

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Макаренко Елена Николаевна

Должность: Ректор

Дата подписания: 28.03.2023 15:02:51

Уникальный программный ключ:

c098bc0c1041cb2a4cf926cf171d6715d99a6ae00adc8e27b55cbe1e2dbd7c78

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение высшего образования «Ростовский государственный экономический университет (РИНХ)»

УТВЕРЖДАЮ

Директор Института магистратуры

Иванова Е.А.

« 30 » 08 2021 г.

Рабочая программа дисциплины
Модели и алгоритмы интеллектуального анализа данных

Направление 01.04.02 Прикладная математика и информатика
магистерская программа 01.04.02.03 "Искусственный интеллект в цифровой экономике"

Для набора 2021 года

Квалификация
Магистр

КАФЕДРА **Фундаментальная и прикладная математика**

Распределение часов дисциплины по семестрам

Семестр («Курс»-«Семестр по курсу») Неделя	3 (2.1)		Итого	
	уч	лп	уч	лп
Вид занятий	уч	лп	уч	лп
Лекции	10	10	10	10
Лабораторные	10	10	10	10
Итого ауд.	20	20	20	20
Контактная работа	20	20	20	20
Сам. работа	84	84	84	84
Часы на контроль	4	4	4	4
Итого	108	108	108	108

ОСНОВАНИЕ

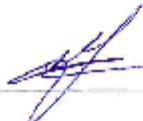
Учебный план утвержден учёным советом вуза от 30.08.2021 протокол № 1.

Программу составил(и): д.ф.-м.н., проф. каф. ФизМ, Сидорова Л.В.

Зав. кафедрой: д.ф.-м.н., доц. Стрюков М.Б.

Методическим советом направления: д.ф.-м.н., доц. Стрюков М.Б.





1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

1.1	Цель изучения дисциплины: демонстрация полного цикла построения и анализа моделей интеллектуального анализа данных, от сбора данных до выбора оптимального решения, а также оценки его качества.
1.2	Задачи изучения дисциплины: научить студентов пользоваться современными аналитическими инструментами и адаптировать их под особенности конкретных задач; освоить основные темы, необходимые в работе с большим массивом данных, поиск структуры в данных, проведение экспериментов, построение выводов, необходимых для решения конкретных прикладных задач.

2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

УК-1: Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий

ОПК-1: Способен решать актуальные задачи фундаментальной и прикладной математики

ОПК-2: Способен совершенствовать и реализовывать новые математические методы решения прикладных задач

ПК-2: Способен разрабатывать и анализировать концептуальные и теоретические модели решаемых научных проблем и задач

В результате освоения дисциплины обучающийся должен:

Знать:
<ul style="list-style-type: none"> - основные методы интеллектуального анализа данных; математические основы методов линейной регрессии, частичных наименьших квадратов, поддерживающих векторов, искусственных нейронных сетей, методов Байеса (соотнесено с индикатором УК-1.1) - новые информационные технологии в гуманитарных областях знаний с использованием средств интеллектуального анализа данных и машинного обучения, компьютерной лингвистики и представления знаний (соотнесено с индикатором ОПК-1.1) - основные подходы к анализу данных и способы их модификации (соотнесено с индикатором ОПК-2.1) - методы интеллектуального анализа данных, необходимых для построения моделей решаемых научных задач (соотнесено с индикатором ПК-2.1)
Уметь:
<ul style="list-style-type: none"> - строить различные регрессионные и классификационные модели с использованием методов интеллектуального анализа данных; подбирать оптимальные значения подгоночных параметров модели (соотнесено с индикатором УК-1.2) - использовать новые информационные технологии в гуманитарных областях знаний с использованием средств интеллектуального анализа данных и машинного обучения, компьютерной лингвистики и представления знаний (соотнесено с индикатором ОПК-1.2) - настраивать алгоритмы под особенности прикладных задач (соотнесено с индикатором ОПК-2.2) - применять методы интеллектуального анализа данных, необходимых для построения моделей решаемых научных задач (соотнесено с индикатором ПК-2.2)
Владеть:
<ul style="list-style-type: none"> - методами интерпретировать необходимые данные для формирования суждений по соответствующим социальным, научным и этическим проблемам (соотнесено с индикатором УК-1.3) - Методами применения новых информационных технологий в гуманитарных областях знаний с использованием средств интеллектуального анализа данных и машинного обучения, компьютерной лингвистики и представления знаний (соотнесено с индикатором ОПК-1.3) - программными инструментами, позволяющими гибкую подстройку алгоритма (соотнесено с индикатором ОПК-2.3) - навыками применения методов интеллектуального анализа данных, необходимых для построения моделей решаемых научных задач (соотнесено с индикатором ПК-2.3)

3. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Компетенции	Литература
	Раздел 1. Интеллектуальный анализ данных				
1.1	Тема 1. Интеллектуальный анализ данных (Data Mining. Сопоставление и сравнение понятий "информация", "данные", "знание". Свойства знаний. Сравнение статистик, машинного обучения и Data Mining. Классификация задач Data Mining. Сферы применения Data Mining. /Лек/	3	2	УК-1 ОПК-1 ОПК-2 ПК-2	Л1.1 Л1.2 Л1.3 Л2.1 Л2.2

1.2	Тема 1. Интеллектуальный анализ данных (Data Mining). Сопоставление и сравнение понятий "информация", "данные", "знание". Свойства знаний. Сравнение статистик, машинного обучения и Data Mining. Классификация задач Data Mining. Сферы применения Data Mining. Применение Loginom /Лаб/	3	2	УК-1 ОПК-1 ОПК-2 ПК-2	Л1.1 Л1.2 Л1.3Л2.1 Л2.2
1.3	Тема 2. Интеллектуальный анализ данных (Data Mining). Предсказательные и описательные модели Data Mining. Процесс Data Mining: анализ предметной области, постановка задачи, подготовка данных, построение модели, проверка и оценка моделей, выбор модели, применение модели. /Лек/	3	2	УК-1 ОПК-1 ОПК-2 ПК-2	Л1.1 Л1.2 Л1.3Л2.1 Л2.2
1.4	Тема 2. Интеллектуальный анализ данных (Data Mining). Предсказательные и описательные модели Data Mining. Процесс Data Mining: анализ предметной области, постановка задачи, подготовка данных, построение модели, проверка и оценка моделей, выбор модели, применение модели. Применение Loginom /Лаб/	3	2	УК-1 ОПК-1 ОПК-2 ПК-2	Л1.1 Л1.2 Л1.3Л2.1 Л2.2
1.5	Тема 1. Интеллектуальный анализ данных (Data Mining). Предсказательные и описательные модели Data Mining. Процесс Data Mining: анализ предметной области, постановка задачи, подготовка данных, построение модели, проверка и оценка моделей, выбор модели, применение модели. /Ср/	3	14	УК-1 ОПК-1 ОПК-2 ПК-2	Л1.1 Л1.2 Л1.3Л2.1 Л2.2
1.6	Тема 3. "Задачи и методы интеллектуального анализа данных". Классификация. Регрессия. Методы представления результатов анализа. Прогнозирование и визуализация. Методы прогнозирования и классификации: алгоритм построения 1R-алгоритм; метод Naïve Bayes; деревья решений, метод опорных векторов, метод «ближайшего соседа», нейронные сети. Поиск ассоциативных правил. Характеристики ассоциативных правил: поддержка, достоверность и улучшение. Алгоритм Apriori. Методы кластерного анализа. Меры близости в алгоритмах кластеризации. /Лек/	3	2	УК-1 ОПК-1 ОПК-2 ПК-2	Л1.1 Л1.2 Л1.3Л2.1 Л2.2

1.7	Тема 3. "Задачи и методы интеллектуального анализа данных".Классификация. Регрессия. Методы представления результатов анализа. Прогнозирование и визуализация. Методы прогнозирования и классификации: алгоритм построения 1R-алгоритм; метод Naïve Bayes; деревья решений, метод опорных векторов, метод «ближайшего соседа», нейронные сети. Поиск ассоциативных правил. Характеристики ассоциативных правил: поддержка, достоверность и улучшение. Алгоритм Apriori. Методы кластерного анализа. Меры близости в алгоритмах кластеризации. /Лаб/	3	2	УК-1 ОПК-1 ОПК-2 ПК-2	Л1.1 Л1.2 Л1.3Л2.1 Л2.2
1.8	Тема 2. "Задачи и методы интеллектуального анализа данных".Классификация. Регрессия. Методы представления результатов анализа. Прогнозирование и визуализация. Методы прогнозирования и классификации: алгоритм построения 1R-алгоритм; метод Naïve Bayes; деревья решений, метод опорных векторов, метод «ближайшего соседа», нейронные сети. Поиск ассоциативных правил. Характеристики ассоциативных правил: поддержка, достоверность и улучшение. Алгоритм Apriori. Методы кластерного анализа. Меры близости в алгоритмах кластеризации. /Ср/	3	22	УК-1 ОПК-1 ОПК-2 ПК-2	Л1.1 Л1.2 Л1.3Л2.1 Л2.2
	Раздел 2. "Инструменты Data Mining"				
2.1	Тема 4. "Инструменты Data Mining". Стандарт CWM: назначение, структура и состав. Стандарт CRISP: структура, фазы и задачи. Стандарт PMML. Стандарт OLE DB для Data Mining. Поставщики Data Mining. Классификация инструментов. /Лек/	3	2	УК-1 ОПК-1 ОПК-2 ПК-2	Л1.1 Л1.2 Л1.3Л2.1 Л2.2
2.2	Тема 5. "Программное обеспечение" Программное обеспечение Data Mining для поиска ассоциативных правил. Программное обеспечение для решения задач кластеризации и сегментации. Программное обеспечение для решения задач классификации. Программное обеспечение Data Mining для решения задач оценивания и прогнозирования. /Лек/	3	2	УК-1 ОПК-1 ОПК-2 ПК-2	Л1.1 Л1.2 Л1.3Л2.1 Л2.2
2.3	Тема 4. "Инструменты Data Mining. Loginom". Программное обеспечение Data Mining для поиска ассоциативных правил. Программное обеспечение для решения задач кластеризации и сегментации. Программное обеспечение для решения задач классификации. Программное обеспечение Data Mining для решения задач оценивания и прогнозирования. Применение Loginom /Лаб/	3	2	УК-1 ОПК-1 ОПК-2 ПК-2	Л1.1 Л1.2 Л1.3Л2.1 Л2.2

2.4	Тема 5. "Инструменты Data Mining. Phython". Программное обеспечение Data Mining для поиска ассоциативных правил. Программное обеспечение для решения задач кластеризации и сегментации. Программное обеспечение для решения задач классификации. Программное обеспечение Data Mining для решения задач оценивания и прогнозирования. Применение Phython /Лаб/	3	2	УК-1 ОПК-1 ОПК-2 ПК-2	Л1.1 Л1.2 Л1.3Л2.1 Л2.2
2.5	Тема 3. "Инструменты Data Mining". Стандарт CWM: назначение, структура и состав. Стандарт CRISP: структура, фазы и задачи. Стандарт PMML. Стандарт OLE DB для Data Mining. Поставщики Data Mining. Классификация инструментов. Программное обеспечение Data Mining для поиска ассоциативных правил. Программное обеспечение для решения задач кластеризации и сегментации. Программное обеспечение для решения задач классификации. Программное обеспечение Data Mining для решения задач оценивания и прогнозирования. /Ср/	3	48	УК-1 ОПК-1 ОПК-2 ПК-2	Л1.1 Л1.2 Л1.3Л2.1 Л2.2
2.6	/Зачёт/	3	4	УК-1 ОПК-1 ОПК-2 ПК-2	Л1.1 Л1.2 Л1.3Л2.1 Л2.2

4. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

Структура и содержание фонда оценочных средств для проведения текущей и промежуточной аттестации представлены в Приложении 1 к рабочей программе дисциплины.

5. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

5.1. Основная литература

	Авторы, составители	Заглавие	Издательство, год	Колич-во
Л1.1	Пальмов, С. В.	Интеллектуальный анализ данных: учебное пособие	Самара: Поволжский государственный университет телекоммуникаций и информатики, 2017	http://www.iprbookshop.ru/75376.html неограниченный доступ для зарегистрированных пользователей
Л1.2	Пол Дейтел, Харви Дейтел	Python: Искусственный интеллект, большие данные и облачные вычисления	Санкт-Петербург: Питер, 2021	https://ibooks.ru/reading.php?short=1&productid=371701 неограниченный доступ для зарегистрированных пользователей
Л1.3	Замятин, А. В.	Интеллектуальный анализ данных: учебное пособие	Томск: Издательский Дом Томского государственного университета, 2020	https://www.iprbookshop.ru/116889.html неограниченный доступ для зарегистрированных пользователей

5.2. Дополнительная литература

	Авторы, составители	Заглавие	Издательство, год	Колич-во
Л2.1	Алексеев, В. В., Ивановский, М. А., Елисеев, А. И., Громов, Ю. Ю., Губсков, Ю. А.	Интеллектуальные информационные системы и технологии их построения: учебное пособие	Тамбов: Тамбовский государственный технический университет, ЭБС АСВ, 2021	https://www.iprbookshop.ru/123026.html неограниченный доступ для зарегистрированных пользователей

	Авторы, составители	Заглавие	Издательство, год	Колич-во
Л2.2		Студент. Аспирант. Исследователь: всероссийский научный журнал: журнал	Владивосток: Эксперт-Наука, 2021	https://biblioclub.ru/index.php?page=book&id=685682 неограниченный доступ для зарегистрированных пользователей

5.3 Профессиональные базы данных и информационные справочные системы

Консультант +

Официальная документация библиотеки pybrain. - <http://pybrain.org/docs/index.html>

Образовательная платформа по Python - <http://pybrain.org/docs/index.html>

сайт федеральной государственной службы статистики - <https://rosstat.gov.ru/databases>

5.4. Перечень программного обеспечения

Loginom

Python

5.5. Учебно-методические материалы для студентов с ограниченными возможностями здоровья

При необходимости по заявлению обучающегося с ограниченными возможностями здоровья учебно-методические материалы предоставляются в формах, адаптированных к ограничениям здоровья и восприятия информации. Для лиц с нарушениями зрения: в форме аудиофайла; в печатной форме увеличенным шрифтом. Для лиц с нарушениями слуха: в форме электронного документа; в печатной форме. Для лиц с нарушениями опорно-двигательного аппарата: в форме электронного документа; в печатной форме.

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Помещения для всех видов работ, предусмотренных учебным планом, укомплектованы необходимой специализированной учебной мебелью и техническими средствами обучения:

- столы, стулья;

- персональный компьютер / ноутбук (переносной);

- проектор;

- экран / интерактивная доска.

7. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)

Методические указания по освоению дисциплины представлены в Приложении 2 к рабочей программе дисциплины.

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

1 Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания

1.1 Показатели и критерии оценивания компетенций:

ЗУН, составляющие компетенцию	Показатели оценивания	Критерии оценивания	Средства оценивания
УК-1:Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, выработать стратегию действий			
З. основные методы интеллектуального анализа данных; математические основы методов линейной регрессии, частичных наименьших квадратов, поддерживающих векторов, искусственных нейронных сетей, методов Байеса	изучает основную и дополнительную литературу, содержащую материал об основных понятиях инструментальных средствах и математических методах, используемых при решении профессиональных задач, для подготовки к зачету, и устному опросу	полнота и содержательность ответа на зачете, устном опросе, соответствие ответов материалу, содержащемуся в изученной литературе	УО (Раздел 1 в. 1-15 Раздел 2 в. 1-11) 3 (1-26)
У. строить различные регрессионные и классификационные модели с использованием методов интеллектуального анализа данных; подбирать оптимальные значения подгоночных параметров модели	решение практико-ориентированных и практических заданий: применение различных методов с использованием Loginom, составление программ на Phython с использованием базовых конструкций, условий, циклов, массивов	правильность решения заданий на применение различных методов с использованием Loginom, составление программ на Phython с использованием базовых конструкций, условий, циклов, массивов	ЛЗ (1-5) ПОЗ (1-4)

<p>В.</p> <p>методами интерпретировать необходимые данные для формирования суждений по соответствующим социальным, научным и этическим проблемам</p>	<p>решение практико-ориентированных и практических заданий:</p> <p>применение различных методов с использованием Loginom, составление программ на Phython с использованием базовых конструкций, условий, циклов, массивов)</p>	<p>обоснованность применения методов для:</p> <p>решения заданий с использованием Loginom,</p> <p>решения заданий на составление программ на Phython с использованием базовых конструкций, условий, циклов, массивов)</p>	<p>ЛЗ (1-5)</p> <p>ПОЗ (1-4)</p>
------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------

Способен решать актуальные задачи фундаментальной и прикладной математики

<p>З.</p> <p>новые информационные технологии в гуманитарных областях знаний с использованием средств интеллектуального анализа данных и машинного обучения, компьютерной лингвистики и представления</p>	<p>изучает основную и дополнительную литературу, содержащую материал об основных понятиях инструментальных средствах и математических методах, используемых при решении профессиональных задач, для подготовки к зачету, и устному опросу</p>	<p>полнота и содержательность ответа на зачете, устном опросе, соответствие ответов материалу, содержащемуся в изученной литературе</p>	<p>УО (Раздел 1 в. 1-15 Раздел 2 в. 1-11)</p> <p>З (1-26)</p>
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------

<p>У.</p> <p>использовать новые информационные технологии в гуманитарных областях знаний с использованием средств интеллектуального анализа данных и машинного обучения, компьютерной лингвистики и представления</p>	<p>решение практико-ориентированных и практических заданий:</p> <p>применение различных методов с использованием Loginom,</p> <p>составление программ на Phython с использованием базовых конструкций, условий, циклов, массивов</p>	<p>правильность решения заданий на применение различных методов с использованием Loginom,</p> <p>составление программ на Phython с использованием базовых конструкций, условий, циклов, массивов</p>	<p>ЛЗ (1-5)</p> <p>ПОЗ (1-4)</p>
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------

знаний			
В. Методами применения новых информационных технологии в гуманитарных областях знаний с использованием средств интеллектуального анализа данных и машинного обучения, компьютерной лингвистики и представления знаний	решение практико-ориентированных и практических заданий: применение различных методов с использованием Loginom, составление программ на Python с использованием базовых конструкций, условий, циклов, массивов)	обоснованность применения методов для: решения заданий с использованием Loginom, решения заданий на составление программ на Python с использованием базовых конструкций, условий, циклов, массивов)	ЛЗ (1-5) ПОЗ (1-4)
ОПК-2:Способен совершенствовать и реализовывать новые математические методы решения прикладных задач			
З. основные подходы к анализу данных и способы их модификации	изучает основную и дополнительную литературу, содержащую материал об основных понятиях инструментальных средствах и математических методах, используемых при решении профессиональных задач, для подготовки к зачету, и устному опросу	полнота и содержательность ответа на зачете, устном опросе, соответствие ответов материалу, содержащемуся в изученной литературе	УО (Раздел 1 в. 1-15 Раздел 2 в. 1-11) З (1-26)
У. настраивать алгоритмы под особенности прикладных задач	решение практико-ориентированных и практических заданий: применение различных методов с использованием Loginom, составление программ на Python с использованием базовых конструкций, условий, циклов, массивов	правильность решения заданий на применение различных методов с использованием Loginom, составление программ на Python с использованием базовых конструкций, условий, циклов,	ЛЗ (1-5) ПОЗ (1-4)

		массивов	
В. программными инструментами, позволяющими гибкую подстройку	решение практико-ориентированных и практических заданий: применение различных методов с использованием Loginom, составление программ на Phython с использованием базовых конструкций, условий, циклов, массивов)	обоснованность применения методов для: решения заданий с использованием Loginom, решения заданий на составление программ на Phython с использованием базовых конструкций, условий, циклов, массивов)	ЛЗ (1-5) ПОЗ (1-4)

ПК-2: Способен разрабатывать и анализировать концептуальные и теоретические модели решаемых научных проблем и задач

3. методы интеллектуального анализа данных, необходимых для построения моделей решаемых научных	изучает основную и дополнительную литературу, содержащую материал библиотек Phython и математических методах, используемых при решении профессиональных задач, для подготовки к зачету, и устному опросу	полнота и содержательность ответа на зачете, устном опросе, соответствие ответов материалу, содержащемуся в изученной литературе	УО (Раздел 1 в. 1-15 Раздел 2 в. 1-11) З (5-26)
У. применять методы интеллектуального анализа данных, необходимых для построения моделей решаемых научных задач	решение практико-ориентированных и практических заданий: применение составление программ на Phython с использованием библиотек	правильность решения заданий на составление программ на Phython с использованием библиотек	ЛЗ (1-5) ПОЗ (1-4)
В. навыками применения методов интеллектуального анализа данных, необходимых для построения моделей решаемых научных	решение практико-ориентированных и практических заданий: составление программ на Phython с использованием библиотек (NumPy, Pandas, matplotlib, PyBriant)	обоснованность применения методов для: решения заданий на составление программ на Phython с использованием библиотек (NumPy, Pandas,	ЛЗ (1-5) ПОЗ (1-4)

задач		matplotlib, PyBria)	
-------	--	---------------------	--

З – вопросы к зачету, ЛЗ-лабораторное задание, ПОЗ-практико-ориентированное задание к зачету, УО- устный опрос

1.2 Шкалы оценивания:

Текущий контроль успеваемости и промежуточная аттестация осуществляется в рамках накопительной балльно-рейтинговой системы в 100-балльной шкале:

50-100 баллов (зачет)

0-49 баллов (незачет)

2 Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

В разделе приводятся типовые варианты оценочных средств: вопросы к зачету, практико-ориентированные задания к зачету, лабораторные задания, вопросы для устного опроса

Вопросы к зачету

1. Понятие Интеллектуального анализа данных (Data Mining). Data Mining как часть рынка интеллектуальных технологий
2. Интеллектуальный анализ данных. Отличия от других видов анализа данных
3. Типы наборов данных. Форматы хранения данных. Метаданные.
4. Особенности использования данных при интеллектуальном анализе данных. Данные, информация и знания.
6. Классификация задач Data Mining. Сферы применения Data Mining.
7. Процесс Data Mining
8. Задачи интеллектуального анализа данных. Классификация и кластеризация
9. Задачи интеллектуального анализа данных. Прогнозирование и визуализация
10. Методы прогнозирования и классификации: деревья решений
11. Методы прогнозирования и классификации: метод опорных векторов

12. Методы прогнозирования и классификации: метод «ближайшего соседа»
13. Методы прогнозирования и классификации: нейронные сети.
14. Методы поиска ассоциативных правил
15. Байесова классификация
16. Поставщики Data Mining. Классификация инструментов.
17. Программное обеспечение Data Mining для поиска ассоциативных правил.
18. Программное обеспечение для решения задач кластеризации и сегментации.
19. Программное обеспечение для решения задач классификации.
20. Программное обеспечение Data Mining для решения задач оценивания и прогнозирования
21. Средства извлечения данных
22. Хранилища данных
23. Сферы применения Data Mining
24. Очистка данных. Инструменты очистки данных
25. Стандарты Data Mining
26. Интеграция Data Mining и OLAP

Практико-ориентированные задания к зачету

Задание 1 (Loginom)

На основе представленного набора данных построить модель логистической регрессии (Примечание: в Сценарии создать два узла: Логистическая регрессия и Нейросеть (классификация)). Построить Куб с матрицей ошибок прогноза.

Задание 2 (Phyton)

Вклад в банке составляет X рублей. Ежегодно он увеличивается на P процентов, после чего дробная часть копеек отбрасывается. Определите, через сколько лет вклад составит не менее Y рублей. Напишите программу, которая по данным числам X , Y , P определяет, сколько лет пройдет, прежде чем сумма достигнет значения Y .

Задание 3 (Phyton)

Оценить экономическую деятельность нескольких предприятий. Известны названия предприятий, значения планового объема розничного товарооборота и значения фактического объема розничного товарооборота. Требуется определить:

1. процент выполнения плана каждым предприятием

2. количество предприятий, недовыполнивших план
3. наибольший плановый товарооборот
4. упорядочить предприятия по возрастанию планового товарооборота

Задание 4. (Phyton)

На основании представленных данных написать программу для прогнозирования цен на недвижимость с графическим выводом данных (примечание: использовать библиотеки `pumpry` и `matplotlib`)

Критерии оценивания:

- 50-100 баллов (зачет) – изложенный материал фактически верен, наличие глубоких исчерпывающих знаний; правильные, уверенные действия по применению полученных знаний на практике, грамотное и логически стройное изложение материала при ответе; практико-ориентированное задание выполнено правильно и прокомментировано; наличие твердых и достаточно полных знаний, правильные действия по применению знаний на практике, четкое изложение материала, допускаются отдельные логические и стилистические погрешности, неуверенность и неточность ответов на дополнительные и наводящие вопросы; практико-ориентированное задание выполнено правильно, но не прокомментировано; при неполном ответе на вопросы; затрудняется ответить на дополнительные вопросы; практико-ориентированное задание выполнено с ошибками и отсутствуют комментарии;
- 0-49 баллов (незачет) – ответы не связаны с вопросами, наличие грубых ошибок в ответе, непонимание сущности излагаемого вопроса, неумение применять знания на практике, неуверенность и неточность ответов на дополнительные и наводящие вопросы; практико-ориентированное задание не выполнено.

Лабораторные задания

Раздел 1. Интеллектуальный анализ данных

Тема 1. Интеллектуальный анализ данных (Data Mining) (10 баллов).

Сопоставление и сравнение понятий "информация", "данные", "знание".

Свойства знаний. Сравнение статистик, машинного обучения и Data Mining.

Классификация задач Data Mining.

Сферы применения Data Mining.

Применение Logitom

Критерии оценивания:

10 б. – задание выполнено верно;

7-9 б. – при выполнении задания были допущены неточности, не влияющие на результат;

4-6 б. – при выполнении задания были допущены ошибки;

3-1 б. – при выполнении задания были допущены существенные ошибки.

0 б. – задание не выполнено.

Тема 2. Интеллектуальный анализ данных (Data Mining) (15 баллов).

Предсказательные и описательные модели Data Mining.

Процесс Data Mining: анализ предметной области, постановка задачи, подготовка данных, построение модели, проверка и оценка моделей, выбор модели, применение модели.

Применение Logitom

Тема 3. "Задачи и методы интеллектуального анализа данных"(15 баллов).

Классификация. Регрессия. Методы представления результатов анализа.

Прогнозирование и визуализация.

Методы прогнозирования и классификации: алгоритм построения

1R-алгоритм; метод Naive Bayes; деревья решений, метод опорных векторов, метод «ближайшего соседа», нейронные сети.

Поиск ассоциативных правил.

Характеристики ассоциативных правил: поддержка, достоверность и улучшение.

Алгоритм Apriori.

Методы кластерного анализа. Меры близости в алгоритмах кластеризации.

Критерии оценивания:

15 б. – задание выполнено верно;

10-14 б. – при выполнении задания были допущены неточности, не влияющие на результат;

4-9 б. – при выполнении задания были допущены ошибки;

3-1 б. – при выполнении задания были допущены существенные ошибки.

0 б. – задание не выполнено.

Максимальное количество баллов по лабораторным заданиям раздела 1 – 40 б.

Раздел 2. "Инструменты Data Mining"

Тема 4. "Инструменты Data Mining. Logitom". (15 баллов).

Программное обеспечение Data Mining для поиска ассоциативных правил.

Программное обеспечение для решения задач кластеризации и сегментации.

Программное обеспечение для решения задач классификации. Программное обеспечение Data Mining для решения задач оценивания и прогнозирования. Применение Logitom

Тема 5. "Инструменты Data Mining. Phyton". (15 баллов).

Программное обеспечение Data Mining для поиска ассоциативных правил.

Программное обеспечение для решения задач кластеризации и сегментации.

Программное обеспечение для решения задач классификации. Программное обеспечение Data Mining для решения задач оценивания и прогнозирования. Применение Phyton

Критерии оценивания:

15 б. – задание выполнено верно;

10-14 б. – при выполнении задания были допущены неточности, не влияющие на результат;

4-9 б. – при выполнении задания были допущены ошибки;

3-1 б. – при выполнении задания были допущены существенные ошибки.

0 б. – задание не выполнено.

Максимальное количество баллов по лабораторным заданиям раздела 2 – 30 б.

Максимальное количество баллов по лабораторным заданиям - 70

Перечень вопросов для устного опроса

Раздел 1. Интеллектуальный анализ данных

1. Понятие Интеллектуального анализа данных (Data Mining). Data Mining как часть рынка интеллектуальных технологий
2. Интеллектуальный анализ данных. Отличия от других видов анализа данных
3. Типы наборов данных. Форматы хранения данных. Метаданные.
4. Особенности использования данных при интеллектуальном анализе данных. Данные, информация и знания.
6. Классификация задач Data Mining. Сферы применения Data Mining.
7. Процесс Data Mining
8. Задачи интеллектуального анализа данных. Классификация и кластеризация
9. Задачи интеллектуального анализа данных. Прогнозирование и визуализация
10. Методы прогнозирования и классификации: деревья решений
11. Методы прогнозирования и классификации: метод опорных векторов
12. Методы прогнозирования и классификации: метод «ближайшего соседа»
13. Методы прогнозирования и классификации: нейронные сети.
14. Методы поиска ассоциативных правил
15. Байесова классификация

Максимальное количество баллов по разделу 1 – 16 б

Раздел 2. "Инструменты Data Mining"

1. Поставщики Data Mining. Классификация инструментов.
2. Программное обеспечение Data Mining для поиска ассоциативных правил.
3. Программное обеспечение для решения задач кластеризации и сегментации.
4. Программное обеспечение для решения задач классификации.
5. Программное обеспечение Data Mining для решения задач оценивания и прогнозирования
6. Средства извлечения данных
7. Хранилища данных
8. Сферы применения Data Mining
9. Очистка данных. Инструменты очистки данных
10. Стандарты Data Mining
11. Интеграция Data Mining и OLAP

Максимальное количество баллов по разделу 2 – 14 б.

Критерии оценивания:

Для каждого вопроса:

2 балла - дан полный, развернутый ответ на поставленный вопрос, изложение материала при ответе – грамотное и логически стройное;

1 балл - дан неполный ответ на поставленный вопрос

0 баллов - обучающийся не владеет материалом по заданному вопросу.

Максимальное количество баллов по устному опросу – 30

3 Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

Процедуры оценивания включают в себя текущий контроль и промежуточную аттестацию.

Текущий контроль успеваемости проводится с использованием оценочных средств, представленных в п. 2 данного приложения. Результаты текущего контроля доводятся до сведения студентов до промежуточной аттестации.

Промежуточная аттестация проводится в форме зачета 2 семестр

Зачет проводится по окончании теоретического обучения до начала экзаменационной сессии в письменном виде. Количество вопросов в задании – 3 (2 теоретических вопроса и 1 практико-ориентированное задание к зачету). Проверка ответов и объявление результатов производится в день зачета. Результаты аттестации заносятся в зачетную ведомость и зачетную книжку студента. Студенты, не прошедшие промежуточную аттестацию по графику сессии, должны ликвидировать задолженность в установленном порядке.

МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Учебным планом предусмотрены следующие виды занятий:

- лекции;
- лабораторные занятия.

В ходе лабораторных работ развиваются навыки применения математических методов, выбора инструментальных средств для обработки и анализа экономических данных в профессиональной деятельности

При подготовке к лабораторным занятиям каждый студент должен:

- изучить рекомендованную учебную литературу;

В процессе подготовки к лабораторным занятиям студенты могут воспользоваться консультациями преподавателя.

Теоретические вопросы должны быть изучены студентами в ходе самостоятельной работы. Контроль самостоятельной работы студентов над учебной программой курса осуществляется методом устного опроса и выполнения лабораторных заданий. В ходе самостоятельной работы каждый студент обязан прочитать основную и по возможности дополнительную литературу по изучаемой теме. Выделить непонятные термины, найти их значение в литературе.

Для подготовки к занятиям, текущему контролю и промежуточной аттестации студенты могут воспользоваться электронно-библиотечными системами. Также обучающиеся могут взять на дом необходимую литературу на абонементе университетской библиотеки или воспользоваться читальными залами.