

**Федеральное государственное образовательное бюджетное
учреждение высшего образования
«Финансовый университет при Правительстве Российской Федерации»
(Финуниверситет)**

Владикавказский филиал Финуниверситета

Кафедра «Корпоративные инфокоммуникационные системы»

УТВЕРЖДАЮ
Директор филиала
Т.А. Хубаев

« 30 » _____ июня _____ 2023 г.

М.А. Ковалева

ТЕХНОЛОГИИ МАШИННОГО ОБУЧЕНИЯ

Оценочные материалы для проведения текущего контроля успеваемости и
промежуточной аттестации студентов,
обучающихся по направлению подготовки
38.03.05 «Бизнес-информатика»,
образовательная программа «Цифровая трансформация
управления бизнесом»,
профиль «ИТ менеджмент в бизнесе»

*Одобрено заседанием кафедры «Корпоративные
инфокоммуникационные системы»
(протокол № 12 от «22» июня 2023 г..)*

Владикавказ 2023

Содержание

1. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине.....	3
2. Показатели и критерии оценивания компетенций на различных этапах их формирования, описание шкал оценивания	4
3. Оценочный лист результатов обучения к дисциплине	6
4. Контрольные материалы, необходимые для оценки индикаторов достижения компетенций, умений и знаний	7

1. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине

Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (умения и знания) соотнесенные с индикаторами достижения компетенции
ПКН-3	Способность применять аналитические системы и консультировать по вопросам разработки и развития аналитических систем работы с данными	1. Применяет аналитические системы работы с данными.	1.Знать принципы проведения научных экспериментов в области анализа больших данных и машинного обучения. Уметь проводить научные эксперименты и оценивать результаты исследований в области анализа больших данных и машинного обучения.
		2. Проводит анализ рынка аналитических систем работы с данными.	2.Знать методы проектирования интеллектуальных информационных систем с точки зрения применения в них технологий машинного обучения. Уметь создавать модели машинного обучения для решения прикладных задач и построения интеллектуальных информационных систем
		3. Консультирует по вопросам применения аналитических систем работы с данными	3. Знать сферы применения технологий машинного обучения в различных информационных системах. Уметь: применять инструментальные и программные средства машинного обучения для различных источников данных и БД.
ПКП-3	Способность предлагать различные варианты инфраструктурных решений для поддержки ИТ/ИС	1. Анализирует текущий уровень инфраструктурных решений предприятия/организации	1.Знать современные компьютерные технологии анализа данных и машинного обучения, современные проблемы и методы прикладной информатики и современное состояние научно-технического развития ИКТ. Уметь исследовать современные проблемы и методами прикладной

			информатики с учетом современного состояния научно-технического развития ИКТ, приобретать новые знания путем применения компьютерных технологий анализа данных и машинного обучения.
		2. Формирует и обосновывает варианты технологического слоя архитектуры предприятия/организации.	2.Знать Подходы к использованию современных компьютерных технологий машинного обучения для задач прикладной области. Уметь внедрять модели машинного обучения в прикладные информационные системы.

2. Показатели и критерии оценивания компетенций на различных этапах их формирования, описание шкал оценивания

Уровень сформированности компетенций			
«Минимальный уровень не достигнут» Компетенции не сформированы.	«Минимальный уровень» Компетенции сформированы.	«Средний уровень» Компетенции сформированы.	«Высокий уровень» Компетенции сформированы.
Знания отсутствуют, умения и навыки не сформированы.	Сформированы базовые структуры знаний. Умения фрагментарны и носят репродуктивный характер. Демонстрируется низкий уровень самостоятельности практического навыка.	Знания обширные, системные. Умения носят репродуктивный характер, применяются к решению типовых заданий. Демонстрируется достаточный уровень самостоятельности и устойчивого практического навыка.	Знания твердые, аргументированные, всесторонние. Умения успешно применяются к решению как типовых, так и нестандартных творческих заданий. Демонстрируется высокий уровень самостоятельности, высокая адаптивность практического навыка
Описание критериев оценивания			
Обучающийся демонстрирует: - существенные пробелы в знаниях	Обучающийся демонстрирует: - знания теоретического	Обучающийся демонстрирует: - знание и понимание	Обучающийся демонстрирует: - глубокие, всесторонние и

учебного материала; - допускаются принципиальные ошибки при ответе на основные вопросы билета, отсутствует знание и понимание основных понятий и категорий; - непонимание сущности дополнительных вопросов в рамках заданий билета; - отсутствие умения выполнять практические задания, предусмотренные программой дисциплины; - отсутствие готовности (способности) к дискуссии и низкую степень контактности.	материала; - неполные ответы на основные вопросы, ошибки в ответе, недостаточное понимание сущности излагаемых вопросов; - неуверенные и неточные ответы на дополнительные вопросы; - недостаточное владение литературой, рекомендованной программой дисциплины; - умение без грубых ошибок решать практические задания, которые следует выполнить.	основных вопросов контролируемого объема программного материала; - твердые знания теоретического материала. - способность устанавливать и объяснять связь практики и теории, выявлять противоречия, проблемы и тенденции развития; - правильные и конкретные, без грубых ошибок, ответы на поставленные вопросы; - умение решать практические задания, которые следует выполнить; - владение основной литературой, рекомендованной программой дисциплины; - наличие собственной обоснованной позиции по обсуждаемым вопросам. Возможны незначительные оговорки и неточности в раскрытии отдельных положений вопросов билета, присутствует неуверенность в ответах на	аргументированные знания программного материала; - полное понимание сущности и взаимосвязи рассматриваемых процессов и явлений, точное знание основных понятий в рамках обсуждаемых заданий; - способность устанавливать и объяснять связь практики и теории; - логически последовательные, содержательные, конкретные и исчерпывающие ответы на все задания билета, а также дополнительные вопросы экзаменатора; - умение решать практические задания; - свободное использование в ответах на вопросы материалов рекомендованной основной и дополнительной литературы.
---	---	---	--

Оценка «неудовлетворительно »	Оценка «удовлетворительно » / «зачтено»	Оценка «хорошо» / «зачтено»	Оценка «отлично» / «зачтено»
--	--	--	---

3. Оценочный лист результатов обучения к дисциплине

Примерные виды работ обучающегося, формирующие текущий контроль по дисциплине

№ п/п	Вид учебной деятельности	Баллы	Максимум за семестр (модуль)
1.	Ведение конспекта лекции / семинара и работа с ним	0,4	2
2.	Активное вовлечение в интерактивный процесс	0,6	3
3.	Ответы на вопросы на семинарском занятии	2	8
4.	Реферативный обзор дополнительного материала по теме	0,6	3
5.	Подготовка доклада по заданной теме	0,6	3
6.	Решение задач с последующим обсуждением их результатов	0,6	6
7.	Собеседование по материалу и обсуждение результатов	0,8	4
8.	Выполнение домашней контрольной (расчетно-аналитической) работы	-	10
9.	Посещение аудиторных занятий	0,2	1
	Всего за семестр (модуль)		40

Критерии бальной оценки включают качество подготовки обучающихся к семинарским занятиям, выполнения различных видов самостоятельной работы (при этом результаты регулярно проводимого экспресс-контроля успеваемости обучающихся должны составлять не менее 50% бальной оценки), а также посещение аудиторных занятий (не более 15% бальной оценки).

При подведении итогов текущего контроля успеваемости в середине семестра его результаты в 20-бальной оценке вносятся в ведомости текущего контроля успеваемости и использованием ТЭКУ (в филиалах – в ведомости текущего контроля успеваемости) и на соответствующей странице предметно-группового журнала преподавателям, ведущим аудиторные занятия.

Обучающийся, получивший от 7 до 20 баллов, считается аттестованным, получивший от 0 до 6 баллов – не аттестованным.

Результаты текущего контроля успеваемости за вторую половину семестра / в модуле преподаватели, ведущие аудиторные занятия, сообщают

обучающимся на последнем занятии и вносят в ведомости (в колонку «баллы, работа в семестре») в день проведения зачета/экзамена.

Итоговая сумма баллов, полученная обучающимся по результатам текущего контроля успеваемости в семестре/модуле и промежуточной аттестации, преобразуется в пятибалльную систему оценок.

Порядок перевода 100-бальной оценки в пятибалльную

100-бальная система	5-бальная система
86-100	отлично
70-85	хорошо
50-69	удовлетворительно
менее 50	неудовлетворительно
50-100	зачтено
менее 50	не зачтено

4. Контрольные материалы, необходимые для оценки индикаторов достижения компетенций, умений и знаний

Текущий контроль

Тема 1. Технологии машинного обучения

Вопросы для подготовки к устному опросу, докладов презентаций ПКН-3(1,2,3), ПКП-3(1,2)

1. Обучение без учителя. Метод главных компонент. Кластеризация
2. Стохастический градиентный спуск. Онлайн-обучение
3. Основы анализа временных рядов

Контрольные вопросы по теме семинарского занятия ПКН-3(1,2,3), ПКП-3(1,2)

Понятия жизненного цикла и модели жизненного цикла.

Общие принципы объектно-ориентированного проектирования.

Этапы объектно-ориентированного проектирования систем

Практико-ориентированная задача 1. ПКН-3(1,2,3), ПКП-3(1,2)

Для набора данных [“This is an error message”, “This is an info message”, “This is a success message”, “Consider this a warning”] вывести количество элементов, содержащих информацию об ошибке или предупреждение..

Тема 2. Реализация решений обработки больших данных

Вопросы для подготовки к устному опросу, докладов презентаций
ПКН-3(1,2,3), ПКП-3(1,2)

1. Построение моделей машинного обучения
2. Разработка системы потоковой аналитики мошеннических транзакций в банкоматах.

Контрольные вопросы по теме семинарского занятия ПКН-3(1,2,3), ПКП-3(1,2)

Первичный анализ данных с библиотекой Pandas

Визуальный анализ данных с библиотеками Seaborn и Matplotlib

Практико-ориентированная задача 1. ПКН-3(1,2,3), ПКП-3(1,2)

Для набора данных с 10 столбцами вещественных чисел `df = pd.DataFrame (np.random.random(size=(5, 10)), columns=list('abcdefghij'))` вывести наименование столбца с наименьшей суммой элементов.

Тема 3. Методы обучения

Вопросы для подготовки к устному опросу, докладов презентаций
ПКН-3(1,2,3), ПКП-3(1,2)

Метрические методы классификации

Гипотеза компактности и непрерывности

Метод ближайших соседей (kNN) и его обобщения.

Контрольные вопросы по теме семинарского занятия ПКН-3(1,2,3), ПКП-3(1,2)

1. Метод ближайших соседей (kNN), деревья решений,
2. Линейная регрессия, Восстановление регрессии, Регуляризация L1, L2.

Практико-ориентированная задача 1. ПКН-3(1,2,3), ПКП-3(1,2)

Таблица данных содержит столбец ключей 'keys' и столбец целочисленных значений 'vals': `df = pd.DataFrame({'keys': list('aaabbcaabcccbbc'), 'vals': [12, 345, 3, 1, 45, 14, 4, 52, 54, 23, 235, 21, 57, 3, 87]})`. Для каждого ключа найдите сумму его значений..

Тема 4. Логистическая регрессия и логические методы классификации

Вопросы для подготовки к устному опросу, докладов презентаций ПКН-3(1,2,3), ПКП-3(1,2)

4. Решающие списки и деревья.
5. Решающее правило.
6. Алгоритм синтеза списка.
7. Решающее дерево и алгоритм ID3
8. Редукция решающих деревьев.
9. Решающий лес

Контрольные вопросы по теме семинарского занятия ПКН-3(1,2,3), ПКП-3(1,2)

1. Логистическая регрессия.
2. Принцип максимума правдоподобия и логарифмическая функция потерь. Пример прикладной задачи: кредитный скоринг. Бинаризация признаков.
3. Скоринговые карты и оценивание вероятности дефолта. Риск кредитного портфеля банка. Жадный алгоритм синтеза списка.

Решающее дерево. Редукция решающих деревьев: предредукция и постредукция. Преобразование решающего дерева в решающий список. Чередующиеся решающие деревья

Практико-ориентированная задача 1. ПКН-3(1,2,3), ПКП-3(1,2)

Реализуйте алгоритмы построения дерева с критерием информационного выигрыша и критерием Джини и определению класса по

мажоритарному классу в листе. Найдите оптимальную глубину дерева в обоих случаях (в отрезке 2-10)..

Тема 5. Методы кластеризации

Вопросы для подготовки к устному опросу, докладов презентаций ПКН-3(1,2,3), ПКП-3(1,2)

10. Классификация, деревья решений и метод ближайших соседей
11. Линейные модели классификации и регрессии
12. Логистическая регрессия и случайный лес

Контрольные вопросы по теме семинарского занятия ПКН-3(1,2,3), ПКП-3(1,2)

Постановка задачи кластеризации. Примеры прикладных задач.

Типы кластерных структур.

Статистические алгоритмы: EM-алгоритм и Алгоритм k-средних (k-means).

Алгоритм построения дендрограммы.

Определение числа кластеров.

Практико-ориентированная задача 1. ПКН-3(1,2,3), ПКП-3(1,2)

Примените метод SVM (например, из библиотеки sklearn) для дата-сета blobs2. Визуализируйте результат (разбиение плоскости и опорные вектора) при разных вариантах ядер (линейное; полиномиальное степеней 2,3,5; RBF).

Тема 6. Нейронные сети в машинном обучении

Вопросы для подготовки к устному опросу, докладов презентаций ПКН-3(1,2,3), ПКП-3(1,2)

Место технологии Больших данных (Big Data) в цифровой экономике.

Архитектура обращения с Большими данными.

Угрозы и риски использования Больших данных.

Контрольные вопросы по теме семинарского занятия ПКН-3(1,2,3), ПКП-3(1,2)

1. Модель МакКаллока-Питтса как линейный классификатор. Функции активации. Проблема полноты.
2. Нейронная сеть Кохонена. Конкурентное обучение, стратегии WTA и WTM.
3. Самоорганизующаяся карта Кохонена.
4. Применение для визуального анализа данных. Искусство интерпретации карт Кохонена.

Практико-ориентированная задача 1. ПКН-3(1,2,3), ПКП-3(1,2)

Реализуйте алгоритм логистической регрессии со стохастическим градиентным спуском, обучите его на датасете `spambase_old` (train) и проверьте на датасете `spambase_new` (val). Получите ROC кривые для вариантов без нормировки и с нормировкой признаков.

Темы для выполнения контрольной работы

1. Исходный набор данных содержит 1000 записей, каждая запись представляет человека, который берет кредит в банке. **ПКН-3(1,2,3), ПКП-3(1,2)**

Набор данных содержит следующие атрибуты:

- • возраст (числовой);
- • пол (текст);
- • работа (числовой: 0 - неквалифицированный и нерезидент, 1 - неквалифицированный и резидент, 2 - квалифицированный, 3 - высококвалифицированный);
- • жилье (текст);
- • характеристика сберегательных и текущих счетов (числовой);
- • сумма кредита (числовой);
- • срок (числовой);

- • назначение кредита

Кредитный риск имеет бинарную классификацию: good, bad.

Задание: создать классификатор кредитного риска.

2. В файле `data.RData` содержатся результаты обучения и прогнозирования трех разных подходов МО. Какой стоит выбрать и по какой причине? **ПКН-3(1,2,3), ПКП-3(1,2)**
3. В файле `fictData.RData` содержатся индикаторные данные и результаты применения машинного обучения. Сравните подходы с помощью F1 критерия. **ПКН-3(1,2,3), ПКП-3(1,2)**
4. В файле `klast.RData` содержатся данные для кластерного анализа. Проведите процедуру K-mean кластеризации по фактору 1 и 2 с количеством кластеров равных 5. Как изменится распределение выбранных кластеров при добавлении третьего фактора в анализ? **ПКН-3(1,2,3), ПКП-3(1,2)**
5. . Искусственный нейрон имеет 2 входа с весами 1, 2 соответственно. Ко входному сигналу прибавляется число 1 и применяется сигмоидная функция. Чему будет равно значение на выходе из нейрона, если на первый вход было подано число 1, а на второй вход было подано число 1? **ПКН-3(1,2,3), ПКП-3(1,2)**
6. Написать программу реализующую функционал искусственного нейрона **ПКН-3(1,2,3), ПКП-3(1,2)**
7. Выполнить обучение линейных моделей с использованием функций библиотеки `scikit-learn`. Выполнить оценку метрик модели на тренировочных и проверочных данных. Изменить гиперпараметры модели для улучшения результата. **ПКН-3(1,2,3), ПКП-3(1,2)**
8. Необходимо разработать программу для обучения линейной регрессии. Разработать модификацию программы для обучения логистической регрессии **ПКН-3(1,2,3), ПКП-3(1,2)**

9. Необходимо разработать программу для реализации метода главных компонент. Разработать программу для реализации алгоритма кластеризации к-средних. **ПКН-3(1,2,3), ПКП-3(1,2)**
10. Выполнить анализ данных с использованием деревьев решений и ансамблевых методов. Выделить значимые параметры. **ПКН-3(1,2,3), ПКП-3(1,2)**
11. Провести первичный анализ данных с помощью Pandas. Выполнить визуализацию параметров набора данных. **ПКН-3(1,2,3), ПКП-3(1,2)**
12. Разработайте интеллектуальный алгоритм для следующей задачи: есть n игровых автоматов, в каждом из которых фиксирована вероятность выигрыша. Цель агента - найти слот-машину с наибольшим ожидаемым выигрышем и всегда выбирать именно ее. **ПКН-3(1,2,3), ПКП-3(1,2)**
13. Реализуйте алгоритмы построения дерева с критерием информационного выигрыша и критерием Джини и определению класса по мажоритарному классу в листе. Найдите оптимальную глубину дерева в обоих случаях (в отрезке 2-10). **ПКН-3(1,2,3), ПКП-3(1,2)**
14. Разработать скрипт на Python, позволяющий задавать с использованием функций принадлежности лингвистическую шкалу для решения задачи по варианту Количество оценок в шкале и параметры функций принадлежности должны задаваться по умолчанию, если в заданном пользователем файле не указано иное. Осуществить графическое отображение функций принадлежности меток шкалы, выделив каждую отдельным цветом. Предусмотреть загрузку файла с данными и оценку их по построенной шкале. **ПКН-3(1,2,3), ПКП-3(1,2)**
15. Реализовать скрипт на языке Python, позволяющий загружать четкий временной ряд величины из задачи по лингвистическим шкалам (задача «Лингвистические шкалы»). С помощью построенной в предыдущей работе шкалы (задача «Лингвистические шкалы»)

построить из четкого временного ряда временной ряд нечетких значений и нечетких тенденций. Осуществить графическое отображение нечетких меток шкалы и нечетких тенденций, выделив каждую отдельным цветом. **ПKN-3(1,2,3), ПКП-3(1,2**

16. Кластеризация и лингвистические шкалы. Язык реализации Python. Построить шкалу из задачи по лингвистическим шкалам (работа «Лингвистические шкалы») с помощью алгоритма fcm-клас- 274 теризации. Количество кластеров задается по умолчанию, но может редактироваться пользователем посредством загрузки из файла. Каждый кластер – метка на шкале с присвоенным ему лингвистическим значением либо по умолчанию, либо пользователем. Осуществить графическое отображение меток шкалы, выделив каждую отдельным цветом. **ПKN-3(1,2,3), ПКП-3(1,2**
17. Прогнозирование значения временного ряда с помощью нейронной сети. Реализовать прогноз следующего значения четкого временного ряда из работы «Нечеткие временные ряды и ряды нечетких тенденций» с помощью нейронной сети, но НЕ рекуррентного типа. Графически отобразить реальный ряд и прогнозное значение, а также проиллюстрировать результат тестового прогноза. **ПKN-3(1,2,3), ПКП-3(1,2**
18. Прогнозирование значения временного ряда с помощью модифицированного метода Сонга. Реализовать прогноз следующего значения временного ряда из работы «Нечеткие временные ряды и ряды нечетких тенденций» с помощью метода, описанного в пункте «Пример моделирования временного ряда в нечетком подходе». Графически отобразить реальный ряд и прогнозное значение, а также проиллюстрировать результат тестового прогноза. **ПKN-3(1,2,3), ПКП-3(1,2**

19. F-преобразование. Для временного ряда из работы «Нечеткие временные ряды и ряды нечетких тенденций» выделить тренд методом F-преобразования. **ПKN-3(1,2,3), ПКП-3(1,2)**
- 20.. На языке Python разработайте скрипт, который с помощью генетического алгоритма решает следующую задачу. Дано N наименований продуктов, для каждого из которых известно m характеристик. Необходимо получить самый дешевый рацион из k наименований, удовлетворяющий заданным медицинским нормам для каждой из m характеристик. **ПKN-3(1,2,3), ПКП-3(1,2)**
21. На языке Python разработайте скрипт, который с помощью генетического алгоритма решает следующую задачу. Дано n пунктов производства продуктов и k городов, которые в них нуждаются. Каждый город может потребить x продуктов, а каждый пункт произвести y продуктов. Необходимо получить оптимальный маршрут, так, чтобы все города получили нужный им объем продуктов без сильного его превышения, а транспортные расходы были минимальными. **ПKN-3(1,2,3), ПКП-3(1,2)**
- 22.3. На языке Python разработайте скрипт, который с помощью генетического алгоритма решает следующую задачу. Дано N наименований продуктов, для каждого из которых известно m характеристик. Необходимо получить самый лучший по характеристикам рацион из k наименований, удовлетворяющий заданным ценовым рамкам. Лучшим считается рацион с минимальным отклонением от нормы. **ПKN-3(1,2,3), ПКП-3(1,2)**
23. На языке Python разработайте скрипт, который с помощью генетического алгоритма решает следующую задачу. Дано n пунктов производства продуктов и k городов, которые в них нуждаются. Каждый город может потребить x продуктов, а каждый пункт произвести y продуктов. Необходимо получить оптимальный маршрут, так, чтобы все города получили нужный им объем продуктов с

минимальным его превышением, а транспортные расходы укладывались в определенные рамки. **ПКН-3(1,2,3), ПКП-3(1,2)**

24. На языке Python разработайте скрипт, который с помощью генетического алгоритма решает следующую задачу. Дано N полей и k культур для посева. Для каждого поля известна характеристика урожайности каждой из k культур, а для каждой культуры – его закупочная стоимость. Необходимо получить самый лучший урожай за наименьшую стоимость. **ПКН-3(1,2,3), ПКП-3(1,2)**

Промежуточная аттестация

Вопросы для подготовки к зачету

1. Особенности типов данных языка Python и библиотеки NumPy **ПКН-3(1,2,3), ПКП-3(1,2)**
2. Массивы библиотеки NumPy **ПКН-3(1,2,3), ПКП-3(1,2)**
3. Сравнения, маски и булева логика NumPy **ПКН-3(1,2,3), ПКП-3(1,2)**
4. «Прихотливая» индексация NumPy **ПКН-3(1,2,3), ПКП-3(1,2)**
5. Объект Series библиотеки Pandas **ПКН-3(1,2,3), ПКП-3(1,2)**
6. Объект DataFrame библиотеки Pandas **ПКН-3(1,2,3), ПКП-3(1,2)**
7. Объект Index библиотеки Pandas **ПКН-3(1,2,3), ПКП-3(1,2)**
8. Обработка отсутствующих данных в Pandas **ПКН-3(1,2,3), ПКП-3(1,2)**
9. Методы индексации в Pandas. **ПКН-3(1,2,3), ПКП-3(1,2)**
10. Визуальные способы оценки совокупности данных: гистограмма, полигон частот, диаграмма размаха, скрипичная диаграмма, теплокарта. **ПКН-3(1,2,3), ПКП-3(1,2)**
11. Специфические формы представления: деревья, наборы правил, наборы карт, диаграммы рассеяния. **ПКН-3(1,2,3), ПКП-3(1,2)**

12. Задачи машинного обучения. ПКН-3(1,2,3), ПКП-3(1,2)
13. Принципы обучения с учителем и без учителя. ПКН-3(1,2,3), ПКП-3(1,2)
14. Оценка качества результатов обучения. ПКН-3(1,2,3), ПКП-3(1,2)
15. Наивная байесовская классификация. ПКН-3(1,2,3), ПКП-3(1,2)
16. Простая линейная регрессия. ПКН-3(1,2,3), ПКП-3(1,2)
17. Регрессия по комбинации базисных функций. ПКН-3(1,2,3), ПКП-3(1,2)
18. Регуляризация базисных функций. ПКН-3(1,2,3), ПКП-3(1,2)
19. Метод опорных векторов. ПКН-3(1,2,3), ПКП-3(1,2)
20. Метод опорных векторов для решения нелинейных задач классификации. ПКН-3(1,2,3), ПКП-3(1,2)
21. Деревья принятия решений. ПКН-3(1,2,3), ПКП-3(1,2)
22. Случайные леса. ПКН-3(1,2,3), ПКП-3(1,2)
23. Метод главных компонент. ПКН-3(1,2,3), ПКП-3(1,2)
24. Понятие больших данных. Характеристики больших данных ПКН-3(1,2,3), ПКП-3(1,2)
25. Экосистема Hadoop. Принципы построения отличительные особенности ПКН-3(1,2,3), ПКП-3(1,2)
26. Парадигма MapReduce. ПКН-3(1,2,3), ПКП-3(1,2)
27. Общая структура кластера Hadoop. ПКН-3(1,2,3), ПКП-3(1,2)
28. Принципы распределенного хранения данных в Hadoop. ПКН-3(1,2,3), ПКП-3(1,2)
29. Принципы организации распределенных вычислений в Hadoop. ПКН-3(1,2,3), ПКП-3(1,2)
30. Компоненты экосистемы Hadoop. ПКН-3(1,2,3), ПКП-3(1,2)
31. Файловая система HDFS. Общие понятия, особенности. ПКН-3(1,2,3), ПКП-3(1,2)
32. Особенности типов данных языка Python и библиотеки NumPy. ПКН-3(1,2,3), ПКП-3(1,2)

- 33. Массивы библиотеки NumPy. ПКН-3(1,2,3), ПКП-3(1,2)
- 34. Сравнения, маски и булева логика NumPy. ПКН-3(1,2,3), ПКП-3(1,2)
- 35. «Прихотливая» индексация NumPy. ПКН-3(1,2,3), ПКП-3(1,2)
- 36. Объект Series библиотеки Pandas. ПКН-3(1,2,3), ПКП-3(1,2)
- 37. Объект DataFrame библиотеки Pandas. ПКН-3(1,2,3), ПКП-3(1,2)
- 38. Объект Index библиотеки Pandas. ПКН-3(1,2,3), ПКП-3(1,2)
- 39. Обработка отсутствующих данных в Pandas. ПКН-3(1,2,3), ПКП-3(1,2)
- 40. Разновидности аналитических оценок. ПКН-3(1,2,3), ПКП-3(1,2)
- 41. Критерий AIC. ПКН-3(1,2,3), ПКП-3(1,2)
- 42. Байесовский информационный критерий (BIC).
- 43. Статистические критерии: коэффициент детерминации, критерий Фишера, анализ регрессионных остатков. ПКН-3(1,2,3), ПКП-3(1,2)
- 44. Сложность задачи отбора признаков. ПКН-3(1,2,3), ПКП-3(1,2)