

<p>Документ подписан простой электронной подписью  Информация о владельце:  ФИО: Таскаев Сергей Валерьевич  Должность: Ректор  Дата подписания: 04.07.2024 00:51:26  Уникальный программный ключ:  09192488137655050640430788872155</p>	<p>МИНИСТЕРСТВО НАУКИ РОССИИ  Федеральное государственное бюджетное образовательное учреждение высшего образования  «Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)</p> <p>Рабочая программа дисциплины "Методы машинного обучения (научный семинар)" по направлению подготовки (специальности) 02.03.02 "Фундаментальная информатика и информационные технологии" направленности (профилю) Математические и алгоритмические основы интеллектуальных систем ФГБОУ ВО «ЧелГУ»</p>	<p>стр. 1</p>
---	---	---------------

**Рабочая программа дисциплины (модуля)\***  
**Методы машинного обучения (научный семинар)**

Направление подготовки (специальность)

02.03.02 Фундаментальная информатика и информационные технологии

Направленность (профиль)

Математические и алгоритмические основы интеллектуальных систем

Присваиваемая квалификация (степень)

бакалавр

Форма обучения

очная

Год(ы) набора 2024

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

Челябинск 2024 г.



## 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 курса – познакомить студентов с наиболее популярными методами и подходами к работе с данными. Задачи курса: изучить основные методы и модели для работы с данными; изучить основы нейронных сетей; научиться обрабатывать данные, выбирать и анализировать параметры качества для конкретной задачи, проверять и оценивать модели.
Результаты изучения дисциплины направлены на достижение следующих индикаторов:
УК-1.1. Выполняет поиск информации, определяет критерии системного анализа поставленных задач.
УК-1.2. Использует критический анализ, систематизацию и обобщение информации для решения поставленных задач.
ПК-1.1. Обладает знаниями о методологии и этапах выполнения научно-исследовательской работы.
ПК-1.2. Демонстрирует умения: обрабатывать и анализировать научно-техническую информацию и результаты исследований.
ПК-1.3. Имеет практический опыт (навыки): научной аргументации при анализе объекта научной и профессиональной деятельности.

### 2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

Цикл (раздел) ОПОП:	К.М.01.ДВ.03.02
<b>2.1 Требования к предварительной подготовке обучающегося:</b>	
При изучении этой дисциплины обучающимся понадобятся знания следующих дисциплин:	
Математические методы обработки изображений (научный семинар)	
Современные компьютерные технологии (научный семинар)	
Введение в программирование на языке Python	
Введение в анализ информационных технологий	
<b>2.2 Дисциплины и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:</b>	
Дисциплина логически связана с дисциплинами	
Интеллектуальные системы	
Подготовка к процедуре защиты и защита выпускной квалификационной работы	

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

<b>УК-1: Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач</b>
<b>Знать:</b>
как выполнять поиск информации, определяет критерии системного анализа поставленных задач;
<b>Уметь:</b>
использовать критический анализ, систематизацию и обобщение информации для решения поставленных задач;
<b>Владеть:</b>
владеть навыками системного анализа и поиска информации.
<b>ПК-1: Способность проводить под научным руководством локальные научно-исследовательские и опытно-конструкторские разработки на основе существующих методов в конкретной области профессиональной деятельности</b>
<b>Знать:</b>
методологии и этапы выполнения научно-исследовательской работы;
<b>Уметь:</b>
обрабатывать и анализировать научно-техническую информацию и результаты исследований;



**Владеть:**

научной аргументации при анализе объекта научной и профессиональной деятельности.

**В результате освоения дисциплины обучающийся должен**

<b>3.1</b>	<b>Знать:</b>
3.1.1	- как выполнять поиск информации, определяет критерии системного анализа поставленных задач;
3.1.2	- методологии и этапы выполнения научно-исследовательской работы.
<b>3.2</b>	<b>Уметь:</b>
3.2.1	- использовать критический анализ, систематизацию и обобщение информации для решения поставленных задач;
3.2.2	- обрабатывать и анализировать научно-техническую информацию и результаты исследований.
<b>3.3</b>	<b>Владеть:</b>
3.3.1	- владеть навыками системного анализа и поиска информации;
3.3.2	- научной аргументации при анализе объекта научной и профессиональной деятельности.

**4. ОБЪЕМ ДИСЦИПЛИНЫ (МОДУЛЯ)**

Общая трудоемкость	<b>5 ЗЕТ</b>
Часов по учебному плану : 180	Виды контроля в семестрах: экзамены 8
в том числе :	
аудиторные занятия : 40	
самостоятельная работа : 95,8	
часов на контроль : 36	
контактная работа: 48,2	
ИКР: 8,2	

**5. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)**

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Литература
	<b>Раздел 1. Обработка данных.</b>			
1.1	Типы данных. Методы обработки данных различного типа. Основные библиотеки для обработки и визуализации данных. /Лек/	8	2	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2 Э2
1.2	Подготовка изображений. Обработка и визуализация статистических и текстовых данных. /Лаб/	8	2	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2
1.3	Поиск и обработка данных согласно индивидуального задания. /Ср/	8	32	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2 Э1
	<b>Раздел 2. Методы классификации.</b>			
2.1	Обучение с учителем. Метод k-ближайших соседей (K-Nearest Neighbors); Метод опорных векторов (Support Vector Machines); Классификатор дерева решений (Decision Tree Classifier) / /Лек/	8	4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2 Э2
2.2	Случайный лес (Random Forests); Наивный байесовский метод (Naive Bayes); Линейный дискриминантный анализ (Linear Discriminant Analysis); Логистическая регрессия (Logistic Regression); /Лек/	8	6	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2 Э2
2.3	Метрические методы классификации. /Лаб/	8	4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2



2.4	Логические методы классификации. /Лаб/	8	4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2
2.5	Построение пайплайна для одномерной и многомерной регрессии. /Лаб/	8	2	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2
2.6	Выполнить классификацию двумя методами для набора данных индивидуального задания. /Ср/	8	32	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2 Э1
<b>Раздел 3. Методы кластеризации.</b>				
3.1	Обучение без учителя. K-means Affinity Propagation Спектральная кластеризация /Лек/	8	4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2 Э1
3.2	Агломеративная кластеризация Метрики качества кластеризации /Лек/	8	4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2 Э2
3.3	Модель кластеризации методом K-средних. Исследование различных метрик. /Лаб/	8	4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2
3.4	Проверка индивидуального задания. /Лаб/	8	4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2
3.5	Подготовка отчета и презентации по индивидуальному заданию. /Ср/	8	31,8	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2 Э1
<b>Раздел 4. Иная контактная работа.</b>				
4.1	Консультации, текущий контроль. /ИКР/	8	8,2	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2

## 6. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

### 6.1. Перечень видов оценочных средств

Виды оценочных средств для текущего контроля:

- комплекты заданий лабораторных работ.
- индивидуальное задание.
- теоретические вопросы к экзамену.

### 6.2. Типовые контрольные задания и иные материалы для текущей аттестации

Пример лабораторного задания

Учебная задача

Необходимо организовать подготовку данных для построения модели (допустим модели классификации). В качестве данных выбран набор данных об ирисах Фишера.

Данный набор данных предназначен для построения модели классификации. Данные о 150 экземплярах ириса (рис. 1.1), по 50 экземпляров из трёх видов – Ирис щетинистый (*Iris setosa*), Ирис виргинский (*Iris virginica*) и Ирис разноцветный (*Iris versicolor*). Для каждого экземпляра измерялись четыре характеристики (в сантиметрах):

- 1) длина наружной доли околоцветника (sepal length);
- 2) ширина наружной доли околоцветника (sepal width);
- 3) длина внутренней доли околоцветника (petal length);
- 4) ширина внутренней доли околоцветника (petal width).

На основании этого набора данных требуется построить правило классификации, определяющее вид растения по данным измерений. Это задача многоклассовой классификации, так как имеется три класса – три вида ириса.

Провести анализ данных классификации по графическому представлению.



Пример индивидуального задания.

1. Подберите набор данных на ресурсе <https://www.kaggle.com/> и согласуйте свой выбор с преподавателем. Студент может предложить синтезированный набор данных.
2. Проведите первичный анализ данных. Особое внимание следует уделить графическому представлению распределений признаков, визуализации взаимосвязей, позволяющие судить о наборе данных. Построение графиков желательно произвести по нескольким проекциям. При анализе данных использовать как можно более разнообразие типы графиков.

### 6.3. Типовые контрольные вопросы и задания для промежуточной аттестации

Контрольные вопросы в приложенном файле

### 6.4. Критерии оценивания

Схема подведения итогов промежуточной аттестации в форме экзамена в 6 семестре.

Баллы, полученные за отдельные задания экзамена (тестирование, теоретический вопрос, практическое задание) суммируются.

Виды заданий	Баллы
Тест	0-20
Теоретический вопрос	0-40
Практическое задание	0-40
Всего:	0-100

0-49 баллов – неудовлетворительно;  
50-69 баллов - удовлетворительно;  
70-89 баллов - хорошо;  
90-100 баллов - отлично.

При подведении итогов промежуточной аттестации учитываются результаты текущей аттестации.

Виды заданий	Баллы
Лабораторный практикум (1-18)	0-50
Индивидуальное задание	0-50

0-49 баллов - неудовлетворительно (2);  
50-69 баллов - удовлетворительно (3);  
70-89 баллов - хорошо (4);  
90-100 баллов - отлично (5).

Полученные за текущую аттестацию баллы могут либо повысить, либо понизить отметку за экзамен за счет вычисления среднего балла за экзамен и текущую аттестацию.

0-49 баллов - неудовлетворительно (2);  
50-69 баллов - удовлетворительно (3);  
70-89 баллов - хорошо (4);  
90-100 баллов - отлично (5).

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

### 7.1. Рекомендуемая литература

#### 7.1.1. Основная литература

	Авторы,	Заглавие	Издательство,	Ресурс
Л1.1	Вохминцев А. В., Ботов Д. С., Шицелов А. В., Петриченко Ю. В.	Машинное обучение. Нейронные сети: практикум ( <a href="http://library.csu.ru/rbooks2/view2?code=local/007890/007890">http://library.csu.ru/rbooks2/view2?code=local/007890/007890</a> )	Челябинск : Издательство Челябинского государственног о университета, 2022	ЭБС



	Авторы,	Заглавие	Издательство,	Ресурс
Л1.2	Шицелов А. В., Вохминцев А. В., Ботов Д. С., Петриченко Ю. В.	Машинное обучение и интеллектуальный анализ данных: практикум ( <a href="http://library.csu.ru/rbooks2/view2?code=local/007889/007889">http://library.csu.ru/rbooks2/view2?code=local/007889/007889</a> )	Челябинск : Издательство Челябинского государственно о университета, 2022	ЭБС
Л1.3	Титов А. Н., Тазиева Р. Ф.	Python. Обработка данных: учебно-методическое пособие ( <a href="https://biblioclub.ru/index.php?page=book&amp;id=702252">https://biblioclub.ru/index.php?page=book&amp;id=702252</a> )	Казань : Казанский национальный исследовательск ий технологически й университет (КНИТУ), 2022	ЭБС
Л1.4	Бугырский Е. Ю., Цехановский В. В., Жукова Н. А., Баймуратов И. Р., Куликов И. А.	Машинное обучение: учебник ( <a href="https://biblioclub.ru/index.php?page=book&amp;id=701807">https://biblioclub.ru/index.php?page=book&amp;id=701807</a> )	Москва : Директ -Медиа, 2023	ЭБС
Л1.5	Маккинни У.	Python и анализ данных. Первичная обработка данных с применением pandas, NumPy и Jupiter ( <a href="https://e.lanbook.com/book/348086">https://e.lanbook.com/book/348086</a> )	Москва : ДМК Пресс, 2023	ЭБС

#### 7.1.2. Дополнительная литература

	Авторы,	Заглавие	Издательство,	Ресурс
Л2.1	Мэрфи К. П.	Вероятностное машинное обучение. Введение ( <a href="https://e.lanbook.com/book/314891">https://e.lanbook.com/book/314891</a> )	Москва : ДМК Пресс, 2022	ЭБС
Л2.2	Топольский Н. Г., Вилисов В. Я., Топольский Н. Г.	Методы, модели и алгоритмы в системах безопасности: машинное обучение, робототехника, страхование, риски, контроль: монография ( <a href="https://biblioclub.ru/index.php?page=book&amp;id=709190">https://biblioclub.ru/index.php?page=book&amp;id=709190</a> )	Москва : Риор, 2021	ЭБС

#### 7.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"

Э1	Kaggle — платформа для соревнований по Data Science от Google. Свободный доступ. <a href="https://www.kaggle.com/">https://www.kaggle.com/</a>			
Э2	Введение в машинное обучение. Яндекс школа. Лектор Воронцов К.В. Свободный доступ. <a href="https://www.youtube.com/playlist?list=PLJOzdkh8T5krxc4HsHbB8g8f0hu7973fK">https://www.youtube.com/playlist?list=PLJOzdkh8T5krxc4HsHbB8g8f0hu7973fK</a> <a href="https://www.youtube.com/playlist?list=PLJOzdkh8T5krxc4HsHbB8g8f0hu7973fK">https://www.youtube.com/playlist?list=PLJOzdkh8T5krxc4HsHbB8g8f0hu7973fK</a>			

#### 7.3 Перечень информационных технологий

##### 7.3.1 Программное обеспечение

LMS Moodle

Python

PostgreSQL

##### 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, лабораторных занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы.

Учебные аудитории укомплектованы специализированной мебелью.

Для проведения занятий лекционного типа используется переносное и / или стационарное мультимедийное оборудование (экран, ноутбук или десктоп, проектор).

Для обеспечения тематической иллюстрации занятий лекционного типа в образовательном процессе используются цифровые образовательные ресурсы (мультимедийные презентации по всем темам программы).



Для проведения лабораторных работ и самостоятельной работы используется компьютерный класс, объединённых в локальную компьютерную сеть с выходом в Интернет и доступом в электронную информационно-образовательную среду университета, с установленным программным обеспечением.

При изучении дисциплины используется программное обеспечение, указанное в п. 7.3.1.

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

Учебный курс строится таким образом, чтобы способствовать созданию у обучающегося понятийно–теоретической базы, развитию умения практического решения задач, умения работать со справочной литературой.

Для успешного усвоения материала студенту необходимо получить достаточное количество баллов по следующим формам обучения:

1. Лекционная форма, которая предполагает посещение лекций.
2. Практическая форма занятий предполагает выполнение лабораторных работ, использование справочной литературы.
3. Самостоятельная форма работы предполагает изучение теоретических вопросов, выполнение практических заданий. Для их выполнения обучающемуся необходимо использование и изучение литературы по заданной теме.

В каждом семестре студенты выполняют цикл лабораторных работ, которые служат основой для сдачи зачета или допуска к экзамену.

По окончании шестого семестра студенты сдают экзамен. Сдача экзамена предполагает выполнение теста. В тест входят вопросы и задачи.

Примеры индивидуальных заданий, образец итогового задания, образцы вопросов теста к экзамену прилагаются в разделе ФОС.

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-

## 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 голо информационных технологий, предоставляемых Ресурсным учебно-методическим центром по обучению инвалидов и лиц с ограниченными возможностями здоровья ЧелГУ по запросу обучающегося.

1. Мобильные специальные технические средства для лиц с нарушениями зрения: портативный компьютер с вводом/выводом шрифтом Брайля с синтезатором речи «E1Braille-W14J G2»; ноутбуки с программной экранного доступа NVDA; электронные увеличители для удаленного просмотра; видеоувеличители портативные; тифлоплеер; цифровые диктофоны.

2. Мобильные специальные технические средства для лиц с нарушениями слуха: система свободного звукового поля со встроенной совместимостью с FM-устройствами; радиоклассы «Сонет-PCM» с передатчиком, заушным индуктором и индукционной петлей; система информационная для слабослышащих переносная «Исток» А2 со встроенным плеером – звуковым информатором; документ-камера; программируемые слуховые аппараты индивидуального пользования.

3. Ассистивные информационные технологии: программное обеспечение экранного доступа с синтезом речи NVDA; программы экранного увеличения; программы речевого синтеза для компьютеров и ноутбуков; программы речевого синтеза для мобильных устройств; экранная клавиатура; экранная лупа.

При необходимости для обучающихся с нарушениями зрения на рабочих местах для проведения практических или лабораторных занятий устанавливается специальное программное обеспечение (программа речевой навигации



NVDA, речевые синтезаторы, экранные лупы).

В учебные аудитории обеспечивается беспрепятственный доступ для обучающихся инвалидов и обучающихся 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 к электронным источникам, представленным в форме электронного документа в фонде научной библиотеки ЧелГУ или электронно-библиотечных системах, с помощью специальных технических и программных средств (рабочее место для незрячего пользователя с программным обеспечением экранного доступа с синтезом речи NVDA, рабочее место с компьютерным роллером и клавиатурой Clevy с большими кнопками и с разделяющей клавиши накладкой).

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

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа,
- в форме аудиофайла,
- в печатной форме шрифтом Брайля.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа,
- в форме аудиофайла.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

Для инвалидов и лиц с ограниченными возможностями здоровья освоение дисциплины может быть частично или полностью осуществлено с использованием дистанционных образовательных технологий (Moodle, Adobe Connect Pro и пр.).

В освоении дисциплины инвалидами и лицами с ограниченными возможностями здоровья используется индивидуальная работа. Под индивидуальной работой подразумевается две формы взаимодействия с преподавателем: индивидуальная учебная работа (консультации), т.е. дополнительное разъяснение учебного материала и углубленное изучение материала с теми обучающимися, которые в этом заинтересованы, и индивидуальная воспитательная работа. Индивидуальные консультации направлены на индивидуализацию обучения и установлению воспитательного контакта между преподавателем и обучающимся инвалидом или обучающимся с ограниченными возможностями здоровья.

При проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине обеспечивается выполнение следующих дополнительных требований в зависимости от индивидуальных особенностей, обучающихся:

- а) инструкция по порядку проведения процедуры оценивания предоставляется в доступной форме (устно, в письменной форме, в письменной форме шрифтом Брайля, устно с использованием услуг сурдопереводчика);
- б) доступная форма предоставления заданий оценочных средств (в печатной форме, в печатной форме увеличенным шрифтом, в печатной форме шрифтом Брайля, в форме электронного документа, задания зачитываются ассистентом, задания предоставляются с использованием сурдоперевода);
- в) доступная форма предоставления 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 технических средств, необходимых им в связи с их индивидуальными особенностями. Эти средства могут быть предоставлены ЧелГУ или могут использоваться собственные технические средства. 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.

Проведение процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья допускается с использованием дистанционных образовательных технологий.

## Вопросы к экзамену

### Основные определения

- 1 Основные определения в машинном обучении: объект, целевая функция, признак, модель, обучающая выборка, функционал качества, обучение, переобучение.
- 2 Задачи машинного обучения - обучение с учителем, без учителя. Задачи регрессии и классификации. Задачи снижения размерности и кластеризации.
- 3 Типы признаков в машинном обучении. Приведите примеры различных признаков.
- 4 Определение ROC-кривой.

### Метрические классификаторы

- 5 Метод  $k$  ближайших соседей в задаче классификации.
- 6 Методы отбора признаков. Жадный метод.
- 7 Определение отступа в метрических алгоритмах классификации. Алгоритм Condensed Nearest Neighbor.
- 8 Метод  $k$  ближайших соседей в задаче регрессии.
- 9 Обобщение метода  $k$  ближайших соседей через взвешенный учет объектов. Ядерная оценка плотности.
- 10 Проклятие размерности. Зависимость метода ближайших соседей от масштабирования признаков. Способы стандартизации признаков.

### Кластеризация

- 11 Постановка задачи кластеризации. Цели кластеризации. Типы кластерных структур. Чувствительность к нормировке и масштабированию признаков.
- 12 Метод  $k$  средних. Особенности метода.

**13** Степени свободы метода  $k$  средних. Метод  $k$ -means++. Метод Xmeans.

**14** Метод распространения близости.

**15** Графовые алгоритмы кластеризации.

**16** Алгоритм Ланса-Уильямса.

### **Деревья принятия решений**

**17** Логическая закономерность. Интерпретируемость и информативность.

**18** Решающий список. Достоинства и недостатки.

**19** Структура решающего дерева, метод спуска по дереву в общем случае. Достоинства и недостатки решающих деревьев.

**20** Подрезание решающих деревьев.

**21** Небрежные решающие деревья.

**22** Деревья принятия решений в задаче регрессии.

### **Байесовские методы**

**23** Вероятностная постановка задачи классификации. Функция правдоподобия и априорная вероятность.

**24** Функционал среднего риска. Общая формула байесовского классификатора.

**25** Наивный байесовский классификатор.

**26** Восстановление плотности распределения по выборке.

**27** Аддитивное сглаживание для байесова классификатора.

## Линейные классификаторы

- 28 Модель МакКаллока-Питтса
- 29 Обобщённая модель линейного классификатора. Определение отступа. Минимизация эмпирического риска.
- 30 Метод градиентного спуска. Выбор величины шага.
- 31  $L_2$  регуляризация.
- 32 Метод стохастического градиента. Недостатки метода SG и как с ними бороться.

## Способность к обобщению

- 33 Внутренний и внешний функционал качества. Кросс-валидация.
- 34 Критерий непротиворечивости моделей.
- 35 Аналитическая оценка вероятности переобучения. Схема использования.
- 36 Неравенство Бернштейна-Хёфдинга в применении к задаче выбора модели.
- 37 Дихотомии. Функция роста. Точка разрыва.
- 38 Оценка на максимальное число дихотомий.

## Нейронные сети

- 39 Представимость функций в виде нейросети.
- 40 Метод обратного распространения ошибок. Основные недостатки и способы их устранения.
- 41 Выбор начального приближения в градиентных методах настройки нейронных сетей. Функции активации.
- 42 Устройство свёрточной нейросети.

43 Нейронные сети для задачи регрессии.

## **Метод опорных векторов**

44 Постановка задачи SVM.

45 Регуляризация в задаче SVM.

46 Двойственная задача SVM.

47 Ядерный алгоритм SVM.

48 Представление метода опорных векторов в виде нейронной сети.

49 Метод SVR для задачи регрессии.

## **Линейная регрессия**

50 Постановка задачи многомерной линейной регрессии. Матричная запись.

51 Использование сингулярного разложения для решения задачи наименьших квадратов.

52 Проблема «мультиколлинеарности» в задачах многомерной линейной регрессии.

53 Гребневая регрессия. Регуляризация Лассо.

54 Нелинейная регрессия. Метод Ньютона-Гаусса.

55 Задача уменьшения размерности. Метод главных компонент.

## **Анализ смещения и разброса**

56 Постановка задачи анализа смещения и разброса. Качество обучения в зависимости от пространства моделей.

57 Внутренний и внешний функционал качества. Средний метод.

**58** Определение смещения и разброса.

**59** Кривые обучения.

### **Ансамбли**

**60** Определение композиции алгоритмов. Типы композиций.

**61** Взвешенное голосование. Бустинг. Алгоритм AdaBoost.

**62** Простое голосование. Бэггинг и метод случайных подпространств.

**63** Случайный лес. Стэкинг.

