

**Федеральное государственное бюджетное образовательное  
учреждение высшего образования  
«РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА  
И ГОСУДАРСТВЕННОЙ СЛУЖБЫ  
ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»**

---

**ИНСТИТУТ ЭКОНОМИКИ, МАТЕМАТИКИ И ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ**  
кафедра Эконометрики и математической экономики

**УТВЕРЖДЕНА**  
на заседании кафедры Системного анализа  
и информатики  
Протокол от «03» июня 2020 г. № 10

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

**Б1.О.04 «Машинное обучение»**

по направлению подготовки 38.04.01 «Экономика»

направленность «Системы больших данных в экономике»

квалификация магистр

очная форма обучения

Год набора - 2021

Москва, 2021 г.

**Автор—составитель:**

д.ф.м.н. Сенько О.В.

**Заведующий кафедрой**

Системного анализа и информатики к.т.н. Маруев С.А.

## СОДЕРЖАНИЕ

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы.....	4
2. Объем и место дисциплины в структуре образовательной программы.....	4
3. Содержание и структура дисциплины.....	4
4. Материалы текущего контроля успеваемости обучающихся и фонд оценочных средств промежуточной аттестации по дисциплине.....	5
5. Методические указания для обучающихся по освоению дисциплины .....	8
6. Учебная литература и ресурсы информационно-телекоммуникационной сети «Интернет», учебно-методическое обеспечение самостоятельной работы обучающихся по дисциплине .....	8
6.1. Основная литература .....	8
6.2. Дополнительная литература .....	8
6.3. Учебно-методическое обеспечение самостоятельной работы .....	8
6.4. Нормативные правовые документы .....	8
6.5. Интернет-ресурсы .....	9
6.6. Иные источники .....	9
7. Материально-техническая база, информационные технологии, программное обеспечение и информационные справочные системы .....	9

## 1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения программы

1.1. Дисциплина Б1.О.04 «Машинное обучение» обеспечивает овладение следующими компетенциями:

Код компетенции	Наименование компетенции	Код этапа освоения компетенции	Наименование этапа освоения компетенции
ОПК ОС-3	Способен применять продвинутое инструментальные методы экономического анализа в прикладных и (или) фундаментальных исследованиях	ОПК ОС-3.2	способен использовать методы машинного обучения при решении прикладных экономических задач
ОПК ОС-5	Способен использовать современные информационные технологии и программные средства при решении практических и (или) исследовательских задач	ОПК ОС-5.2	способен использовать современные программные средства для решения прикладных задач

1.2. В результате освоения дисциплины у студентов должны быть сформированы:

ОТФ/ТФ профессиональные действия	Код этапа освоения компетенции	Результаты обучения
ведение аналитической деятельности в области экономики	ОПК ОС-3.2	на уровне знаний: знание основных алгоритмов машинного обучения и областей их применения на уровне умений: уметь отбирать алгоритмы машинного обучения для решения поставленных задач на уровне навыков: осуществлять поиск оптимального алгоритма машинного обучения, используя кроссвалидацию и решетчатый поиск
	ОПК ОС-5.2	на уровне знаний: знает синтаксис Python на уровне умений: уметь настраивать алгоритмы машинного обучения для решения поставленных задач

## 2. Объем и место дисциплины в структуре ОП ВО

Общий объем дисциплины 4 ЗЕ (144 ак.ч./108 астр.ч.). В целях планирования, при расчете структуры дисциплины применены академические часы (ак.ч.).

Количество академических часов, выделенных на занятия лекционного типа – 32 ак.ч., занятия семинарского типа – 32 ак.ч.; на самостоятельную работу обучающихся – 42 ак.ч. и промежуточную аттестацию – 36 ак.ч.

Место дисциплины в структуре образовательной программы:

- дисциплина Б1.О.04 «Машинное обучение» изучается во 2 семестре на 1 курсе;
- дисциплина реализуется после изучения дисциплин:

Математические инструменты в экономических исследованиях

- дисциплина может реализоваться частично или полностью с применением ЭО и/или ДОТ. Учебные материалы дисциплины размещаются по адресу [lms.ranepa.ru](https://lms.ranepa.ru)

форма промежуточной аттестации – экзамен.

### 3. Содержание и структура дисциплины

№ п/п	Наименование тем (разделов)	Объем дисциплины (модуля), ак. час./ час.						Форма текущего контроля успеваемости*, промежуточной аттестации
		Всего	Контактная работа обучающихся с преподавателем по видам учебных занятий**				СР	
			Л	ЛР	ПЗ	КСР		
Очная форма обучения								
Тема 1	Введение в машинное обучение	20	8		4		8	ДЗ
Тема 2	Классификация (обучение с учителем)	32	8		12		12	ДЗ
Тема 3	Регрессия (обучение с учителем)	32	8		12		12	ДЗ
Тема 4	Кластерный анализ. Обучение без учителя.	22	8		4		10	ДЗ
		2						консультация
Промежуточная аттестация		36						Экзамен
Всего:		144/108	32/24		32/24		42/31,5	

Примечание\* – формы текущего контроля успеваемости: домашнее задание (ДЗ)

Примечание \*\*: в рамках указанной контактной работы с обучающимися учебные занятия могут проводиться с использованием ДОТ и/или ЭО

### Содержание дисциплины

#### Тема 1. Введение в машинное обучение

Основные задачи, решаемые машинным обучением. Обобщающая способность. Задачи распознавания и регрессионного анализа. Методы оценки эффективности алгоритмов машинного обучения. Кросс-валидация.

#### Тема 2. Классификация (распознавание), обучение с учителем.

Область применения. Метрики, оценивающие эффективность решения задач классификации. ROC-анализ. Методы машинного обучения, используемые для решения задач распознавания байесовские классификаторы, логистическая регрессия, метод k-ближайших соседей, деревья решений, случайный лес, адаптивный бустинг, градиентный бустинг, метод опорных векторов, нейронные сети. Предварительная подготовка данных. Снижение размерности. Работа с категориальными переменными. Подбор оптимальных параметров обучения.

#### Тема 3. Регрессия (обучение с учителем).

Область применения. Метрики, оценивающие эффективность решения задач регрессионного анализа. Поиск оптимальных регрессий с помощью метода наименьших квадратов. Методы, основанные на регуляризации. Эластичная сеть Регрессионные деревья и леса.

#### Тема 4. Обучение без учителя.

Область применения. Алгоритмы кластеризации. Метрики качества кластеризации.

### 4. Материалы текущего контроля успеваемости обучающихся и фонд оценочных средств промежуточной аттестации по дисциплине

#### **4.1. Формы и методы текущего контроля успеваемости обучающихся и промежуточной аттестации.**

**4.1.1. В ходе реализации дисциплины Б1.О.04 «Машинное обучение» используются следующие методы текущего контроля успеваемости обучающихся:**

Тема (раздел)	Методы текущего контроля успеваемости
Тема 1	Домашнее задание 1
Тема 2	Домашнее задание 2
Тема 3	Домашнее задание 3
Тема 4	Домашнее задание 4

**4.1.2. Экзамен проводится с применением следующих методов (средств):** экзамен проводится в виде устного собеседования по результатам выполнения мини-кейса.

#### **4.2. Материалы текущего контроля успеваемости обучающихся**

##### **Типовые оценочные материалы по теме 1**

Домашнее задание 1. Датасет breast cancer

Решить задачу распознавания с использованием метода логистическая регрессия и оценить точность в режиме кросс-валидации

##### **Типовые оценочные материалы по теме 2**

Домашнее задание 2. Датасет

Решить задачу распознавания с использованием нескольких методов и построить соответствующие ROC- кривые

##### **Типовые оценочные материалы по теме 3**

Домашнее задание 3. Датасет city bikes, forecast apple.

Решить задачу регрессионного анализа с отбором признаков.

##### **Типовые оценочные материалы по теме 4**

Домашнее задание 4. Провести кластеризацию с помощью метода k-средних и иерархической кластеризации

#### **4.3. Оценочные средства для промежуточной аттестации.**

**4.3.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы. Показатели и критерии оценивания компетенций с учетом этапа их формирования**

Код компетенции	Наименование компетенции	Код этапа освоения компетенции	Наименование этапа освоения компетенции
ОПК ОС-3	Способен применять продвинутое инструментальные методы экономического анализа в прикладных и (или) фундаментальных исследованиях	ОПК ОС-3.2	способен использовать методы машинного обучения при решении прикладных экономических задач
ОПК ОС-5	Способен использовать современные информационные технологии и программные средства при решении практических и (или) исследовательских задач	ОПК ОС-5.2	способен использовать современные программные средства для решения прикладных задач

Этап освоения компетенции	Показатель оценивания	Критерий оценивания
ОПК ОС-3.2	способен использовать методы машинного обучения при решении прикладных экономических задач	демонстрирует понимание методов машинного обучения, умеет отбирать алгоритм для решения поставленных задач
ОПК ОС-5.2	способен решать задачу машинного обучения	задача решена верно, код оптимален и работоспособен. настроен оптимальный алгоритм машинного обучения

### 4.3.2 Типовые оценочные средства

Мини-кейс. Обработать исходные данные и построить модель. Исходные данные предоставляются преподавателем. Решение о результативности принимается на основе контрольной выборки преподавателя. Пример. Задача прогнозирования отказов при выплате по банковским кредитам.

#### Шкала оценивания.

Оценка определяется по формуле:

$$I = \left[ \frac{\sum_{i=1}^n B_i}{n} + B_{mc} \right] * 0.5 \quad (1)$$

где  $n$  – домашних заданий,  $B_i$  – оценка за домашнее задание по 10-ти бальной шкале,  $I$  – итоговое количество баллов.

$B_{mc}$  -оценка за мини кейс по 10-ти бальной шкале

10-бальная шкала	Традиционная шкала	«Зачтено»/«Не зачтено»	Определение
10	Отлично	Зачтено	Поставленная задача решена верно, код оптимален и работоспособен. Анализ решения проведен в полном объеме.
9	Отлично	Зачтено	Поставленная задача решена верно, код частично оптимален и работоспособен. Анализ решения проведен в полном объеме.
8	Отлично	Зачтено	Поставленная задача решена верно, код работоспособен. Анализ решения проведен в полном объеме.
7	Хорошо	Зачтено	Поставленная задача решена верно, код работоспособен. Анализ решения проведен в достаточном объеме.
6	Хорошо	Зачтено	Поставленная задача решена верно, код работоспособен. Проведен анализ решения.
5	Удовлетворительно	Зачтено	Поставленная задача решена верно, код работоспособен. Проведен неполный анализ решения.
4	Удовлетворительно	Зачтено	Поставленная задача решена верно, код работоспособен. Проведен неполный анализ решения с частичными ошибками.
3	Неудовлетворительно	Не зачтено	Поставленная задача решена с неверным ответом.
2	Неудовлетворительно	Не зачтено	Поставленная задача не решена, представленный студентом код частично работоспособен
1	Неудовлетворительно	Не зачтено	Поставленная задача не решена, представленный студентом код не работоспособен.
0	Неудовлетворительно	Не зачтено	Поставленная задача не выполнена, код не представлен

### 4.4. Методические материалы

Экзамен проводится в компьютерной аудитории, с установленным пакетом Anaconda

с Python 3, с доступом в локальную сеть Академии и глобальную сеть Интернет. Обучающейся решает задание кейса, выбирает единственную модель и проверяет ее на контрольной выборке преподавателя на ложные решения, с обязательным построением матрицы ошибок. При проверке корректура исходной модели не допускается.

Обучающейся имеет право пользоваться сетью Интернет при решении кейса. Обучающийся обязан объяснить каждую строчку кода.

Время на выполнение кейса 1 час. Результаты сохраняются на рабочем столе. Время на проверку контрольной выборкой не более 15 мин. Результаты сохраняются на рабочем столе. Обучающиеся покидают аудиторию, и по установленной очереди заходят в аудиторию для защиты кейса. Время на защиту кейса не более 7 мин.

В случае, если дисциплина полностью или частично проводилась с применением технологий электронного обучения и/или дистанционных технологий, экзамен может производиться с использованием системы СДО Академии и применением прокторинга.

## **5. Методические указания для обучающихся по освоению дисциплины**

Кейсы и домашние задания для самостоятельной подготовки могут быть опубликованы:

1. В блоге преподавателя на сайте <http://economy.ranepa.ru/prepodavатели/>;
2. В группе ВКонтакте;
3. Высланы на общую почту группы

Конкретный способ коммуникации со студентами определяется преподавателем.

Для подготовки к практическим занятиям необходимо ознакомиться с рекомендованной литературой, повторить материал предыдущих практических занятий. При возникновении вопросов обратиться к преподавателю по электронной почте с указанием конкретной проблемы и (или) прийти к преподавателю на консультацию в установленное время.

Для подготовки письменных домашних заданий необходимо ознакомиться с рекомендованной литературой, повторить материал предыдущих практических занятий.

В течении курса преподаватель вправе предлагать студентам дополнительные задания повышенной сложности для начисления дополнительных баллов. Правила выполнения данных заданий и начисления баллов объявляются преподавателем индивидуально для каждого задания повышенной сложности.

Самостоятельная работа студентом осуществляется для закрепления изученного материала после практических занятий, для выполнения домашних заданий, для изучения дополнительных материалов.

## **6. Учебная литература и ресурсы информационно- телекоммуникационной сети "Интернет", включая перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине**

### **6.1. Основная литература.**

1. Сенько О.В., Цифровые методы диагностики и диагностирования процессов. НИТУ МиСиС, 2016-86 с
2. Андреас Мюллер, Сара Гвидо. Введение в машинное обучение с помощью Python. Руководство для специалистов по работе с данными. М.: Диалектика, 2017 – 480 с

6.3. Учебно-методическое обеспечение самостоятельной работы.  
Отдельное обеспечение не предусмотрено.

6.4. Нормативные правовые документы.



Не предусмотрены.

6.5. Интернет-ресурсы.

<http://www.machinelearning.ru/>  
<https://www.anaconda.com/download/>  
<http://scikit-learn.org>  
<https://pandas.pydata.org>  
<https://seaborn.pydata.org>  
<https://matplotlib.org>  
<http://lms.ranepa.ru>

6.6. Иные источники.

Не предусмотрены.

**7. Материально-техническая база, информационные технологии,  
программное обеспечение и информационные справочные системы**

1. Компьютерный класс
2. Доступ в интернет и локальную сеть Академии
3. Проекционное оборудование
4. Программное обеспечение:
  - Windows/Linux/Mac OS
  - Дистрибутив Anaconda