

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ  
УЧРЕЖДЕНИЕ ИНКЛЮЗИВНОГО ВЫСШЕГО ОБРАЗОВАНИЯ

**«МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ГУМАНИТАРНО  
ЭКОНОМИЧЕСКИЙ УНИВЕРСИТЕТ»**

---

Факультет Прикладная математика и информатика  
Кафедра Информационных технологий и прикладной математики

«Утверждаю»

Зав. кафедрой 

«26» августа 2020 г.

**ФОНД ОЦЕНОЧНЫХ СРЕДСТВ  
ПО ДИСЦИПЛИНЕ**

**«Методы машинного обучения»**

образовательная программа направления подготовки  
09.03.03 Прикладная информатика  
Блок Б1.В.ДВ.05.02 «Дисциплины (модули)», часть, формируемая  
участниками образовательных отношений, дисциплины (модули) по выбору

**Профиль подготовки**

Прикладная информатика в биоинформационных технологиях

Квалификация (степень) выпускника

Бакалавр

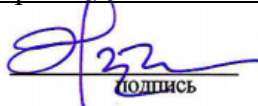
Форма обучения очная

Курс 3,4 семестр 6,7

Москва

2020

Составитель / составители: МГГЭУ, доцент кафедры информационных технологий и прикладной математики

  
подпись

место работы, занимаемая должность

Никольский А.Е. «22» августа 2020 г.  
Ф.И.О. Дата

Рецензент: МГГЭУ, профессор кафедры информационных технологий и прикладной математики

  
подпись

место работы, занимаемая должность

Истомина Т.В. «23» августа 2020 г.  
Ф.И.О. Дата

Согласовано:

*Представитель работодателя или объединения работодателей*

Генеральный директор, АО «Микропроцессорные системы», к.т.н.

(должность, место работы)

  
подпись

Демидов Л.Н. «26» августа 2020 г.  
Ф.И.О. Дата

Фонд оценочных средств рассмотрен и одобрен на заседании кафедры Информационных технологий и прикладной математики (протокол № 1 от «26» августа 2020 г.)

/Зав. кафедрой ИТиПМ/  Петрунина Е.В. «26» августа 2020 г.

подпись

Ф.И.О.

Дата

Дополнения и изменения, внесенные в фонд оценочных средств, утверждены на заседании кафедры \_\_\_\_\_,

протокол № \_\_\_\_ от « \_\_\_\_ » \_\_\_\_\_ 20\_\_ г.

Заведующий кафедрой \_\_\_\_\_ / Ф.И.О/

Дополнения и изменения, внесенные в фонд оценочных средств, утверждены на заседании кафедры \_\_\_\_\_,

протокол № \_\_\_\_ от « \_\_\_\_ » \_\_\_\_\_ 20\_\_ г.

Заведующий кафедрой \_\_\_\_\_ / Ф.И.О/

Дополнения и изменения, внесенные в фонд оценочных средств, утверждены на заседании кафедры \_\_\_\_\_,

протокол № \_\_\_\_ от « \_\_\_\_ » \_\_\_\_\_ 20\_\_ г.

Заведующий кафедрой \_\_\_\_\_ / Ф.И.О/

## Содержание

1. Паспорт фонда оценочных средств.....
2. Перечень оценочных средств.....
3. Описание показателей и критериев оценивания компетенций.....
4. Методические материалы, определяющие процедуры оценивания результатов обучения, характеризующих этапы формирования компетенций.....
5. Материалы для проведения текущего контроля и промежуточной аттестации.....
- ...

## 1. ПАСПОРТ ФОНДА ОЦЕНОЧНЫХ СРЕДСТВ

по дисциплине «Методы машинного обучения»

Оценочные средства составляются в соответствии с рабочей программой дисциплины и представляют собой совокупность контрольно-измерительных материалов ( типовые задачи ( задания), контрольные работы, тесты и др.), предназначенных для измерения уровня достижения обучающимися установленных результатов обучения.

Оценочные средства используются при проведении текущего контроля успеваемости и промежуточной аттестации.

Таблица 1 - Перечень компетенций, формируемых в процессе освоения дисциплины

Код компетенции	Наименование результата обучения
ПК-10	<p>Способен применять системный подход и математические методы в формализации решения прикладных задач</p> <p>ПК-10.1. Знает базовые положения фундаментальных разделов системного анализа и математики в объеме, необходимом для обработки информации и анализа данных в прикладной области; принципы и методы проведения исследований в области информационных систем и технологий; техники планирования и проведения вычислительного эксперимента.</p> <p>ПК-10.2. Умеет формулировать и доказывать наиболее важные результаты в прикладных областях; применять численные методы для решения прикладных задач; программно реализовать вычислительный эксперимент посредством языков программирования или с использованием специализированных пакетов прикладных программ; разрабатывать алгоритмы решения конкретных задач.</p> <p>ПК-10.3. Владеет навыками постановки задачи; навыками работы с библиографическими источниками информации; навыками решения поставленных задач в предметной области в рамках выбранного профиля.</p>

Конечными результатами освоения дисциплины являются сформированные когнитивные дескрипторы «знать», «уметь», «владеть», расписанные по отдельным компетенциям. Формирование дескрипторов происходит в течение всего семестра по этапам в рамках контактной работы, включающей различные виды занятий и самостоятельной работы, с применением различных форм и методов обучения (табл.2).

Таблица 2 - Формирование компетенций в процессе изучения дисциплины:

Код компетенции	Уровень освоения компетенций	Индикаторы достижения компетенций	Вид учебных занятий <sup>1</sup> , работы, формы и методы обучения, способствующие формированию и развитию компетенций <sup>2</sup>	Контролируемые разделы и темы дисциплины <sup>3</sup>	Оценочные средства, используемые для оценки уровня сформированности компетенции <sup>4</sup>
ПК-10		<i>Знает</i>			
	Недостаточный уровень	ПК-10. Студент не способен применять системный подход и математические методы в формализации решения прикладных задач. Не знает фундаментальных разделов кибернетики; принципов и методов проведения исследований в области кибернетики, а также не знает, что представляют собой экзоскелеты: назначение и классификация экзоскелетов, существующие конструкции экзоскелетов, а также микророботы.	Лекционные и практические занятия, самостоятельная работа обучающихся, подготовка и сдача промежуточной аттестации.	Раздел 1. Понятийный аппарат и методологическая база методов машинного обучения. Раздел 2. Основные технологии, используемые в методах машинного обучения. Раздел 3. Анализ многомерных данных. Корреляционные и причинно-следственные связи Раздел 4. Графовые методы. Нейронные сети.	Текущий контроль – устный опрос, контрольная работа, тестирование, защита отчетов по практическим работам
	Базовый	ПК-10.1. Студент имеет	Лекционные и	Раздел 1. Понятийный аппарат	Текущий контроль –

<sup>1</sup> Лекционные занятия, практические занятия, лабораторные занятия, самостоятельная работа...

<sup>2</sup> Необходимо указать активные и интерактивные методы обучения (например, интерактивная лекция, работа в малых группах, методы мозгового штурма и т.д.), способствующие развитию у обучающихся навыков командной работы, межличностной коммуникации, принятия решений, лидерских качеств.

<sup>3</sup> Наименование темы (раздела) берется из рабочей программы дисциплины.

<sup>4</sup> Оценочное средство должно выбираться с учетом запланированных результатов освоения дисциплины, например:

«Знать» – собеседование, коллоквиум, тест...

«Уметь», «Владеть» – индивидуальный или групповой проект, кейс-задача, деловая (ролевая)

игра, портфолио...

	уровень	несистематизированные знания о фундаментальных разделах кибернетики; принципах и методах проведения исследований в области кибернетики. Показывает поверхностные знания о экзоскелетах и микророботах.	практические занятия, самостоятельная работа обучающихся, подготовка и сдача промежуточной аттестации.	и методологическая база методов машинного обучения. Раздел 2. Основные технологии, используемые в методах машинного обучения. Раздел 3. Анализ многомерных данных. Корреляционные и причинно-следственные связи. Раздел 4. Графовые методы. Нейронные сети.	устный опрос, контрольная работа, тестирование, защита отчетов по практическим работам
	Средний уровень	ПК-10.1. Студент знает основное содержание материала дисциплины. Знает фундаментальные разделы кибернетики; принципы и методы проведения исследований в области кибернетики, а также знает, что представляют собой экзоскелеты: назначение и классификация экзоскелетов, существующие конструкции экзоскелетов, а также микророботы, но допускает незначительные ошибки	Лекционные и практические занятия, работа в малых группах, интерактивная лекция, дискуссия, самостоятельная работа обучающихся, подготовка и сдача промежуточной аттестации.	Раздел 1. Понятийный аппарат и методологическая база методов машинного обучения. Раздел 2. Основные технологии, используемые в методах машинного обучения. Раздел 3. Анализ многомерных данных. Корреляционные и причинно-следственные связи. Раздел 4. Графовые методы. Нейронные сети.	Текущий контроль – устный опрос, контрольная работа, тестирование, защита отчетов по практическим работам

Высокий уровень	ПК-10.1. Студент знает, понимает, выделяет главные положения в изученном материале и способен дать краткую характеристику основным идеям проработанного материала дисциплины. Знает фундаментальные разделы кибернетики; принципы и методы проведения исследований в области кибернетики, а также знает, что представляют собой экзоскелеты: назначение и классификация экзоскелетов, существующие конструкции экзоскелетов, а также микророботы.	Лекционные и практические занятия, работа в малых группах, интерактивная лекция, дискуссия, самостоятельная работа обучающихся, подготовка и сдача промежуточной аттестации.	Раздел 1. Понятийный аппарат и методологическая база методов машинного обучения. Раздел 2. Основные технологии, используемые в методах машинного обучения. Раздел 3. Анализ многомерных данных. Корреляционные и причинно-следственные связи. Раздел 4. Графовые методы. Нейронные сети.	Текущий контроль – устный опрос, контрольная работа, тестирование, защита отчетов по практическим работам
	<i>Умеет</i>			
Базовый уровень	ПК-10.2. Студент умеет формулировать и доказывать наиболее важные результаты в кибернетических областях; разрабатывать алгоритмы решения конкретных задач, умеет разрабатывать требования к системе управления многозвенным экзоскелетонным комплексом, алгоритмы управления роботами, как	Лекционные и практические занятия, работа в малых группах, интерактивная лекция, дискуссия, самостоятельная работа обучающихся, подготовка и сдача промежуточной аттестации.	Раздел 1. Понятийный аппарат и методологическая база методов машинного обучения. Раздел 2. Основные технологии, используемые в методах машинного обучения. Раздел 3. Анализ многомерных данных. Корреляционные и причинно-следственные связи. Раздел 4. Графовые методы. Нейронные сети.	Текущий контроль – устный опрос, контрольная работа, тестирование, защита отчетов по практическим работам

		многоагентными системами, но допускает ошибки.			
Средний уровень	ПК-10.2 Студент умеет формулировать и доказывать наиболее важные результаты в кибернетических областях; разрабатывать алгоритмы решения конкретных задач, умеет разрабатывать требования к системе управления многозвенным экзоскелетонным комплексом, алгоритмы управления роботами, как многоагентными системами.	Лекционные и практические занятия, работа в малых группах, интерактивная лекция, дискуссия, самостоятельная работа обучающихся, подготовка и сдача промежуточной аттестации.	Раздел 1. Понятийный аппарат и методологическая база методов машинного обучения. Раздел 2. Основные технологии, используемые в методах машинного обучения. Раздел 3. Анализ многомерных данных. Корреляционные и причинно-следственные связи. Раздел 4. Графовые методы. Нейронные сети.	Текущий контроль – устный опрос, контрольная работа, тестирование, защита отчетов по практическим работам	
Высокий уровень	ПК-10.2. Студент умеет самостоятельно, безошибочно формулировать и доказывать наиболее важные результаты в кибернетических областях; разрабатывать алгоритмы решения конкретных задач, умеет разрабатывать требования к системе управления многозвенным экзоскелетонным комплексом, алгоритмы управления роботами, как многоагентными системами	Лекционные и практические занятия, работа в малых группах, интерактивная лекция, дискуссия, самостоятельная работа обучающихся, подготовка и сдача промежуточной аттестации.	Раздел 1. Понятийный аппарат и методологическая база методов машинного обучения. Раздел 2. Основные технологии, используемые в методах машинного обучения. Раздел 3. Анализ многомерных данных. Корреляционные и причинно-следственные связи. Раздел 4. Графовые методы. Нейронные сети.	Текущий контроль – устный опрос, контрольная работа, тестирование, защита отчетов по практическим работам	
		<i>Владеет</i>			



Базовый уровень	ПК-10.3. Студент на базовом уровне владеет навыками решения поставленных задач в предметной области в рамках выбранного профиля; методами разработки экспертных систем по использованию робототехнических средств, проектирования АСУ с использованием системного анализа, математического и имитационного моделирования для систем различного назначения.	Лекционные и практические занятия, работа в малых группах, интерактивная лекция, дискуссия, самостоятельная работа обучающихся, подготовка и сдача промежуточной аттестации.	Раздел 1. Понятийный аппарат и методологическая база методов машинного обучения. Раздел 2. Основные технологии, используемые в методах машинного обучения. Раздел 3. Анализ многомерных данных. Корреляционные и причинно-следственные связи. Раздел 4. Графовые методы. Нейронные сети.	Текущий контроль – устный опрос, контрольная работа, тестирование, защита отчетов по практическим работам
Средний уровень	ПК-10.3. Студент владеет знаниями всего изученного материала. Владеет навыками решения поставленных задач в предметной области в рамках выбранного профиля; методами разработки экспертных систем по использованию робототехнических средств, проектирования АСУ с использованием системного анализа, математического и имитационного моделирования для систем различного назначения.	Лекционные и практические занятия, работа в малых группах, интерактивная лекция, дискуссия, самостоятельная работа обучающихся, подготовка и сдача промежуточной аттестации.	Раздел 1. Понятийный аппарат и методологическая база методов машинного обучения. Раздел 2. Основные технологии, используемые в методах машинного обучения. Раздел 3. Анализ многомерных данных. Корреляционные и причинно-следственные связи. Раздел 4. Графовые методы. Нейронные сети.	Текущий контроль – устный опрос, контрольная работа, тестирование, защита отчетов по практическим работам
Высокий	ПК-10.3. Студент владеет	Лекционные и	Раздел 1. Понятийный аппарат	Текущий контроль –

	уровень	<p>концептуально-понятийным аппаратом.          Владеет на высоком уровне навыками решения поставленных задач в предметной области в рамках выбранного профиля; методами разработки экспертных систем по использованию робототехнических средств, проектирования АСУ с использованием системного анализа, математического и имитационного моделирования для систем различного назначения.</p>	<p>практические занятия, работа в малых группах, интерактивная лекция, дискуссия, самостоятельная работа обучающихся, подготовка и сдача промежуточной аттестации.</p>	<p>и методологическая база методов машинного обучения.          Раздел 2. Основные технологии, используемые в методах машинного обучения.          Раздел 3. Анализ многомерных данных. Корреляционные и причинно-следственные связи          Раздел 4. Графовые методы. Нейронные сети.</p>	<p>устный опрос, контрольная работа, тестирование, защита отчетов по практическим работам</p>
--	---------	---	--	--	---

## 2. ПЕРЕЧЕНЬ ОЦЕНОЧНЫХ СРЕДСТВ<sup>5</sup>

Таблица 3

№	Наименование оценочного средства	Характеристика оценочного средства	Представление оценочного средства в ФОС
1.	Опрос	Средство контроля усвоения учебного материала темы, раздела или разделов дисциплины, организованное как учебное занятие в виде собеседования преподавателя с обучающимися.	Вопросы по темам/разделам дисциплины
2.	Контрольная работа, защита отчетов по практическим работам	<p>Различают задачи и задания:</p> <p>а) репродуктивного уровня, позволяющие оценивать и диагностировать знание фактического материала (базовые понятия, алгоритмы, факты) и умение правильно использовать специальные термины и понятия, узнавание объектов изучения в рамках определенного раздела дисциплины;</p> <p>б) реконструктивного уровня, позволяющие оценивать и диагностировать умения синтезировать, анализировать, обобщать фактический и теоретический материал с формулированием конкретных выводов, установлением причинно-следственных связей;</p> <p>в) творческого уровня, позволяющие оценивать и диагностировать умения, интегрировать знания различных областей, аргументировать собственную точку зрения.</p>	Комплект разноуровневых задач (заданий)
3.	Тестирование	Средство, позволяющее оценить уровень знаний обучающегося путем выбора им одного из нескольких вариантов ответов на поставленный вопрос. Возможно использование тестовых вопросов, предусматривающих ввод обучающимся короткого и однозначного ответа на поставленный вопрос.	Тестовые задания

<sup>5</sup> Указываются оценочные средства, применяемые в ходе реализации рабочей программы данной дисциплины.

### **3. ОПИСАНИЕ ПОКАЗАТЕЛЕЙ И КРИТЕРИЕВ ОЦЕНИВАНИЯ КОМПЕТЕНЦИЙ**

Оценивание результатов обучения по дисциплине «Математика» осуществляется в соответствии с Положением о текущем контроле успеваемости и промежуточной аттестации обучающихся.

Предусмотрены следующие виды контроля: текущий контроль (осуществление контроля всех видов аудиторной и внеаудиторной деятельности обучающегося с целью получения первичной информации о ходе усвоения отдельных элементов содержания дисциплины) и промежуточная аттестация (оценивается уровень и качество подготовки по дисциплине в целом).

Показатели и критерии оценивания компетенций, формируемых в процессе освоения данной дисциплины, описаны в табл. 4.

Таблица 4.

Код компетенции	Уровень освоения компетенции	Индикаторы достижения компетенции	Критерии оценивания результатов обучения
ПК-10		Знает	
	Недостаточный уровень Оценка «неудовлетворительно».	ПК-10.1.	<i>Не знает значительной части материала курса, не способен самостоятельно выделять главные положения в изученном материале дисциплины.</i>
	Базовый уровень Оценка «удовлетворительно».	ПК-10.1.	<i>Знает не менее 50 % основного материала курса, однако испытывает затруднения в его применении.</i>
	Средний уровень Оценка «хорошо».	ПК-10.1.	<i>Знает основную часть материала курса, способен применить изученный материал на практике, испытывает незначительные затруднения в решении задач.</i>
	Высокий уровень Оценка «отлично».	ПК-10.1.	<i>Показывает глубокое знание и понимание материала, способен применить изученный материал на практике.</i>
		Умеет	
	Базовый уровень	ПК-10.2.	<i>Умеет воспроизвести не менее 50 % основного материала курса, однако испытывает затруднения при решении практических задач.</i>
	Средний уровень	ПК-10.2.	<i>Умеет решать стандартные профессиональные задачи с применением полученных знаний, испытывает незначительные затруднения в решении задач.</i>
	Высокий уровень	ПК-10.2.	<i>Умеет решать стандартные профессиональные задачи с применением полученных знаний, показывает глубокое знание и понимание материала, способен решить задачу при изменении формулировки.</i>
		Владеет	
Базовый уровень	ПК-10.3.	<i>Студент владеет основными навыками теоретического и практического применения методов аналитической геометрии, линейной алгебры и математического анализа. Имеет несистематизированные знания основных разделов дисциплины.</i>	
Средний уровень	ПК-10.3.	<i>Студент владеет знаниями всего изученного материала, владеет навыками теоретического и практического применения методов аналитической геометрии, линейной алгебры и математического анализа. Испытывает незначительные</i>	

			<i>затруднения в решении задач.</i>
	Высокий уровень	<i>ПК-10.3.</i>	<i>Свободно владеет навыками теоретического и практического применения методов аналитической геометрии, линейной алгебры и математического анализа, показывает глубокое знание и понимание изученного материала. Студент владеет концептуально-понятийным аппаратом, научным языком и терминологией профессиональной деятельности.</i>

## **4. Методические материалы, определяющие процедуры оценивания результатов обучения**

### **Задания в форме устного опроса:**

Устный опрос используется для текущего контроля успеваемости обучающихся по дисциплине в качестве проверки результатов освоения материала. Каждому студенту выдается свой собственный, узко сформулированный вопрос. Ответ должен быть четким и кратким, содержащим все основные характеристики описываемого понятия. В своем ответе студент должен показать умения прослеживать причинно-следственные связи и навыки рассуждений и доказательства.

### **Задания в форме практических работ. Комплект разноуровневых задач (заданий)**

Практическая работа представляет собой контрольное мероприятие по учебному материалу каждой темы (раздела) дисциплины, состоящее в индивидуальном выполнении обучающимся практических заданий для оценки полученных знаний, умений и владений компетенциями, формируемыми по данной дисциплине.

Выполнение практических работ является средством текущего контроля успеваемости обучающихся по дисциплине и может включать в себя следующие типы заданий: задания типового вида и задания творческого характера, по результатам выполнения практических заданий обучающие оформляют отчеты, содержащие анализ полученных результатов и выводы.

### **Тестовые задания. Задания в форме тестирования**

Тест представляет собой контрольное мероприятие по учебному материалу каждой темы (раздела) дисциплины, состоящее в выполнении обучающимся системы стандартизированных заданий, которая позволяет автоматизировать процедуру измерения уровня знаний и умений обучающегося.

Тестирование является средством текущего контроля успеваемости обучающихся по дисциплине и может включать в себя следующие типы заданий: задание с единственным выбором ответа из предложенных вариантов, задание на определение верных и неверных суждений; задание с множественным выбором ответов.

В каждом задании необходимо выбрать все правильные ответы.

## **5. Материалы для проведения текущего контроля и промежуточной аттестации**

### **Задания в форме опроса**

**Раздел 1. Понятийный аппарат и методологическая база методов машинного обучения.**

- 1) Основные понятия и определения в методах машинного обучения.
- 2) История развития методов машинного обучения.
- 3) Методологическая база методов машинного обучения.

**Раздел 2. Основные технологии, используемые в методах машинного обучения.**

- 1) Типы и способы представления методов машинного обучения
- 2) Базовые алгоритмы решения задач машинного обучения
- 3) Основные программно-информационные ресурсы методов машинного обучения

### **Раздел 3. Анализ многомерных данных. Корреляционные и причинно-следственные связи**

- 1) Анализ многомерных данных.
- 2) Корреляционные и причинно-следственные связи. Корреляция признаков и структура данных.
- 3) Регрессия. Метод наименьших квадратов.
- 4) Теорема Гаусса-Маркова. Обобщенный метод наименьших квадратов.
- 5) Многомерная регрессия. Особенности построения регрессии по многомерным данным.
- 6) Множественная линейная регрессия, ее преимущества и недостатки.
- 7) Кластеризация. Кластеризация как классификация без учителя.
- 8) Меры сходства и меры различия образов.
- 9) Метод K средних.
- 10) Метод ISODATA.
- 11) Метод FOREL.

### **Раздел 4. Графовые методы. Нейронные сети.**

- 1) Графовые методы. Иерархическая кластеризация.
- 2) Агломеративные и разделяющие алгоритмы кластеризации. Дендрограммы.
- 3) Нейронные сети. Предпосылки возникновения нейросетей.
- 4) Перцептрон Розенблатта. Многослойный перцептрон.
- 5) Карты Кохонена. Сети Хопфилда.
- 6) Методы обучения нейросетей.
- 7) Метод опорных векторов.
- 8) Машинное обучение и теория Вапника-Червоненкиса.
- 9) Принцип структурной минимизации риска.
- 10) Метод опорных векторов.
- 11) Политика назначения штрафов.

Контролируемые компетенции: ПК-10

*Оценка компетенций осуществляется в соответствии с таблицей 4.*

### **Задания в форме практических работ. Комплект разноуровневых задач (заданий)**

Практическая работа 1: «Автоматизация конструирования признаков»

Цель работы: приобрести практический опыт автоматизации конструирования признаков.

Задание: найти оптимальное представление набора данных с помощью метода главных компонент (МГК) и критериев отбора компонент.

Шаги:

1 Сгенерировать набор данных произвольной размерности  $d$ , например, методом «sklearn.datasets.make\_blobs».



2 Получить дисперсии значений компонент. Это можно сделать с помощью класса «sklearn.decomposition.PCA» и атрибута «explained\_variance\_».

3 Определить пороговые значения дисперсии компонент с помощью:

- правила Кайзера;
- правила сломанной трости;
- «метода локтя».

4 Построить диаграммы, например, с помощью метода «matplotlib.pyplot.bar», отметить на диаграмме пороговые значения по каждому критерию.

5 Применить МГК с количеством компонент, полученным на основе одного из критериев.

Отчет должен содержать следующие файлы:

1 Сгенерированный набор данных в формате .csv.

2 Диаграмма, полученная на шаге 4

3 Преобразованный набор данных, полученный на шаге 5, в формате .csv.

Практическая работа 2: «Автоматизация выбора модели и оптимизации гиперпараметров»

Цель работы: приобрести практический опыт автоматизации классификации.

Задание: разработать автоматизированный классификатор.

Шаги:

1 Сгенерировать набор данных для классификации, например, методом «sklearn.datasets.make\_blobs».

2 Выбрать не менее 3 алгоритмов классификации, например, из библиотеки «sklearn».

3 Выбрать целевую функцию. Если была использована библиотека «sklearn», то функции можно выбрать из модулю «sklearn.metrics».

4 Для каждого выбранного классификатора реализовать автоматическую оптимизацию гиперпараметров методами поиска по решетке и случайного поиска с кросс-валидацией. При использовании «sklearn» можно воспользоваться методами «model\_selection.GridSearchCV» и «model\_selection.RandomizedSearchCV».

5 Реализовать автоматизированный выбор классификатора из сформированного набора с оптимизированными гиперпараметрами на основе выбранной целевой функции. 6 Применить выбранный классификатор.

Отчет должен содержать следующие файлы:

1 Сгенерированный набор данных в формате .csv.

2 Текстовый файл, в котором указаны:

- выбранные классификаторы;
- выбранная целевая функция;
- список полученных в результате оптимизации значений гиперпараметров для каждого классификатора и значения целевой функции.

3 Результат классификации в формате .csv.

Контролируемые компетенции: ПК-10

*Оценка компетенций осуществляется в соответствии с таблицей 4.*

**Тестовые задания. Задания в форме тестирования**

1 Нейронная сеть, в которой каждый нейрон в узле решетки связан только с ближайшими нейронами, называется:

- a. Слабосвязная нейронная сеть,
- b. Циклическая нейронная сеть,
- c. Многослойная нейронная сеть,
- d. Полносвязная нейронная сеть.

2 Нейронная сеть, в которой каждый нейрон связан со всеми другими нейронами и все выходные сигналы подаются всем нейронам, называется:

- a. Полносвязная нейронная сеть,
- b. Циклическая нейронная сеть,
- c. Слоисто-циклическая нейронная сеть,
- d. Слабосвязная нейронная сеть.

3 Нейронная сеть, в которой нейроны образуют слои, в которых они не взаимодействуют

друг с другом, а только нейронами последующего, по ходу распространения информации, слоя называются:

- a. Многослойная нейронная сеть,
- b. Слабосвязная нейронная сеть,
- c. Слоисто-циклическая нейронная сеть,
- d. Полносвязная нейронная сеть.

4 Нейронная сеть, в которой слои замкнуты в кольцо, причем все слои равноправны и могут,

как получать выходные сигналы, так и отдавать их, называются:

- a. Многослойная нейронная сеть,
- b. Циклическая нейронная сеть,
- c. Слоисто-циклическая нейронная сеть,
- d. Полносвязная нейронная сеть.

5 Нейронная сеть называется конвергирующей, если:

- a. Общее число входных нейронов больше, чем число выходных нейронов,
- b. Общее число входных нейронов меньше, чем число выходных нейронов,
- c. Общее число входных нейронов равно числу выходных нейронов,
- d. Общее число выходных нейронов больше, чем число выходных нейронов.

6 Нейронная сеть называется дивергирующей, если:

- a. Общее число входных нейронов больше, чем число выходных нейронов,
- b. Общее число входных нейронов меньше, чем число выходных нейронов,
- c. Общее число входных нейронов равно числу выходных нейронов,
- d. Общее число выходных нейронов меньше, чем число выходных нейронов.

7 Нейронная сеть называется симметричной, если:

- a. Общее число входных нейронов больше, чем число выходных нейронов,
- b. Общее число входных нейронов меньше, чем число выходных нейронов,
- c. Общее число входных нейронов равно числу выходных нейронов,
- d. Общее число выходных нейронов равно числу выходных нейронов.

8 Как называется нейронная сеть, в которой есть хотя бы один слой, выходные сигналы с

которого поступают на этот же слой или на один из предыдущих слоев.

- a. Рекуррентная нейронная сеть,
- b. Нейронная сеть прямого распространения,
- c. Нейронная сеть встречного распространения,
- d. Рециркуляционная нейронная сеть.

9 Как называется нейронная сеть, в которой выходной сигнал передается от слоя к слою

только в направлении от входного слоя сети к выходному.

- a. Рекуррентная нейронная сеть,
- b. Нейронная сеть прямого распространения,
- c. Нейронная сеть с обратными связями,
- d. Рециркуляционная нейронная сеть.

10 Как называется входная связь нейрона?

- a. Аксон,
- b. Синапс,
- c. Входной слой,
- d. Выходной слой.

11 Как называется выходная связь нейрона?

- a. Аксон,
- b. Синапс,
- c. Выходной слой,
- d. Входной слой.

12 Какая функция даёт оценку качества обучения нейронной сети:

- a. Функция ошибки,
- b. Функция активации,
- c. Обе функции,
- d. Функция принадлежности.

13 Какой вид обучения нейронной сети предполагает явное задание целевого вектора,

представляющего собой требуемый выход?

- a. Обучение с учителем,
- b. Обучение без учителя,
- c. Смешанное обучение,
- d. Конкурентное обучение.

14 Обучение нейронной сети с учителем предполагает, что:

- a. Для каждого входного вектора известен требуемый выход,
- b. В качестве обучающих примеров используются только входные значения,
- c. Неизвестны выходы сети, но известна критическая оценка правильности сети,
- d. Существует обучающая программа.

15 Какой вид обучения нейронной сети предполагает, что в качестве обучающих примеров

сети используются только входные значения

- a. Обучение с учителем,
- b. Обучение без учителя,
- c. Смешанное обучение,
- d. Дельта-правило.

16 Смешанное обучение нейронной сети предполагает, что:

- a. Для каждого входного вектора известен требуемый выход,
- b. В качестве обучающих примеров используются только входные значения,
- c. Часть весов определяется обучением «с учителем», другая часть самообучением,
- d. Конкурентное обучение.

17 Какой вид обучения наиболее употребим для нейронной сети типа персептрон?

- a. Обучение с учителем,
- b. Обучение без учителя,
- c. Смешанное обучение,
- d. Конкурентное обучение.

18. Какой вид обучения характерен для нейронной сети встречного распространения?

- a. Обучение с учителем,
- b. Обучение без учителя,

- c. Смешанное обучение,
  - d. Дельта-правило.
- 19 Сеть Хопфилда является:
- a. Сетью с обратной связью,
  - b. Прямонаправленной сетью,
  - c. Сетью обратного распространения,
  - d. Рециркуляционной сетью.
- 20 Персептрон является:
- a. Сетью с обратной связью,
  - b. Прямонаправленной сетью,
  - c. Сетью обратного распространения,
  - d. Рециркуляционной сетью.

1.	B
2.	D
3.	A
4.	C
5.	C
6.	B
7.	A
8.	D
9.	B
10.	A
11.	B
12.	D
13.	A
14.	C
15.	B
16.	B
17.	A
18.	D
19.	B
20.	A

### Тематика рефератов

1. Работа с типами данных в языке Python.
2. Введение в массивы библиотеки NumPy.
3. Выполнение вычислений над массивами библиотеки NumPy.
4. Операции над данными в библиотеке Pandas.
5. Визуализация с помощью библиотеки Matplotlib.
6. Библиотека Scikit-Learn.
7. Смеси Гауссовых распределений.
8. Ядерная оценка плотности распределения.
9. Метод опорных векторов. Оптимальная разделяющая гиперплоскость.
10. Случаи линейной разделимости и отсутствия линейной разделимости. Кусочно-линейная функция потерь.
11. Задача квадратичного программирования и двойственная задача.
12. Понятие опорных векторов. Линейные методы классификации.
13. Градиентные методы. Линейный классификатор, связь с методом максимума правдоподобия.

14. Метод стохастического градиента и частные случаи: адаптивный линейный элемент ADALINE, персептрон Розенблатта, правило Хэбба.
  15. Метрические методы классификации. Метод ближайших соседей и его обобщения.
  16. Постановка задач обучения по прецедентам.
  17. Типы задач: классификация, регрессия, прогнозирование, кластеризация.
- Примеры прикладных задач.
18. Основные понятия: модель алгоритмов, метод обучения, функция потерь и функционал качества.
  19. Методика экспериментального исследования и сравнения алгоритмов на модельных и реальных данных.
  20. Полигон алгоритмов классификации.
  21. CRISP-DM — межотраслевой стандарт ведения проектов интеллектуального анализа данных.
  22. Понятие логической закономерности. Решающие списки и деревья. Объединение в решающие леса.
  23. Сингулярное разложение, метод главных компонент.
  24. Наивная байесовская классификация.
  25. Машинное обучение с учителем и обучение без учителя.

### Контролируемые компетенции: ПК-10

*Оценка компетенций осуществляется в соответствии с таблицей 4.*

#### **Вопросы к экзамену**

1. Основные понятия. Определение предмета машинного обучения. Примеры задач и областей приложения. Образы и признаки.
2. Типы задач предсказания. Регрессия. Таксономия. Классификация. Типы ошибок классификации. Обобщающая способность классификатора.
3. Принцип минимизации эмпирического риска. Недообучение. Переобучение. Статистический, нейросетевой и структурно-лингвистический подходы к распознаванию образов.
4. Структура типичной системы распознавания образов. Цикл построения системы распознавания образов.
5. Классификация. Общие принципы. Этапы классификации. Алгоритмы обучения классификаторов с учителем и без учителя. Дискриминантный анализ. Геометрическая интерпретация задачи классификации.
6. Проективный подход. Метрики в пространстве признаков. Евклидово расстояние. Расстояние Махаланобиса. Ошибки первого и второго рода. Чувствительность и избирательность.
7. Кривая мощности критерия классификации. ROC-кривые. Проверка классификатора. Проверка тестовой выборкой. Перекрестная проверка. Оценка информативности признаков.
8. Основные методы машинного обучения. Байесовская классификация. Условная вероятность. Формула полной вероятности. Формула Байеса. Статистическое распознавание образов. Задача классификации спама. Критерий отношения правдоподобия. Байесовский риск. Критерий Байеса.
9. Критерий максимального правдоподобия. Многоклассовые байесовские классификаторы. Байесовские классификаторы для нормально распределенных классов при различной структуре матрицы ковариации
10. Оценивание функций распределения. Параметрическое оценивание. Метод максимума правдоподобия. Байесовское оценивание. Непараметрическое оценивание. Распознавание рукописных цифр с помощью наивного байесовского

11. Деревья решений. Основные понятия. Классы решаемых задач: описание данных, классификация, регрессия. Общий алгоритм построения дерева решений. Критерии выбора наилучшего атрибута: прирост информации, относительный прирост информации, индекс Гини.

12. Правила останова разбиения дерева. Обрезание дерева. Алгоритм ID3. Переобучение деревьев решений. Обработка непрерывных атрибутов. Обучение на данных с пропусками. Программное обеспечение для построения деревьев решений. Распознавание спамовых писем с помощью деревьев решений

13. Анализ многомерных данных. Корреляционные и причинно-следственные связи. Корреляция признаков и структура данных.

14. Метод главных компонент как декомпозиция матрицы данных. Матрица счетов. Матрица нагрузок. Матрица ошибок. Объясненная и остаточная вариация в данных. Графическая интерпретация метода главных компонент. Критерии выбора количества главных компонент. Понижение размерности признакового пространства методом главных компонент

15. Регрессия. Метод наименьших квадратов. Теорема Гаусса-Маркова. Обобщенный метод наименьших квадратов. Рекурсивный метод наименьших квадратов. Анализ регрессионных остатков.

16. Многомерная регрессия. Особенности построения регрессии по многомерным данным. Множественная линейная регрессия, ее преимущества и недостатки.

17. Кластеризация. Кластеризация как классификация без учителя. Меры сходства и меры различия образов. Метод K средних. Метод ISODATA. Метод FOREL.

18. Графовые методы. Иерархическая кластеризация. Агломеративные и разделяющие алгоритмы кластеризации. Дендрограммы.

19. Нейронные сети. Предпосылки возникновения нейросетей. Перцептрон Розенблатта. Многослойный перцептрон. Карты Кохонена. Сети Хопфилда. Методы обучения нейросетей. Метод опорных векторов.

20. Машинное обучение и теория Вапника-Червоненкиса. Принцип структурной минимизации риска. Метод опорных векторов. Политика назначения штрафов.

**Контролируемые компетенции: ПК-10**

*Оценка компетенций осуществляется в соответствии с таблицей 4.*