

*Федеральное государственное автономное образовательное учреждение высшего образования
«Российский университет дружбы народов»*

Факультет физико-математических и естественных наук

Рекомендовано МССН
38.00.00 «Экономика и управление»,
подгруппа 4 «Бизнес-информатика»

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Наименование дисциплины

Методы машинного обучения

Рекомендуется для направления подготовки

38.03.05 Бизнес-информатика

Квалификация (степень) выпускника

бакалавр

1. Цели и задачи дисциплины: Настоящий курс ставит своей целью ознакомление обучающихся с задачами, возникающими в области машинного обучения (Machine Learning), и методами их решения, которые помогут выявлять, формализовывать и успешно решать практические задачи машинного обучения, возникающие в процессе профессиональной деятельности.

В ходе изучения дисциплины перед обучающимися ставятся следующие задачи:

- изучение основных постановок задач и принципов машинного обучения;
- получение представления о методах и алгоритмах решения задач машинного обучения;
- освоение основных методов и алгоритмов классификации и регрессии;
- освоение основных методов и алгоритмов построения деревьев решений;
- получение представления о задачах и принципах глубокого обучения;
- получение представления о задачах и принципах обучения с подкреплением.

2. Место дисциплины в структуре ОП ВО:

Дисциплина «Методы машинного обучения» относится к *части, формируемой участниками образовательных отношений*, блока 1 учебного плана.

В таблице № 1 приведены предшествующие и последующие дисциплины, направленные на формирование компетенций дисциплины в соответствии с матрицей компетенций ОС ВО.

Таблица № 1

Предшествующие и последующие дисциплины, направленные на формирование компетенций

№ п/п	Шифр и наименование компетенции	Предшествующие дисциплины	Последующие дисциплины (группы дисциплин)
Универсальные компетенции			
-	УК-12	Анализ данных	Интеллектуальный анализ данных
Общепрофессиональные компетенции			
	ОПК-7	Основы программирования Технология программирования Алгоритмы и анализ сложности Java и его приложения	Интеллектуальный анализ данных, Модели на гиперграфах
Профессиональные компетенции (вид профессиональной деятельности: <i>научно-исследовательская</i>)			
	ПК-3	Основы программирования Технология программирования Алгоритмы и анализ сложности Java и его приложения	Интеллектуальный анализ данных
Профессионально-специализированные компетенции специализации			
-	-	-	-

3. Требования к результатам освоения дисциплины:

Процесс изучения дисциплины направлен на формирование следующих компетенций: УК-12; ОПК-7; ПК-3

УК-12 Способен: искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных

УК-12.1 Способен: искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных

ОПК-7 Способен использовать цифровые технологии и методы в профессиональной деятельности в области бизнес-информатики для: изучения и моделирования объектов профессиональной деятельности, анализа данных, представления информации и пр.

- ОПК-7.1 Знает базовые принципы цифровых технологий и методов, необходимых в профессиональной деятельности в области бизнес-информатики для: изучения и моделирования объектов профессиональной деятельности, анализа данных, представления информации и пр.
- ОПК-7.2 Умеет применять необходимые в профессиональной деятельности цифровые технологии и методы в области бизнес-информатики для: изучения и моделирования объектов профессиональной деятельности, анализа данных, представления информации и пр.
- ОПК-7.3 Владеет необходимыми в профессиональной деятельности технологиями и методами в области бизнес-информатики для: изучения и моделирования объектов профессиональной деятельности, анализа данных, представления информации и пр.

ПК-3 Способен выполнять работы и управлять работами по созданию (модификации) и сопровождению ИС, автоматизирующих задачи организационного управления и бизнес-процессы

- ПК-3.3 Знает основы программирования; современные объектно-ориентированные языки программирования; современные структурные языки программирования; языки современных бизнес-приложений
- ПК-3.4 Умеет кодировать на языках программирования; тестировать результаты кодирования

В результате изучения дисциплины студент должен:

Знать: принципы машинного обучения; основные задачи и методы машинного обучения; возможности современных и перспективных библиотек и фреймворков машинного обучения.

Уметь: формулировать задачи машинного обучения; выбирать адекватные методы и алгоритмы их решения; оценивать качество получаемых моделей; выбирать адекватные программные и технические средства построения моделей машинного обучения.

Владеть: технологиями разработки алгоритмов и программными системами построения моделей машинного обучения; средствами автоматизации обучения моделей и оценки их качества.

4. Объем дисциплины и виды учебной работы

Общая трудоемкость дисциплины составляет 4 зачетные единицы.

Вид учебной работы	Всего часов	Семестры
		Семестр 6, модуль 12 (С)
Аудиторные занятия (всего)	54	54
<i>Лекции</i>	18	18
<i>Практические занятия (ПЗ)</i>	-	-
<i>Семинары (С)</i>	-	-
<i>Лабораторные работы (ЛР)</i>	36	36
Самостоятельная работа (всего)	90	90
Общая трудоемкость, час.	144	144
зач. ед.	4	4

5. Содержание дисциплины

5.1. Содержание разделов дисциплины

№ п/п	Наименование раздела дисциплины	Содержание раздела (темы)
1.	Введение в машинное обучение	Тема 1. Основные определения и постановки задач. Виды обучения: с учителем, без учителя, с подкреплением. Основные типы задач: задача классификации, задача регрессии, задача кластеризации. Основные проблемы машинного обучения (недостаточный объем обучающей выборки, пропуски в данных, переобучение).
2.	Задачи регрессии	Тема 1. Метод наименьших квадратов. Измерение ошибки в задачах. Многомерная регрессия, проблема мультиколлинеарности. Регрессия, линейная по параметрам, полиномиальная регрессия. Решение проблемы переобучения: L1-регуляризация (Lasso), L2-регуляризация (гребневая регрессия), эластичная сеть.
3.	Задачи классификации	Тема 1. Линейная модель классификации. Логистическая регрессия как линейный классификатор. Функция потерь (ошибок классификации). Матрица ошибок классификации. Метрики качества классификации. Метод ближайших соседей (kNN). Использование наивной байесовской модели.
4.	Деревья решений	Тема 1. Этапы построения дерева решений, выбор критерия точности прогноза. Метрики ветвления на основе прироста информации, нормализованного прироста информации, индекса Джини. Правила разбиения. Механизм отсечения дерева. Критерии останова алгоритма. Переобучение решающих деревьев. Случайный лес. Обучение случайного леса. Достоинства и недостатки случайного леса.
5.	Глубокое обучение	Тема 1. Основы нейронных сетей. Архитектура нейронной сети. Сети прямого распространения (многослойный перцептрон). Обучение нейронной сети градиентными методами. Функция потерь (loss function). Функции активации (сигмоида, гиперболический тангенс, ReLU). Алгоритм обратного распространения (backpropagation algorithm).

№ п/п	Наименование раздела дисциплины	Содержание раздела (темы)
6.	Обучение с подкреплением	Тема 1. Основы обучения с подкреплением. Постановка задачи обучения с подкреплением. Взаимодействие агент–окружающая среда. Цели и вознаграждения. Марковские процессы принятия решения. Функции ценности. Динамическое программирование. Оценка и улучшение политики.

5.2. Разделы дисциплин и виды занятий

№ п/п	Наименование раздела дисциплины	Лек-ции	Прак-зан.	Лаб. зан.	Сем. зан.	СР С	Всего час.
1.	Введение в машинное обучение	2	-	4	-	10	16
2.	Задачи регрессии	4	-	8	-	20	32
3.	Задачи классификации	2	-	4	-	10	16
4.	Деревья решений	2	-	4	-	10	16
5.	Глубокое обучение	4	-	8	-	20	32
6.	Обучение с подкреплением	4	-	8	-	20	32
	ВСЕГО:	18	-	36	-	90	144

6. Лабораторный практикум

№ п/п	№ раздела	Наименование лабораторных работ	Трудо-емкость (час.)
1.	1	Применение градиентного спуска для оптимизации и визуализация данных	4
2.	2	Решение задачи прогнозирования при помощи линейной и нелинейной регрессии.	8
3.	3	Решение задачи классификации данных при помощи байесовских методов и классификации по ближайшим соседям.	4
4.	4	Решение задачи классификации и регрессии при помощи обучения деревьев решений.	4
5.	5	Построение решений дифференциальных уравнений при помощи нейронных сетей прямого распространения.	8
6.	6	Моделирование управления динамической системой при помощи обучения с подкреплением	8

7. Практические занятия (семинары)

Не предусмотрены

8. Материально-техническое обеспечение дисциплины:

Учебная аудитория с ПК и проектором для проведения учебных занятий (в том числе для практического и лекционного типов занятий, индивидуальных и групповых консультаций, текущего контроля и промежуточной аттестации).

Компьютерные (дисплейные) классы с доступом к сети Интернет и электронно-образовательной среде Университета для выполнения обучающимися заданий лабораторного практикума, выполнения обучающимися самостоятельной работы и проведения компьютерного тестирования обучающихся (при необходимости).

9. Информационное обеспечение дисциплины

а) программное обеспечение

- ОС Windows, MS Office (программа корпоративного лицензирования (Microsoft Subscription) Enrollment for Education Solutions), браузер Firefox (лицензия MPL-2.0) или браузер Chrome (лицензия Google Chrome Terms of Service); Adobe Reader (Adobe Software License Agreement)
- программное обеспечение со свободной лицензией: Dev-C++ (лицензия GNU GPL)
- ОС Linux, офисный пакет LibreOffice (лицензия MPL-2.0), ПО для просмотра pdf (например, evince (лицензия GPL-2+ CC-BY-SA-3.0)), dev-lang/perl (лицензия Artistic GPL-1+), dev-lang/python (лицензия PSF-2), numpy (лицензия NumPy license), sympy (лицензия The 3-Clause BSD License), matplotlib (лицензия Python Software Foundation License), dev-lang/R (лицензия || (GPL-2 GPL-3) LGPL-2.1)

б) базы данных, информационно-справочные и поисковые системы:

1. Сайт библиотеки РУДН <http://lib.rudn.ru/>
2. ТУИС <http://esystem.pfur.ru/>

10. Учебно-методическое обеспечение дисциплины:

а) основная литература:

1. Воронцов К.В. Машинное обучение. НОУ Интуит, 2015 // [Электронный ресурс] URL: <https://www.intuit.ru/studies/courses/13844/1241/info>, режим доступа: свободный.

б) дополнительная литература:

1. Тарков М. Нейрокомпьютерные системы. НОУ Интуит, 2006 // [Электронный ресурс] URL: <https://www.intuit.ru/studies/courses/61/61/info>, режим доступа: свободный.

11. Методические указания для обучающихся по освоению дисциплины

Учебным планом на изучение дисциплины отводится 1 семестр. В течение семестра выполняются лабораторные работы, домашние задания и проводятся контрольные мероприятия. В конце семестра производится итоговый контроль знаний - экзамен.

11.1 Методические указания по самостоятельному освоению теоретического материала по дисциплине

Лекционный материал дисциплины охватывает темы, указанные в разделе 5.1 программы дисциплины. В ТУИС (<http://esystem.pfur.ru>) по темам лекций размещены учебные материалы. Рекомендуется по указанным темам в дополнение к учебным материалам ТУИС изучить материалы, указанные в п. 10 программы дисциплины.

11.2 Методические указания по выполнению лабораторных работ

Задания по лабораторным работам выполняются индивидуально каждым студентом в соответствии с календарным планом. По результатам выполнения каждой лабораторной работы студентом готовится отчет.

11.3. Методические указания по подготовке к контрольным мероприятиям

Контрольные мероприятия по дисциплине проводятся в форме контрольных работ и/или оценки результатов выполнения лабораторных работ. Итоговый контроль в форме опроса проводится по темам всех разделов дисциплины. Материалы для подготовки к промежуточному и итоговому контролю размещены в соответствующем разделе ТУИС (<http://esystem.pfur.ru>).

12. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине

ФОС по дисциплине представлен в приложении к данной программе.
Программа составлена в соответствии с требованиями ОС ВО РУДН.

Разработчик:

доцент кафедры
информационных технологий, к.ф.-м.н.



С.Г. Шорохов

Зав. кафедрой информационных
технологий, д.ф.-м.н.



Ю.Н. Орлов

Руководитель программы

Заведующий кафедрой
прикладной информатики
и теории вероятностей, д.т.н., проф.



К.Е. Самуйлов

*Федеральное государственное автономное образовательное учреждение высшего образования
«Российский университет дружбы народов»*

Факультет физико-математических и естественных наук

Кафедра информационных технологий

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

ПО УЧЕБНОЙ ДИСЦИПЛИНЕ

Методы машинного обучения

Рекомендуется для направления подготовки

38.03.05 Бизнес-информатика

Квалификация (степень) выпускника

Бакалавр

Паспорт фонда оценочных средств по дисциплине

Дисциплина: Методы машинного обучения

Направление: 38.03.05 Бизнес-информатика

Код контролируемой компетенции или ее части	Контролируемый раздел дисциплины	Контролируемая тема дисциплины	ФОСы (формы контроля уровня освоения ООП)			Баллы темы	Баллы раздела
			Аудиторная работа	СРС	Экзамен/Зачет		
			Выполнение ЛР	Выполнение ДЗ			
УК-12 ОПК-7 ПК-3	Раздел 1: Введение в машинное обучение	Тема 1. Основные определения и постановки задач. Виды обучения: с учителем, без учителя, с подкреплением. Основные типы задач: задача классификации, задача регрессии, задача кластеризации. Основные проблемы машинного обучения (недостаточный объем обучающей выборки, пропуски в данных, переобучение).	10		3	12	13
	Раздел 2: Задачи регрессии	Тема 1. Метод наименьших квадратов. Измерение ошибки в задачах. Многомерная регрессия, проблема мультиколлинеарности. Регрессия, линейная по параметрам, полиномиальная регрессия. Решение проблемы переобучения: L1-регуляризация (Lasso), L2-регуляризация (гребневая регрессия), эластичная сеть.	10		3	13	13
	Раздел 3: Задачи классификации	Тема 1. Линейная модель классификации. Логистическая регрессия как линейный классификатор. Функция потерь (ошибок классификации). Матрица ошибок классификации. Метрики качества классификации. Метод ближайших соседей (kNN). Использование наивной байесовской модели.	10		3	13	13
	Раздел 4: Деревья решений	Тема 1. Этапы построения дерева решений, выбор критерия точности прогноза. Метрики ветвления на основе прироста информации, нормализованного прироста информации, индекса Джини. Правила разбиения. Механизм отсечения дерева. Критерии останова алгоритма. Переобучение решающих деревьев. Случайный лес. Обучение случайного леса. Достоинства и недостатки случайного леса.	10		3	12	13

	Раздел 5: Глубокое обучение	Тема 1. Основы нейронных сетей. Архитектура нейронной сети. Сети прямого распространения (многослойный перцептрон). Обучение нейронной сети градиентными методами. Функция потерь (loss function). Функции активации (сигмоида, гиперболический тангенс, ReLU). Алгоритм обратного распространения (backpropagation algorithm).	20		4	24	24
	Раздел 6: Обучение с подкреплением	Тема 1. Основы обучения с подкреплением. Постановка задачи обучения с подкреплением. Взаимодействие агент–окружающая среда. Цели и вознаграждения. Марковские процессы принятия решения. Функции ценности. Динамическое программирование. Оценка и улучшение политики.	20		4	24	24
		ИТОГО:	80		20	100	100

Процесс изучения дисциплины направлен на формирование следующих компетенций:

УК-12; ОПК-7; ПК-3

УК-12 Способен: искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных

УК-12.1 Способен: искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных

ОПК-7 Способен использовать цифровые технологии и методы в профессиональной деятельности в области бизнес-информатики для: изучения и моделирования объектов профессиональной деятельности, анализа данных, представления информации и пр.

- ОПК-7.1 Знает базовые принципы цифровых технологий и методов, необходимых в профессиональной деятельности в области бизнес-информатики для: изучения и моделирования объектов профессиональной деятельности, анализа данных, представления информации и пр.
- ОПК-7.2 Умеет применять необходимые в профессиональной деятельности цифровые технологии и методы в области бизнес-информатики для: изучения и моделирования объектов профессиональной деятельности, анализа данных, представления информации и пр.
- ОПК-7.3 Владеет необходимыми в профессиональной деятельности технологиями и методами в области бизнес-информатики для: изучения и моделирования объектов профессиональной деятельности, анализа данных, представления информации и пр.

ПК-3 Способен выполнять работы и управлять работами по созданию (модификации) и сопровождению ИС, автоматизирующих задачи организационного управления и бизнес-процессы

- ПК-3.3 Знает основы программирования; современные объектно-ориентированные языки программирования; современные структурные языки программирования; языки современных бизнес-приложений
- ПК-3.4 Умеет кодировать на языках программирования; тестировать результаты кодирования

Балльно-рейтинговая система оценки уровня знаний

Таблица соответствия баллов и оценок

Баллы БРС	Традиционные оценки РФ	Оценки ECTS
95 - 100	5	A
86 - 94		B
69 - 85	4	C
61 - 68	3	D
51 - 60		E
31 - 50	2	FX
0 - 30		F
51-100	Зачет	Passed

Правила применения БРС

1. Раздел (тема) учебной дисциплины считаются освоенными, если студент набрал более 50 % от возможного числа баллов по этому разделу (теме).
2. Студент не может быть аттестован по дисциплине, если он не освоил все темы и разделы дисциплины, указанные в сводной оценочной таблице дисциплины.
3. По решению преподавателя и с согласия студентов, не освоивших отдельные разделы (темы) изучаемой дисциплины, в течение учебного семестра могут быть повторно проведены мероприятия текущего контроля успеваемости или выданы дополнительные учебные задания по этим темам или разделам. При этом студентам за данную работу засчитывается минимально возможный положительный балл (51 % от максимального балла).
4. При выполнении студентом дополнительных учебных заданий или повторного прохождения мероприятий текущего контроля полученные им баллы засчитываются за конкретные темы. Итоговая сумма баллов не может превышать максимального количества баллов, установленного по данным темам (в соответствии с приказом Ректора № 564 от 20.06.2013). По решению преподавателя предыдущие баллы, полученные студентом по учебным заданиям, могут быть аннулированы.
5. График проведения мероприятий текущего контроля успеваемости формируется в соответствии с календарным планом курса. Студенты обязаны сдавать все задания в сроки, установленные преподавателем.
6. Время, которое отводится студенту на выполнение мероприятий текущего контроля успеваемости, устанавливается преподавателем. По завершение отведенного времени студент должен сдать работу преподавателю, вне зависимости от того, завершена она или нет.
7. Использование источников (в том числе конспектов лекций и лабораторных работ) во время выполнения контрольных мероприятий возможно только с разрешения преподавателя.

8. Отсрочка в прохождении мероприятий текущего контроля успеваемости считается уважительной только в случае болезни студента, что подтверждается наличием у него медицинской справки, заверенной круглой печатью в поликлинике № 25, предоставляемой преподавателю не позднее двух недель после выздоровления. В этом случае выполнение контрольных мероприятий осуществляется после выздоровления студента в срок, назначенный преподавателем. В противном случае, отсутствие студента на контрольном мероприятии признается не уважительным.
9. Студент допускается к итоговому контролю знаний с любым количеством баллов, набранных в семестре.
10. Итоговый контроль знаний оценивается из 20 баллов независимо от числа баллов за семестр.
11. Если в итоге за семестр студент получил менее 31 балла, то ему выставляется оценка F и студент должен повторить эту дисциплину в установленном порядке. Если же в итоге студент получил 31-50 баллов (т. е. FX), то студенту разрешается добор необходимого (до 51) количества баллов путем повторного одноразового выполнения предусмотренных контрольных мероприятий, при этом по усмотрению преподавателя аннулируются соответствующие предыдущие результаты. Ликвидация задолженностей проводится в период с 07.02 по 28.02 (с 07.09 по 28.09) по согласованию с деканатом.

Примерный перечень оценочных средств

п/ п	Наименование оценочного средства	Краткая характеристика оценочного средства	Представление оценочного средства в фонде
<i>Аудиторная работа</i>			
1	Лабораторная работа	Система практических заданий, направленных на формирование практических навыков у обучающихся	Фонд практических заданий
4	Экзамен	Оценка работы студента в течение семестра (года, всего срока обучения и др.) и призван выявить уровень, прочность и систематичность полученных им теоретических и практических знаний, приобретения навыков самостоятельной работы, развития творческого мышления, умение синтезировать полученные знания и применять их в решении практических задач.	Примеры заданий/вопросов, пример экзаменационного билета
<i>Самостоятельная работа</i>			
1	Подготовка отчетов по результатам выполнения лабораторных работ	Форма проверки качества выполнения студентами лабораторных работ в соответствии с утвержденной программой.	Фонд практических заданий

Учебным планом на изучение дисциплины отводится один семестр. В дисциплине предусмотрены лекции, лабораторный практикум, контрольные мероприятия по проверке отчетов по лабораторным работам. В конце семестра проводится итоговый контроль знаний.

Оценивание результатов освоения дисциплины производится в соответствии с балльно-рейтинговой системой. По дисциплине предусмотрен экзамен.

Критерии оценки по дисциплине

95-100 баллов:

- полное и своевременное выполнение на высоком уровне лабораторных работ с оформлением отчетов, успешное прохождение контрольных мероприятий, предусмотренных программой курса;
- систематизированное, глубокое и полное освоение навыков и компетенций по всем разделам программы дисциплины;
- использование научной терминологии, стилистически грамотное, логически правильное изложение ответов на вопросы, умение делать обоснованные выводы;
- безупречное владение программным обеспечением, умение эффективно использовать его в постановке и решении научных и профессиональных задач;
- выраженная способность самостоятельно и творчески решать поставленные задачи;
- полная самостоятельность и творческий подход при изложении материала по программе дисциплины;
- полное и глубокое усвоение основной и дополнительной литературы, рекомендованной программой дисциплины и преподавателем.

86- 94 балла:

- полное и своевременное выполнение на хорошем уровне лабораторных работ с оформлением отчетов, успешное прохождение контрольных мероприятий, предусмотренных программой курса;
- систематизированное, глубокое и полное освоение навыков и компетенций по всем разделам программы дисциплины;
- использование научной терминологии, стилистически грамотное, логически правильное изложение ответа на вопросы, умение делать обоснованные выводы;
- хорошее владение программным обеспечением, умение эффективно использовать его в постановке и решении научных и профессиональных задач;
- способность самостоятельно решать поставленные задачи в нестандартных производственных ситуациях;
- усвоение основной и дополнительной литературы, нормативных и законодательных актов, рекомендованных программой дисциплины и преподавателем.

69-85 баллов:

- своевременное выполнение на хорошем уровне лабораторных работ с оформлением отчетов, прохождение контрольных мероприятий, предусмотренных программой курса;
- хороший уровень культуры исполнения лабораторных работ;
- систематизированное и полное освоение навыков и компетенций по всем разделам программы дисциплины;

- владение программным обеспечением, умение использовать его в постановке и решении научных и профессиональных задач;
- способность самостоятельно решать проблемы в рамках программы дисциплины;
- усвоение основной литературы;

51-68 баллов:

- выполнение на удовлетворительном уровне лабораторных работ с оформлением отчетов, прохождение контрольных мероприятий, предусмотренных программой курса;
- систематизированное и полное освоение навыков и компетенций по всем разделам программы дисциплины;
- удовлетворительное владение программным обеспечением, умение использовать его в постановке и решении научных и профессиональных задач;
- способность решать проблемы в рамках программы дисциплины;
- удовлетворительное усвоение основной литературы;

31 - 50 баллов – НЕ ЗАЧТЕНО:

- не выполнение, несвоевременное выполнение или выполнение на неудовлетворительном уровне лабораторных работ, не прохождение контрольных мероприятий, предусмотренных программой курса;
- недостаточно полный объем навыков и компетенции в рамках программы дисциплины;
- неумение использовать в практической деятельности научной терминологии, изложение ответа на вопросы с существенными стилистическими и логическими ошибками;
- слабое владение программным обеспечением по разделам программы дисциплины, некомпетентность в решении стандартных (типовых) производственных задач;
- способность решать проблемы в рамках программы дисциплины;
- удовлетворительное усвоение основной литературы;

0-30 баллов, НЕ ЗАЧТЕНО:

- отсутствие умений, навыков, знаний и компетенции в рамках программы дисциплины;
- невыполнение лабораторных заданий, не прохождение контрольных мероприятий, предусмотренных программой курса; отказ от ответов по программе дисциплины;
- игнорирование занятий по дисциплине по неуважительной причине.

Экзаменационные вопросы

Дисциплина Методы машинного обучения

1. Виды машинного обучения (с учителем, без учителя, с подкреплением).
2. Основные задачи машинного обучения: задача классификации, задача регрессии, задача кластеризации.
3. Линейная модель классификации. Логистическая регрессия как линейный классификатор.
4. Байесовский классификатор. Теорема Байеса. Оценка априорной вероятности класса.
5. Байесовский классификатор. Параметрический подход для числовых признаков. Алгоритм байесовской классификации. Пример.
6. Байесовский классификатор. Классификация категориальных признаков. Пример.
7. Наивный байесовский классификатор. Алгоритм наивной байесовской классификации. Пример.
8. Наивный байесовский классификатор. Классификация категориальных признаков. Пример.
9. Метод К ближайших соседей. Пример.
10. Классификатор дерева решений. Рекурсивные разбиения. Гиперплоскости. Чистота области.
11. Алгоритм построения дерева принятия решений.
12. Оценка разбиения: энтропия, информационный выигрыш, индекс Джини.
13. Оценка разбиения для числовых признаков. Алгоритм оценки числовых признаков.
14. Оценка разбиения для категориальных признаков. Алгоритм оценки категориальных признаков.
15. Линейный дискриминантный анализ. Проекция на прямую. Оптимальный линейный дискриминант.
16. Линейный дискриминантный анализ. Линейный дискриминант Фишера. Алгоритм линейного дискриминанта.
17. Оценка классификации. Меры качества классификации: доля ошибок, точность.
18. Меры оценки качества классификации на основе таблицы сопряженности. Точность и полнота класса. F-мера.
19. Регрессионная модель. Линейная регрессия. Одномерный случай. Нелинейная регрессия.
20. Оценка качества регрессии. Коэффициент детерминации.
21. Метод опорных векторов. Разделяющая гиперплоскость. Зазор и опорные векторы. Каноническая гиперплоскость.
22. Метод опорных векторов. Линейный и разделимый случай. Классификатор метода опорных векторов.
23. Метод опорных векторов с мягким зазором. Прямая задача оптимизации. Оптимизация методом Ньютона.
24. Метод опорных векторов с мягким зазором. Двойственная задача оптимизации. Градиентный подъем и стохастический градиентный подъем.
25. Основы нейронных сетей. Модели нейронов.
26. Методы обучения отдельного нейрона.
27. Вычислительные способности отдельного нейрона.
28. Классификация нейронных сетей.
29. Аналитический метод обучения нейронных сетей.
30. Метод обратного распространения ошибки.

31. Параметры и гиперпараметры нейронной сети.
32. Недообучение и переобучение нейронной сети.
33. Регуляризация в глубоком обучении.
34. L1 и L2-регуляризация весов нейронной сети.
35. Аугментация выборки в глубоком обучении.
36. Понятие Dropout в глубоком обучении.
37. Постановка задачи обучения с подкреплением в виде марковского процесса принятия решений. Примеры.
38. Дилемма exploration/exploitation и подходы к её решению в обучении с подкреплением.
39. Понятие on-policy и off-policy алгоритмов обучения с подкреплением. Примеры.
40. Уравнения Беллмана для функций ценности. Связь функций ценности между собой.

Критерии оценки итогового тестирования

Итоговое тестирование оценивается в соответствии с БРС и паспортом ФОС. Проверяется правильность и полнота ответов на вопросы экзаменационного билета.

Наименование лабораторных работ

1. Применение градиентного спуска для оптимизации и визуализация данных
2. Решение задачи прогнозирования при помощи линейной и нелинейной регрессии.
3. Решение задачи классификации данных при помощи байесовских методов и классификации по ближайшим соседям.
4. Решение задачи классификации и регрессии при помощи обучения деревьев решений.
5. Построение решений дифференциальных уравнений при помощи нейронных сетей прямого распространения.
6. Моделирование управления динамической системой при помощи обучения с подкреплением

Критерии оценки выполнения лабораторных работ

Оценивается полнота выполнения работы, оформление результатов, наличие примеров использования, полнота ответов на контрольные вопросы, если это предусмотрено заданием.