

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Волков В.В.

Должность: Ректор

Дата подписания: 14.05.2023 18:37:05

Уникальный программный ключ:

ed68fd4b85b778e0f0b1bfea5dbc56cf4148f1229917e799a70e3191f90051f

**Автономная некоммерческая образовательная организация высшего образования  
«Европейский университет в Санкт-Петербурге»**

**Факультет экономики**

УТВЕРЖДАЮ:



Ректор

В.В. Волков

«03 мая 2021 г.»

03 мая 2021 г.

Протокол Ученого Совета

№ 2 от 03 мая 2021 г.

Рабочая программа дисциплины  
**Основы машинного обучения**

образовательная программа  
направление подготовки  
**38.04.01 Экономика**

направленность (профиль)  
**«Исследовательская экономика»**  
программа подготовки – магистратура

язык обучения – русский  
форма обучения - очная

квалификация (степень) выпускника  
**Магистр**

**Санкт-Петербург**

**Автор:**

Тушканова О.Н., кандидат технических наук, доцент факультета социологии и философии АНООВО «ЕУСПб»

**Рецензент:**

Соколов М.В., кандидат экономических наук, старший научный сотрудник факультета экономики АНООВО «ЕУСПб»

Рабочая программа дисциплины «**Основы машинного обучения**», входящей в образовательную программу уровня магистратуры «Исследовательская экономика», утверждена на заседании Совета факультета экономики

Протокол заседания № 9 от 26 февраля 2021 года.

**АННОТАЦИЯ РАБОЧЕЙ ПРОГРАММЫ ДИСЦИПЛИНЫ**  
**«Основы машинного обучения»**  
**ФТД.В.05**

Дисциплина **«Основы машинного обучения»** является факультативной дисциплиной основной профессиональной образовательной программы высшего образования «Исследовательская экономика» по направлению подготовки 38.04.01 Экономика.

Дисциплина реализуется на факультете экономики.

Дисциплина нацелена на формирование следующих компетенций, которыми должен обладать выпускник, освоивший программу магистратуры:

профессиональных компетенций (ПК):

— способен работать с различными источниками статистической информации, разрабатывать и совершенствовать методологии сбора и обработки статистических данных, проводить статистические расчеты с применением соответствующих математических методов и информационных технологий, а также последующую аналитическую работу с полученными данными (ПК-5);

— способен осуществлять деятельность, направленную на решение задач аналитического характера, предполагающих выбор и многообразие актуальных способов решения задач, проводить теоретические обобщения научных данных в соответствии с задачами исследования (ПК-6).

Целью освоения дисциплины **«Основы машинного обучения»** является ознакомление с теоретическими основами алгоритмов машинного обучения, типами задач анализа данных и методов анализ данных. В ходе изучения дисциплины обучающийся учится применять алгоритмы машинного обучения в соответствии с поставленной задачей анализа данных; реализовывать программы и применять современные методы и библиотеки машинного обучения.

Преподавание дисциплины предусматривает следующие формы организации учебного процесса: лекции, практические занятия, самостоятельная работа магистранта.

Преподавание дисциплины предусматривает следующие формы организации учебного процесса: лекции, практические занятия, самостоятельная работа магистранта.

Программой дисциплины предусмотрены следующие виды контроля: текущий контроль успеваемости, промежуточный контроль в форме зачета с оценкой (в конце 8 модуля).

Общая трудоемкость освоения дисциплины составляет 3 зачетных единицы, 108 часов.

Программой дисциплины предусмотрены: 14 лекционных часов, 14 часов практических занятий, 71 час самостоятельной работы магистранта, 9 часов промежуточного контроля.

## Содержание

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ .....	5
2. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ .....	5
3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ .....	6
4. ОБЪЕМ ДИСЦИПЛИНЫ .....	6
5. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ .....	7
5.1 Содержание дисциплины .....	7
5.2 Структура дисциплины .....	8
6. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ .....	9
6.1 Общие положения.....	9
6.2 Рекомендации по распределению учебного времени по видам самостоятельной работы и разделам дисциплины.....	9
6.3 Перечень основных вопросов по изучаемым темам для самостоятельной работы обучающихся по дисциплине.....	11
6.4 Перечень литературы для самостоятельной работы обучающегося.....	11
6.5 Перечень учебно-методического обеспечения для самостоятельной работы.....	11
7. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ .....	12
7.1 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации .....	12
7.2 Контрольные задания для текущей аттестации .....	14
7.3 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации .....	14
7.4 Типовые задания к промежуточной аттестации .....	16
7.4.1. Перечень вопросов для подготовки устному ответу на вопросы.....	16
7.5 Средства оценки индикаторов достижения компетенций .....	17
8. ОСНОВНАЯ И ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА .....	17
8.1. Основная литература.....	17
8.2. Дополнительная литература .....	17
9. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ, ИСПОЛЬЗУЕМЫЕ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА .....	18
9.1 Программное обеспечение.....	18
9.2 Перечень информационно-справочных систем и профессиональных баз данных информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины: .....	18
9.3 Лицензионные электронные ресурсы библиотеки Университета.....	19
9.4 Электронная информационно-образовательная среда Университета.....	19
10. МАТЕРИАЛЬНО-ТЕХНИЧЕСКАЯ БАЗА, НЕОБХОДИМАЯ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА .....	20
ПРИЛОЖЕНИЕ 1 .....	21

## 1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

**Цель обучения:** освоения дисциплины «Машинное обучение: введение» — ознакомление с теоретическими основами алгоритмов машинного обучения, типами задач анализа данных и методов анализ данных.

### Задачи обучения:

- применение алгоритма машинного обучения в соответствии с поставленной задачей анализа данных;
- реализация программ и применение современных методов и библиотек машинного обучения.

## 2. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ

В результате изучения учебной дисциплины обучающийся должен овладеть следующими компетенциями: профессиональными (ПК). Планируемые результаты формирования компетенций и индикаторы их достижения в результате освоения дисциплины представлены в Таблице 1.

Таблица 1

**Планируемые результаты освоения дисциплины, соотнесенные с индикаторами достижения компетенций обучающихся**

Код и наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (знать, уметь, владеть)
ПК-5 Способен работать с различными источниками статистической информации, разрабатывать и совершенствовать методологии сбора и обработки статистических данных, проводить статистические расчеты с применением соответствующих математических методов и информационных технологий, а также последующую аналитическую работу с полученными данными	ИД.ПК-5.1. Работает с различными источниками статистической информации ИД.ПК-5.2. Производит экспериментальные статистические расчеты с применением информационных технологий ИД.ПК-5.3. Анализирует данные на основе методов математической статистики ИД.ПК-5.4. Разрабатывает и совершенствует системы статистических показателей и методик их расчета	Знать: российские и международные методологические положения и стандарты, методические подходы к проведению статистических расчетов и анализу, статистические пакеты прикладных программ, актуальные научные публикации по статистике, в том числе зарубежные, источники статистической информации З (ПК-5)
		Уметь: производить статистические расчеты с применением соответствующих математических методов и информационных технологий, а также последующую аналитическую работу с полученными данными, применять статистические пакеты прикладных программ У (ПК-5)
		Владеть: навыками применения традиционных и инновационных методов статистического анализа В (ПК-5)
ПК-6 Способен осуществлять деятельность, направленную на решение задач аналитического характера, предполагающих выбор и	ИД.ПК-6.1. Осуществляет выбор актуальных способов решения экономических задач ИД.ПК-6.2. Проводит теоретические обобщения научных данных в профессиональной области ИД.ПК-6.3. Применяет комплексный подход по решению	Знать: методологический аппарат современной экономики, аналитические приемы отбора методов и моделей в соответствии с целями и задачами исследования, аналитический подход к теоретическому обобщению научных данных З (ПК-6)
		Уметь:

Код и наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (знать, уметь, владеть)
многообразие актуальных способов решения задач, проводить теоретические обобщения научных данных в соответствии с задачами исследования	профессиональных задач в области экономики и смежных областях	осуществлять комплекс действий по выбору актуальных способов решения экономических задач У (ПК-6)
		Владеть: навыками решения задач аналитического характера в рамках научных исследования в области экономики и смежных отраслей В (ПК-6)

В результате освоения дисциплины магистрант должен:

**ЗНАТЬ:**

- методы сбора и обработки данных;
- основы машинного обучения;
- особенности применения различных теоретико-методологических концепций с использованием технологий прикладного анализа данных;

- современные методы машинного обучения;

- библиотеки машинного обучения.

**УМЕТЬ:**

- правильно ставить задачи анализа данных;
- уметь применять алгоритмы машинного обучения;
- правильно оформлять и представлять результаты работы.

**ВЛАДЕТЬ:**

- навыками применение алгоритма машинного обучения в соответствии с поставленной задачей анализа данных;

- навыками реализация программ и применение современных методов и библиотек машинного обучения;

- способностью к самостоятельному обучению новым методам исследования, к изменению научного и научно-производственного профиля своей профессиональной деятельности.

### 3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина «**Основы машинного обучения**» является факультативной дисциплиной образовательной программы «Исследовательская экономика». Код дисциплины по Учебному плану ФТД.В.05. Курс читается в восьмом модуле, форма промежуточной аттестации – зачет с оценкой.

Успешное освоение дисциплины предполагает наличие у обучающегося знаний, умений и навыков, полученных в ходе изучения дисциплины:

Б1.В.05 Введение в R для экономистов;

ФТД.В.02 Язык программирования Python.

Знания, умения и навыки, полученные при освоении данной дисциплины, применяются магистрантами в процессе изучения различных дисциплин, а также прохождения учебной и производственной практик.

### 4. ОБЪЕМ ДИСЦИПЛИНЫ

Общая трудоемкость освоения дисциплины составляет 3 (три) зачетных единицы, 108 часов.

Таблица 2

Типы учебных занятий и самостоятельная работа		Объем дисциплины												
		Всего	Модуль											
			1	2	3	4	5	6	7	8	9	10		
<i>Очная форма обучения</i>														
<b>Контактная работа обучающихся с преподавателем в соответствии с УП:</b>		<b>28</b>	-	-	-	-	-	-	-	-	-	<b>28</b>	-	-
лекционного типа (Лек)		14	-	-	-	-	-	-	-	-	-	14	-	-
лабораторные занятия (Лаб)		-	-	-	-	-	-	-	-	-	-	-	-	-
практические занятия (Пр)		14	-	-	-	-	-	-	-	-	-	14	-	-
<b>Самостоятельная работа обучающихся (СР)</b>		<b>71</b>	-	-	-	-	-	-	-	-	-	<b>71</b>	-	-
<b>Промежуточная аттестация</b>	<b>форма</b>	<b>Зачет с оценкой</b>	-	-	-	-	-	-	-	-	-	<b>Зачет с оценкой</b>	-	-
	<b>час.</b>	<b>9</b>	-	-	-	-	-	-	-	-	-	<b>9</b>	-	-
<b>Общая трудоемкость (час. / з.е.)</b>		<b>108/3</b>	-	-	-	-	-	-	-	-	-	<b>108/3</b>	-	-

## 5. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ

Содержание дисциплины соотносится с планируемыми результатами обучения по дисциплине: через задачи, формируемые компетенции и их компоненты (знания, умения, навыки – далее ЗУВ) по средствам индикаторов достижения компетенций в соответствии с Таблицей 3.

### 5.1 Содержание дисциплины

Таблица 3

Содержание дисциплины					
№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Индикаторы компетенций (в соот. с Таблицей 1)	Коды ЗУВ (в соот. с Таблицей 1)
1	Типы задач машинного обучения	Предмет и задачи машинного обучения и анализа данных. Основные принципы, задачи и подходы, использование в различных областях науки и индустрии. Основные этапы эволюции алгоритмов машинного обучения.	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	3 (ПК-5) У (ПК-5) В (ПК-5) 3 (ПК-6) У (ПК-6) В (ПК-6)
2	Метрические классификаторы	Общий вид метрического классификатора. Алгоритм К ближайших соседей. Алгоритмы отбора эталонов.	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	3 (ПК-5) У (ПК-5) В (ПК-5) 3 (ПК-6) У (ПК-6) В (ПК-6)
3	Алгоритмы кластеризации	Алгоритмы кластеризации с фиксированным количеством кластеров. Алгоритмы кластеризации по плотности. Иерархическая кластеризация.	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	3 (ПК-5) У (ПК-5) В (ПК-5) 3 (ПК-6) У (ПК-6) В (ПК-6)
4	Деревья решений	Правила и анализ качества (точность, полнота). Анализ с помощью ROC кривой. Алгоритм построения деревьев решений. Критерий информационного выигрыша и	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1.	3 (ПК-5) У (ПК-5) В (ПК-5) 3 (ПК-6) У (ПК-6)

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Индикаторы компетенций (в соот. с Таблицей 1)	Коды ЗУВ (в соот. с Таблицей 1)
		критерий Джини. Леса решающих деревьев.		ИД.ПК-6.2. ИД.ПК-6.3.	В (ПК-6)
5	Линейные классификаторы	Перцептрон и разделяющая гиперплоскость. Переход в пространство повышенной размерности. Метод опорных векторов.	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	3 (ПК-5) У (ПК-5) В (ПК-5) 3 (ПК-6) У (ПК-6) В (ПК-6)
6	Нейронные сети и глубокое обучение	Логистическая регрессия. Градиентный спуск. Нейронные сети и алгоритм обратного распространения градиента. Глубокое обучение, свертки и пулинг.	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	3 (ПК-5) У (ПК-5) В (ПК-5) 3 (ПК-6) У (ПК-6) В (ПК-6)
7	Регрессионный анализ	Линейная регрессия. Полиномиальная регрессия. Смещение и дисперсия. Гребневая регрессия.	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	3 (ПК-5) У (ПК-5) В (ПК-5) 3 (ПК-6) У (ПК-6) В (ПК-6)
8	Ансамблевые методы	Голосование. Бутстраппинг. Бустинг, адаптивный бустинг, градиентный бустинг.	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	3 (ПК-5) У (ПК-5) В (ПК-5) 3 (ПК-6) У (ПК-6) В (ПК-6)
9	Стохастический поиск	Монте-Карло поиск. Алгоритм симулированного отжига. Генетический алгоритм.	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	3 (ПК-5) У (ПК-5) В (ПК-5) 3 (ПК-6) У (ПК-6) В (ПК-6)

## 5.2 Структура дисциплины

Таблица 4

### Структура дисциплины

№ п/п	Наименование тем (разделов)	Объем дисциплины, час.					Форма текущего контроля успеваемости*, промежуточной аттестации
		Всего	Контактная работа обучающихся с преподавателем по типам учебных занятий в соответствии с УП			СР	
			Лек	Лаб	Пр		
<b>Очная форма обучения</b>							
Тема 1	Типы задач машинного обучения	10	2	–	1	7	ПЗ
Тема 2	Метрические классификаторы	12	2	–	2	8	ПЗ
Тема 3	Алгоритмы кластеризации	12	2	–	2	8	ПЗ
Тема 4	Деревья решений	11	1	–	2	8	ПЗ
Тема 5	Линейные классификаторы	12	2	–	2	8	ПЗ



№ п/п	Наименование тем (разделов)	Объем дисциплины, час.					Форма текущего контроля успеваемости*, промежуточной аттестации
		Всего	Контактная работа обучающихся с преподавателем по типам учебных занятий в соответствии с УП			СР	
			Лек	Лаб	Пр		
<b>Очная форма обучения</b>							
Тема 6	Нейронные сети и глубокое обучение	11	2	–	1	8	ПЗ
Тема 7	Регрессионный анализ	11	1	–	2	8	ПЗ
Тема 8	Ансамблевые методы	12	2	–	2	8	ПЗ
Тема 9	Стохастический поиск	11	2	–	1	8	ПЗ
Промежуточная аттестация		9	-	-	-	-	Зачет с оценкой
<b>Всего:</b>		<b>108</b>	<b>14</b>	<b>-</b>	<b>14</b>	<b>71</b>	<b>9</b>

\*Примечание: формы текущего контроля успеваемости: практическое задание (ПЗ).

## 6. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

### 6.1 Общие положения

Знания и навыки, полученные в результате лекций и семинарских занятий, закрепляются и развиваются в результате повторения материала, усвоенного в аудитории, путем чтения текстов и исследовательской литературы (из списков основной и дополнительной литературы) и их анализа.

Самостоятельная работа является важнейшей частью процесса высшего образования. Ее следует осознанно организовать, выделив для этого необходимое время и соответствующим образом организовав рабочее пространство. Важнейшим элементом самостоятельной работы является проработка материалов прошедших занятий (анализ конспектов, чтение рекомендованной литературы) и подготовка к следующим лекциям/практическим (семинарским) занятиям. Литературу, рекомендованную в программе курса, следует, по возможности, читать в течение всего семестра, концентрируясь на обусловленных программой курса темах.

Существенную часть самостоятельной работы магистранта представляет самостоятельное изучение вспомогательных учебно-методических изданий, лекционных конспектов, интернет-ресурсов и пр. Подготовка к практическим занятиям является важной формой работы магистранта. Самостоятельная работа может вестись как индивидуально, так и при содействии преподавателя.

### 6.2 Рекомендации по распределению учебного времени по видам самостоятельной работы и разделам дисциплины

#### Тема 1. Типы задач машинного обучения

1.1. Повторение пройденного на лекциях и на практических занятиях материала – 2 часа.

1.2. Самостоятельная работа с рекомендованной литературой, поиск ответов на возникшие в ходе подготовки вопросы – 5 часов. Итого: 7 часов.

#### Тема 2. Метрические классификаторы

2.1. Повторение пройденного на лекциях и на практических занятиях материала – 4 часа.

2.2. Самостоятельная работа с рекомендованной литературой, поиск ответов на возникшие в ходе подготовки вопросы – 4 часа.

Итого: 8 часов.

### **Тема 3. Алгоритмы кластеризации**

3.1. Повторение пройденного на лекциях и на практических занятиях материала – 4 часа.

3.2. Самостоятельная работа с рекомендованной литературой, поиск ответов на возникшие в ходе подготовки вопросы – 4 часа.

Итого: 8 часов.

### **Тема 4. Деревья решений**

4.1. Повторение пройденного на лекциях и на практических занятиях материала – 4 часа.

4.2. Самостоятельная работа с рекомендованной литературой, поиск ответов на возникшие в ходе подготовки вопросы – 4 часа.

Итого: 8 часов.

### **Тема 5. Линейные классификаторы**

5.1. Повторение пройденного на лекциях и на практических занятиях материала – 4 часа.

5.2. Самостоятельная работа с рекомендованной литературой, поиск ответов на возникшие в ходе подготовки вопросы – 4 часа.

Итого: 4 часов.

### **Тема 6. Нейронные сети и глубокое обучение**

6.1. Повторение пройденного на лекциях и на практических занятиях материала – 4 часа.

6.2. Самостоятельная работа с рекомендованной литературой, поиск ответов на возникшие в ходе подготовки вопросы – 4 часа.

Итого: 8 часов.

### **Тема 7. Регрессионный анализ**

7.1. Повторение пройденного на лекциях и на практических занятиях материала – 4 часа.

7.2. Самостоятельная работа с рекомендованной литературой, поиск ответов на возникшие в ходе подготовки вопросы – 4 часа.

Итого: 8 часов.

### **Тема 8. Ансамблевые методы**

8.1. Повторение пройденного на лекциях и на практических занятиях материала – 4 часа.

8.2. Самостоятельная работа с рекомендованной литературой, поиск ответов на возникшие в ходе подготовки вопросы – 4 часов.

Итого: 8 часов.

### **Тема 9. Стохастический поиск**

9.1. Повторение пройденного на лекциях и на практических занятиях материала – 4 часа.

9.2. Самостоятельная работа с рекомендованной литературой, поиск ответов на возникшие в ходе подготовки вопросы – 4 часа.

Итого: 8 часов.

### 6.3 Перечень основных вопросов по изучаемым темам для самостоятельной работы обучающихся по дисциплине

1. Препроцессинг. Масштабирование. Нормировка. Полиномиальные признаки. One-hot encoding.
2. Логистическая регрессия. Градиентный спуск.
3. Пороговые условия. Эффективность по Парето. Precision-Recall и ROC кривые. AUC.
4. Ансамблевые методы регрессии. RANSAC. Theil-Sen. Huber.
5. Перцептрон. Перцептрон с карманом.
6. Метод опорных векторов. Постановка задачи. Формулировка и решение двойственной задачи. Типы опорных векторов. Ядра.
7. Гипотезы и дихотомии. Функция роста. Точка поломки. Доказательство полиномиальности функции роста в присутствии точки поломки.
8. Деревья решений. Информационный выигрыш, критерий Джини. Регуляризация деревьев. Небрежные решающие деревья.
9. Байесовский классификатор. Типы оценки распределений признаков (Gaussian, Bernoulli, Multinomial). EM алгоритм.
10. Нейронные сети. Перцептрон Розенблатта. Функции активации. Обратное распространение градиента. Softmax.
11. Стохастическая оптимизация. Hill Climb. Отжиг. Генетический алгоритм

### 6.4 Перечень литературы для самостоятельной работы обучающегося:

1. Кук, Д. Машинное обучение с использованием библиотеки H2O / Д. Кук; пер. с англ. А.Б. Огурцова. - Москва: ДМК Пресс, 2018. - 250 с. - ISBN 978-5-97060-508-0. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1028135>
2. Обработка изображений с помощью OpenCV / Глория Буэно Гарсия [и др.]; пер. с англ. А.А. Слинкина. - Москва: ДМК Пресс, 2016. - 210 с. - ISBN 978-5-970(0)-387-1. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1028080>
3. Груздев, А.В. Прогнозное моделирование в IBM SPSS Statistics и R: Метод деревьев решений / А.В. Груздев. - Москва: ДМК Пресс, 2016. - 278 с. - ISBN 978-5-97060-456-4. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1028064>
4. Горелов, В.И. Анализ статистических данных: практикум: [16+] / В.И. Горелов, Т.Н. Ледащева; Российская международная академия туризма. – Москва: Университетская книга, 2015. – 120 с.: ил. – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=574944>

### 6.5 Перечень учебно-методического обеспечения для самостоятельной работы

Для обеспечения самостоятельной работы магистрантов по дисциплине «**Основы машинного обучения**» разработано учебно-методическое обеспечение в составе:

1. Контрольные задания для подготовки к процедурам текущего контроля (п. 7.2 Рабочей программы).
2. Типовые задания для подготовки к промежуточной аттестации (п. 7.4 Рабочей программы).
3. Рекомендуемые основная, дополнительная литература, Интернет-ресурсы и справочные системы (п. 8, 9 Рабочей программы).
4. Рабочая программа дисциплины размещена в электронной информационно-образовательной среде Университета на электронном учебно-методическом ресурсе АНООВО «ЕУСПб» — образовательном портале LMS Sakai — Sakai@EU.

## 7. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

### 7.1 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации

Информация о содержании и процедуре текущего контроля успеваемости, методике оценивания знаний, умений и навыков обучающегося в ходе текущего контроля доводятся научно-педагогическими работниками Университета до сведения обучающегося на первом занятии по данной дисциплине.

Текущий контроль предусматривает подготовку магистрантов к каждому занятию, активное слушание на лекциях, выполнение практических заданий. Магистрант должен присутствовать на лекциях, отвечать на поставленные вопросы, показывая, что прочитал разбираемую литературу, представлять содержательные реплики по темам обсуждения.

Текущий контроль проводится в форме оценивания выполнения практических заданий, демонстрирующих степень знакомства магистрантов с дополнительной литературой.

Таблица 5

#### Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации

Наименование тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. с Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
Типы задач машинного обучения	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 1	зачтено/ не зачтено
Метрические классификаторы	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 2	зачтено/ не зачтено
Алгоритмы кластеризации	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 3	зачтено/ не зачтено
Деревья решений	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 4	зачтено/ не зачтено
Линейные классификаторы	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 5	зачтено/ не зачтено

Наименование тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. с Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
Нейронные сети и глубокое обучение	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 6	зачтено/ не зачтено
Регрессионный анализ	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 7	зачтено/ не зачтено
Ансамблевые методы	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 8	зачтено/ не зачтено
Стохастический поиск	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 9	зачтено/ не зачтено

При освоении дисциплины опросы оцениваются с использованием бинарной системы, а домашние задания – с использованием 100-балльной шкалы с последующим переводом в бинарную систему для получения результатов текущего контроля, фиксирующих ход образовательного процесса, согласно Таблице 6.

Таблица 6

### Критерии оценивания

Формы текущего контроля успеваемости	Описание	Показатели	Количество баллов по 100-балльной шкале	Результаты текущего контроля
Практическое задание	Выполнение домашних заданий подразумевает письменное изложение ответов на теоретические вопросы и решений поставленных преподавателем задач. Предполагается, что ход решения задач сопровождается подробными комментариями обучающегося	Обучающийся правильно выполняет задания и отвечает на все поставленные в задании вопросы.	81–100	зачтено
		Обучающийся выполняет предложенные задания и отвечает на все поставленные в задании вопросы, но допускает при этом несущественные ошибки.	61–80	
		Обучающийся предлагает правильную идеологию решения задач.	41–60	

Формы текущего контроля успеваемости	Описание	Показатели	Количество баллов по 100-балльной шкале	Результаты текущего контроля
		Обучающийся испытывает затруднения при выполнении предложенных заданий.	0–40	не зачтено

## 7.2 Контрольные задания для текущей аттестации

### Примерный материал для практических заданий

#### Практическое задание 1:

- Реализуйте алгоритм kNN классификации по k ближайшим соседям, используя простое евклидовое расстояние.
- Реализуйте алгоритм k-means для кластеризации на 2-4 кластера.
- Реализуйте алгоритм DBSCAN, найдите параметры для кластеризации на 4 кластера.

#### Практическое задание 2:

- Реализуйте алгоритмы построения дерева с критерием информационного выигрыша и критерием Джини и определению класса по мажоритарному классу в листе. Найдите оптимальную глубину дерева в обоих случаях (в отрезке 2-10).
- Примените метод SVM (например, из библиотеки sklearn) для датасета blobs2. Визуализируйте результат (разбиение плоскости и опорные вектора) при разных вариантах ядер (линейное; полиномиальное степеней 2,3,5; RBF).
- Реализуйте алгоритм логистической регрессии со стохастическим градиентным спуском, обучите его на датасете spambase\_old (train) и проверьте на датасете spambase\_new (val). Получите ROC кривые для вариантов без нормировки и с нормировкой признаков.

#### Практическое задание 3:

- Реализуйте алгоритм линейной регрессии, и полиномиальной регрессии (для датасета noisysine – степеней от 2 до 5, для датасета hydrodynamics – степени 2) без регуляризации.
- Реализуйте алгоритм гребневой регрессии и найдите оптимальный параметр регуляризации для случаев из задачи 1.
- Найдите максимум функции с помощью алгоритма кросс-энтропийного поиска, изображая распределение на каждом шаге.
- Найдите лучший путь в задаче коммивояжера с помощью алгоритма отжига.

## 7.3 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации

Форма промежуточной аттестации – **зачет с оценкой**, при выставлении которого учитываются результаты текущего контроля успеваемости обучающегося по дисциплине и результаты устного ответа на вопросы, выраженные в 100-балльной шкале.

Перед зачетом с оценкой проводятся консультация, на которой преподаватель отвечает на вопросы обучающихся.

Критерии оценивания письменной зачетной работы приведены в таблице 7.

**Критерии оценивания письменной зачетной работы**

Вид промежуточной аттестации	Показатели	Количество баллов
Устный ответ на вопросы	<p>Магистрант дает ответы на вопросы билета, для которых характерно:</p> <ul style="list-style-type: none"> <li>• глубокое усвоение программного материала,</li> <li>• изложение его исчерпывающе, последовательно, четко,</li> <li>• умение делать обоснованные выводы,</li> <li>• соблюдение норм устной и письменной литературной речи;</li> </ul>	81–100
	<p>Магистрант верно отвечает на вопрос, указанный в билете, при условии, что ответ на вопрос характеризуется отсутствием серьезных, значимых неточностей, при следующих характеристиках ответа:</p> <ul style="list-style-type: none"> <li>• твердое знание материала курса,</li> <li>• последовательное изложение материала,</li> <li>• знание теоретических положений без обоснованной их аргументации,</li> <li>• соблюдение норм устной и письменной литературной речи;</li> </ul>	61–80
	<p>Магистрант представляет правильный ответ на теоретический вопрос, указанный в билете, при условии, что ответ на вопрос характеризуется значительными неточностями, при следующих параметрах ответа:</p> <ul style="list-style-type: none"> <li>• знание основного материала, но владение им не в полном объеме,</li> <li>• допущение существенных неточностей, недостаточно правильных формулировок,</li> <li>• допущение нарушения логической последовательности в изложении материала,</li> <li>• наличие нарушений норм литературной устной и письменной речи.</li> </ul>	41–60
	<p>Магистрант представляет ответ на вопрос билета, свидетельствующий о некомпетентности магистранта, при следующих параметрах ответа:</p> <ul style="list-style-type: none"> <li>• незнание значительной части программного материала,</li> <li>• наличие существенных ошибок в определениях, формулировках, понимании теоретических положений;</li> <li>• бессистемность при ответе на поставленный вопрос,</li> <li>• отсутствие в ответе логически корректного анализа, аргументации, классификации,</li> <li>• наличие нарушений норм устной и письменной литературной речи.</li> </ul>	0–40

Максимальная оценка по итогам освоения дисциплины составляет 100 баллов (50% оценки составляют результаты практических заданий по темам курса, 50% – результаты устного ответа на вопросы). При вычислении оценки G обучающегося по итогам освоения дисциплины используется расчетная формула

$$G = 0.5R + 0.5F,$$

после чего проводится процедура округления до целого. Здесь R – количество баллов за выполнение практических заданий, F – количество баллов, полученных за письменную зачетную работу.

В результате промежуточного контроля знаний студенты получают аттестацию по дисциплине. На основании оценки обучающегося по итогам освоения дисциплины, выраженной в 100-балльной шкале, выставляется **зачет с оценкой** в соответствии с Таблицей 8.

Таблица 8

**Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации**

Форма промежуточной аттестации/вид промежуточной аттестации	Коды компетенций	Индикаторы компетенций (в соот. с Таблицей 1)	Коды ЗУВ (в соот. с Таблицей 1)	Оценка по итогам освоения дисциплины (в 100-балльной шкале)	Результаты текущего контроля
зачет с оценкой / устный ответ на вопросы	ПК-5 ПК-6	ИД.ПК-5.1.	З (ПК-5)	81–100	Зачтено, отлично
		ИД.ПК-5.2.	У (ПК-5)	61–80	Зачтено, хорошо
		ИД.ПК-5.3. ИД.ПК-5.4.	В (ПК-5) З (ПК-6)	41–60	Зачтено, удовлетворительно
		ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	У (ПК-6) В (ПК-6)	0–40	Не зачтено, неудовлетворительно

Результаты промежуточной аттестации по дисциплине во втором модуле, выраженные в оценках «зачтено, удовлетворительно», «зачтено, хорошо», «зачтено, отлично», свидетельствуют о сформированности у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Исследовательская экономика» по направлению подготовки 38.04.01 Экономика (уровень магистратуры).

Результат промежуточной аттестации по дисциплине во втором модуле, выраженный в оценке «не зачтено, неудовлетворительно», свидетельствует об отсутствии или критическом уровне сформированности у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Исследовательская экономика» по направлению подготовки 38.04.01 Экономика (уровень магистратуры).

## 7.4 Типовые задания к промежуточной аттестации

### 7.4.1. Перечень вопросов для подготовки устному ответу на вопросы

1. Препроцессинг. Масштабирование. Нормировка. Полиномиальные признаки. One-hot encoding.
2. Кластеризация. kMeans, MeanShift, DBSCAN, Affinity Propagation.
3. Смещение и дисперсия (bias and variance). Понятие средней гипотезы.
4. Ансамблевые методы. Soft and Hard Voting. Bagging. Случайные леса. AdaBoost.
5. Типы обучения: с учителем, без учителя, с подкреплением, с частичным участием учителя, активное обучение.
6. Бустинг деревьев решений.
7. Ошибка внутри и вне выборки. Ошибка обобщения. Неравенство Хёфдинга. Валидация и кросс-валидация.
8. Линейная регрессия. Полиномиальная регрессия. Гребневая регрессия.
9. Размерность Вапника-Червоненкиса. Размерность Вапника-Червоненкиса для перцептрона.
10. Логистическая регрессия. Градиентный спуск.



## 7.5 Средства оценки индикаторов достижения компетенций

Таблица 9

### Средства оценки индикаторов достижения компетенций

Коды компетенций	Индикаторы компетенций (в соот.с Таблицей 1)	Средства оценки (в соот. с Таблицами 5, 7)
ПК-5	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4.	практическое задание (ПЗ), устный ответ на вопросы (УО)
ПК-6	ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	практическое задание (ПЗ), устный ответ на вопросы (УО)

Таблица 10

### Описание средств оценки индикаторов достижения компетенций

Средства оценки (в соот. с Таблицами 5, 7)	Рекомендованный план выполнения работы
Практическое задание	Магистрант в ходе подготовки и выполнения домашних заданий по темам дисциплины, показывает способность совершать следующий набор профессиональных действий, получивший развитие в рамках данной дисциплины: 1. Применяет математические методы и информационные технологии, проводит статистические расчеты, работает с различными источниками статистической информации, ведет аналитическую работу с данными, разрабатывает и совершенствует методологии сбора и обработки статистических данных 2. Решает задачи аналитического характера, в том числе осуществляет выбор актуальных способов решения задач, проводит теоретические обобщения научных данных в соответствии с задачами исследования
Устный ответ на вопросы	Магистрант в ходе подготовки и ответа на вопросы по темам дисциплины, показывает способность совершать следующий набор профессиональных действий, получивший развитие в рамках данной дисциплины: 1. Применяет математические методы и информационные технологии, проводит статистические расчеты, работает с различными источниками статистической информации, ведет аналитическую работу с данными, разрабатывает и совершенствует методологии сбора и обработки статистических данных 2. Решает задачи аналитического характера, в том числе осуществляет выбор актуальных способов решения задач, проводит теоретические обобщения научных данных в соответствии с задачами исследования

## 8. ОСНОВНАЯ И ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА

### 8.1. Основная литература

1. Кук, Д. Машинное обучение с использованием библиотеки H2O / Д. Кук; пер. с англ. А.Б. Огурцова. - Москва: ДМК Пресс, 2018. - 250 с. - ISBN 978-5-97060-508-0. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1028135>
2. Обработка изображений с помощью OpenCV / Глория Буэно Гарсия [и др.] ; пер. с англ. А.А. Слинкина. - Москва: ДМК Пресс, 2016. - 210 с. - ISBN 978-5-970(0)-387-1. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1028080>

### 8.2. Дополнительная литература

1. Груздев, А.В. Прогнозное моделирование в IBM SPSS Statistics и R: Метод деревьев решений / А.В. Груздев. - Москва: ДМК Пресс, 2016. - 278 с. - ISBN 978-5-97060-456-4. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1028064>
2. Горелов, В.И. Анализ статистических данных: практикум: [16+] / В.И. Горелов, Т.Н. Ледашева; Российская международная академия туризма. – Москва: Университетская книга, 2015. – 120 с.: ил. – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=574944>.

## **9. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ, ИСПОЛЬЗУЕМЫЕ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА**

### **9.1 Программное обеспечение**

При осуществлении образовательного процесса магистрантами и профессорско-преподавательским составом используется следующее лицензионное программное обеспечение:

1. OS Microsoft Windows (OVS OS Platform)
2. MS Office (OVS Office Platform)
3. Adobe Acrobat Professional 11.0 MLP AOO License RU
4. Adobe CS5.5 Design Standart Win IE EDU CLP
5. ABBYY FineReader 11 Corporate Edition
6. ABBYY Lingvo x5
7. Adobe Photoshop Extended CS6 13.0 MLP AOO License RU
8. Adobe Acrobat Reader DC /Pro – бесплатно
9. Google Chrome – бесплатно
10. Opera – бесплатно
11. Mozilla – бесплатно
12. VLC – бесплатно
13. R — бесплатно

### **9.2 Перечень информационно-справочных систем и профессиональных баз данных информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины:**

#### **Информационно-справочные системы**

1. Гарант.Ру. Информационно-правовой портал: <http://www.garant.ru>
2. Информационная система «Единое окно доступа к образовательным ресурсам»: <http://window.edu.ru/>
3. Открытое образование. Ассоциация «Национальная платформа открытого образования»: <http://npoad.ru>
4. Официальная Россия. Сервер органов государственной власти Российской Федерации: <http://www.gov.ru>
5. Официальный интернет-портал правовой информации. Государственная система правовой информации: <http://pravo.gov.ru>
6. Правовой сайт КонсультантПлюс: <http://www.consultant.ru/sys>
7. Российское образование. Федеральный портал: <http://www.edu.ru>

#### **Профессиональные базы данных информационно-телекоммуникационной сети «Интернет»:**

1. Google. Книги: <https://books.google.com>
2. Internet Archive: <https://archive.org>
3. Koob.ru. Электронная библиотека «Куб»: <http://www.koob.ru/philosophy/>
4. Библиотека Гумер – гуманитарные науки: <http://www.gumer.info>
5. Библиотека Ихтика [ihtik.lib.ru]: <http://ihtik.lib.ru/>
6. Докусфера — Российская национальная библиотека: <http://leb.nlr.ru>
7. ЕНИП — Электронная библиотека «Научное наследие России»: <http://e-heritage.ru/>
8. Интелрос. Интеллектуальная Россия: <http://www.intelros.ru/>
9. Национальная электронная библиотека НЭБ: <http://www.rusneb.ru>
10. Президентская библиотека: <http://www.prlib.ru>
11. Российская государственная библиотека: <http://www.rsl.ru/>
12. Российская национальная библиотека: <http://www.nlr.ru/poisk/>

### 9.3 Лицензионные электронные ресурсы библиотеки Университета

#### Профессиональные базы данных:

1. **Cambridge University Press** — полнотекстовая коллекция журналов издательства Cambridge University Press: <https://www.cambridge.org/>
2. **East View** – 100 ведущих российских журналов по гуманитарным наукам (архив и текущая подписка): [https://dlib.eastview.com/browse](https://dlib.eastview.com/browse;);
3. **Ebook Central** коллекция электронных книг **Academic Complete** библиотеки компании ProQuest — **Ebook Central** — более 140 тыс. электронных научных книг крупнейших издательств мира: <https://ebookcentral.proquest.com/>;
4. **EBSCO** – научные журналы, справочники, полнотекстовые и многопрофильные базы данных: <http://search.ebscohost.com/>;
5. **eLIBRARY.RU** — Российский информационно-аналитический портал в области науки, технологии, медицины и образования, содержащий рефераты и полные тексты научных статей и публикаций, наукометрическая база данных: <http://elibrary.ru>;
6. **JSTOR** – полнотекстовая база данных междисциплинарного характера, включающая более тысячи научных журналов по гуманитарным, социальным наукам и математике с их первого выпуска: <http://www.jstor.org/>;
7. **Oxford Reference Online** — словари издательства Oxford University Press - <http://www.oxfordreference.com/>;
8. **Oxford University Press** — полнотекстовая коллекция журналов издательства Oxford University Press (текущая подписка и архив): <http://www.oxfordjournals.org/en/>;
9. **Project MUSE Standard Collection** — полные тексты более чем 300 журналов по гуманитарным наукам зарубежных научных издательств: <http://muse.jhu.edu/>;
10. **ProQuest Dissertations & Theses** — база диссертаций и дипломных работ: <http://search.proquest.com/pqdtglobal/dissertations>;
11. **Sage** — полнотекстовая коллекция журналов издательства Sage (текущая подписка и архив): <http://online.sagepub.com/>;
12. **SCOPUS** – реферативная наукометрическая база данных: <https://www.scopus.com/>;
13. **Taylor&Francis** – полнотекстовая коллекция журналов издательства Taylor&Francis (текущая подписка и архив) – <http://www.tandfonline.com/>;
14. **Web of Science** — реферативная наукометрическая база данных: <http://apps.webofknowledge.com>;
15. **Университетская информационная система РОССИЯ** — база электронных ресурсов для учебных программ и исследовательских проектов в области социально-гуманитарных наук: <http://www.uisrussia.msu.ru/>;
16. Электронные журналы по подписке (текущие номера научных зарубежных журналов)

#### Электронные библиотечные системы:

1. **Znanium.com** – Электронная библиотечная система (ЭБС) – <http://znanium.com/>;
2. Университетская библиотека онлайн – Электронная библиотечная система (ЭБС) – <http://biblioclub.ru/>

### 9.4 Электронная информационно-образовательная среда Университета

Образовательный процесс по дисциплине поддерживается средствами электронной информационно-образовательной среды Университета, которая включает в себя электронный учебно-методический ресурс АНООВО «ЕУСПб» — образовательный портал LMS Sakai — Sakai@EU, лицензионные электронные ресурсы библиотеки Университета, официальный сайт Университета (Европейский университет в Санкт-Петербурге [<https://eu.spb.ru>]), локальную сеть и корпоративную электронную почту Университета, и обеспечивает:

- доступ к учебным планам, рабочим программам дисциплин (модулей), практик и к изданиям электронных библиотечных систем и электронным образовательным ресурсам, указанным в рабочих программах;
- фиксацию хода образовательного процесса, результатов промежуточной аттестации и результатов освоения основной образовательной программы;
- формирование электронного портфолио обучающегося, в том числе сохранение работ обучающегося, рецензий и оценок за эти работы со стороны любых участников образовательного процесса;
- взаимодействие между участниками образовательного процесса, в том числе синхронное и (или) асинхронное взаимодействие посредством сети «Интернет» (электронной почты и т.д.).

Каждый обучающийся в течение всего периода обучения обеспечен индивидуальным неограниченным доступом к электронным ресурсам библиотеки Университета, содержащей издания учебной, учебно-методической и иной литературы по изучаемой дисциплине.

## **10. МАТЕРИАЛЬНО-ТЕХНИЧЕСКАЯ БАЗА, НЕОБХОДИМАЯ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА**

В ходе реализации образовательного процесса используются специализированные многофункциональные аудитории для проведения занятий лекционного типа, занятий семинарского типа (практических занятий, лабораторных работ), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, укомплектованные специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Проведение занятий лекционного типа обеспечивается демонстрационным оборудованием.

Помещения для самостоятельной работы оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду организации.

**Для лиц с ограниченными возможностями здоровья и инвалидов** предоставляется возможность присутствия в аудитории вместе с ними ассистента (помощника). Для слабовидящих предоставляется возможность увеличения текста на экране ПК. Для самостоятельной работы лиц с ограниченными возможностями здоровья в помещении для самостоятельной работы организовано одно место (ПК) с возможностями бесконтактного ввода информации и управления компьютером (специализированное лицензионное программное обеспечение – Camera Mouse, веб камера). Библиотека университета предоставляет удаленный доступ к электронным ресурсам библиотеки Университета с возможностями для слабовидящих увеличения текста на экране ПК. Лица с ограниченными возможностями здоровья могут при необходимости воспользоваться имеющимся в университете креслом-коляской. В учебном корпусе имеется адаптированный лифт. На первом этаже оборудован специализированный туалет. У входа в здание университета для инвалидов оборудована специальная кнопка, входная среда обеспечена информационной доской о режиме работы университета, выполненной рельефно-точечным тактильным шрифтом (азбука Брайля).

**ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ПО ДИСЦИПЛИНЕ  
«Основы машинного обучения»**

## ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

### 1. Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации

Информация о содержании и процедуре текущего контроля успеваемости, методике оценивания знаний, умений и навыков обучающегося в ходе текущего контроля доводятся научно-педагогическими работниками Университета до сведения обучающегося на первом занятии по данной дисциплине.

Текущий контроль предусматривает подготовку магистрантов к каждому занятию, активное слушание на лекциях, выполнение практических заданий. Магистрант должен присутствовать на лекциях, отвечать на поставленные вопросы, показывая, что прочитал разбираемую литературу, представлять содержательные реплики по темам обсуждения.

Текущий контроль проводится в форме оценивания выполнения практических заданий, демонстрирующих степень знакомства магистрантов с дополнительной литературой.

Таблица 1

#### Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации

Наименование тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. с Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
Типы задач машинного обучения	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 1	зачтено/ не зачтено
Метрические классификаторы	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 2	зачтено/ не зачтено
Алгоритмы кластеризации	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 3	зачтено/ не зачтено
Деревья решений	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 4	зачтено/ не зачтено
Линейные классификаторы	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 5	зачтено/ не зачтено

Наименование тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. с Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
Нейронные сети и глубокое обучение	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 6	зачтено/ не зачтено
Регрессионный анализ	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 7	зачтено/ не зачтено
Ансамблевые методы	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 8	зачтено/ не зачтено
Стохастический поиск	ПК-5 ПК-6	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 9	зачтено/ не зачтено

При освоении дисциплины опросы оцениваются с использованием бинарной системы, а домашние задания – с использованием 100-балльной шкалы с последующим переводом в бинарную систему для получения результатов текущего контроля, фиксирующих ход образовательного процесса, согласно Таблице 2.

Таблица 2

### Критерии оценивания

Формы текущего контроля успеваемости	Описание	Показатели	Количество баллов по 100-балльной шкале	Результаты текущего контроля
Практическое задание	Выполнение домашних заданий подразумевает письменное изложение ответов на теоретические вопросы и решений поставленных преподавателем задач. Предполагается, что ход решения задач сопровождается подробными комментариями обучающегося	Обучающийся правильно выполняет задания и отвечает на все поставленные в задании вопросы.	81–100	зачтено
		Обучающийся выполняет предложенные задания и отвечает на все поставленные в задании вопросы, но допускает при этом несущественные ошибки.	61–80	
		Обучающийся предлагает правильную идеологию решения задач.	41–60	

Формы текущего контроля успеваемости	Описание	Показатели	Количество баллов по 100-балльной шкале	Результаты текущего контроля
		Обучающийся испытывает затруднения при выполнении предложенных заданий.	0–40	не зачтено

## 2 Контрольные задания для текущей аттестации

### Материал для практических заданий

#### Практическое задание 1:

- Реализуйте алгоритм kNN классификации по k ближайшим соседям, используя простое евклидовое расстояние.
- Реализуйте алгоритм k-means для кластеризации на 2-4 кластера.
- Реализуйте алгоритм DBSCAN, найдите параметры для кластеризации на 4 кластера.

#### Практическое задание 2:

- Реализуйте алгоритмы построения дерева с критерием информационного выигрыша и критерием Джини и определению класса по мажоритарному классу в листе. Найдите оптимальную глубину дерева в обоих случаях (в отрезке 2-10).
- Примените метод SVM (например, из библиотеки sklearn) для датасета blobs2. Визуализируйте результат (разбиение плоскости и опорные вектора) при разных вариантах ядер (линейное; полиномиальное степеней 2,3,5; RBF).
- Реализуйте алгоритм логистической регрессии со стохастическим градиентным спуском, обучите его на датасете spambase\_old (train) и проверьте на датасете spambase\_new (val). Получите ROC кривые для вариантов без нормировки и с нормировкой признаков.

#### Практическое задание 3:

- Реализуйте алгоритм линейной регрессии, и полиномиальной регрессии (для датасета noisysine – степеней от 2 до 5, для датасета hydrodynamics – степени 2) без регуляризации.
- Реализуйте алгоритм гребневой регрессии и найдите оптимальный параметр регуляризации для случаев из задачи 1.
- Найдите максимум функции с помощью алгоритма кросс-энтропийного поиска, изображая распределение на каждом шаге.
- Найдите лучший путь в задаче коммивояжера с помощью алгоритма отжига.

## 3 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации

Форма промежуточной аттестации – **зачет с оценкой**, при выставлении которого учитываются результаты текущего контроля успеваемости обучающегося по дисциплине и результаты устного ответа на вопросы, выраженные в 100-балльной шкале.

Перед зачетом с оценкой проводятся консультация, на которой преподаватель отвечает на вопросы обучающихся.

Критерии оценивания письменной зачетной работы приведены в таблице 3.



**Критерии оценивания письменной зачетной работы**

Вид промежуточной аттестации	Показатели	Количество баллов
Устный ответ на вопросы	<p>Магистрант дает ответы на вопросы билета, для которых характерно:</p> <ul style="list-style-type: none"> <li>• глубокое усвоение программного материала,</li> <li>• изложение его исчерпывающе, последовательно, четко,</li> <li>• умение делать обоснованные выводы,</li> <li>• соблюдение норм устной и письменной литературной речи;</li> </ul>	81–100
	<p>Магистрант верно отвечает на вопрос, указанный в билете, при условии, что ответ на вопрос характеризуется отсутствием серьезных, значимых неточностей, при следующих характеристиках ответа:</p> <ul style="list-style-type: none"> <li>• твердое знание материала курса,</li> <li>• последовательное изложение материала,</li> <li>• знание теоретических положений без обоснованной их аргументации,</li> <li>• соблюдение норм устной и письменной литературной речи;</li> </ul>	61–80
	<p>Магистрант представляет правильный ответ на теоретический вопрос, указанный в билете, при условии, что ответ на вопрос характеризуется значительными неточностями, при следующих параметрах ответа:</p> <ul style="list-style-type: none"> <li>• знание основного материала, но владение им не в полном объеме,</li> <li>• допущение существенных неточностей, недостаточно правильных формулировок,</li> <li>• допущение нарушения логической последовательности в изложении материала,</li> <li>• наличие нарушений норм литературной устной и письменной речи.</li> </ul>	41–60
	<p>Магистрант представляет ответ на вопрос билета, свидетельствующий о некомпетентности магистранта, при следующих параметрах ответа:</p> <ul style="list-style-type: none"> <li>• незнание значительной части программного материала,</li> <li>• наличие существенных ошибок в определениях, формулировках, понимании теоретических положений;</li> <li>• бессистемность при ответе на поставленный вопрос,</li> <li>• отсутствие в ответе логически корректного анализа, аргументации, классификации,</li> <li>• наличие нарушений норм устной и письменной литературной речи.</li> </ul>	0–40

Максимальная оценка по итогам освоения дисциплины составляет 100 баллов (50% оценки составляют результаты практических заданий по темам курса, 50% – результаты устного ответа на вопросы). При вычислении оценки G обучающегося по итогам освоения дисциплины используется расчетная формула

$$G = 0.5R + 0.5F,$$

после чего проводится процедура округления до целого. Здесь R – количество баллов за выполнение практических заданий, F – количество баллов, полученных за письменную зачетную работу.

В результате промежуточного контроля знаний студенты получают аттестацию по дисциплине. На основании оценки обучающегося по итогам освоения дисциплины, выраженной в 100-балльной шкале, выставляется **зачет с оценкой** в соответствии с Таблицей 4.

Таблица 4

**Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации**

Форма промежуточной аттестации/вид промежуточной аттестации	Коды компетенций	Индикаторы компетенций (в соот. с Таблицей 1)	Коды ЗУВ (в соот. с Таблицей 1)	Оценка по итогам освоения дисциплины (в 100-балльной шкале)	Результаты текущего контроля
зачет с оценкой / устный ответ на вопросы	ПК-5 ПК-6	ИД.ПК-5.1.	З (ПК-5)	81–100	Зачтено, отлично
		ИД.ПК-5.2.	У (ПК-5)	61–80	Зачтено, хорошо
		ИД.ПК-5.3.	В (ПК-5)	41–60	Зачтено, удовлетворительно
		ИД.ПК-5.4.	З (ПК-6)	0–40	Не зачтено, неудовлетворительно
		ИД.ПК-6.1.	У (ПК-6)		
		ИД.ПК-6.2.	В (ПК-6)		
		ИД.ПК-6.3.			

Результаты промежуточной аттестации по дисциплине во втором модуле, выраженные в оценках «зачтено, удовлетворительно», «зачтено, хорошо», «зачтено, отлично», свидетельствуют о сформированности у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Исследовательская экономика» по направлению подготовки 38.04.01 Экономика (уровень магистратуры).

Результат промежуточной аттестации по дисциплине во втором модуле, выраженный в оценке «не зачтено, неудовлетворительно», свидетельствует об отсутствии или критическом уровне сформированности у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Исследовательская экономика» по направлению подготовки 38.04.01 Экономика (уровень магистратуры).

#### 4 Задания к промежуточной аттестации

##### 4.1. Перечень вопросов для подготовки устному ответу на вопросы

1. Препроцессинг. Масштабирование. Нормировка. Полиномиальные признаки. One-hot encoding.
2. Кластеризация. kMeans, MeanShift, DBSCAN, Affinity Propagation.
3. Смещение и дисперсия (bias and variance). Понятие средней гипотезы.
4. Ансамблевые методы. Soft and Hard Voting. Bagging. Случайные леса. AdaBoost.
5. Типы обучения: с учителем, без учителя, с подкреплением, с частичным участием учителя, активное обучение.
6. Бустинг деревьев решений.
7. Ошибка внутри и вне выборки. Ошибка обобщения. Неравенство Хёфдинга. Валидация и кросс-валидация.
8. Линейная регрессия. Полиномиальная регрессия. Гребневая регрессия.
9. Размерность Вапника-Червоненкиса. Размерность Вапника-Червоненкиса для перцептрона.
10. Логистическая регрессия. Градиентный спуск.
11. Пороговые условия. Эффективность по Парето. Precision-Recall и ROC кривые. AUC.

12. Ансамблевые методы регрессии. RANSAC. Theil-Sen. Huber.
13. Перцептрон. Перцептрон с карманом.
14. Метод опорных векторов. Постановка задачи. Формулировка и решение двойственной задачи. Типы опорных векторов. Ядра.
15. Гипотезы и дихотомии. Функция роста. Точка поломки. Доказательство полиномиальности функции роста в присутствии точки поломки.
16. Деревья решений. Информационный выигрыш, критерий Джини. Регуляризация деревьев. Небрежные решающие деревья.
17. Байесовский классификатор. Типы оценки распределений признаков (Gaussian, Bernoulli, Multinomial). EM алгоритм.
18. Нейронные сети. Перцептрон Розенблатта. Функции активации. Обратное распространение градиента. Softmax.
19. Стохастическая оптимизация. Hill Climb. Отжиг. Генетический алгоритм.
20. Метрические классификаторы. kNN. WkNN. Отбор эталонов. DROP5. Kdtree

## 5 Средства оценки индикаторов достижения компетенций

Таблица 5

### Средства оценки индикаторов достижения компетенций

Коды компетенций	Индикаторы компетенций (в соот.с Таблицей 1)	Средства оценки (в соот. с Таблицами 5, 7)
ПК-5	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4.	практическое задание (ПЗ), устный ответ на вопросы (УО)
ПК-6	ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3.	практическое задание (ПЗ), устный ответ на вопросы (УО)

Таблица 6

### Описание средств оценки индикаторов достижения компетенций

Средства оценки (в соот. с Таблицами 5, 7)	Рекомендованный план выполнения работы
Практическое задание	Магистрант в ходе подготовки и выполнения домашних заданий по темам дисциплины, показывает способность совершать следующий набор профессиональных действий, получивший развитие в рамках данной дисциплины: 1. Применяет математические методы и информационные технологии, проводит статистические расчеты, работает с различными источниками статистической информации, ведет аналитическую работу с данными, разрабатывает и совершенствует методологии сбора и обработки статистических данных 2. Решает задачи аналитического характера, в том числе осуществляет выбор актуальных способов решения задач, проводит теоретические обобщения научных данных в соответствии с задачами исследования
Устный ответ на вопросы	Магистрант в ходе подготовки и ответа на вопросы по темам дисциплины, показывает способность совершать следующий набор профессиональных действий, получивший развитие в рамках данной дисциплины: 1. Применяет математические методы и информационные технологии, проводит статистические расчеты, работает с различными источниками статистической информации, ведет аналитическую работу с данными, разрабатывает и совершенствует методологии сбора и обработки статистических данных 2. Решает задачи аналитического характера, в том числе осуществляет выбор актуальных способов решения задач, проводит теоретические обобщения научных данных в соответствии с задачами исследования