть:



4. ОБЪЕМ ДИСЦИПЛИНЫ (МОДУЛЯ)

Общая трудоемкость	4 ЗЕТ
Часов по учебному плану : 144 в том числе : аудиторные занятия : 48 самостоятельная работа : 51 часов на контроль : 36 контактная работа: 57 ИКР: 9	Виды контроля в семестрах: экзамены 1

5. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Литература
	Раздел 1. Введение в машинное обучение. Задача линейной регрессии			
1.1	Вводная лекция. Примеры задач машинного обучения с учителем и без. Одномерная линейная регрессия и метод максимального правдоподобия. Функция потерь, метод градиентного спуска. Множественная линейная регрессия. Нормализация признаков. Построение нелинейных моделей. /Лек/	1	4	Л1.1 Л1.2 Л1.3Л2.1 Л2.2 Э1 Э2
1.2	Реализация метода одномерной линейной регрессии для решения задачи прогнозирования прибыли при открытии нового филиала сети ресторанов. Реализация метода множественной линейной регрессии для решения задачи предсказания цены на дом при известных значениях площади дома и числе комнат. Библиотека sklearn и ее методы решения задач линейной регрессии. /Пр/	1	6	Л1.1 Л1.2 Л1.3Л2.1 Л2.2 Э1 Э2
	Раздел 2. Задача классификации. Логистическая регрессия. Проблема переобучения. Регуляризация			
2.1	Задача бинарной классификации. Логистическая регрессия. Сигмоида и логлосс. Задача множественной классификации. Проблема переобучения. Регуляризация. Гребневая регрессия. Лассо. /Лек/	1	4	Л1.1 Л1.2 Л1.3Л2.1 Л2.2 Э1 Э2
2.2	Решение задач бинарной классификации методом логистической регрессии для случая линейно разделимых классов. Решение задач бинарной классификации методом логистической регрессии для случая линейно неразделимых классов. Использование логистической регрессии для решения задачи множественной классификации. Распознавание рукописных цифр от 0 до 9 /Пр/	1	8	Л1.1 Л1.2 Л1.3Л2.1 Л2.2 Э1 Э2
	Раздел 3. Метод kNN, деревья решений и ансамблевые методы.			
3.1	Метод ближайших соседей kNN и его модификации. Деревья решений. Ансамбли деревьев решений. /Лек/	1	2	Л1.1 Л1.2 Л1.3Л2.1 Л2.2 Э1 Э2
3.2	Классификация ирисов Фишера с помощью метода kNN. Использование ансамблевых моделей на основе деревьев решений. /Пр/	1	4	Л1.1 Л1.2 Л1.3Л2.1 Л2.2 Э1 Э2
	Раздел 4. Метод опорных векторов			



4.1	Метод опорных векторов. Оптимальная разделяющая гиперплоскость. Зазор между классами. Функции ядра (kernel trick). Нелинейный SVM. /Лек/	1	2	Л1.1 Л1.2 Л1.3Л2.1 Л2.2 Э1 Э2
4.2	Настройка параметров SVM. Использование SVM для решения задачи бинарной классификации. Построение классификатора спама на основе SVM. /Пр/	1	6	Л1.1 Л1.2 Л1.3Л2.1 Л2.2 Э1 Э2
Раздел 5. Методы понижения размерности и визуализации данных				
5.1	Методы понижения размерности данных - метод главных компонент (PCA) и метод t-SNE. Сжатие и визуализация данных. /Лек/	1	2	Л1.1 Л1.2 Л1.3Л2.1 Л2.2 Э1 Э2
5.2	Использование PCA для уменьшения размерности данных. Метод t-SNE. /Пр/	1	4	Л1.1 Л1.2 Л1.3Л2.1 Л2.2 Э1 Э2
Раздел 6. Дополнительные вопросы обучения моделей машинного обучения				
6.1	Организация надежной валидации (dataset split, cross-validation), анализ learning curves. Метрики качества моделей, отбор признаков. /Лек/	1	2	Л1.1 Л1.2 Л1.3Л2.1 Л2.2 Э1 Э2
6.2	Валидация модели машинного обучения. Изучение метрик качества в задаче с несбалансированными классами. /Пр/	1	4	Л1.1 Л1.2 Л1.3Л2.1 Л2.2 Э1 Э2
Раздел 7. Самостоятельная работа				
7.1	Подготовка к промежуточным тестам /Ср/	1	12	Л1.1 Л1.2 Л1.3Л2.1 Л2.2 Э1 Э2
7.2	Изучение тем и проблем, не выносимых на лекции и практические занятия /Ср/	1	14	Л1.1 Л1.2 Л1.3Л2.1 Л2.2 Э1 Э2
7.3	Подготовка к итоговому тесту /Ср/	1	14	Л1.1 Л1.2 Л1.3Л2.1 Л2.2 Э1 Э2
7.4	Подготовка к защите практических работ /Ср/	1	11	Л1.1 Л1.2 Л1.3Л2.1 Л2.2 Э1 Э2
Раздел 8. Иная контактная работа				
8.1	Индивидуальные консультации, текущий контроль /ИКР/	1	9	Л1.1 Л1.2 Л1.3Л2.1 Л2.2 Э1 Э2

6. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

6.1. Перечень видов оценочных средств

Компьютерный тест. Практическая работа

6.2. Типовые контрольные задания и иные материалы для текущей аттестации

На графиках представлена зависимость значений целевой функции от числа эпох обучения. Какой график из представленных на рисунке соответствует случаю успешного обучения модели линейной регрессии?[целевая функция - число эпох]

- A) ни один
- B) верхний
- C) нижний
- D) оба

ANSWER: A

От скольких переменных зависит целевая функция в задаче одномерной линейной регрессии?



- A) 2
B) 1
C) n, где n - число обучающих элементов
D) n+1, где n - число обучающих элементов
ANSWER: A

Количество баллов равно числу правильных ответов.
Продолжительность тестирования 20 мин.

Практическая работа: Классификация ирисов Фишера с помощью метода kNN. Использование ансамблевых моделей на основе деревьев решений.

Максимальное количество баллов - 12
12 баллов: задание полностью выполнено
1-11: задание выполнено частично
0: задание не выполнено

6.3. Типовые контрольные вопросы и задания для промежуточной аттестации

Испытывают два независимо работающих элемента. Длительность времени безотказной работы $F_1(t)=1-e^{-0,02t}$, $F_2(t)=1-e^{-0,05t}$ Найти вероятность того, что за 6 часов оба элемента откажут.

- A) 0,03
B) 0,02
C) 0,56
D) 0,4

ANSWER: A

Как выглядит выход $h(x)$ модели линейной регрессии?

- A) $h(x)=\theta_0+\theta_1 x$
B) $h(x)=\text{sigmoid}(\theta_0+\theta_1 x)$
C) $h(x)=\begin{cases} -1, & x < 0, \\ 1, & x \geq 0 \end{cases}$
D) $h(x)=\frac{1}{2m} \sum_{i=1}^n (\theta_0+\theta_1 x_i - y_i)^2$

ANSWER: A

На автозавод поступили двигатели от трех моторных заводов. От первого завода поступило 10 двигателей, от второго – 6, от третьего – 4. Вероятности безотказной работы этих двигателей в течении гарантийного срока соответственно равны 0,9; 0,8; 0,7. Какова вероятность того, что установленный на машине двигатель будет работать без дефектов в течении гарантийного срока?

- A) 0,83
B) 0,78
C) 0,87
D) 0,56

ANSWER: A

6.4. Критерии оценивания

На экзамене происходит оценивание учебной деятельности обучающихся по дисциплине на основе полученных оценок за контрольно-рейтинговые мероприятия текущего контроля и промежуточной аттестации. При оценивании результатов учебной деятельности обучающегося по дисциплине используется балльно-рейтинговая система оценивания результатов учебной деятельности обучающихся.

Отлично: Величина рейтинга обучающегося по дисциплине 85...100 %

Хорошо: Величина рейтинга обучающегося по дисциплине 75...84 %

Удовлетворительно: Величина рейтинга обучающегося по дисциплине 60...74 %

Неудовлетворительно: Величина рейтинга обучающегося по дисциплине 0...59 %.

Допускается выставление оценки на основе текущего рейтинга (автоматом).

7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

7.1. Рекомендуемая литература

7.1.1. Основная литература

Авторы, составители	Заглавие	Издательство, год	Ресурс
---------------------	----------	-------------------	--------



	Авторы, составители	Заглавие	Издательство, год	Ресурс
Л1.1	Флах П.	Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных (http://e.lanbook.com/books/element.php?p11_id=69955)	Москва : ДМК Пресс, 2015	ЭБС
Л1.2	Андрейчиков А.В., Андрейчикова О.Н.	Интеллектуальные информационные системы и методы искусственного интеллекта: учебник (https://znanium.ru/catalog/document?id=439338)	Москва : ООО "Научно- издательский центр ИНФРА- М", 2024	ЭБС
Л1.3	Вишневский В.М., Ефросинин Д.В.	Теория очередей и машинное обучение: монография (https://znanium.ru/catalog/document?id=456847)	Москва : ООО "Научно- издательский центр ИНФРА- М", 2024	ЭБС

7.1.2. Дополнительная литература

	Авторы, составители	Заглавие	Издательство, год	Ресурс
Л2.1	Рашка С.	Python и машинное обучение: крайне необходимое пособие по новейшей предсказательной аналитике, обязательное для более глубокого понимания методологии машинного обучения (https://e.lanbook.com/book/100905)	Москва : ДМК Пресс, 2017	ЭБС
Л2.2	Шарден Б., Массарон Л., Боскетти А.	Крупномасштабное машинное обучение вместе с Python (https://e.lanbook.com/book/105836)	Москва : ДМК Пресс, 2018	ЭБС

7.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"

Э1	Лань [Электронный ресурс] : электронно-библиотечная система (ЭБС) / издательство Лань http://e.lanbook.com
Э2	Znanium.com [Электронный ресурс] : электронно-библиотечная система (ЭБС) / Науч. электрон. б-ка http://znanium.com/

7.3 Перечень информационных технологий

7.3.1 Программное обеспечение

LMS Moodle

Python

7.3.2 Профессиональные базы данных и информационно-справочные системы

8. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Для реализации дисциплины используются учебные аудитории для проведения занятий лекционного типа, занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы.

Учебные аудитории укомплектованы специализированной мебелью и техническими средствами обучения: доска, парты, мультимедийное и аудиооборудование.

Для проведения занятий лекционного типа предлагаются наборы демонстрационного оборудования и учебно-наглядных пособий: цифровые образовательные ресурсы, а также используется переносное и / или стационарное мультимедийное оборудование (экран, ноутбук, проектор, колонки).

Для семинарских занятий используются аудитории оснащенные обычной доской, партами, переносным мультимедийным и аудиооборудованием (в случае необходимости).

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с подключением к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду университета.

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

Успешное изучение курса требует от обучающихся посещения лекций, активной работы на семинарах, выполнения всех учебных заданий преподавателя, ознакомления с базовыми учебниками, основной и дополнительной литературой. Запись лекции – одна из форм активной работы студентов, требующая навыков и умения кратко, схематично, последовательно и логично фиксировать основные положения, выводы, обобщения, формулировки.



Последующая работа над текстом лекции воскрешает в памяти ее содержание, позволяет развивать экономическое мышление. В конце лекции преподаватель оставляет время (5 минут) для того, чтобы студенты имели возможность задать уточняющие вопросы по изучаемому материалу.

Важным методом обучения является самостоятельная работа студентов с учебно-методическими материалами, научной литературой. При изучении дисциплины необходимо изучить вопросы, которые преподаватель вынес на самостоятельное изучение, быть готовым к обсуждению этих вопросов.

К промежуточной аттестации необходимо готовиться целенаправленно, регулярно, систематически и с первых дней обучения по данной дисциплине. После этого у обучающегося должно сформироваться четкое представление об объеме и характере знаний и умений, которыми надо будет овладеть по дисциплине. Систематическое выполнение учебной работы на лекциях и семинарских занятиях позволит успешно освоить дисциплину.

В случае применения при обучении дисциплины электронного обучения, дистанционных образовательных технологий общение обучающихся и преподавателя осуществляется в режиме реального времени (онлайн-лекции (вебинары), чаты, видео-конференции и др.) или отложенного времени (система дистанционного обучения Moodle, форумы, электронная почта и др.).

Большую часть времени обучающиеся самостоятельно работают с учебно-методическими материалами. Студенты имеют возможность консультироваться с преподавателем по всем вопросам, возникающим в ходе самостоятельной работы посредством электронной почты, социальных сетей и т.п.

Доступ обучающегося к учебным ресурсам в режиме отложенного времени, самостоятельной работы осуществляется через сеть Интернет в удобном для него месте, времени и темпе.

При обучении лиц с ограниченными возможностями здоровья электронное обучение, дистанционные образовательные технологии предусматривают возможность приема-передачи информации в доступных для них формах.

Реализация дисциплины с применением электронного обучения, дистанционных образовательных технологий (далее – ЭО, ДОТ) осуществляется на основании «Положения о реализации основных и дополнительных образовательных программ с применением электронного обучения и дистанционных образовательных технологий в федеральном государственном бюджетном образовательном учреждении высшего образования «Челябинский государственный университет», «Положения о порядке зачета обучающимися по основным профессиональным образовательным программам высшего образования в ФГБОУ ВО «ЧелГУ» результатов освоения в организациях, осуществляющих образовательную деятельность, учебных предметов, курсов, дисциплин (модулей), практик, дополнительных образовательных программ» посредством электронной информационно-образовательной среды ФГБОУ ВО «ЧелГУ». В исключительных случаях (форс-мажор и т.п.) при реализации образовательной деятельности с применением ЭО, ДОТ могут применять компоненты, не входящие в перечень электронной информационно-образовательной среды.

10. СПЕЦИАЛЬНЫЕ УСЛОВИЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ ОБУЧАЮЩИМИСЯ С ИНВАЛИДНОСТЬЮ И ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ

Освоение дисциплины инвалидами и лицами с ограниченными возможностями здоровья осуществляется с использованием специальных технических средств и информационных технологий, предоставляемых Ресурсным учебно-методическим центром по обучению инвалидов и лиц с ограниченными возможностями здоровья ЧелГУ по запросу обучающегося (мобильные специальные технические средства для лиц с нарушениями зрения и с нарушением слуха, ассистивные информационные технологии).

При необходимости для обучающихся с нарушениями зрения на рабочих местах для проведения практических или лабораторных занятий устанавливается специальное программное обеспечение (программа речевой навигации, речевые синтезаторы, экранные лупы).

В учебные аудитории обеспечивается беспрепятственный доступ для обучающихся с инвалидностью и с ограниченными возможностями здоровья. В каждой аудитории, где обучаются инвалиды и лица с ограниченными возможностями здоровья, предусматривается соответствующее количество мест для обучающихся с учетом нарушений их здоровья.

Для освоения дисциплины инвалидам и лицам с ограниченными возможностями здоровья предоставляется доступ к печатным источникам, имеющимся в научной библиотеке ЧелГУ, с помощью специальных технических средств; доступ с помощью специальных технических и программных средств к электронным источникам, представленным в форме электронного документа в фонде научной библиотеки ЧелГУ или электронно-библиотечных системах.

Учебно-методические материалы для обучающихся из числа инвалидов и лиц с ограниченными возможностями здоровья предоставляются в формах, адаптированных к ограничениям их здоровья и особенностям восприятия информации.

Для инвалидов и лиц с ограниченными возможностями здоровья освоение дисциплины может быть частично или полностью осуществлено с использованием дистанционных образовательных технологий.

При проведении промежуточной аттестации по дисциплине обучающимся с инвалидностью и с ограниченными возможностями здоровья обеспечивается по их заявлению предоставление в доступной форме в зависимости от их



индивидуальных особенностей инструкции о порядке проведения промежуточной аттестации, оценочных средств и возможности ответов на задания (письменно на бумаге, набор ответов на компьютере, письменно шрифтом Брайля, с использованием услуг ассистента, устно).

При проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья предусматривается использование предоставленных ЧелГУ или собственных технических средств, необходимых им в связи с их индивидуальными особенностями. При необходимости инвалидам и лицам с ограниченными возможностями здоровья предоставляется дополнительное время для подготовки ответа на задания, процедура оценивания результатов обучения по дисциплине может проводиться в несколько этапов.