

Документ подписан простой электронной подписью

Информация о владельце: МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФИО: Силин Яков Петрович

Должность: Ректор

Дата подписания: 10.07.2024 15:08:03

Уникальный программный код:

24f866be2aca16484036841850905116056

ФГБОУ ВО «Уральский государственный экономический университет»

Одобрена

на заседании кафедры

26.11.2023 г.

протокол № 4

Зав. кафедрой Назаров Д.М.

Утверждена

Советом по учебно-методическим
вопросам и качеству образования

13 декабря 2023 г.

протокол № 4

Председатель

Карх Д.А.

(подпись)

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Наименование дисциплины	Интеллектуальный анализ данных
Направление подготовки	09.04.03 Прикладная информатика
Профиль	Цифровая бизнес-аналитика
Форма обучения	очная
Год набора	2024

Разработана:
Доцент, к.х.н
Галактионов А.Д.

Профессор, д.э.н.
Назаров Д.М.

Екатеринбург
2023 г.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	3
1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ	3
2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП	3
3. ОБЪЕМ ДИСЦИПЛИНЫ	3
4. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ОПОП	3
5. ТЕМАТИЧЕСКИЙ ПЛАН	6
6. ФОРМЫ ТЕКУЩЕГО КОНТРОЛЯ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ШКАЛЫ ОЦЕНИВАНИЯ	7
7. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ	9
8. ОСОБЕННОСТИ ОРГАНИЗАЦИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ ДЛЯ ЛИЦ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ	12
9. ПЕРЕЧЕНЬ ОСНОВНОЙ И ДОПОЛНИТЕЛЬНОЙ УЧЕБНОЙ ЛИТЕРАТУРЫ, НЕОБХОДИМОЙ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ	12
10. ПЕРЕЧЕНЬ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, ВКЛЮЧАЯ ПЕРЕЧЕНЬ ЛИЦЕНЗИОННОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ И ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ, ОНЛАЙН КУРСОВ, ИСПОЛЬЗУЕМЫХ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ	13
11. ОПИСАНИЕ МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЙ БАЗЫ, НЕОБХОДИМОЙ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ	14

ВВЕДЕНИЕ

Рабочая программа дисциплины является частью основной профессиональной образовательной программы высшего образования - программы магистратуры, разработанной в соответствии с ФГОС ВО

ФГОС ВО	Федеральный государственный образовательный стандарт высшего образования - магистратура по направлению подготовки 09.04.03 Прикладная информатика (приказ Минобрнауки России от 19.09.2017 г. № 916)
ПС	

1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целью освоения дисциплины «Интеллектуальный анализ данных» является формирование у студентов целостного представления о современных принципах аналитической деятельности и инструментах интеллектуального анализа данных в контексте реализации перехода в цифровую экономику (экономику данных).

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

Дисциплина относится к вариативной части учебного плана.

3. ОБЪЕМ ДИСЦИПЛИНЫ

Промежуточная аттестация	Часов					З.е.
	Всего за семестр	Контактная работа (по уч.зан.)			Самостоятельная работа в том числе подготовка контрольных и курсовых	
		Всего	Лекции	Лабораторные		
Семестр 1						
Зачет	144	24	8	16	120	4
Семестр 2						
Экзамен, Курсовая работа	144	28	8	20	89	4
	288	52	16	36	209	8

4. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ОПОП

В результате освоения ОПОП у выпускника должны быть сформированы компетенции, установленные в соответствии ФГОС ВО.

Шифр и наименование компетенции	Индикаторы достижения компетенций
производственно-технологический	
ПК-4 Планирование, организация и контроль аналитических работ в ИТ-проекте	ИД-1.ПК-4 Знать Теория обучения Английский язык Методы планирования проектных работ Теория управления группой Теория управления Управление изменениями в системах План работ по разработке требований к системе

<p>ПК-4 Планирование, организация и контроль аналитических работ в ИТ-проекте</p>	<p>ИД-2.ПК-4 Уметь</p> <ul style="list-style-type: none">Создавать учебно-методические материалыПланировать проектные работыВыбирать методики и шаблоныПроводить совещанияРазрешать конфликтыКонтролировать состояние работЗаполнять формы отчета
---	---

<p>ПК-4 Планирование, организация и контроль аналитических работ в ИТ-проекте</p>	<p>ИД-3.ПК-4 Иметь практический опыт:</p> <ul style="list-style-type: none"> Исследование и изучение мировых практик выполнения аналитических работ Выявление проблем и сложностей в существующих практиках выполнения аналитических работ в организации Разработка рекомендаций по изменению практик Описание методик выполнения аналитических работ Апробация методик на выбранных проектах и их доработка Выявление потребителей требований и их интересов Определение источников информации для требований Выбор методов разработки требований Выбор типов и атрибутов требований Выбор шаблонов документов требований Составление и согласование перечня поставок требований Достижение договоренностей с потребителями требований о методах и процедуре приемки требований Создание графика поставок требований Определение состава работ по разработке требований Определение требований к компетенциям исполнителей разных работ по созданию требований Определение графика контрольных мероприятий по аналитическим работам Определение кандидатов на исполнение отдельных аналитических работ Достижение договоренностей с владельцами ресурсов об их выделении на аналитические работы Постановка задач на разработку планов аналитических работ по отдельным частям системы Интегрирование планов аналитических работ по отдельным частям системы Передача и согласование плана аналитических работ с менеджером проекта Определение состава аналитической группы проекта Знакомство аналитической группы Представление и обсуждение плана аналитических работ Распределение ролей и аналитических работ по участникам аналитической группы проекта Ответы на вопросы и предложения участников аналитической группы проекта Достижение соглашений с владельцами ресурсов о выделении ресурсов для выполнения аналитических работ в проекте Сбор информации о состоянии аналитических работ в проекте Анализ соответствия фактического состояния работ плановому Определение причин отклонений от планов Выявление проблемных ситуаций в ходе работ Разработка мероприятий по компенсации отклонений Проведение коррекции планов аналитических работ Разрешение проблемных ситуаций в ходе аналитических работ Оценка соответствия состояния аналитических работ плановому Описание состояния аналитических работ в формате отчета Передача отчетности о состояниях аналитических работ руководителю проекта
---	--

ПК-1 Обоснование подходов, используемых в бизнес-анализе	ИД-1.ПК-1 Знать: Определение подхода к проведению бизнес-анализа Определение подхода к работе с информацией бизнес-анализа Определение подхода к работе с заинтересованными сторонами Определение подхода к разработке различных типов требований Определение подхода к работе с изменениями различных типов требований Определение подхода к оценке эффективности работы по бизнес-анализу
	ИД-2.ПК-1 Уметь: Планировать, организовывать и проводить встречи и обсуждения с заинтересованными сторонами Выявлять, регистрировать, анализировать и классифицировать риски и разрабатывать комплекс мероприятий по их минимизации Оформлять результаты бизнес-анализа в соответствии с выбранными подходами Применять информационные технологии в объеме, необходимом для целей бизнес-анализа
	ИД-3.ПК-1 Иметь практический опыт: Определение подхода к проведению бизнес-анализа Определение подхода к работе с информацией бизнес-анализа Определение подхода к работе с заинтересованными сторонами Определение подхода к разработке различных типов требований Определение подхода к работе с изменениями различных типов требований Определение подхода к оценке эффективности работы по бизнес-анализу

5. ТЕМАТИЧЕСКИЙ ПЛАН

Тема	Часов						
	Наименование темы	Всего часов	Контактная работа (по уч.зан.)			Самост. работа	Контроль самостоятельной работы
			Лекции	Лабораторные	Практические занятия		
Семестр 1		144					
Тема 1.	Интеллектуальный анализ данных (BDA). Основные понятия. ПК-1: Обоснование подходов, используемых в бизнес-анализе ПК-4: Планирование, организация и контроль аналитических работ в ИТ-проекте	72	4	8		60	
Тема 2.	Основные алгоритмы интеллектуального анализа данных. ПК-1: Обоснование подходов, используемых в бизнес-анализе ПК-4: Планирование, организация и контроль аналитических работ в ИТ-проекте	72	4	8		60	
Семестр 2		117					

Тема 3.	Основные инструментальные средства интеллектуального анализа данных.ПК-1: Обоснование подходов, используемых в бизнес-анализе ПК-4: Планирование, организация и контроль аналитических работ в ИТ-проекте	42	4	8		30	
Тема 4.	Понятие о принципах управления аналитической деятельностью на основе интеллектуального анализа данных.ПК-1: Обоснование подходов, используемых в бизнес-анализе ПК-4: Планирование, организация и контроль аналитических работ в ИТ-проекте	38	2	6		30	
Тема 5.	Комплексная работа по интеллектуальному анализу данных в цифровой экономике.ПК-1: Обоснование подходов, используемых в бизнес-анализе ПК-4: Планирование, организация и контроль аналитических работ в ИТ-проекте	37	2	6		29	

6. ФОРМЫ ТЕКУЩЕГО КОНТРОЛЯ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ШКАЛЫ ОЦЕНИВАНИЯ

Раздел/Тема	Вид оценочного средства	Описание оценочного средства	Критерии оценивания
Текущий контроль (Приложение 4)			
Тема 1	Тест (Приложение 4)	Тест состоит из 6 вопросов с выбором ответа	5
Тема 2	Контрольная работа (Приложение 4)	Контрольная работа состоит из 5 заданий с открытым ответом	10
Тема 3	Тест (Приложение 4)	Тест состоит из 6 вопросов с выбором ответа	5
Тема 4-5	Ситуационная задача (Приложение 4)	Задача на вычисление переменной по формуле	10
Промежуточная аттестация (Приложение 5)			
1 семестр (За)	Тест (Приложение 5)	30 тестовых вопросов с открытым ответом	100 баллов
2 семестр (Эк)	Экзаменационные билеты (Приложение 5)	20 билетов. Каждый билет содержит 2 теоретических вопроса и 1 практическое задание	100 баллов
2 семестр (КР)	Курсовая работа (Приложение 3, Приложение 7)	Перечень тем курсовой работы (Приложение 3), Методические указания к выполнению курсовой работы (Приложение 7)	100 баллов

ОПИСАНИЕ ШКАЛ ОЦЕНИВАНИЯ

Показатель оценки освоения ОПОП формируется на основе объединения текущего контроля и промежуточной аттестации обучающегося.

Показатель рейтинга по каждой дисциплине выражается в процентах, который показывает уровень подготовки студента.

Текущий контроль. Используется 100-балльная система оценивания. Оценка работы студента в течение семестра осуществляется преподавателем в соответствии с разработанной им системой оценки учебных достижений в процессе обучения по данной дисциплине.

В рабочих программах дисциплин и практик закреплены виды текущего контроля, планируемые результаты контрольных мероприятий и критерии оценки учебных достижений.

В течение семестра преподавателем проводится не менее 3-х контрольных мероприятий, по оценке деятельности студента. Если посещения занятий по дисциплине включены в рейтинг, то данный показатель составляет не более 20% от максимального количества баллов по дисциплине.

Промежуточная аттестация. Используется 5-балльная система оценивания. Оценка работы студента по окончании дисциплины (части дисциплины) осуществляется преподавателем в соответствии с разработанной им системой оценки достижений студента в процессе обучения по данной дисциплине. Промежуточная аттестация также проводится по окончании формирования компетенций.

Порядок перевода рейтинга, предусмотренных системой оценивания, по дисциплине, в пятибалльную систему.

Высокий уровень – 100% - 70% - отлично, хорошо.

Средний уровень – 69% - 50% - удовлетворительно.

Показатель оценки	По 5-балльной системе	Характеристика показателя
100% - 85%	отлично	обладают теоретическими знаниями в полном объеме, понимают, самостоятельно умеют применять, исследовать, идентифицировать, анализировать, систематизировать, распределять по категориям, рассчитать показатели, классифицировать, разрабатывать модели, алгоритмизировать, управлять, организовать, планировать процессы исследования, осуществлять оценку результатов на высоком уровне
84% - 70%	хорошо	обладают теоретическими знаниями в полном объеме, понимают, самостоятельно умеют применять, исследовать, идентифицировать, анализировать, систематизировать, распределять по категориям, рассчитать показатели, классифицировать, разрабатывать модели, алгоритмизировать, управлять, организовать, планировать процессы исследования, осуществлять оценку результатов. Могут быть допущены недочеты, исправленные студентом самостоятельно в процессе работы (ответа и т.д.)
69% - 50%	удовлетворительно	обладают общими теоретическими знаниями, умеют применять, исследовать, идентифицировать, анализировать, систематизировать, распределять по категориям, рассчитать показатели, классифицировать, разрабатывать модели, алгоритмизировать, управлять, организовать, планировать процессы исследования, осуществлять оценку результатов на среднем уровне. Допускаются ошибки, которые студент затрудняется исправить самостоятельно.
49 % и менее	неудовлетворительно	обладают не полным объемом общих теоретическими знаниями, не умеют самостоятельно применять, исследовать, идентифицировать, анализировать, систематизировать, распределять по категориям, рассчитать показатели, классифицировать, разрабатывать модели, алгоритмизировать, управлять, организовать, планировать процессы исследования, осуществлять оценку результатов. Не сформированы умения и навыки для решения профессиональных задач
100% - 50%	зачтено	характеристика показателя соответствует «отлично», «хорошо», «удовлетворительно»
49 % и менее	не зачтено	характеристика показателя соответствует «неудовлетворительно»

7. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

7.1. Содержание лекций

<p>Тема 1. Интеллектуальный анализ данных (BDA). Основные понятия. ПК-1: Обоснование подходов, используемых в бизнес-анализе ПК-4: Планирование, организация и контроль аналитических работ в ИТ-проекте</p> <p>Определение больших данных. Характеристики и описание больших данных.</p>
<p>Тема 2. Основные алгоритмы интеллектуального анализа данных.ПК-1: Обоснование подходов, используемых в бизнес-анализе ПК-4: Планирование, организация и контроль аналитических работ в ИТ-проекте</p> <p>Методы аналитики данных. Кластеризация методом K-средних.Ассоциативные правила.</p>
<p>Тема 3. Основные инструментальные средства интеллектуального анализа данных.ПК-1: Обоснование подходов, используемых в бизнес-анализе ПК-4: Планирование, организация и контроль аналитических работ в ИТ-проекте</p> <p>Надстройка интеллектуального анализа данных в Excel.</p>
<p>Тема 4. Понятие о принципах управления аналитической деятельностью на основе интеллектуального анализа данных.ПК-1: Обоснование подходов, используемых в бизнес-анализе ПК-4: Планирование, организация и контроль аналитических работ в ИТ-проекте</p> <p>Жизненный цикл аналитики данных. Роли для успешного аналитического проекта.</p>
<p>Тема 5. Комплексная работа по интеллектуальному анализу данных в цифровой экономике.ПК-1: Обоснование подходов, используемых в бизнес-анализе ПК-4: Планирование, организация и контроль аналитических работ в ИТ-проекте</p> <p>Сценарии использования и алгоритмы реализации изученных методов. Диагностика, которая наиболее часто используется для оценки эффективности метода.</p>

7.2 Содержание практических занятий и лабораторных работ

<p>Тема 1. Интеллектуальный анализ данных (BDA). Основные понятия. ПК-1: Обоснование подходов, используемых в бизнес-анализе ПК-4: Планирование, организация и контроль аналитических работ в ИТ-проекте</p> <p>Воздействие неструктурированных данных на аналитику больших данных. Возможности для аналитиков в хранилищах данных.</p>
<p>Тема 2. Основные алгоритмы интеллектуального анализа данных.ПК-1: Обоснование подходов, используемых в бизнес-анализе ПК-4: Планирование, организация и контроль аналитических работ в ИТ-проекте</p> <p>Линейная регрессия.Логистическая регрессия. Статистические методы в аналитике. Наивный байесовский классификатор.Дерево решений.</p>
<p>Тема 3. Основные инструментальные средства интеллектуального анализа данных.ПК-1: Обоснование подходов, используемых в бизнес-анализе ПК-4: Планирование, организация и контроль аналитических работ в ИТ-проекте</p> <p>Язык R. Основные понятия и правила.</p>
<p>Тема 4. Понятие о принципах управления аналитической деятельностью на основе интеллектуального анализа данных.ПК-1: Обоснование подходов, используемых в бизнес-анализе ПК-4: Планирование, организация и контроль аналитических работ в ИТ-проекте</p> <p>Учебный пример для применения жизненного цикла аналитики данных.</p>

Тема 5. Комплексная работа по интеллектуальному анализу данных в цифровой экономике.ПК-1: Обоснование подходов, используемых в бизнес-анализе ПК-4: Планирование, организация и контроль аналитических работ в ИТ-проекте

Алгоритм выбора метода. Анализ достоинств и недостатков изученных методов интеллектуального анализа данных.

7.3. Содержание самостоятельной работы

Тема 1. Интеллектуальный анализ данных (BDA). Основные понятия. ПК-1: Обоснование подходов, используемых в бизнес-анализе ПК-4: Планирование, организация и контроль аналитических работ в ИТ-проекте

Примеры эффективного использования больших данных.

Тема 2. Основные алгоритмы интеллектуального анализа данных.ПК-1: Обоснование подходов, используемых в бизнес-анализе ПК-4: Планирование, организация и контроль аналитических работ в ИТ-проекте

Анализ временных рядов. Анализ текста.

Тема 3. Основные инструментальные средства интеллектуального анализа данных.ПК-1: Обоснование подходов, используемых в бизнес-анализе ПК-4: Планирование, организация и контроль аналитических работ в ИТ-проекте

Технология представления данных QlikView. Основы технологии SAP.

Тема 4. Понятие о принципах управления аналитической деятельностью на основе интеллектуального анализа данных.ПК-1: Обоснование подходов, используемых в бизнес-анализе ПК-4: Планирование, организация и контроль аналитических работ в ИТ-проекте

Формулировка бизнес-проблемы в качестве аналитической проблемы.

Тема 5. Комплексная работа по интеллектуальному анализу данных в цифровой экономике.ПК-1: Обоснование подходов, используемых в бизнес-анализе ПК-4: Планирование, организация и контроль аналитических работ в ИТ-проекте

Применение технологий интеллектуального анализа данных в аналитической деятельности организации.

7.3.1. Примерные вопросы для самостоятельной подготовки к зачету/экзамену
Приложение 1

7.3.2. Практические задания по дисциплине для самостоятельной подготовки к зачету/экзамену
Приложение 2

7.3.3. Перечень курсовых работ
Приложение 3

7.4. Электронное портфолио обучающегося
Курсовые работы размещаются на <https://portfolio.usue.ru/>

7.5. Методические рекомендации по выполнению контрольной работы
Материалы не предусмотрены

7.6 Методические рекомендации по выполнению курсовой работы
Приложение 7

8. ОСОБЕННОСТИ ОРГАНИЗАЦИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ ДЛЯ ЛИЦ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ

По заявлению студента

В целях доступности освоения программы для лиц с ограниченными возможностями здоровья при необходимости кафедра обеспечивает следующие условия:

- особый порядок освоения дисциплины, с учетом состояния их здоровья;
- электронные образовательные ресурсы по дисциплине в формах, адаптированных к ограничениям их здоровья;
- изучение дисциплины по индивидуальному учебному плану (вне зависимости от формы обучения);
- электронное обучение и дистанционные образовательные технологии, которые предусматривают возможности приема-передачи информации в доступных для них формах.
- доступ (удаленный доступ), к современным профессиональным базам данных и информационным справочным системам, состав которых определен РПД.

9. ПЕРЕЧЕНЬ ОСНОВНОЙ И ДОПОЛНИТЕЛЬНОЙ УЧЕБНОЙ ЛИТЕРАТУРЫ, НЕОБХОДИМОЙ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Сайт библиотеки УрГЭУ

<http://lib.usue.ru/>

Основная литература:

1. Богданов Е.П. Интеллектуальный анализ данных [Электронный ресурс]: учебное пособие. - Волгоград: ФГБОУ ВПО Волгоградский государственный аграрный университет, 2019. - 112 – Режим доступа: <https://znanium.ru/catalog/product/1087885>

2. Назаров Д.М., Рыжкина Д.А. Интеллектуальные средства бизнес-аналитики [Электронный ресурс]: учебник. - , 2022. - 241 с – Режим доступа: <https://book.ru/book/941734>

3. Кулаичев А.П. Методы и средства комплексного статистического анализа данных [Электронный ресурс]: учебное пособие. - Москва: ООО "Научно-издательский центр ИНФРА-М", 2022. - 484 – Режим доступа: <https://znanium.ru/catalog/product/1815604>

4. Назаров, Д. М., Бегичева, С. В., Ковтун, Д. Б., Назаров, А. Д. Data Science и интеллектуальный анализ данных [Электронный ресурс]: учебное пособие. - Москва: Ай Пи Ар Медиа, 2023. - 304 – Режим доступа: <https://www.iprbookshop.ru/127201.html>

Дополнительная литература:

1. Затонский А. В. Информационные технологии: разработка информационных моделей и систем: учебное пособие для студентов вузов, обучающихся по направлению 23.01.00 "Информатика и вычислительная техника". - Москва: РИОР: ИНФРА-М, 2014. - 344 с.

2. Гурвиц Дж. С., Ньюджент А. Ф., Халпер Ф., Кауфман М. А. Просто о больших данных: перевод с английского. - Москва: Сбербанк: [Эксмо], 2015. - 395 с.

10. ПЕРЕЧЕНЬ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, ВКЛЮЧАЯ ПЕРЕЧЕНЬ ЛИЦЕНЗИОННОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ И ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ, ОНЛАЙН КУРСОВ, ИСПОЛЬЗУЕМЫХ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ

Перечень лицензионного программного обеспечения:

Microsoft Office 2016. Договор № 52/223-ПО/2020 от 13.04.2020, Акт № Tr000523459 от 14.10.2020 Срок действия лицензии - Без ограничения срока.

МойОфис стандартный. Соглашение № СК-281 от 7 июня 2017. Дата заключения - 07.06.2017. Срок действия лицензии - без ограничения срока.

R Studio (среда для языка программирования R). Лицензия GNU Affero General Public License v3. Срок действия лицензии - без ограничения срока.

Язык программирования R. Лицензия GNU GPL 2. Срок действия лицензии - без ограничения срока.

Перечень информационных справочных систем, ресурсов информационно-телекоммуникационной сети «Интернет»:

ИАД

<https://www.intuit.ru/studies/courses/2312/612/info>

11. ОПИСАНИЕ МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЙ БАЗЫ, НЕОБХОДИМОЙ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ

Реализация учебной дисциплины осуществляется с использованием материально-технической базы УрГЭУ, обеспечивающей проведение всех видов учебных занятий и научно-исследовательской и самостоятельной работы обучающихся:

Специальные помещения представляют собой учебные аудитории для проведения всех видов занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации.

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

Все помещения укомплектованы специализированной мебелью и оснащены мультимедийным оборудованием спецоборудованием (информационно-телекоммуникационным, иным компьютерным), доступом к информационно-поисковым, справочно-правовым системам, электронным библиотечным системам, базам данных действующего законодательства, иным информационным ресурсам служащими для представления учебной информации большой аудитории.

Для проведения занятий лекционного типа презентации и другие учебно-наглядные пособия, обеспечивающие тематические иллюстрации.

7.3.1. Примерные вопросы для самостоятельной подготовки к зачету/экзамену К зачету

1. Назовите три характеристики больших данных
2. Что такое аналитическая песочница?
3. Поясните, в чем различие между BI и DS
4. Охарактеризовать задачи текущей аналитической архитектуры для аналитика по обработке данных.
5. Каковы основные наборы навыков и поведенческих характеристик для аналитика по обработке данных?
6. В какой фазе вы планируете инвестировать большую часть вашего времени проекта и почему? Где было бы ожидаемо потратить минимум времени?
7. Каковы преимущества ведения экспериментальной программы до полномасштабного внедрения новой аналитической методологии? Обсудите это в контексте мини-кейса.
8. Какие инструменты будут использоваться в следующих фазах, и для каких видов сценариев?
9. Какие фазы жизненного цикла вам нужны, чтобы ориентироваться на выявлении способов сделать его?
10. Вернемся к нашему примеру дисперсионного анализа (ANOVA), приведенному ранее. Что Вы думаете по этому поводу? Есть ли разница практически в различии первого и второго предложения? Следует ли нам прибегать к одному из них? Если да, и затраты бы составляли US \$25 для каждого offer1 и US \$10 для offer2, вы все равно бы приняли такое решение?
11. В примере производства дисководов на фабрике, предположим, что вы бы проверили ее на наличие проблем в процессе производства. Как вы бы обосновали свое решение с финансовой точки зрения.
12. Почему кластеризацию методом K-средних считают неконтролируемой машиной, изучающей алгоритм?
13. Как использовать парные диаграммы, чтобы оценить эффективность кластеризации?
14. Опишите 4 шага алгоритма кластеризации методом K-средних.
15. Как использовать WSS, чтобы определить значение K?
16. Какова наиболее распространенная мера расстояния, используемого алгоритмами кластеризации методом K-средних?
17. Что такое «свойство Apriori» и какое значение оно имеет в алгоритме Apriori?
18. Назовите три распространенных прецедента для алгоритмов, основанных на ассоциативных правилах.
19. В чём разница между Lift и Leverage? Как с помощью соотношения Lift можно оценить качество обнаруженных правил?
20. Дайте определение понятиям «поддержка» и «достоверность».
21. Как определить эффективность сгенерированных правил с помощью контрольных данных?
22. Что такое логит, как с его использованием вычисляются вероятности класса?
23. Как используется кривая ROC для диагностики эффективности логистической регрессионной модели?
24. Что такое псевдо R2 и что он измеряет в логистической регрессионной модели?
25. Как вы опишите задачу из бинарного класса?
26. Сравните и сопоставьте линейные и логистические методы регрессии.
27. Что такое информационный выигрыш?
28. При каких условиях энтропия имеет максимальное и минимальное значения?
29. Назовите три сценария использования деревьев решений.

30. Что такое слабые модели, и как они применяются в групповых методах?
31. Когда результаты вычисления зависят от многих переменных, мы получаем переобученную модель с глубокими деревьями и большим набором данных. Почему?
32. Какой классификатор вы бы порекомендовали в следующих случаях:
Если вы работаете с многомерными данными
Если нелинейные переменные или разрывы переменных влияют на результат вычисления.

К экзамену

1. Основные понятия искусственного интеллекта
2. Интеллектуальные информационные системы и их эволюция. Общие понятия и определения.
3. Направления развития искусственного интеллекта Основные виды интеллектуальных информационных систем. Экономические экспертные информационные системы.
4. Свойства и состав интеллектуальной системы Обобщенная структура построения интеллектуальных информационных систем.
5. Информационная модель деятельности специалиста и место систем искусственного интеллекта в этой деятельности.
6. Основные требования к эффективности и надежности интеллектуальных информационных систем.
7. Современные технологии проектирования интеллектуальных информационных систем.
8. Свойства и состав интеллектуальной системы
9. Жизненный цикл систем искусственного интеллекта Основные этапы проектирования интеллектуальных информационных систем.
10. Понятие знания, их классификация. Формы и методы представления знаний.
11. Модели представления знаний. Логическая модель представления знаний.
12. Модели представления знаний. Фреймовая модель
13. Модели представления знаний. Продукционная модель
14. Системы управления базами знаний. Общие принципы организации. Методы интеграции баз данных и знаний.
15. Методы организации рассуждений в ИИС. Рассуждения на основе дедукции, индукции, аналогии.
16. Понятие экспертной системы. Особенности ЭС. Основные типы задач
17. Принципы классификации ЭС
18. Структура экспертной системы.
19. Методы организации баз знаний интеллектуальных информационных систем.
20. Разработка экспертных систем.
21. Инструментальные средства разработки экспертных систем
22. Перспективы развития интеллектуальных информационных систем.
23. Методы приобретения знаний, используемые в ИИС.
24. Методы извлечения знаний из данных.
25. Понятие нейронной сети. Архитектура нейронных систем. Типы решаемых задач
26. Этапы построения нейронной сети
27. Алгоритмы обучения нейронных сетей.
28. Современные ИИС поддержки принятия управленческих решений.
29. Критерии отбора набора стратегий ИИС поддержки принятия решений.
30. Понятие интеллектуального анализа данных. Функции систем бизнес-аналитики
31. Основные понятия интеллектуального анализа
32. Классификация методов Data Mining
33. Основы анализа данных

7.3.2. Практические задания по дисциплине для самостоятельной подготовки к зачету/экзамену

ЗАДАНИЯ ПО ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЕ

09.04.03 Прикладная информатика

Дисциплина: Интеллектуальный анализ данных

Компетенции ПК-1, ПК-4

ПК-1 Обоснование подходов, используемых в бизнес-анализе

ПК-4 Планирование, организация и контроль аналитических работ в ИТ-проекте

Задания закрытого типа

1. Что такое машинное обучение?

- a) процесс передачи знаний от человека к компьютеру
- b) процесс автоматического извлечения информации из данных
- c) процесс создания искусственного интеллекта
- d) процесс передачи бизнес-процессов

Ответ: б) процесс автоматического извлечения информации из данных

2. Какой тип машинного обучения находит зависимость между входными и выходными данными и использует их для предсказания новых данных?

- a) а) обучение с учителем
- b) б) обучение без учителя
- c) в) обучение с подкреплением
- d) г) регрессия

Ответ: а) обучение с учителем

3. Что такое переобучение в машинном обучении?

- a) процесс улучшения качества модели
- b) процесс ухудшения качества модели
- c) процесс обучения на слишком большое количество данных
- d) процесс создания модели, которая хорошо работает только на обучающей выборке, но плохо на новых данных

Ответ: d) процесс создания модели, которая хорошо работает только на обучающей выборке, но плохо на новых данных

4. Алгоритм машинного обучения находит структуру в структурах с использованием нескольких слоев нейронов?

- a) SVM

- b) k-ближайшие соседи
- c) решающие деревья
- d) нейронные сети

Ответ: г) нейронные сети

5. Какой тип машинного обучения не требует выходных данных и находит структуру в данных?

- a) а) обучение с учителем
- b) б) обучение без учителя
- c) в) обучение с подкреплением
- d) г) регрессия

Ответ: б) обучение без учителя

6. Какие из обнаруженных методов относятся к методам обучения с учителем?

- a) k-ближайшие соседи
- b) линейная регрессия
- c) k-средние
- d) PCA

Ответ: а) k-ближайших соседей и б) линейная регрессия

7. Что такое кластеризация в машинном обучении?

- a) а) процесс поиска в данных
- b) б) процесс поиска оптимальных параметров
- c) в) процесс разбиения данных на группы по схожести
- d) г) процесс выбора моделей признаков, на

8. Что такое машинное обучение?

- a) Процесс, в ходе которого компьютерные системы улучшают свою производительность
- b) Обучение машинам, как выполнять задачи, без явного программирования
- b) Разработка аппаратного обеспечения для компьютеров
- c) Искусственный интеллект, созданный на основе естественных алгоритмов

Ответ: б

9. Какого типа машинного обучения?

- a) а) Обучение с учителем, безучителем, обучение с подкреплением
- b) б) Обучение с учителем, безучителем, обучение с биоретроспекцией
- c) в) Обучение с учителем, обучение с курсом, обучение с подкреплением
- d) г) Обучение с учителем, безучителем, обучение с обратной связью

Ответ: а

10. Что такое обучение с учителем?

- a) Обучение без экспериментов
- b) Обучение с помощью подкрепления
- c) Обучение на основе завышенных правильных ответов
- d) Обучение на основе градиентного опыта

Ответ: c

11. Что такое обучение без учителя?

- a) Обучение на основе использования правильных ответов
- b) Обучение на основе подкрепления
- c) Обучение без использования примеров с исходными ответами
- d) Обучение на основе градиентного запуска

Ответ: c

12. Что такое кластеризация?

- a) Алгоритм, который используется для обучения без учителя
- b) Алгоритм, который используется для обучения с учителем
- c) Алгоритм, который используется для обучения с подкреплением
- d) Алгоритм, который используется для обучения с биоретроспекцией

Ответ: a

13. Что такое "данные" в контексте ИАД?

- a) Информация, которая может быть обработана компьютером
- b) Случайные цифры и символы
- c) Информация, которая не может быть обработана компьютером
- d) Случайные слова и предложения

Ответ: a

14. Какие методы машинного обучения используются в ИАД?

- a) Кластерный анализ
- b) Линейная алгебра
- c) Теория вероятности
- d) Все вышеперечисленные

Ответ: a, c (в зависимости от того, какая точка зрения на машинное обучение)

15. Какую роль играют алгоритмы в ИАД?

- a) Они позволяют собрать данные
- b) Они позволяют обработать данные

c) Они позволяют принимать решения на основе данных

d) Они не играют никакой роли

Ответ: b, c

Задания открытого типа

1. Что такое выбросы (outliers)? Приведите пример.
2. Что такое корреляция? Приведите пример.
3. Что такое перекрестная проверка (cross-validation)? Приведите пример.
4. Что такое overfitting? Приведите пример.
5. Что такое метрики оценки качества моделей? Приведите пример.
6. Что такое ассоциативные правила? Приведите пример.
7. Что такое решающие деревья? Приведите пример.
8. Что такое метод опорных векторов (SVM)? Приведите пример.
9. Что такое байесовские сети? Приведите пример.
10. Что такое нейронные сети? Приведите пример.
11. Что такое рекуррентные нейронные сети (RNN)? Приведите пример.
12. Что такое сверточные нейронные сети (CNN)? Приведите пример.
13. Что такое глубокое обучение (deep learning)? Приведите пример.
14. Какие методы можно использовать для поиска выбросов в данных? Приведите пример.
15. Какие методы можно использовать для определения наилучшего числа кластеров? Приведите пример.
16. Какие методы можно использовать для извлечения признаков из данных? Приведите пример.
17. Что такое data mining? Приведите пример.
18. Какие типы данных можно использовать в data mining? Приведите пример.
19. Что такое обучение с учителем (supervised learning)? Приведите пример.
20. Что такое обучение без учителя (unsupervised learning)? Приведите пример.
21. Что такое задача кластеризации (clustering)? Приведите пример.
22. Что такое задача классификации (classification)? Приведите пример.
23. Что такое задача регрессии (regression)? Приведите пример.
24. Какие методы можно использовать для отбора признаков (feature selection)? Приведите пример.
25. Что такое метод главных компонент (PCA)? Приведите пример.
26. Что такое кривая обучения (learning curve)? Приведите пример.

27. Что такое метрики качества моделей? Приведите пример.
28. Что такое кросс-валидация (cross-validation)? Приведите пример.
29. Что такое случайный лес (random forest)? Приведите пример.
30. Что такое градиентный бустинг (gradient boosting)? Приведите пример.

7.3.3. Перечень курсовых работ

Список примерных тем для выполнения курсовых работ по дисциплине Интеллектуальный анализ данных.

1. Основные этапы разработки интеллектуальных систем бизнес-анализа;
2. Подготовка исходных данных к машинному обучению;
3. Создание Хранилища данных и организация доступа;
4. Оценка качества данных. Профайлинг и аудит данных;
5. Проблема пропущенных данных и ее возможные решения;
6. Редактирование выбросов и аномальных значений;
7. Поиск и редактирование дубликатов и противоречий;
8. Спектральная обработка ряда;
9. Преобразование Дата/время при подготовке временного ряда;
10. Группировка и разгруппировка данных. Задачи, решаемые группировкой;
11. Квантование данных. Назначение квантования. Параметры квантования, определение границ интервалов и задание меток;
12. Создание новых полей с помощью калькулятора;
13. Слияние наборов данных. Типы объединения;
14. Изменение структуры таблицы. Кросс-таблица и свертка столбцов;
15. Создание репрезентативного множества. Сэмплинг и Разбиение на множества. Методы выборки данных. Размеры обучающего и тестового множества;
16. Подготовка временного ряда к прогнозированию. Скользящее окно. Критерии выбора погружения в ряд;
17. Отбор переменных в модель логистической регрессии. Конечные классы. WoE-анализ. Метрики классов и значимость входных признаков;
18. Применение скриптов в сценариях обработки данных;
19. Групповая обработка данных по похожим сценариям;
20. Применение переменных в сценариях обработки данных;
21. Оценка степени периодичности ряда с помощью метода автокорреляции;
22. Выявление закономерностей между связанными событиями. Ассоциативные правила. Шаблоны покупок. Достоверность и Лифт AP;
23. Задачи, решаемые с помощью кластеризации. Кластеризация k-means, g-means. Область применения;
24. EM-кластеризация. Автоматическое определение кластеров. Оценка на основе логарифмической функции правдоподобия;
25. Кластеризация транзакций. Эвристический алгоритм CLOPE. Глобальная функция стоимости. Область применения кластеризации транзакций;
26. Самоорганизующиеся карты (Self Organizing Maps - SOM). Кластеризации многомерных векторов – алгоритм проецирования с сохранением топологического подобия. Инициализация начальных весов. Область применения карт Кохонена.
27. Линейная регрессия. Коэффициенты регрессии. Область применения;
28. Логистическая регрессия. Прогнозирование бинарной переменной. Коэффициенты регрессии. ROC-анализ. Lift-анализ. Оценка качества модели;
29. Дерево решений. Нормализация полей. Обучающая выборка. Параметры обучения. Значимость атрибутов. Правила. Что-если. Таблица сопряженности;
30. Нейронная сеть. Требования к обучающей и тестовой выборке. Задание структуры нейронной сети. Скрытые слои и активационная функция.

Определение числа связей и переобучение сети. Область применения нейронных сетей.

7.4. Электронное портфолио обучающегося

В электронном портфолио обучающегося по дисциплине размещается

<http://portfolio.usue.ru>

- контрольные работы

- курсовая работа

Приложение 6

Методические рекомендации по выполнению Курсовых работ для магистрантов очно-заочной формы обучения по дисциплине «Интеллектуальный анализ данных».

Цель выполнения Курсовых работ:

- систематизация теоретических знаний и практических навыков, полученных при изучении данного курса;
- применение полученных теоретических знаний и практических навыков при решении прикладных задач;
- стимулирование навыков самостоятельной аналитической работы
- овладение современными методами научного исследования;
- выработка навыков презентации публичной дискуссии и защиты научных идей, предложений и рекомендаций.

Выполненная курсовая работа должна свидетельствовать о способности автора к применению полученных во время учебы теоретических знаний и практических навыков по дисциплине «Интеллектуальный анализ данных». В работе необходимо продемонстрировать способность создавать интеллектуальные системы бизнес-анализа применительно к поставленной задаче.

Согласно предложенной теме студент должен определить инструменты Data Mining необходимые для решения поставленной задачи.

Результаты исследования оформить в виде отчета со следующей структурой:

1. *Введение*. Приводится постановка задачи с указанием цели исследования, исходные данные с указанием источника.
2. *Подготовка данных к анализу*. Приводится порядок работ по очистке и трансформации данных для решения проблем с обоснованием каждого этапа. Создается хранилище данных применительно к существующим бизнес-процессам.
3. *Data Mining*. Приводится порядок работ по применению инструментов Data Mining для решения поставленной задачи. Как правило, в реальных системах применяются несколько инструментов позволяющих получить наилучший результат. В ряде случаев разрабатываются несколько моделей с последующей оценкой их эффективности и достоверности.
4. *Выводы и резюме*. В сжатой форме приводятся основные этапы выполненной работы, делается заключение о работоспособности системы и возможности ее практического применения.
5. Список использованных источников.

Так же по результатам выполнения контрольной работы необходимо:

- 1) Подготовить доклад (длительность 4 минуты)
- 2) Подготовить презентацию к докладу;

Требования к оформлению презентации:

название, авторство, год на первом слайде

на следующих слайдах: цель работы, задачи; на последнем слайде: заключение
оглавление-план

слайды пронумеровать

задать колонтитул (ФИО, год)

Контрольная работа сдает на портал.

При проверке оценивается:

- постановка задачи;
- обоснованность выбора методологии и методов интеллектуального анализа данных применительно к исследуемому бизнес-процессу;
- эффективность применения методов Data Mining;
- понимание работы интеллектуальных методов бизнес-анализа;
- степень самостоятельности выполнения контрольной работы (плагиат).