

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

Владикавказский филиал Финуниверситета

Кафедра «Корпоративные инфокоммуникационные системы»

М.А. Ковалева

ТЕХНОЛОГИИ МАШИННОГО ОБУЧЕНИЯ

Рабочая программа дисциплины

для студентов, обучающихся по направлению подготовки
38.03.05 «Бизнес-информатика»,
образовательная программа «Цифровая трансформация управления бизне-
сом»,
профиль «ИТ- менеджмент в бизнесе»

Владикавказ 2023

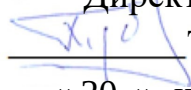
**Федеральное государственное образовательное бюджетное
учреждение высшего образования
«Финансовый университет при Правительстве Российской Федерации»
(Финуниверситет)**

Владикавказский филиал Финуниверситета

Кафедра «Корпоративные инфокоммуникационные системы»

УТВЕРЖДАЮ

Директор филиала

 Т.А. Хубаев

« 30 » июня 2023

М.А. Ковалева

ТЕХНОЛОГИИ МАШИННОГО ОБУЧЕНИЯ

Рабочая программа дисциплины

для студентов, обучающихся по направлению подготовки
38.03.05 «Бизнес-информатика»,
образовательная программа «Цифровая трансформация управления бизне-
сом»,
профиль «ИТ- менеджмент в бизнесе»

*Рекомендовано Ученым советом Владикавказского филиала
Финансового университета
(протокол от «28» июня 2023 г. № 2.)*

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

Владикавказ 2023

Содержание

1. Наименование дисциплины.....	4
2. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине.....	4
3. Место дисциплины в структуре образовательной программы.....	5
4. Объем дисциплины (модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся.....	5
5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий.....	5
5.1. Содержание дисциплины.....	5
5.2. Учебно-тематический план.....	7
5.3. Содержание семинаров, практических занятий.....	8
6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине.....	9
6.1 Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы.....	9
6.2 Перечень вопросов, заданий, тем для подготовки к текущему контролю.....	10
7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине.....	13
8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины.....	16
9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины.....	17
10. Методические указания для обучающихся по освоению дисциплины.....	18

11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем (при необходимости).....	22
11.1. Комплект лицензионного программного обеспечения.....	22
11.2. Современные профессиональные базы данных и информационные справочные системы.....	22
11.3. Сертифицированные программные и аппаратные средства защиты информации.....	22
12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине.....	22

1. Наименование дисциплины

Учебная дисциплина «Технологии машинного обучения».

2. 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.

3. Место дисциплины в структуре образовательной программы

Дисциплина входит в цикл профиля (элективный) части модуль «Сквозные цифровые технологии в бизнесе», формируемой участниками образовательных отношений образовательной программы по направлению подготовки 38.03.05 «Бизнес-информатика», образовательная программа «Цифровая трансформация управления бизнесом», профиль «ИТ- менеджмент в бизнесе».

4. Объем дисциплины (модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся

Вид учебной работы по дисциплине	Всего (в з/е и часах)	Семестр 7 (в часах)
Общая трудоемкость дисциплины	3 /108 час.	108
Контактная работа- Аудиторные занятия	30	30
<i>Лекции</i>	14	14
<i>Семинары, практические занятия</i>	16	16
Самостоятельная работа	78	78
Вид текущего контроля	Контрольная работа	Контрольная работа
Вид промежуточной аттестации	Зачет	Зачет

5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий

5.1. Содержание дисциплины

Тема 1. Технологии машинного обучения

Общая постановка задачи машинного обучения. Обучение с учителем и без учителя, обучение с подкреплением. Задачи классификации, восстановления регрессии, предсказания. Модели алгоритмов. Признаки. Типы признаков. Понятие функционала качества. Вероятностная постановка задачи. Оценка обобщающей способности. Проблема переобучения. Критерии оценки качества работы алгоритмов машинного обучения. ROC-кривые. Примеры практических задач машинного обучения

Тема 2. Реализация решений обработки больших данных

Общая постановка задачи машинного обучения. Информационные технологии для обработки больших данных. Первичный анализ данных с библиотекой Pandas. Визуальный анализ данных с библиотеками Seaborn и Matplotlib

Тема 3. Методы обучения

Обобщенный метрический классификатор. Виды метрик. Метод ближайшего соседа. Алгоритм k-ближайших соседей. Взвешенная версия алгоритма k-ближайших соседей. Метод окна Парзена. Метод потенциальных функций. Понятие эталона. Отступы и классификация объектов. Классификация, деревья решений и метод ближайших соседей

Тема 4. Логистическая регрессия и логические методы классификации

Постановка задач линейной регрессии и линейной классификации. Линейная и многомерная регрессия в библиотеке Scikit-Learn. Линейная регрессия, Lasso и RF-регрессия. Построение и отбор признаков Логистическая регрессия. Принцип максимума правдоподобия и логарифмическая функция потерь. Пример прикладной задачи: кредитный скоринг. Бинаризация признаков. Скоринговые карты и оценивание вероятности дефолта. Риск кредитного портфеля банка. Настройка порога решающего правила по критерию числа

ошибок I и II рода. Кривая ошибок (ROC curve). Алгоритм эффективного построения ROC-кривой. Градиентный метод максимизации площади под ROC-кривой.

Тема 5. Методы кластеризации

Методы кластеризации. Типы кластерных структур. Функционал качества кластеризации. EM-алгоритм. Метод k-средних. Иерархическая кластеризация. Формула Ланса-Уильямса. Быстрая агломеративная кластеризация. Прикладные задачи снижения размерности. Методы снижения размерности и метод главных компонент в библиотеке Scikit-Learn.

Тема 6. Нейронные сети в машинном обучении

Нейронные сети и Deep Learning. Нейронные сети для классификации 3. Нейронные сети для прогнозирования: feedforward nets. Нейронные сети для прогнозирования: LSTM. Сверточные (convolution) сети.

5.2. Учебно-тематический план

№ п/ п	Наименование тем (разделов) дисциплины	Трудоёмкость в часах					Формы теку- щего контроля успеваемости
		Всего	Контактная работа - Аудиторная работа			Самостоя- тельная работа	
			Общая, в т.ч.:	Лек- ции	Семинары, практиче- ские заня- тия		
1.	Тема 1. Технологи- и машинного обучения	14	4	2	2	10	Дискуссия, об- суждение, вы- полнение зада- ний
2.	Тема 2. Реализация решений обра- ботки больших данных	14	4	2	2	10	Дискуссия, об- суждение, вы- полнение зада- ний
3.	Тема 3. Методы обучения	14	4	2	2	10	Дискуссия, об- суждение, вы- полнение зада- ний
4.	Тема 4. Логистиче- ская регрессия и логические методы классификации	14	4	2	2	10	Дискуссия, об- суждение, вы- полнение зада- ний

5.	Тема 5. Методы кластеризации	14	4	2	2	10	Дискуссия, обсуждение, выполнение заданий
6.	Тема 6. Нейронные сети в машинном обучении	38	10	4	6	28	Дискуссия, обсуждение, выполнение заданий. Защита контрольной работы
В целом по дисциплине		108	30	14	16	78	Согласно учебному плану: контрольная работа
Итого в %		100	27,8	14,8	14,8	72,2	

5.3. Содержание семинаров, практических занятий

Наименование тем (разделов) дисциплины	Перечень вопросов для обсуждения на семинарах, практических занятиях, рекомендуемые источники из разделов 8,9 (указывается раздел и порядковый номер источника)	Формы проведения занятий
Тема 1. Технологии машинного обучения	1. Понятия жизненного цикла и модели жизненного цикла. 2. Общие принципы объектно-ориентированного проектирования. 3. Этапы объектно-ориентированного проектирования систем Рекомендуемые источники: 8: (1,2,3); 9: (1,2,3,4,5,6)	Опрос, доклады (рефераты), презентации, решение практико-ориентированных задач
Тема 2. Реализация решений обработки больших данных	1. Построение моделей машинного обучения 2. Разработка системы потоковой аналитики мошеннических транзакций в банкоматах. Рекомендуемые источники: 8: (1,2,3); 9: (1,2,3,4,5,6)	Опрос, доклады (рефераты), презентации, решение практико-ориентированных задач
Тема 3. Методы обучения	1. Метод ближайших соседей (kNN), деревья решений, 2. Линейная регрессия, Восстановление регрессии, Регуляризация L1, L2. Рекомендуемые источники: 8: (1,2,3); 9: (1,2,3,4,5,6)	Опрос, доклады (рефераты), презентации, решение практико-ориентированных задач
Тема 4. Логистическая регрессия и	1. Логистическая регрессия.	Опрос, доклады (рефераты), презентации, решение

логические методы классификации	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. 4. Жадный алгоритм синтеза списка. 5. Решающее дерево. Редукция решающих деревьев: предредукция и постредукция. Преобразование решающего дерева в решающий список. Чередующиеся решающие деревья Рекомендуемые источники: 8: (1,2,3); 9: (1,2,3,4,5,6)	практико-ориентированных задач
Тема 5. Методы кластеризации	1. Постановка задачи кластеризации. Примеры прикладных задач. 2. Типы кластерных структур. 3. Статистические алгоритмы: EM-алгоритм и Алгоритм k-средних (k-means). 4. Алгоритм построения дендрограммы. 5. Определение числа кластеров. Рекомендуемые источники: 8: (1,2,3); 9: (1,2,3,4,5,6)	Опрос, доклады (рефераты), презентации, решение практико-ориентированных задач
Тема 6. Нейронные сети в машинном обучении	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. Рекомендуемые источники: 8: (1,2,3); 9: (1,2,3,4,5,6)	Опрос, доклады (рефераты), презентации, решение практико-ориентированных задач. Защита контрольной работы

6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине.

6.1. Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы

Наименование тем (разделов) дисциплины	Перечень вопросов, отводимых на самостоятельное освоение	Формы внеаудиторной самостоятельной работы
Тема 1. Технологии машинного обучения	Модели алгоритмов.	Подготовка презентации на занятие Изучение основной и дополнительной литературы, поиск и анализ информации, содержащейся в Интернет-ресурсах.

Тема 2. Реализация решений обработки больших данных	Библиотеки Pandas. Seaborn и Matplotlib	Подготовка презентации на занятие Изучение основной и дополнительной литературы, поиск и анализ информации, содержащейся в Интернет-ресурсах.
Тема 3. Методы обучения	Понятие эталона. Отступы и классификация объектов. Классификация, деревья решений и метод ближайших соседей	Подготовка презентации на занятие Изучение основной и дополнительной литературы, поиск и анализ информации, содержащейся в Интернет-ресурсах.
Тема 4. Логистическая регрессия и логические методы классификации	Линейная регрессия Логистическая регрессия	Подготовка презентации на занятие Изучение основной и дополнительной литературы, поиск и анализ информации, содержащейся в Интернет-ресурсах.
Тема 5. Методы кластеризации	Метод k-средних. Иерархическая кластеризация.	Подготовка презентации на занятие Изучение основной и дополнительной литературы, поиск и анализ информации, содержащейся в Интернет-ресурсах.
Тема 6. Нейронные сети в машинном обучении	Нейронные сети	Подготовка презентации на занятие Изучение основной и дополнительной литературы, поиск и анализ информации, содержащейся в Интернет-ресурсах.

6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю

Примерный перечень тем для подготовки к устному опросу, докладов (презентаций)

1. Место технологии Больших данных (Big Data) в цифровой экономике. Архитектура обращения с Большими данными. Угрозы и риски использования Больших данных.
2. Метрические методы классификации
3. Гипотеза компактности и непрерывности
4. Метод ближайших соседей (kNN) и его обобщения.
5. Логические методы классификации.
6. Решающие списки и деревья.

7. Решающее правило.
8. Алгоритм синтеза списка.
9. Решающее дерево и алгоритм ID3
10. Редукция решающих деревьев.
11. Решающий лес
12. Байесовские классификаторы
13. Метод потенциальных функций, градиентный алгоритм.
14. Отбор эталонных объектов.
15. Первичный анализ данных с библиотекой Pandas
16. Визуальный анализ данных с библиотеками Seaborn и Matplotlib
17. Классификация, деревья решений и метод ближайших соседей
18. Линейные модели классификации и регрессии
19. Логистическая регрессия и случайный лес
20. Линейная регрессия, Lasso и RF-регрессия. Построение и отбор признаков
21. Обучение без учителя. Метод главных компонент. Кластеризация
22. Стохастический градиентный спуск. Онлайн-обучение
23. Основы анализа временных рядов

Примеры практико-ориентированных задач

Задание 1.

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.

Задание 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.

Задание 3.

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.

Задание 4.

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).

Задание 5.

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

Задание 6.

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.

Задание 7.

Модифицируйте модель из задачи 3, заменив последний нейрон на 10 нейронов, и реализовав мультиклассовую классификацию с `softmax` в качестве решающей функции и кросс-энтропией в качестве функции потерь и обучите на подготовленном датасете `mnist`.

Примерное задание для выполнения контрольной работы

Примеры заданий контрольной работы

Исходный набор данных содержит 1000 записей, каждая запись представляет человека, который берет кредит в банке.

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.

Примерные темы для выполнения контрольной работы

1. Задачи машинного обучения.
2. Метрические методы классификации.
3. Метод ближайших соседей (kNN) и его обобщения.
4. Методы классификации.
5. Решающие списки и деревья.
6. Редукция решающих деревьев.
7. Решающий лес.
8. Байесовские классификаторы.
9. Вероятностная постановка задачи классификации.
10. Метод опорных векторов.

«Критерии балльной оценки различных форм текущего контроля успеваемости содержатся в соответствующих методических рекомендациях кафедры «Корпоративные инфокоммуникационные системы».

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

Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов достижения и планируемых результатов обучения по дисциплине содержится в разделе 2 «Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине».

Типовые контрольные задания или иные материалы, необходимые для оценки индикаторов достижения компетенций, умений и знаний

Примерный перечень вопросов для подготовки к зачету

1. Особенности типов данных языка Python и библиотеки NumPy
2. Массивы библиотеки NumPy
3. Сравнения, маски и булева логика NumPy
4. «Прихотливая» индексация NumPy
5. Объект Series библиотеки Pandas
6. Объект DataFrame библиотеки Pandas
7. Объект Index библиотеки Pandas
8. Обработка отсутствующих данных в Pandas
9. Методы индексации в Pandas.
10. Визуальные способы оценки совокупности данных: гистограмма, полигон частот, диаграмма размаха, скрипичная диаграмма, теплокарта.
11. Специфические формы представления: деревья, наборы правил, наборы карт, диаграммы рассеяния.
12. Задачи машинного обучения.
13. Принципы обучения с учителем и без учителя.
14. Оценка качества результатов обучения.
15. Наивная байесовская классификация.
16. Простая линейная регрессия.
17. Регрессия по комбинации базисных функций.

18. Регуляризация базисных функций.
19. Метод опорных векторов.
20. Метод опорных векторов для решения нелинейных задач классификации.

Примеры оценочных средств для проверки индикаторов достижения компетенций, формируемых дисциплиной

Наименование компетенции	Наименование индикатора достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
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.	Вопросы для проведения опроса, подготовки докладов (рефератов), презентаций Место технологии Больших данных (Big Data) в цифровой экономике. Архитектура обращения с Большими данными. Угрозы и риски использования Больших данных. Метрические методы классификации Практическое задание Задача 1 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.
	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.	Вопросы для проведения опроса, подготовки докладов (рефератов), презентаций Гипотеза компактности и непрерывности Метод ближайших соседей (kNN) и

		Уметь создавать модели машинного обучения для решения прикладных задач и построения интеллектуальных информационных систем	его обобщения. Логические методы классификации. Практическое задание-Задача1 Загрузить данные о финансовой активности клиентов банка из файлов в форматах JSON, CSV, TXT для последующей обработки Задача2 Найдите максимум функции с помощью алгоритма кросс-энтропийного поиска, изображая распределение на каждом шаге.
	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.	Вопросы для проведения опроса, подготовки докладов (рефератов), презентаций Построение моделей машинного обучения Практическое задание Задача 1 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).
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	Вопросы для проведения опроса, подготовки докладов (рефератов), презентаций Роль облачных вычислений в обработке и хранении данных, получаемых от IoT-систем Практическое задание Задача1 Для заданного набора данных о клиентах банка на основе применения технологий машинного обучения создать классификатор для кредитного скоринга. Задача2

		развития ИКТ, приобретать новые знания путем применения компьютерных технологий анализа данных и машинного обучения.	Реализуйте алгоритм линейной регрессии, и полиномиальной регрессии (для датасета noisysine – степеней от 2 до 5, для датасета hydrodynamics – степени 2) без регуляризации.
	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.	Вопросы для проведения опроса, подготовки докладов (рефератов), презентаций Решающие списки и деревья. Решающее правило. Алгоритм синтеза списка. Решающее дерево и алгоритм ID3 Редукция решающих деревьев. Решающий лес практическое задание Задание 1 Написать программу реализации алгоритма машинного обучения, позволяющую классифицировать клиентов банка по их финансовой активности. Задание 2 Для набора данных <code>df = pd.DataFrame(np.random.random(size=(5, 3)))</code> вывести список средних значений для каждой строки данных.

8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

Основная литература

1. Кондрашов, Ю.Н. Анализ данных и машинное обучение на платформе MS SQL Server: учебное пособие / Ю.Н. Кондрашов. — Москва: Русайнс, 2020. — 303 с. — Текст : непосредственный.— То же .— ЭБС BOOK.ru. — URL: <https://book.ru/book/933497>— Текст : электронный.
2. Миркин, Б.Г. Введение в анализ данных: учебник и практикум для бакалавриата и магистратуры / Б.Г. Миркин. - Москва: Юрайт, 2019. - 174 с. – Текст : непосредственный. Миркин, Б. Г. Введение в анализ данных

: учебник и практикум / Б. Г. Миркин. — Москва : Юрайт, 2020. — 174 с. — (Высшее образование). — ЭБС Юрайт. — URL: <https://urait.ru/bcode/450262>. — Текст : электронный.

Дополнительная литература

3. Анализ данных : учебник для вузов / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва : Юрайт, 2020. — 490 с. — (Высшее образование).—ЭБС Юрайт. — URL: <https://urait.ru/bcode/> — Текст : электронный.
4. Мартишин, С.А. Базы данных: Работа с распределенными базами данных и файловыми системами на примере MongoDB и HDFS с использованием Node.js, Express.js, Apache Spark и Scala: учебное пособие / С.А. Мартишин, В.Л. Симонов, М.В. Храпченко. — Москва: ИНФРА-М, 2019. — 235 с. — ЭБС ZNANIUM.com. — URL: <http://znanium.com/catalog/product/1018196>. - Текст : электронный.

9. Перечень ресурсов информационно-телекоммуникационной сети

«Интернет», необходимых для освоения дисциплины

1. Электронно-библиотечная система BOOK.RU <http://www.book.ru>
2. Электронно-библиотечная система Znanium <http://www.znanium.com>
3. Электронно-библиотечная система «Университетская библиотека ОНЛАЙН» <http://biblioclub.ru/>
4. Электронно-библиотечная система издательства «ЮРАЙТ» <https://www.biblio-online.ru>
5. Электронная библиотека издательского дома «Гребенников» <https://grebennikon.ru>
6. Электронно-библиотечная система издательства «Лань» <https://e.lanbook.com>

10. Методические указания для обучающихся по освоению дисциплины

Методика освоения дисциплины предусматривает подготовку обучающихся к лекциям, семинарам и практическим занятиям, выполнение студентами самостоятельной внеаудиторной работы, в том числе – контрольной работы.

Рекомендации по подготовке к лекционным занятиям.

Для наиболее полного освоения дисциплины студентам необходимо:

- перед каждой лекцией просматривать рабочую программу дисциплины, ее основные вопросы и рекомендуемую литературу. Это позволит сэкономить время на записывание основных вопросов темы;
- к лекциям готовить материал по теме;
- перед очередной лекцией просматривать материалы предыдущих, чтобы освоение материала не оставляло пробелов.

Рекомендации по подготовке к семинарам, практическим занятиям.

Студентам следует:

- проработать теоретический материал к занятию по рекомендованным литературным источникам и лекциям;
- использовать при подготовке к занятию нормативно-правовые документы, научные публикации, информационный материал, рекомендуемый преподавателем;
- перед занятиями задать вопросы по невыясненным в ходе самостоятельной подготовки темам или отдельным положениям;
- в ходе занятия давать четкие и исчерпывающие ответы на вопросы;
- на занятии демонстрировать понимание обсуждаемых тем и вопросов.

Студентам, по различным причинам пропустившим занятия, необходимо перед очередным занятием отработать пропущенный материал, подготовив его самостоятельно.

Методические рекомендации по выполнению различных форм самостоятельной работы

Самостоятельная работа содержит в себе различные виды и формы работ. В ходе изучения дисциплины предусмотрены следующие формы:

- опрос;
- подготовка докладов (рефератов), презентаций;
- выполнение контрольной работы.
- подготовка к зачету.

Самостоятельная работа студентов включает в себя выполнение различного рода занятий, которые ориентированы на более глубокое усвоение материала изучаемой дисциплины. По каждой теме учебной дисциплины студентам предлагается перечень заданий для самостоятельной работы.

К выполнению заданий для самостоятельной работы предъявляются следующие требования: задания должны выполняться самостоятельно и представляться в установленный срок, а также должны соответствовать установленным требованиям по оформлению.

Студентам следует:

- руководствоваться графиком самостоятельной работы, определенным РПД;
- выполнять все плановые задания, выдаваемые преподавателем для самостоятельного выполнения, разбирать на занятиях и консультациях неясные вопросы;
- прорабатывать соответствующие теоретические и практические разделы дисциплины, фиксируя неясные фрагменты для их обсуждения на консультации.

Методические рекомендации по выполнению контрольной работы

Контрольная работа – вид самостоятельной практической и исследовательской работы студентов, с целью углубления и закрепления теоретических знаний. Для выполнения контрольной работы по дисциплине необходимо использовать учебно-методические материалы, подготовленные кафедрой.

В процессе выполнения контрольной работы студенту рекомендуется

выполнить следующие виды работ:

- составить план контрольной работы;
- отобрать источники, собрать и проанализировать информацию по проблеме;
- систематизировать и проанализировать собранную информацию по научной проблеме;
- представить проведенный анализ с собственными выводами и предложениями.

Контрольная работа выполняется студентом под руководством преподавателя кафедры «Корпоративные инфокоммуникационные системы» самостоятельно. Тему контрольной работы студент выбирает из предлагаемого примерного перечня.

Контрольная работа имеет следующую структуру:

1. Титульный лист.
2. Содержание.
3. Введение (с обоснованием выбора темы).
4. Текстовое изложение материала.
5. Заключение с выводами по всей работе.
6. Список использованных источников.
7. Приложения (при необходимости)

Введение - суть и обоснование выбора данной темы, состоит из ряда компонентов, связанных логически и стилистически.

Текстовое изложение материала (основная часть) - теоретические основы выбранной проблемы и изложение основного вопроса. Данная часть предполагает развитие аргументации и анализа, а также обоснование их, исходя из имеющихся данных, других аргументов и позиций по этому вопросу.

Заключение - обобщения и аргументированные выводы по теме контрольной работы с указанием области ее применения и т.д. Оно подытоживает контрольную работу или еще раз вносит пояснения, подкрепляет смысл и значение изложенного в основной части материала.

К тексту прилагается оформленный по требованиям список использованных источников. При составлении списка в перечень включаются только те источники, которые действительно были использованы при подготовке к написанию работы.

Приложения - иллюстративный материал (схемы, диаграммы, рисунки и др.).

Требования к оформлению контрольной работы

Контрольная работа выполняется на компьютере (гарнитура Times New Roman, шрифт 14) через 1,5 интервала с полями: верхнее, нижнее - 2; правое - 3; левое - 1,5. Отступ первой строки абзаца - 1,25. Должна быть нумерация страниц. Таблицы и рисунки встраиваются в текст работы и тоже нумеруются.

Титульный лист является первой страницей и заполняется по строго определенным правилам.

Список использованных источников составляет одну из частей работы, отражающей самостоятельную творческую работу автора, позволяет судить о степени фундаментальности данной работы. Общее оформление списка использованных источников для контрольной работы аналогично оформлению списка использованной литературы для реферата. В тексте обязательны ссылки на литературные источники, лучше всего постраничные.

К тексту контрольной работы может прилагаться иллюстративный материал (схемы, диаграммы, рисунки, таблицы и др.), оформленный в виде приложений.

11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем (при необходимости)

11.1 Комплект лицензионного программного обеспечения:

1) Антивирусная защита Kaspersky Security для виртуальных и облачных сред;

2) Windows, Microsoft Office.

3) Дистрибутив языка Python 3.11 (или более поздней версии) Anaconda

3

11.2 Современные профессиональные базы данных, и информационные справочные системы

1. Информационно-правовая система «Гарант»: <https://www.garant.ru>

2. Электронная энциклопедия: <http://ru.wikipedia.org/wiki/Wiki>

3. Большая Российская энциклопедия: <https://bigenc.ru/>

11.3 Сертифицированные программные и аппаратные средства защиты информации

Не используются

12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

Учебная аудитория для проведения учебных занятий всех видов, предусмотренных программой магистратуры, оснащенная оборудованием и техническими средствами обучения.

362002, Республика Северная Осетия-Алания, г. Владикавказ, ул. Молодежная, д. 7, ауд. № 47.

Специализированная мебель:

Стол (учительский) – 1 шт.

Стол компьютерный – 18 шт.

Стол (студенческий) двухместный – 6 шт.

Стулья – 26 шт.

Доска для маркера – 1 шт.

Шкаф для документов – 1 шт.

Технические средства обучения:

Компьютер в сборе – 19 шт.,

Мультимедиа-проектор – 1 шт.,

Экран настенный – 1 шт.,

Перечень лицензионного программного обеспечения:

1) Программы для ЭВМ Kaspersky Security для виртуальных и облачных сред;

2) Продукты компании Microsoft, включая ОС Windows Office.

Учебная аудитория оснащена компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду Финансового университета.

Помещения для самостоятельной работы обучающихся:

362002, Республика Северная Осетия-Алания, г. Владикавказ, ул. Молодежная, д.7, каб. № 72.

Специализированная мебель:

Стол (учительский) – 1 шт.

Стол (студенческий) двухместный – 13 шт.

Стулья – 27 шт.

Доска меловая – 1 шт.

Шкаф для документов – 1 шт.

Технические средства обучения:

Компьютер в сборе – 10 шт.,

Мультимедиа-проектор – 1 шт.,

Экран настенный – 1 шт.

Перечень лицензионного программного обеспечения:

1) Программы для ЭВМ Kaspersky Security для виртуальных и облачных сред;

2) Продукты компании Microsoft, включая ОС Windows Office.

362002, Республика Северная Осетия-Алания, г. Владикавказ, ул. Молодежная, д.7, каб. № 55. Читальный зал.

Специализированная мебель:

Столы – 20 шт.

Стулья – 40 шт.

Шкаф для книг – 4 шт.

Стеллажи книжные – 13 шт.

Стеллажи выставочные – 4 шт.

Компьютер в сборе – 6 шт.

Телевизор – 1 шт.

Перечень лицензионного программного обеспечения:

1) Программы для ЭВМ Kaspersky Security для виртуальных и облачных сред;

2) Продукты компании Microsoft, включая ОС Windows Office.

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