



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное автономное образовательное учреждение  
высшего образования  
«Дальневосточный федеральный университет»  
(ДФУ)

**ИНСТИТУТ МАТЕМАТИКИ И КОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ (ШКОЛА)**  
**«СОГЛАСОВАНО»**

Руководитель ОП

Артемяева И.Л.

**«УТВЕРЖДАЮ»**

И.о. директора департамента

Смагин С.В.

«26» января 2022 г.



**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

Технология машинного обучения

**Направление подготовки 09.03.04 Программная инженерия**

**Форма подготовки очная**

курс 4 семестр 7

лекции 16 час.

практические занятия 00 час.

лабораторные работы 34 час.

в том числе с использованием МАО лек. 0 /пр. 0 /лаб. 18 час.

всего часов аудиторной нагрузки 50 час.

в том числе с использованием МАО 18 час.

самостоятельная работа 58 час.

в том числе на подготовку к экзамену 36 час.

контрольные работы (количество) не предусмотрены

курсовая работа / курсовой проект не предусмотрены

зачет \_ семестр

экзамен 7 семестр

Рабочая программа составлена в соответствии с требованиями Федерального государственного образовательного стандарта высшего образования по направлению подготовки 09.03.04 Программная инженерия, утвержденный приказом Министерства образования и науки Российской Федерации от 19 сентября 2017 г. № 920.

Рабочая программа обсуждена на заседании департамента Программной инженерии и искусственного интеллекта, протокол № 1.1 от «24» января 2022 г.

И.о. директора департамента Программной инженерии и искусственного интеллекта С.В. Смагин, к.т.н.

Составитель: доцент кафедры ПММУиПО Смагин С.В., к.т.н.

Владивосток  
2022

**Оборотная сторона титульного листа РПУД**

**I. Рабочая программа пересмотрена на заседании кафедры:**

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

Заведующий кафедрой \_\_\_\_\_  
(подпись) (И.О. Фамилия)

**II. Рабочая программа пересмотрена на заседании кафедры:**

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

Заведующий кафедрой \_\_\_\_\_  
(подпись) (И.О. Фамилия)

**III. Рабочая программа пересмотрена на заседании кафедры:**

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

Заведующий кафедрой \_\_\_\_\_  
(подпись) (И.О. Фамилия)

**IV. Рабочая программа пересмотрена на заседании кафедры:**

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

Заведующий кафедрой \_\_\_\_\_  
(подпись) (И.О. Фамилия)

## ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

**Цель:** познакомить студентов с прикладными задачами обучения по прецедентам, изучить основные методы решения этих задач и базовые алгоритмы, реализующие эти методы.

### Задачи:

- Ознакомление с наукой о данных (Data Science), большими данными (Big Data), а также с существующими направлениями исследований и областями применения машинного обучения (Machine Learning).
- Изучение базовых понятий и терминов машинного обучения.
- Рассмотрение основных типов и примеров прикладных задач.
- Проведение обзора современных приложений машинного обучения.
- Изучение основных групп методов и базовых алгоритмов решения задач классификации, регрессии, прогнозирования, кластеризации.
- Реализация и проведение исследования одного из рассмотренных в курсе алгоритмов на основе обучающих выборок реальных данных.

Планируемые результаты обучения по данной дисциплине (знания, умения, владения), соотнесенные с планируемыми результатами освоения образовательной программы.

Универсальные компетенции выпускников и индикаторы их достижения:

Тип задач	Код и наименование профессиональной компетенции (результат освоения)	Код и наименование индикатора достижения компетенции
Проектный	ПК-6. Способность использовать навыки моделирования, анализа и формальных методов конструирования программного обеспечения	ПК-6.1. демонстрирует знание основ моделирования и формальных методов конструирования программного обеспечения.
		ПК-6.2. использует формальные методы конструирования программного обеспечения.
		ПК-6.3. применяет методы формализации и моделирования программного обеспечения.

Производственно-технологический	ПК-10. Способность использовать различные технологии разработки программного обеспечения	ПК-10.1. демонстрирует знание современных технологий разработки ПО (структурное, объектно-ориентированное).
		ПК-10.2. использует структурное и объектно-ориентированное проектирование при разработке ПО.
		ПК-10.3. применяет современные технологии разработки ПО.

<b>Код и наименование индикатора достижения компетенции</b>	<b>Наименование показателя оценивания (результата обучения по дисциплине)</b>
ПК-6.1. демонстрирует знание основ моделирования и формальных методов конструирования программного обеспечения	Знает основные модели информационных технологий, используемых при создании программ на языке высокого уровня, процесс создания программ, стандарты языков программирования.
	Умеет использовать основные модели информационных технологий при создании программ на языке высокого уровня, общие принципы построения алгоритмов, основные алгоритмические конструкции.
	Владеет основными моделями информационных технологий, используемых при создании программ на языке высокого уровня.
ПК-6.2. использует формальные методы конструирования программного обеспечения	Знает методы использования информационных технологий при подготовке документации и текстовых редакторов при разработке программ и документации.
	Умеет использовать информационные технологии при подготовке документации и разработке программ.
	Владеет средствами, предоставляемыми информационными технологиями для подготовки документации и разработке программ.
ПК-6.3. применяет методы формализации и моделирования программного обеспечения	Знает методы формализации и моделирования программного обеспечения.
	Умеет разрабатывать программы, использующие данные в произвольном формате.
	Владеет навыками разработки формальных моделей программного обеспечения.
ПК-10.1. демонстрирует знание современных технологий разработки ПО (структурное, объектно-ориентированное)	Знает ключевые понятия структурной и объектно-ориентированной парадигм программирования.
	Умеет производить анализ предметной области в рамках современных технологий разработки ПО (структурной, объектно-ориентированной).

	Владеет навыками использования современных технологий разработки ПО (структурной, объектно-ориентированной).
ПК-10.2. использует структурное и объектно-ориентированное проектирование при разработке ПО	Знает нотацию моделей структурного (DFD) и объектно-ориентированного (диаграмма классов, диаграмма перехода состояний объекта, трасса событий) проектирования.
	Умеет производить функциональную декомпозицию системы, переходить от контекстной диаграммы к спецификациям процессов (при структурном проектировании); производить объектно-ориентированный анализ (переходить от объектов к классам объектов, определять отношения между классами, поведение объектов, состояния объектов и переходы от одного состояния к другому).
	Владеет навыками построения структурных и объектно-ориентированных моделей проектирования при разработке ПО.
ПК-10.3. применяет современные технологии разработки ПО	Знает современные технологии разработки ПО, алгоритмы решения прикладных задач на основе типовых структур алгоритмов, методы организации сложных структур данных.
	Умеет применять современные технологии разработки ПО.
	Владеет синтаксисом и семантикой основных конструкций языка программирования высокого уровня, основанного на структурной и объектно-ориентированной технологии.

Для формирования вышеуказанных компетенций в рамках дисциплины «Методы машинного обучения» применяются следующие методы активного/интерактивного обучения: метод круглого стола, метод проектов, дискуссия, анализ конкретных ситуаций, ситуационный анализ.

#### 1. Трудоёмкость дисциплины и видов учебных занятий по дисциплине

Общая трудоёмкость дисциплины составляет 3 зачётные единицы (108 академических часов).

(1 зачетная единица соответствует 36 академическим часам)

Видами учебных занятий и работы обучающегося по дисциплине могут являться:

Обозначение	Виды учебных занятий и работы обучающегося
Лек	Лекции
Лаб	Лабораторные работы
СР	Самостоятельная работа обучающегося в период теоретического обучения
Контроль	Самостоятельная работа обучающегося и контактная работа обучающегося с преподавателем в период промежуточной аттестации

## Структура дисциплины:

Форма обучения – очная.

№	Наименование раздела дисциплины	Семестр	Количество часов по видам учебных занятий и работы обучающегося						Формы промежуточной аттестации		
			Лек	Лаб	Пр	ОК	СР	Контроль			
1	Тема 1. Введение, цели и задачи курса, основные термины	7	2					22	36	УО-1 собеседование	
2	Тема 2. Современная концепция анализа данных, разбор основных типов прикладных задач	7	1								
3	Тема 3. Формальные определения терминов, введение в проблему переобучение	7	1								УО-1 собеседование
4	Тема 4. Метрические методы классификации	7	2								УО-1 собеседование
5	Тема 5. Логические методы классификации	7	2								УО-1 собеседование
6	Тема 6. Линейные методы классификации. Метод стохастического градиента	7	2								УО-1 собеседование
7	Тема 7. Линейные методы классификации. Метод опорных векторов	7	2								УО-1 собеседование
8	Тема 8. Байесовские методы классификации	7	2								УО-1 собеседование
9	Тема 9. Методы кластеризации и частичного обучения	7	2								УО-1 собеседование
10	Лабораторная работа №1. Организация экспериментов по машинному обучению	7		4					ПР-6 лабораторная работа, ПР-9 индивидуальный проект		
11	Лабораторная работа №2. Метрические методы классификации	7		6					ПР-6 лабораторная работа, ПР-9 индивидуальный проект		
12	Лабораторная работа №3. Логические методы классификации	7		6					ПР-6 лабораторная работа,		

									ПР-9 индивидуальны й проект
1 3	Лабораторная работа №4. Линейные методы классификации	7		6					ПР-6 лабораторная работа, ПР-9 индивидуальны й проект
1 4	Лабораторная работа №5. Байесовские методы классификации	7		6					ПР-6 лабораторная работа, ПР-9 индивидуальны й проект
1 5	Лабораторная работа №6. Сравнение методов классификации	7		6					ПР-6 лабораторная работа, ПР-9 индивидуальны й проект
	Итого:	16		34			22	36	

## **I. СТРУКТУРА И СОДЕРЖАНИЕ ТЕОРЕТИЧЕСКОЙ ЧАСТИ КУРСА**

### **Лекционный материал (16 час.)**

#### **Тема 1. Введение, цели и задачи курса, основные термины (2 час.)**

Рассматриваются понятия: модель предметной области, задача обучения по прецедентам, алгоритм обучения, решающее правило, критерий качества обучения, обучающая и контрольная выборки, модельные и реальные данные, объекты, признаки, свойства, классы, кластеры, типы шкал (бинарные, номинальные, порядковые, количественные).

#### **Тема 2. Современная концепция анализа данных, разбор основных типов прикладных задач (1 час.)**

Рассматриваются типы задач обучения с учителем: классификация, регрессия, ранжирование, прогнозирование. Рассматриваются типы задач обучения без учителя: кластеризация, поиск ассоциативных правил, фильтрация выбросов, сокращение размерности, заполнение пропущенных значений. Каждый тип задач сопровождается одним или несколькими детально расписанными примерами.

#### **Тема 3. Формальные определения терминов, введение в проблему переобучение (1 час.)**

Рассматриваются формальные (математические) определения терминов: модель алгоритмов, алгоритм обучения, этапы обучения, ошибка, функция потерь и функционал качества, принцип минимизации эмпирического риска, обобщающая способность, скользящий контроль, проблема переобучения.

#### **Тема 4. Метрические методы классификации (2 час.)**

Рассматриваются понятия: гипотеза компактности, метрика, метод k ближайших соседей KNN и его обобщения, окно Парзена ParzenWindows и потенциальные функции, отбор эталонов (понятие отступа, алгоритм отбора эталонных объектов STOLP), отбор признаков и оптимизация метрики (задача выбора метрики, жадный алгоритм отбора признаков, полный скользящий контроль CCV).

#### **Тема 5. Логические методы классификации (2 час.)**

Рассматриваются понятия: закономерность, информативность, интерпретируемость, решающие деревья (алгоритм ID3, небрежные решающие деревья ODT, бинаризация вещественного признака).

#### **Тема 6. Линейные методы классификации. Метод стохастического градиента (2 час.)**

Рассматриваются понятия: минимизация эмпирического риска для градиентных методов обучения, линейный классификатор, математическая модель нейрона, метод стохастического градиента SG.

#### **Тема 7. Линейные методы классификации. Метод опорных векторов (2 час.)**

Рассматриваются понятия: принцип оптимальной разделяющей гиперплоскости, двойственная задача, понятие опорного вектора, обобщения линейного SVM (ядра и спрямляющие пространства, нейронные сети и SVM, обзор регуляризаторов для SVM).

#### **Тема 8. Байесовские методы классификации (2 час.)**

Рассматриваются понятия: оптимальный байесовский классификатор, вероятностная постановка задачи классификации, задача восстановления плотности распределения, наивный байесовский классификатор NaiveBayes.

#### **Тема 9. Методы кластеризации и частичного обучения (2 час.)**

Рассматриваются понятия: цели кластеризации и частичного обучения, некорректность задачи кластеризации, типы кластерных структур, проблема чувствительности к выбору метрики, графовые методы кластеризации (алгоритм выделения связанных компонент, алгоритм ФОРЭЛ, функционалы

качества кластеризации), иерархическая кластеризация (таксономия) (агломеративная иерархическая кластеризация, дендрограмма и свойство монотонности, свойства сжатия, растяжения и редуktivности).

## **II. СТРУКТУРА И СОДЕРЖАНИЕ ПРАКТИЧЕСКОЙ ЧАСТИ КУРСА**

### **Практические занятия (0 час.)**

Не предусмотрено.

### **Лабораторные работы (34 час.)**

**Лабораторная работа №1. Организация экспериментов по машинному обучению (4 час.)**

Подбор обучающей выборки для проведения экспериментов с сайтов <https://archive.ics.uci.edu/ml/index.php> и <https://www.kaggle.com/datasets>.

**Лабораторная работа №2. Метрические методы классификации (6 час.)**

Настройка значений параметров методов KNN и ParzenWindows для исследования заданной обучающей выборки с UCI ML Repository в среде Google Colaboratory.

**Лабораторная работа №3. Логические методы классификации (6 час.)**

Настройка значений параметров методов ID3 и ODT для исследования заданной обучающей выборки с UCI ML Repository в среде Google Colaboratory.

**Лабораторная работа №4. Линейные методы классификации (6 час.)**

Настройка значений параметров методов SG и SVM для исследования заданной обучающей выборки с UCI ML Repository в среде Google Colaboratory.

**Лабораторная работа №5. Байесовские методы классификации (6 час.)**

Настройка значений параметров метода NaiveBayes для исследования заданной обучающей выборки с UCI ML Repository в среде Google Colaboratory.

**Лабораторная работа №6. Сравнение методов классификации (6 час.)**

Сравнение метрических, логических, линейных и байесовских методов классификации по критерию минимума ошибок на контрольной выборке.

### **III. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ**

Учебно-методическое обеспечение самостоятельной работы обучающихся по дисциплине включает в себя:

- план-график выполнения самостоятельной работы по дисциплине, в том числе примерные нормы времени на выполнение по каждому заданию;
- требования к представлению и оформлению результатов самостоятельной работы;
- критерии оценки выполнения самостоятельной работы.

#### **План-график выполнения самостоятельной работы по дисциплине**

<b>№ п/п</b>	<b>Дата/сроки выполнения</b>	<b>Вид самостоятельной работы</b>	<b>Примерные нормы времени на выполнение</b>	<b>Форма контроля</b>
1.	7 неделя обучения	Формулировка задачи, выбор предметной области	1	УО-1, ПР-6
2.	8 неделя обучения	Разработка модели предметной области	1	УО-1, ПР-6
3.	9 неделя обучения	Исследование свойств модели предметной области	2	УО-1, ПР-6
4.	10 неделя обучения	Выделение набора информативных признаков	2	УО-1, ПР-6
5.	11 неделя обучения	Постановка задачи машинного обучения	2	УО-1, ПР-6
6.	12 неделя обучения	Выбор метода машинного обучения и анализ его свойств	2	УО-1, ПР-6
7.	13 неделя обучения	Формирование спецификации программного средства, в котором будет реализован выбранный метод машинного обучения	2	УО-1, ПР-6
8.	14 неделя обучения	Выбор/генерация обучающей выборки	2	УО-1, ПР-6

9.	15 неделя обучения	Реализация выбранного метода машинного обучения	4	УО-1, ПР-6
10.	17 неделя обучения	Оценка свойств метода	4	УО-1, ПР-9
	17-18 недели	Подготовка к экзамену	36	Экзамен
	Итого		58	

### **Рекомендации по планированию и организации времени, отведенного на изучение дисциплины**

Освоение дисциплины следует начинать с изучения рабочей учебной программы, которая содержит основные требования к знаниям, умениям и навыкам. Обязательно следует учитывать рекомендации преподавателя, данные в ходе установочных занятий. Затем – приступить к изучению отдельных разделов и тем в порядке, предусмотренном программой.

Получив представление об основном содержании раздела, темы, необходимо изучить материал с помощью рекомендуемой основной литературы. Целесообразно составить краткий конспект или схему, отображающую смысл и связи основных понятий данного раздела и включенных в него тем. Обязательно следует записывать возникшие вопросы, на которые не удалось ответить самостоятельно.

Подготовку к началу обучения включает несколько необходимых пунктов:

1) Необходимо создать для себя рациональный и эмоционально достаточный уровень мотивации к последовательному и планомерному изучению дисциплины.

2) Необходимо изучить список рекомендованной основной и дополнительной литературы и убедиться в её наличии у себя дома или в библиотеке в бумажном или электронном виде.

3) Желательно в самом начале периода обучения возможно тщательнее спланировать время, отводимое на работу с источниками и литературой по дисциплине, представить этот план в наглядной форме (график работы с датами) и в дальнейшем его придерживаться, не допуская срывов графика индивидуальной работы и «аврала» в предсессионный период. Пренебрежение этим пунктом приводит к переутомлению и резкому снижению качества усвоения учебного материала.

## **Рекомендации по работе с литературой**

1) Всю учебную литературу желательно изучать «под конспект». Цель написания конспекта по дисциплине – сформировать навыки по поиску, отбору, анализу и формулированию учебного материала.

2) Написание конспекта должно быть творческим – нужно не переписывать текст из источников, но пытаться кратко излагать своими словами содержание ответа, при этом максимально структурируя конспект, используя символы и условные обозначения.

3) При написании конспекта каждый новый вопрос начинается с нового листа, для каждого экзаменационного вопроса отводится 1-2 страницы конспекта.

5) При работе над конспектом обязательно выявляются и отмечаются трудные для самостоятельного изучения вопросы, с которыми уместно обратиться к преподавателю при посещении установочных лекций и консультаций, либо в индивидуальном порядке.

6) При чтении учебной и научной литературы всегда следить за точным и полным пониманием значения терминов и содержания понятий, используемых в тексте.

## **Требования к представлению и оформлению результатов самостоятельной работы**

Самостоятельная работа включает в себя повторение теоретического и практического материала дисциплины, заслушиваемого и конспектируемого в ходе аудиторных занятий; изучение основной и дополнительной литературы, указанной в рабочей учебной программе дисциплины, самоконтроль ответов на основные проблемные вопросы по темам занятий, самостоятельный повтор действий, осуществляемых в ходе выполнения лабораторных работ, в том числе при работе со специальным программным обеспечением.

Результаты самостоятельной работы представляются и оформляются в виде документации, по теме лабораторной работы.

## **Критерии оценки выполнения самостоятельной работы**

Критерии оценки выполнения самостоятельной работы – правильность ответов на вопросы по темам теоретической части дисциплины, достижение правильного результата при осуществлении собственных действий по лабораторным работам.

## Подготовка к промежуточной аттестации по дисциплине: экзамену

К аттестации допускаются студенты, которые систематически в течение всего семестра посещали и работали на занятиях и показали уверенные знания в ходе выполнении лабораторных работ.

Непосредственная подготовка к аттестации осуществляется по вопросам, представленным в рабочей учебной программе. Тщательно изучите формулировку каждого вопроса, вникните в его суть, составьте план ответа. Обычно план включает в себя:

- определение сущности рассматриваемого вопроса, основных положений, утверждений, определение необходимости их доказательства;
- запись обозначений, формул, необходимых для полного раскрытия вопроса;
- графический материал (таблицы, рисунки, графики), необходимые для раскрытия сущности вопроса;
- роль и значение рассматриваемого материала для практической деятельности, примеры использования в практической деятельности.

### IV. КОНТРОЛЬ ДОСТИЖЕНИЯ ЦЕЛЕЙ КУРСА

№ п/п	Контролируемые разделы/темы дисциплины	Коды и этапы формирования компетенций		Оценочные средства – наименование	
				текущий контроль	промежуточная аттестация
1.	Тема 1. Введение, цели и задачи курса, основные термины	ПК-6	знает	ПР-7 конспект	Экзамен. Вопросы 1-6 (теория)
			умеет владеет	ПР-6 Лабораторные работы № 1,6	
2.	Тема 2. Современная концепция анализа данных, разбор основных типов прикладных задач	ПК-6	знает	ПР-7 конспект	Экзамен. Вопросы 7-12 (теория)
			умеет владеет	ПР-6 Лабораторные работы № 1,6	
3	Тема 3. Формальные определения терминов, введение в проблему переобучение	ПК-6	знает	ПР-7 конспект	Экзамен. Вопросы 1-12 (теория)
			умеет владеет	ПР-6 Лабораторные работы № 1,6	
4		ПК-10	знает	ПР-7 конспект	

	Тема 4. Метрические методы классификации		умеет владеет	ПР-6 Лабораторная работа № 2	Экзамен. Вопросы 1-2 (практика)
5	Тема 5. Логические методы классификации	ПК-10	знает	ПР-7 конспект	Экзамен. Вопросы 3-5 (практика)
			умеет владеет	ПР-6 Лабораторная работа № 3	
6	Тема 6. Линейные методы классификации. Метод стохастического градиента	ПК-10	знает	ПР-7 конспект	Экзамен. Вопрос 6 (практика)
			умеет владеет	ПР-6 Лабораторная работа № 4	
7	Тема 7. Линейные методы классификации. Метод опорных векторов	ПК-10	знает	ПР-7 конспект	Экзамен. Вопрос 7 (практика)
			умеет владеет	ПР-6 Лабораторная работа № 4	
8	Тема 8. Байесовские методы классификации	ПК-10	знает	ПР-7 конспект	Экзамен. Вопросы 8-9 (практика)
			умеет владеет	ПР-6 Лабораторная работа № 5	
9	Тема 9. Методы кластеризации и частичного обучения	ПК-10	знает	ПР-7 конспект	Экзамен. Вопросы 10-12 (практика)

Типовые контрольные задания, методические материалы, определяющие процедуры оценивания знаний, умений и навыков и (или) опыта деятельности, а также критерии и показатели, необходимые для оценки знаний, умений, навыков и характеризующие этапы формирования компетенций в процессе освоения образовательной программы, представлены в разделе VIII.

## **V. СПИСОК УЧЕБНОЙ ЛИТЕРАТУРЫ И ИНФОРМАЦИОННО МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ**

### **Основная литература**

*(электронные и печатные издания)*

1. Флах П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных – М.: ДМК Пресс, 2015. – 400 с. <http://www.studentlibrary.ru/book/ISBN9785970602737.html>  
<https://e.lanbook.com/book/69955>

2. Коэлю Л.П., Ричарт В. Построение систем машинного обучения на языке Python. 2-е издание / пер. с англ. Слинкин А.А. – М.: ДМК Пресс, 2016. – 302 с.: ил. <http://www.studentlibrary.ru/book/ISBN9785970603307.html>

3. Вьюгин В.В. Математические основы машинного обучения и прогнозирования [Электронный ресурс]: учеб. пособие – М.: МЦНМО, 2013. – 304 с. <https://e.lanbook.com/book/56397>

4. Рашка С. Python и машинное обучение: крайне необходимое пособие по новейшей предсказательной аналитике, обязательное для более глубокого понимания методологии машинного обучения [Электронный ресурс]: рук. / С. Рашка; пер. с англ. Логунова А.В. – М.: ДМК Пресс, 2017. – 418 с. <https://e.lanbook.com/book/100905>

5. Кулаичев А.П. Методы и средства комплексного анализа данных: учебное пособие для вузов. – М.: Форум: ИНФРА-М, 2014. 511 с. <https://lib.dvfu.ru:8443/lib/item?id=chamo:795113&theme=FEFU>

#### **Дополнительная литература** (печатные и электронные издания)

1. Загоруйко Н. Г. Прикладные методы анализа данных и знаний. – Новосибирск: ИМ СО РАН, 1999. – 270 с. <https://lib.dvfu.ru:8443/lib/item?id=chamo:10172&theme=FEFU>

2. Кухаренко Б.Г. Интеллектуальные системы и технологии – М.: Московская государственная академия водного транспорта, 2015. – 116 с. <http://www.iprbookshop.ru/47933.html>

3. Дедукция и обобщение в системах принятия решений / В.Н. Вагин. М: Наука, 1988. – 383 с. <https://lib.dvfu.ru:8443/lib/item?id=chamo:55526&theme=FEFU>

4. Давнис В.В., Тинякова В.И., Мокшина С.И., Алексеева А.И. Компьютерные решения задач многомерной статистики. Часть 1. Кластерный и дискриминантный анализ: Учебное пособие. – Воронеж: Изд-во ВГУ, 2005. – 37 с. <http://window.edu.ru/resource/417/40417>

5. Симчера В.М. Методы многомерного анализа статистических данных: учеб. пособие. – М.: Финансы и статистика, 2008. – 400 с.: ил. <http://www.studentlibrary.ru/book/ISBN9785279031849.html>

#### **Перечень ресурсов информационно-телекоммуникационной сети «Интернет»**

1. <http://window.edu.ru/resource/417/40417> Давнис В.В., Тинякова В.И., Мокшина С.И., Алексеева А.И. Компьютерные решения задач многомерной

статистики. Часть 1. Кластерный и дискриминантный анализ: Учебное пособие. – Воронеж: Изд-во ВГУ, 2005. – 37 с.

2. <http://machinelearning.ru/> MachineLearning.ru Профессиональный информационно-аналитический ресурс, посвященный машинному обучению, распознаванию образов и интеллектуальному анализу данных.

3. [http://shad.yandex.ru/lectures/machine\\_learning.xml](http://shad.yandex.ru/lectures/machine_learning.xml) Видеолекции курса «Машинное обучение» Школы анализа данных.

4. <http://www.studentlibrary.ru/book/ISBN9785940575061.html>  
Николенко С.И., Тулупьев А.Л. Самообучающиеся системы. – М.: МЦНМО, 2009. – 288 с.: 24 ил.

5. <http://znanium.com/catalog.php?bookinfo=451186> Информационные аналитические системы: учебник / Т.В. Алексеева, Ю.В. Амириди, В.В. Дик и др.; под ред. В.В. Дика. – М.: МФПУ Синергия, 2013. – 384 с. – (Университетская серия). – ISBN 978-5-4257-0092-6.

### **Перечень информационных технологий и программного обеспечения**

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

1. Google Colaboratory.
2. Microsoft Office (Access, Excel, PowerPoint, Word и т.д.).
3. Microsoft Teams.
4. Open Office.
5. Skype.

При осуществлении образовательного процесса студентами и профессорско-преподавательским составом используются следующие информационно-справочные системы:

1. Научная электронная библиотека eLIBRARY.
2. Электронно-библиотечная система издательства «Лань».
3. Электронная библиотека "Консультант студента".
4. Электронно-библиотечная система IPRbooks.
5. Информационная система "ЕДИНОЕ ОКНО доступа к образовательным ресурсам".
6. Доступ к электронному заказу книг в библиотеке ДВФУ, доступ к нормативным документам ДВФУ, расписанию, рассылке писем.

Лабораторные занятия проводятся в специализированном компьютерном классе. Для проведения практических занятий требуется следующее программное обеспечение: браузеры для доступа в интернет.

## **VI. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ**

Дисциплина «Методы машинного обучения» изучается в следующих организационных формах: лекционное занятие; лабораторные работы; самостоятельное изучение теоретического материала; самостоятельное выполнение индивидуального проекта; индивидуальные и групповые консультации.

Основной формой самостоятельной работы студента является изучение конспекта лекций, их дополнение рекомендованной литературой, выполнение проекта, а также активная работа на лабораторных занятиях.

К прослушиванию лекции следует готовиться, для этого необходимо знать программу курса и рекомендованную литературу. Тогда в процессе лекции легче отделить главное от второстепенного, легче сориентироваться: что записать, что самостоятельно проработать, что является трудным для понимания, а что легко усвоить.

Контроль за выполнением самостоятельной работы студента производится в виде контроля каждого этапа работы, отраженного в документации и защиты проекта.

Студент должен планировать график самостоятельной работы по дисциплине и придерживаться его.

### **Подготовка к экзамену**

К аттестации допускаются студенты, которые систематически в течение всего семестра посещали и работали на занятиях и показали уверенные знания в ходе выполнения лабораторных работ.

Непосредственная подготовка к аттестации осуществляется по вопросам, представленным в рабочей учебной программе. Тщательно изучите формулировку каждого вопроса, вникните в его суть, составьте план ответа. Обычно план включает в себя:

- определение сущности рассматриваемого вопроса, основных положений, утверждений, определение необходимости их доказательства;
- графический материал (таблицы, рисунки, графики), необходимые для раскрытия сущности вопроса;

- роль и значение рассматриваемого материала для практической деятельности, примеры использования в практической деятельности.

## VII. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Перечень материально-технического и программного обеспечения дисциплины приведен в таблице.

### Материально-техническое и программное обеспечение дисциплины

Наименование специальных помещений и помещений для самостоятельной работы	Оснащенность специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения. Реквизиты подтверждающего документа
<p>690922, Приморский край, г. Владивосток, остров Русский, полуостров Саперный, поселок Аякс, 10, Корпус D, аудитория 940/818</p> <p>Учебная аудитория для проведения занятий лекционного типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации</p>	<p>Помещение укомплектовано специализированной учебной мебелью (посадочных мест – 30)</p> <p>Оборудование:</p> <p>ЖК-панель 47", Full HD, LG M4716 CCBA – 1 шт.</p> <p>Доска аудиторная.</p>	<p>1С Предприятия8 (8.2), 7-Zip, ABBYY Lingvo12, Alice 3, Anaconda3, Autodesk, CodeBlocks, CorelDRAW X7, Dia ,Directum4.8 ,DosBox-0.74, Farmanager, Firebird 2.5, FlameRobin, Foxit Reader, Free Pascal, Geany, Ghostscript, Git version 1.9.5-preview20141217, Greenfoot, gsview, Inscapе0.91, Java 8 Update 60 (64-bit), Java development Kit, Kaspersky, Lazarus, LibreOffice4.4, MatLab R2017b, Maxima 5.37.2, Microsoft Expression, Microsoft Office 2013, Microsoft Silverlight, Microsoft Silverlight 5SDK-русский, Microsoft Sistem Center, Microsoft Visual Studio 2012, MikTeX2.9, MySQL, NetBeans, Notepad++, Oracle VM VirtualBox, PascalABC.NET, PostgreSQL 9.4, PTC Mathcad, Putty, PyQt GPL v5.4.1 for Pythonv 3.4, Python2.7(3.4,3.6), QGIS Brighton, RStudio, SAM CoDeC Pack, SharePoint, Strawberry Perl, Tecnomatix, TeXnicCenter, TortoiseSVN, Unity2017.3.1f1, Veusz, Vim8.1, Visual Paradigm CE, Visual Studio 2013, Microsoft Visual C++, Windows Kits, Windows Phone SDK8.1, Xilinx Design Tools, Acrobat Reader DC, Adobe Bridge CS3, Adobe Device Central, CS3,Adobe Extend Script Toolkit 2,Adobe Photoshоpe CS3,DVD-студия Windows, Google Chrome, Internet Explorer, ITMOproctor, Mozilla Firefox 39.0 (x86 ru), Visual Studio Installer, Windows Media Center, WinSCP, Xming 6.9.0.31, ESET Endpoint Security, Firebird 2.5.3.26780 (x64), 7-Zip 9.20 (x64 edition), Adobe Flash Player, Adobe Shockwave Player, ESET Endpoint Antivirus, FOG Service, Hao Zip</p>

<p>690922, Приморский край, г. Владивосток, остров Русский, полуостров Саперный, поселок Аякс, 10, Корпус D, аудитория 733/733а</p> <p>Учебная аудитория для проведения лабораторных занятий</p>	<p>Помещение укомплектовано специализированной учебной мебелью (посадочных мест – 20)</p> <p>Оборудование:</p> <p>ЖК-панель 47", Full HD, LG M4716 CCBA – 1 шт.</p> <p>Доска аудиторная,</p> <p>Моноблок Lenovo C360G-i34164G500UDK с лицензионными программами Microsoft Office 2013 (15 шт.) и аудиовизуальными средствами проектор Panasonic DLPPProjectorPT-D2110XE</p>	<p>1С Предприятия8 (8.2), 7-Zip, ABBYY Lingvo12, Alice 3, Anaconda3, Autodesk, CodeBlocks, CorelDRAW X7, Dia ,Directum4.8 ,DosBox-0.74, Farmanager, Firebird 2.5, FlameRobin, Foxit Reader, Free Pascal, Geany, Ghostscript, Git version 1.9.5-preview20141217, Greenfoot, gsview, Inscapе0.91, Java 8 Update 60 (64-bit), Java development Kit, Kaspersky, Lazarus, LibreOffice4.4, MatLab R2017b, Maxima 5.37.2, Microsoft Expression, Microsoft Office 2013, Microsoft Silverlight, Microsoft Silverlight 5SDK-русский, Microsoft Sistem Center, Microsoft Visual Studio 2012, MikTeX2.9, MySQL, NetBeans, Notepad++, Oracle VM VirtualBox, PascalABC.NET, PostgreSQL 9.4, PTC Mathcad, Putty, PyQt GPL v5.4.1 for Pythonv 3.4, Python2.7(3.4,3.6), QGIS Brighton, RStudio, SAM CoDeC Pack, SharePoint, Strawberry Perl, Tecnomatix, TeXnicCenter, TortoiseSVN, Unity2017.3.1f1, Veusz, Vim8.1, Visual Paradigm CE, Visual Studio 2013, Microsoft Visual C++, Windows Kits, Windows Phone SDK8.1, Xilinx Design Tools, Acrobat Reader DC, Adobe Bridge CS3, Adobe Device Central, CS3,Adobe Extend Script Toolkit 2,Adobe Photoshope CS3,DVD-студия Windows, Google Chrome, Internet Explorer, ITMOproctor, Mozilla Firefox 39.0 (x86 ru), Visual Studio Installer, Windows Media Center, WinSCP, Xming 6.9.0.31, ESET Endpoint Security, Firebird 2.5.3.26780 (x64), 7-Zip 9.20 (x64 edition), Adobe Flash Player, Adobe Shockwave Player, ESET Endpoint Antivirus, FOG Service, Hao Zip</p>
--	---	--

Для проведения учебных занятий по дисциплине, а также для организации самостоятельной работы студентам доступно следующее лабораторное оборудование и специализированные кабинеты, соответствующие действующим санитарным и противопожарным нормам, а также требованиям техники безопасности при проведении учебных и научно-производственных работ.

В целях обеспечения специальных условий обучения инвалидов и лиц с ограниченными возможностями здоровья в ДВФУ все здания оборудованы пандусами, лифтами, подъемниками, специализированными местами, оснащенными туалетными комнатами, табличками информационно-навигационной поддержки.

## VIII. ФОНДЫ ОЦЕНОЧНЫХ СРЕДСТВ

Для дисциплины «Методы и технологии интеллектуализации программных систем» используются следующие оценочные средства:

Устный опрос:

1. Собеседование (УО-1)

Письменные работы:

1. Лабораторная работа (ПР-6)

2. Индивидуальный проект (ПР-9)

**Устный опрос**

Устный опрос позволяет оценить знания и кругозор студента, умение логически построить ответ, владение монологической речью и иные коммуникативные навыки.

Обучающая функция состоит в выявлении деталей, которые по каким-то причинам оказались недостаточно осмысленными в ходе учебных занятий и при подготовке к зачёту.

Собеседование (УО-1) – средство контроля, организованное как специальная беседа преподавателя с обучающимся на темы, связанные с изучаемой дисциплиной, и рассчитанное на выяснение объема знаний обучающегося по определенному разделу, теме, проблеме и т.п.

**Письменные работы**

Письменный ответ приучает к точности, лаконичности, связности изложения мысли. Письменная проверка используется во всех видах контроля и осуществляется как в аудиторной, так и во внеаудиторной работе.

Лабораторная работа (ПР-6) – средство для закрепления и практического освоения материала по определенному разделу.

Индивидуальный проект (ПР-9) – Конечный продукт, получаемый в результате планирования и выполнения комплекса учебных и исследовательских заданий. Позволяет оценить умения обучающихся самостоятельно конструировать свои знания в процессе решения практических задач и проблем, ориентироваться в информационном пространстве и уровень сформированности аналитических, исследовательских навыков, навыков практического и творческого мышления. Может выполняться в индивидуальном порядке или группой обучающихся

**Методические рекомендации, определяющие процедуры  
оценивания результатов освоения дисциплины**

**Оценочные средства для промежуточной аттестации**

Промежуточная аттестация студентов по дисциплине «Методы машинного обучения» проводится в соответствии с локальными нормативными актами ДВФУ и является обязательной. Форма отчётности по дисциплине – экзамен 7 семестр.

### **Оценочные средства для текущей аттестации**

Текущая аттестация студентов по дисциплине проводится в соответствии с локальными нормативными актами ДВФУ и является обязательной.

Текущая аттестация проводится в форме контрольных мероприятий (собеседования, презентации, эссе, лабораторных работ, контрольно-расчетных работ, творческого задания) по оцениванию фактических результатов обучения студентов и осуществляется ведущим преподавателем.

Объектами оценивания выступают:

- учебная дисциплина (активность на занятиях, своевременность выполнения различных видов заданий, посещаемость всех видов занятий по аттестуемой дисциплине);
- степень усвоения теоретических знаний;
- уровень овладения практическими умениями и навыками по всем видам учебной работы;
- результаты самостоятельной работы.

Составляется календарный план контрольных мероприятий по дисциплине. Оценка посещаемости, активности обучающихся на занятиях, своевременность выполнения различных видов заданий ведётся на основе журнала, который ведёт преподаватель в течение учебного семестра.

### **Вопросы для собеседования**

1. Определение основных терминов (модель предметной области, объект, класс, алгоритм обучения, решающее правило, признаковое описание, обучающая и контрольная выборка, модельные и реальные данные).
2. Классификатор, ошибки 1-ого и 2-ого рода, проблема переобучения.
3. Основные типы задач машинного обучения: обучение с учителем.
4. Основные типы задач машинного обучения: обучение без учителя.
5. Примеры задач: классификации, восстановления регрессии, ранжирования, кластеризации.
6. Слабая и сильная постановки задачи соревнования среди алгоритмов классификации.

7. Постановка задачи обучения по прецедентам (объекты и признаки, ответы и типы задач, модель алгоритмов, метод обучения, этап обучения и этап применения, функционалы качества).

### **Методические указания к выполнению проекта**

Целью индивидуального проекта является исследование свойств метода машинного обучения на различных наборах данных. Для достижения этой цели необходимо выбрать один из методов машинного обучения, реализовать его на любом языке программирования (либо воспользоваться существующей реализацией), после чего выбрать не менее 3 наборов данных с сайтов: UCI Machine Learning Repository (<http://archive.ics.uci.edu/ml/index.php>), Kaggle (<https://www.kaggle.com/datasets>) и провести исследование. Для этого нужно разбить каждый набор данных на обучающую и контрольную выборки (в пропорциях 80-20 или 70-30), после чего провести обучение и контроль метода. В результате необходимо сформировать оценку свойств исследуемого метода в зависимости от типа и состава использованных наборов данных – построить сравнительную таблицу с указанием процента ошибок на контрольной выборке. Также нужно определить те значения параметров метода, при которых достигаются наилучшие результаты (наименьший процент ошибок на контрольной выборке).

### **Темы индивидуальных заданий (проектов)**

1. Метрические методы классификации: метод ближайшего соседа, метод к ближайших соседей, метод окна Парзена, метод потенциальных функций.
2. Логические методы классификации: жадный алгоритм построения дерева ID3, достоинства и недостатки решающих деревьев ID3.
3. Байесовские методы классификации: наивный байесовский классификатор, оптимальный байесовский классификатор.
4. Методы кластеризации: графовые методы кластеризации (алгоритм выделения связанных компонент, алгоритм КНП (кратчайший незамкнутый путь), алгоритм ФОРЭЛ (ФОРмальные Элементы)).
5. Методы кластеризации: иерархическая кластеризация (агломеративная иерархическая кластеризация, формула Ланса-Уильямса, визуализация кластерной структуры).

### **Критерии оценки проектов**

- 100-86 баллов выставляется, если студент/группа точно определили содержание и составляющие части задания, умеют аргументировано отвечать на вопросы, связанные с заданием. Продемонстрировано знание и владение навыками самостоятельной исследовательской работы по теме. Фактических ошибок, связанных с пониманием проблемы, нет.

- 85-76 баллов – работа студента/группы характеризуется смысловой цельностью, связностью и последовательностью изложения; допущено не более 1 ошибки при объяснении смысла или содержания проблемы. Продемонстрированы исследовательские умения и навыки. Фактических ошибок, связанных с пониманием проблемы, нет.

- 75-61 балл – проведен достаточно самостоятельный анализ основных этапов и смысловых составляющих проблемы; понимание базовых основ и теоретического обоснования выбранной темы. Привлечены основные источники по рассматриваемой теме. Допущено не более 2 ошибок в смысле или содержании проблемы.

- 60-50 баллов – если работа представляет собой пересказанный или полностью переписанный исходный текст без каких бы то ни было комментариев, анализа. Не раскрыта структура и теоретическая составляющая темы. Допущено три или более трех ошибок смыслового содержания раскрываемой проблемы.

### **Методические рекомендации, определяющие процедуры оценивания результатов освоения дисциплины**

**Текущая аттестация студентов.** Текущая аттестация студентов проводится в соответствии с локальными нормативными актами ДВФУ и является обязательной.

Текущая аттестация проводится в форме защиты проекта и осуществляется ведущим преподавателем.

Объектами оценивания выступают:

- степень усвоения теоретических знаний – оценивается в форме собеседования;
- уровень овладения практическими умениями и навыками – оценивается в форме защиты проекта.

**Промежуточная аттестация студентов.** Промежуточная аттестация студентов проводится в соответствии с локальными нормативными актами ДВФУ и является обязательной.

По дисциплине «Методы машинного обучения» предусмотрен экзамен (устный опрос в форме ответов на вопросы).

### **Текущий контроль**

Текущий контроль предполагает систематическую проверку усвоения учебного материала, сформированности компетенций или их элементов, регулярно осуществляемую на протяжении изучения дисциплины, в соответствии с ее рабочей программой.

Состоит в проверке правильности выполнения заданий по самостоятельной работе. Задание зачтено, если нет ошибок. По текущим ошибкам даются пояснения.

Тесты предназначены для проверки знаний по компетенциям. Проверка достижения умений и навыков по компетенциям проверяется выполнением практических работ и курсовой работы.

### **Промежуточный контроль**

Промежуточный контроль осуществляется в конце семестра и завершает изучение дисциплины. Помогает оценить более крупные совокупности знаний и умений, сформированность определенных профессиональных компетенций по дисциплине. Промежуточный контроль проводится в форме зачета, допуск к зачету возможен для обучающихся, получивших оценку «зачтено» в результате выполнения самостоятельной работы и успешно выполнившие все лабораторные работы.

### **Оценочные средства для текущей аттестации**

1. Определение основных терминов (модель предметной области, объект, класс, алгоритм обучения, решающее правило, признаковое описание, обучающая и контрольная выборка, модельные и реальные данные).
2. Классификатор, ошибки 1-ого и 2-ого рода, проблема переобучения.
3. Основные типы задач машинного обучения: обучение с учителем.
4. Основные типы задач машинного обучения: обучение без учителя.
5. Примеры задач: классификации, восстановления регрессии, ранжирования, кластеризации.
6. Слабая и сильная постановки задачи соревнования мира среди алгоритмов классификации.
7. Постановка задачи обучения по прецедентам (объекты и признаки, ответы и типы задач, модель алгоритмов, метод обучения, этап обучения и этап применения, функционалы качества).

## Темы проектов

1. Метрические методы классификации: метод ближайшего соседа, метод k ближайших соседей, метод окна Парзена, метод потенциальных функций.

2. Логические методы классификации: жадный алгоритм построения дерева ID3, достоинства и недостатки решающих деревьев ID3.

3. Байесовские методы классификации: наивный байесовский классификатор, оптимальный байесовский классификатор.

4. Методы кластеризации: графовые методы кластеризации (алгоритм выделения связанных компонент, алгоритм КНП (кратчайший незамкнутый путь), алгоритм ФОРЭЛ (ФОРмальные Элементы)).

5. Методы кластеризации: иерархическая кластеризация (агломеративная иерархическая кластеризация, формула Ланса-Уильямса, визуализация кластерной структуры).

## Критерии оценки проектов

- 100-86 баллов выставляется, если студент/группа точно определили содержание и составляющие части задания, умеют аргументировано отвечать на вопросы, связанные с заданием. Продемонстрировано знание и владение навыками самостоятельной исследовательской работы по теме. Фактических ошибок, связанных с пониманием проблемы, нет.

- 85-76 – баллов – работа студента/группы характеризуется смысловой цельностью, связностью и последовательностью изложения; допущено не более 1 ошибки при объяснении смысла или содержания проблемы. Продемонстрированы исследовательские умения и навыки. Фактических ошибок, связанных с пониманием проблемы, нет.

- 75-61 балл – проведен достаточно самостоятельный анализ основных этапов и смысловых составляющих проблемы; понимание базовых основ и теоретического обоснования выбранной темы. Привлечены основные источники по рассматриваемой теме. Допущено не более 2 ошибок в смысле или содержании проблемы.

- 60-50 баллов – если работа представляет собой пересказанный или полностью переписанный исходный текст без каких бы то ни было комментариев, анализа. Не раскрыта структура и теоретическая составляющая темы. Допущено три или более трех ошибок смыслового содержания раскрываемой проблемы.

## Шкала оценивания

Менее 60 баллов	незачтено	неудовлетворительно
От 61 до 75 баллов	зачтено	удовлетворительно
От 76 до 85 баллов	зачтено	хорошо
От 86 до 100 баллов	зачтено	отлично

## Оценочные средства для промежуточной аттестации

### Вопросы к экзамену

#### Первый блок (теория):

1. Цель курса, три определения машинного обучения, его специфика. Необходимость проведения экспериментов, области применения.
2. Типы обучения: дедуктивное, индуктивное, комбинированное. Сценарий индуктивного обучения. Критерий (функционал) качества обучения.
3. Модель зависимости, выборка, прецедент, объект, объект, класс, кластер. Data Science, интеллектуальный анализ данных, распознавание образов.
4. Признаковое описание объектов обучающей и контрольной выборок. Типы признаков и критерии их отбора. Модельные и реальные данные.
5. Задача обучения по прецедентам, обобщающая способность, метод и алгоритм обучения, решающее правило, этапы обучения и применения.
6. Ошибка, эмпирический риск, скользящий контроль (кросс-проверка). Вычислительная эффективность. Сведение обучения к оптимизации.
7. Современная концепция анализа данных. Ошибки 1-ого и 2-ого рода. Переобучение, недообучение. Соревнования по машинному обучению.
8. Основные типы задач обучения с учителем (классификация, регрессия, ранжирование, прогнозирование), их постановки и отличия.
9. Основные типы задач обучения без учителя (кластеризация, фильтрация выбросов, поиск ассоциативных правил, сокращение размерности, заполнение пропущенных значений), их постановки и отличия.
10. Частичное обучение, трансдуктивное обучение, динамическое обучение, обучение с подкреплением, активное обучение, метаобучение.
11. Примеры задач: классификации (медицинская диагностика, оценивание заемщиков, предсказание ухода клиента), восстановления регрессии (прогноз потребительского спроса), ранжирования (ранжирование текстов), кластеризации (рубрикация текстов, рекомендации).
12. Оценка качества классификации: TP, TN, FP, FN; ROC-кривая, AUC-ROC. Точность и полнота для двухклассовой и многоклассовой классификации.

## Второй блок (практика):

1. Метрические методы классификации: гипотеза компактности, метрика, метрическое пространство, обобщенный метрический классификатор, метод ближайшего соседа, метод к ближайших соседей.
2. Метрические методы классификации: метод окна Парзена, метод потенциальных функций, понятие отступа, типы объектов (в зависимости от отступа), две реализации алгоритма STOLP.
3. Логические методы классификации: логическая закономерность, интерпретируемость, информативность, часто используемые виды закономерностей, часто используемые критерии информативности, обобщенный алгоритм нахождения информативных закономерностей.
4. Логические методы классификации: определение бинарного решающего дерева, пример решающего дерева, жадный алгоритм построения дерева ID3, достоинства и недостатки решающих деревьев ID3.
5. Логические методы классификации: определение бинарного решающего дерева, редукция дерева, небрежные решающие деревья ODT, задача бинаризации вещественного признака.
6. Линейные методы классификации (метод стохастического градиента): задача построения разделяющей поверхности, линейный классификатор, математическая модель нейрона, алгоритм SG.
7. Линейные методы классификации (метод опорных векторов): кусочно-линейная аппроксимация, задача SVM, оптимальная разделяющая гиперплоскость, достоинства и недостатки метода.
8. Байесовские методы классификации: байесовский классификатор апостериорная вероятность, вероятностная постановка задачи, функционал среднего риска, оптимальный байесовский классификатор.
9. Байесовские методы классификации: разбиение задачи классификации на две подзадачи и алгоритм решение первой из них, три основных подхода к оцениванию плотностей, наивный байесовский классификатор.
10. Методы кластеризации: постановка задачи кластеризации, некорректность задачи кластеризации, цели кластеризации, типы кластерных структур, проблема чувствительности к выбору метрики.
11. Методы кластеризации: качество кластеризации в метрическом и линейном векторном пространстве, метод K-средних (K-means) для кластеризации и частичного обучения, алгоритм DBSCAN и его преимущества.
12. Методы кластеризации: агломеративная иерархическая кластеризация (алгоритм Ланса-Уильямса, формула Ланса-Уильямса и ее частные случаи, визуализация кластерной структуры – дендрограмма).

**Критерии выставления оценки студенту на зачете по дисциплине  
«Методы машинного обучения»**

<b>Баллы (рейтин- говой оценки)</b>	<b>Оценка зачета/ экзамена (стандартная)</b>	<b>Требования к сформированным компетенциям</b>
86-100	«зачтено»/ «отлично»	Оценка «отлично» выставляется студенту, если он глубоко и прочно усвоил программный материал, исчерпывающе, последовательно, четко и логически стройно его излагает, умеет тесно увязывать теорию с практикой, свободно справляется с задачами, вопросами и другими видами применения знаний, причем не затрудняется с ответом при видоизменении заданий, правильно обосновывает принятое решение, владеет разносторонними навыками и приемами выполнения практических задач.
76-85	«зачтено»/ «хорошо»	Оценка «хорошо» выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, не допуская существенных неточностей в ответе на вопрос, правильно применяет теоретические положения при решении практических вопросов и задач, владеет необходимыми навыками и приемами их выполнения.
61-75	«зачтено»/ «удовлетво- рительно»	Оценка «удовлетворительно» выставляется студенту, если он имеет знания только основного материала, но не усвоил его деталей, допускает неточности, недостаточно правильные формулировки, нарушения логической последовательности в изложении программного материала, испытывает затруднения при выполнении практических работ.
0-60	«не зачтено»/ «неудовлетво- рительно»	Оценка «неудовлетворительно» выставляется студенту, который не знает значительной части программного материала, допускает существенные ошибки, неуверенно, с большими затруднениями выполняет практические работы. Как правило, оценка «неудовлетворительно» ставится студентам, которые не могут продолжить обучение без дополнительных занятий по соответствующей дисциплине.