

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Андрей Драгомирович Хлутков
Должность: директор
Дата подписания: 25.06.2023 18:41:06
Уникальный программный ключ:
880f7c07c583b07b775f6604a630281b13ca9fd2

Приложение 6 ОП ВО

**Федеральное государственное бюджетное образовательное
учреждение высшего образования
«РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА
И ГОСУДАРСТВЕННОЙ СЛУЖБЫ
ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»**

Северо-Западный институт управления – филиал РАНХиГС

ФАКУЛЬТЕТ ЭКОНОМИКИ и ФИНАНСОВ
(наименование структурного подразделения (института/факультета/филиала))
Кафедра экономики
(наименование кафедры)

УТВЕРЖДЕНА

Директор СЗИУ РАНХиГС

_____ Хлутков А.Д.

Электронная подпись

**ПРОГРАММА МАГИСТРАТУРЫ
Корпоративная экономика**

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ,
Реализуемой без применения электронного(онлайн)курса**

Б1.В. 07 Анализ данных
(индекс, наименование дисциплины, в соответствии с учебным планом)

38.04.01 Экономика
(код, наименование направления подготовки (специальности))

Очная/Заочная
(форма(ы) обучения)
Год набора: 2023

Санкт-Петербург, 2022 г.

Авторы–составители:

д. воен. н., зав. кафедрой бизнес-информатики Наумов Владимир Николаевич
д.э.н., профессор кафедры бизнес-информатики Куклина Евгения Анатольевна

Заведующий кафедрой экономики, д.э.н., профессор Мисько Олег Николаевич
(наименование кафедры) (ученая степень и(или) ученое звание) (Ф.И.О.)

РПД Б1.В.07 «Анализ данных» одобрена на заседании кафедры экономики. Протокол от (27 мая 2022 года) № (2).

СОДЕРЖАНИЕ

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы
2. Объем и место дисциплины в структуре образовательной программы
3. Содержание и структура дисциплины
4. Материалы текущего контроля успеваемости обучающихся
5. Оценочные материалы промежуточной аттестации по дисциплине
6. Методические материалы для освоения дисциплины
7. Учебная литература и ресурсы информационно-телекоммуникационной сети "Интернет"
 - 7.1. Основная литература
 - 7.2. Дополнительная литература
 - 7.3. Нормативные правовые документы и иная правовая информация
 - 7.4. Интернет-ресурсы
 - 7.5. Иные источники
8. Материально-техническая база, информационные технологии, программное обеспечение и информационные справочные системы

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения программы

1.1. Дисциплина «Анализ данных» обеспечивает овладение следующими компетенциями:

Код компетенции	Наименование компетенции	Код компонента компетенции	Наименование компонента компетенции
УК-2	Способен управлять проектом на всех этапах его жизненного цикла	УК-2.1	Способен оценивать ресурсное обеспечение проекта с учетом последовательности этапов его жизненного цикла
ПКс-1	Способен обосновывать актуальность, теоретическую и практическую значимость избранной темы научного исследования	ПКс-1.1	Способен обосновывать актуальность выбранной темы научного исследования
		ПКс-1.2	Способен обосновывать теоретическую и практическую значимость выбранной темы научного исследования
ПКс-8	Способен разрабатывать варианты управленческих решений и обосновывать их выбор на основе критериев социально-экономической эффективности	ПКс-8.1	Способен определять критерии социально-экономической эффективности и разрабатывает альтернативные варианты управленческих решений
		ПКс-8.2	Способен обосновывать выбор оптимального управленческого решения на основе критериев социально-экономической эффективности

1.2. В результате освоения дисциплины у студентов должны быть сформированы:

ОТФ/ТФ (при наличии профстандарта)/ профессиональные действия	Код этапа освоения компетенции	Результаты обучения
В/02.7 Разработка финансового плана для клиента и целевого инвестиционного портфеля	ПКс-1.1	уровне знаний: - о актуальных направлениях научных исследований в области оценочных процедур на уровне умений: - умением выбрать актуальную тему научного исследования в области оценочной деятельности

		<p>на уровне навыков: - практическими навыками обоснования актуальности выбранной темы научного исследования</p>
	ПКс-1.2	<p>на уровне знаний: методики выполнения научных исследований; - программирования процесса выполнения научной деятельности</p> <p>на уровне умений: - умением определить теоретическую и практическую значимость научного исследования в области оценочной деятельности</p> <p>на уровне навыков: - обоснования теоретической и практической значимости темы научного исследования</p>
В/02.7 Разработка финансового плана для клиента и целевого инвестиционного портфеля	ПКс-8.1	<p>на уровне знаний: - критерии социально-экономической эффективности управленческих решений;</p> <p>на уровне умения: определять критерии социально-экономической эффективности управленческих решений;</p> <p>на уровне владения: - навыками определения и использования критериев социально-экономической эффективности управленческих решений</p>
	ПКс-8.2	<p>на уровне знаний: - методов разработки альтернативных вариантов управленческих решений</p> <p>на уровне умений: - умением разрабатывать альтернативные варианты управленческих решений</p> <p>на уровне навыков: - навыками разработки альтернативных вариантов управленческих решений</p>
	УК-2.1	<p>на уровне знаний: - о способах управления проектом с учетом последовательности этапов его жизненного цикла</p> <p>на уровне умений: - определить способы управления проектом с учетом последовательности этапов его жизненного цикла</p> <p>на уровне навыков: - навыками управления проектом с учетом последовательности этапов его жизненного цикла</p>

2. Объем и место дисциплины в структуре образовательной программы

Дисциплина Б1.В.07 «Анализ данных» преподается в 4 семестре по очной форме обучения и 3 семестре по заочной форме обучения.

Общая трудоемкость Б1.В.07 «Анализ данных» составляет 3 зачётных единиц – 108 ак. часов.

Вид работы	Трудоемкость (очная/заочная форма)
Общая трудоемкость	108
Контактная работа с преподавателем	54 / 10
Лекции	16 / 2
Практические занятия	24 / 4
Лабораторные занятия	12 / 2
Самостоятельная работа	54 / 94
консультация	2/2
Контроль	- / 4
Формы текущего контроля	Контрольная работа, расчетное задание, тестирование
Форма промежуточной аттестации	Зачет с оценкой

Количество академических часов, выделенных на контактную работу с преподавателем для очной формы обучения, 54 ак. часов: лекционные занятия – 16 ак. часа., лабораторные занятия – 12 ак. часов, практические занятия – 24 ак. часа, консультация к зачету с оценкой составляет – 2 ак. часа. Самостоятельная работа составляет 54 ак. часов.

Количество академических часов, выделенных на контактную работу с преподавателем для заочной формы обучения, составляет 10 ак. часов: лекционные занятия – 2 ак. часа., лабораторные занятия – 2 ак. часов, практические занятия – 4 ак. часа, консультация к зачету с оценкой – 2 ак. часа. Самостоятельная работа составляет 94 ак. часов. Контроль – 4 ак. часа.

Форма промежуточной аттестации – зачет с оценкой.

Дисциплина Б1.В.07 «Анализ данных» является теоретическим и методологическим курсом, в процессе изучения которого магистранты знакомятся с основными инструментами анализа данных для изучения экономических систем и явлений.

«Входными» для ее освоения являются знания, умения и навыки, полученные обучающимися в процессе изучения таких дисциплин как «Теория фирмы» (Б1.В.03, 1 курс), «Управленческий анализ (продвинутый уровень)» (Б1.О.04, 1 курс), что обеспечивает успешное освоение профессиональных компетенций. Знания, умения и навыки, полученные при изучении дисциплины, используются студентами при выполнении выпускных квалификационных работ (магистерских диссертаций), а также в дальнейшей практической деятельности.

Дисциплина реализуется с применением дистанционных образовательных технологий (далее – ДОТ). Доступ к системе дистанционных образовательных технологий осуществляется каждым обучающимся самостоятельно с любого устройства на портале: <https://lms.ranepa.ru>.

Пароль и логин к личному кабинету / профилю предоставляется студенту в деканате.

Все формы текущего контроля, проводимые в системе дистанционного обучения, оцениваются в системе дистанционного обучения. Доступ к видео и материалам лекций предоставляется в течение всего семестра. Доступ к каждому виду работ и количество попыток на выполнение задания предоставляется на ограниченное время согласно регламенту дисциплины, опубликованному в СДО. Преподаватель оценивает выполненные обучающимся работы не позднее 10 рабочих дней после окончания срока выполнения.

3 Содержание и структура дисциплины

3.1 Структура дисциплины (очная форма обучения)

№ п/п	Наименование тем	Объем дисциплины, час.					Форма текущего контроля успеваемости, промежуточной аттестации	
		Всего	Контактная работа обучающихся с преподавателем по видам учебных занятий					СР
			Л/Д ОТ	ЛР/ДОТ	ПЗ/Д ОТ	КС Р		
Тема 1	Основы анализа данных. Системы поддержки принятия решений	20	4	2	4	-	10	Т, КР
Тема 2	Предобработка и очистка данных	18	2	2	4	-	10	Т, КР
Тема 3	Классификационный анализ без обучения. Кластерный анализ	20	4	2	4	-	10	Т, КР
Тема 4	Анализ взаимосвязей между переменными. Ассоциативные правила	21	4	3	4	-	10	КР, РЗ / РГЗ
Тема 5	Классификационный анализ с обучением	27	2	3	8	-	14	КР, РЗ / РГЗ
Промежуточная аттестация				-	-	-	-	Зач. с оц.
Всего:		108	16	12	24	2	54	

(заочная форма обучения)

№ п/п	Наименование тем	Объем дисциплины, час.					Форма текущего контроля успеваемости, промежуточной аттестации	
		Всего	Контактная работа обучающихся с преподавателем по видам учебных занятий					СР
			Л/Д ОТ	ЛР/ДОТ	ПЗ/Д ОТ	КС Р		
Тема 1	Основы анализа данных. Системы поддержки принятия решений	20	0,5	0,5	1	-	18	Т, КР
Тема 2	Предобработка и очистка данных	19,5	-	-	0,5	-	19	Т, КР
Тема 3	Классификационный анализ без обучения. Кластерный анализ	20,5	0,5	0,5	0,5	-	19	Т, КР
Тема 4	Анализ взаимосвязей между переменными. Ассоциативные правила	21	0,5	0,5	1	-	19	КР, РЗ / РГЗ
Тема 5	Классификационный анализ с обучением	21	0,5	0,5	1	-	19	КР, РЗ / РГЗ
Промежуточная аттестация		4		-	-	-	-	Зач. с оц.
Всего:		108	2	2	4	2	94	

*Примечание: * – разработчик указывает необходимые формы текущего контроля успеваемости: тест (Т); контрольная работа (КР), расчетное задание (РЗ);*

*** – разработчик указывает необходимые формы промежуточной аттестации: экзамен (экз.).*

3.2 Содержание дисциплины

Тема 1. Основы анализа данных. Системы поддержки принятия решений

Введение. Понятие анализа данных. Задачи систем поддержки принятия решений. OLTP и OLAP-системы. Принципы построения информационных хранилищ. Модели информационных хранилищ. Многомерная модель данных. Правила Кодда. Размерностные модели. MOLAP, ROLAP, HOLAP- системы. Витрины данных. ETL (Extracting Transforming and Loading) – средство извлечения, обработки и загрузки данных. Добыча данных. Добыча данных в управлении качеством. Data Mining. Стандарты Data Mining. Стандарт CWM, CRISP, PMML. Жизненный цикл процесса анализа данных. Классификация методов Data Mining. Модели Data Mining. Понятие данные и знания. Процесс обнаружения знаний. Классификация задач Data Mining. Методы анализа данных. Разведочный анализ данных. Очистка и фильтрация данных. Статистические диаграммы. «Ящичные» диаграммы. Диаграммы «ствол-листья». Задачи классификации и регрессии. Использование статистических пакетов для интеллектуального анализа данных. Понятие бизнес-аналитики. Средства бизнес-аналитики. Средства легкой бизнес-аналитики. QlikView, QlikSense.

Тема 2. Предобработка и очистка данных

Методология KDD. Задачи предобработки данных. Технология ETL. Просмотр данных. Очистка данных. Оценка качества данных. Заполнение пропущенных данных. Аномальные и предельные данные. Использование ящичной диаграммы. Выявление дубликатов и противоречий. Корреляционный анализ. Использование факторного анализа при предобработке данных. Трансформация данных. Квантование. Сэмплинг. Группировка данных.

Тема 3. Классификационный анализ без обучения. Кластерный анализ

Постановка задач кластерного анализа. Определение кластера. Параметры кластера. Меры близости. Метрики кластерного анализа. Базовые алгоритмы кластеризации. Иерархическая кластеризация. Дендограммы. Метод К-средних. Профили кластеров. Взаимосвязь кластерного и регрессионного анализа. Использование пакета Deductor для решения задач кластерного анализа. Кластерный анализ в средствах интеллектуального анализа Microsoft Office.

Тема 4. Анализ взаимосвязей между переменными. Ассоциативные правила

Основные положения непараметрической и нечисловой статистики. Таблицы сопряженности. Таблица сопряженности 2x2. Таблицы флагов и заголовков. Непараметрические и нечисловые критерии. Канонический анализ. Корреляционная матрица. Коэффициенты канонической корреляции. Меры избыточности переменных. Задачи ассоциации. Ассоциативные правила. Поддержка и достоверность ассоциативных правил. Лифт. Алгоритмы построения ассоциативных правил. Рекомендации по генерации правил. Алгоритм a priori. Использование пакета Deductor для построения ассоциативных правил.

Тема 5. Классификационный анализ с обучением

Формулировка задачи классификации. Классификационный анализ с обучением. Деревья решений. Алгоритмы построения деревьев решений. Классификация критериев разбиений. Критерий Gini. Деревья классификации и их свойства. Типы ветвления. Методы и

алгоритмы построения деревьев. Алгоритм CART. Определение прекращения построения дерева классификации. Использование нейронных сетей для решения задач классификации. Карты Кохоннена. Логистическая регрессия. Сравнение результатов классификации различными методами.

Примеры алгоритмов построения деревьев решений. Использование статистических пакетов Deductor, Statistica, Excel для построения деревьев решений.

Промежуточная аттестация может проводиться с использованием ДОТ.

4. Материалы текущего контроля успеваемости обучающихся

4.1. В ходе реализации дисциплины Б1.В.06 «Электронный бизнес» используются следующие методы текущего контроля успеваемости обучающихся:

Очная и заочная форма обучения

Тема (раздел)	Методы текущего контроля успеваемости
Тема 1. Основы анализа данных. Системы поддержки принятия решений	Решение тестовых заданий Контрольная работа
Тема 2. Предобработка и очистка данных	Решение тестовых заданий Контрольная работа
Тема 3. Классификационный анализ без обучения. Кластерный анализ	Решение тестовых заданий Контрольная работа
Тема 4. Анализ взаимосвязей между переменными. Ассоциативные правила	Контрольная работа Расчетное задание
Тема 5. Классификационный анализ с обучением	Контрольная работа Расчетное задание

Экзамен проводится в форме тестирования.

4.2. Типовые материалы текущего контроля успеваемости обучающихся

4.2.1. Типовые материалы текущего контроля успеваемости обучающихся к теме 1.
Основы анализа данных. Системы поддержки принятия решений

Тестовые задания:

ЗАДАНИЕ № 1 (выберите один вариант ответа)

В дискриминантном анализе обучающая выборка используется для ...

ВАРИАНТЫ ОТВЕТОВ:

- | | |
|----------------------------------|------------------------------------|
| 1) Выявления значимых признаков | 2) Выявления аномального измерения |
| 3) Разделения объектов на классы | 4) Выбора вида модели |

ЗАДАНИЕ № 2 (выберите один вариант ответа)

В факторном анализе при n измерениях и k факторах матрица факторных нагрузок имеет размерность ...

ВАРИАНТЫ ОТВЕТОВ:

- | | |
|-----------------|-----------------|
| 1) $n \times n$ | 2) $k \times k$ |
| 3) $n \times k$ | 4) $k \times n$ |

**4.2.2. Типовые материалы текущего контроля успеваемости обучающихся к теме 2.
*Предобработка и очистка данных***

ЗАДАНИЕ № 1 (выберите несколько вариантов ответа)

Метод главных компонент ...

ВАРИАНТЫ ОТВЕТОВ:

- | | |
|---|---|
| 1) Является частным случаем метода факторного анализа | 2) Предназначен для снижения размерности задачи |
| 3) Устраняет проблему коррелированности факторов | 4) Предназначен для классификации |

ЗАДАНИЕ № 2 (выберите один вариант ответа)

Сигмоидальная активизационная функция искусственного нейрона имеет вид...

**4.2.3. Типовые материалы текущего контроля успеваемости обучающихся к теме 3.
*Классификационный анализ без обучения. Кластерный анализ***

Тестовые задания:

ЗАДАНИЕ № 1 (выберите несколько вариантов ответа)

В кластерном анализе используются методы объединения ...

ВАРИАНТЫ ОТВЕТОВ:

- | | |
|--------------------|----------------------|
| 1) Ближнего соседа | 2) Дальнего соседа |
| 3) Среднего соседа | 4) центроидный метод |

ЗАДАНИЕ № 2 (выберите несколько вариантов ответа)

В кластерном анализе для определения близости между кластерами используются метрики ...

ВАРИАНТЫ ОТВЕТОВ:

- | | |
|------------------------------------|---------------------------------|
| 1) Эвклидово расстояние | 2) Куб Эвклидова расстояния |
| 3) Взвешенное эвклидово расстояние | 4) Квадрат Эвклидова расстояния |

**4.2.4. Типовые материалы текущего контроля успеваемости обучающихся к теме 4.
*Анализ взаимосвязей между переменными. Ассоциативные правила***

Расчетное задание:

Ассоциативные правила

Построить ассоциативные правила по имеемым транзакциям. Рассчитать характеристики для каждого правила.

Транзакционная база данных	
TID	Приобретенные покупки
100	ремень, женская сумка, портмоне
200	женская сумка, косметичка
300	женская сумка, ремень, ключница, портмоне
400	дамский зонт, ключница, косметичка
500	ремень, женская сумка, портмоне, ключница
600	косметичка, портмоне
700	ремень, портфель

Деревья решений.

Построить дерево решений по данным, приведенным в таблице.

Рейтинг	Возраст	Уровень Дохода	Образование
0	35	3000	0
0	25	5000	1
0	31	7000	1
1	56	1000	0
1	62	1100	1
1	49	1500	0

4.2.5. Типовые материалы текущего контроля успеваемости обучающихся к теме 5. Классификационный анализ с обучением

Расчетное задание:

Решение задач классификации в *DeductorAcademic*

Решить задачу логистической регрессии. Определить качество построенной модели классификации. Решить данную задачу другими методами классификации, реализованными в *DeductorAcademic*. Сравнить результаты решения задачи классификации с помощью таблицы сопряженности.

Рейтинг	Образование, A1	Доход, A2	Возраст, A3
низкий	высшее	малый	35
низкий	среднее	большой	40
высокий	высшее	большой	30
высокий	высшее	большой	30
низкий	среднее	малый	30
высокий	высшее	малый	35
высокий	высшее	большой	45
высокий	высшее	большой	35

Контрольная работа

Контрольная работа включает пять задач. Шаблоны контрольной работы размещены в файле Excel. К тематике задач относятся: задача очистки данных, иерархическая задача кластерного анализа, решение задачи кластерного анализа методов k-средних, построение ассоциативных правил, построение дерева решений.

Пример задачи. Построить дендограмму, используя Евклидово расстояние и метод "дальнего соседа". Перед построением кластеров выполнить стандартизацию значений атрибутов

Номер объекта	x1	x2
1	3,00	10,00
2	4,00	11,00
3	6,00	10,00
4	10,00	9,00
5	11,00	9,00
6	10,00	7,00

Найти ассоциативные правила, если множества транзакций имеют вид

TID	Предметные наборы			
TID1	зубная паста	крем для бритья	шампунь	
TID2	мыло	дезодорант	шампунь	
TID3	шампунь	дезодорант	лосьен после бритья	шампунь
TID4	крем для бритья	шампунь	дезодорант	лосьен после бритья
TID5	лосьен после бритья	мыло	зубная паста	
TID6	дезодорант	мыло	лосьен после бритья	дезодорант
TID7	дезодорант	шампунь		
TID8	зубная паста	дезодорант	крем для бритья	
TID9	дезодорант	мыло	лосьен после бритья	
TID10	лосьен после бритья	шампунь		

Расчетно-графическое задание (пример).

Использование пакета QlikView для решения задач анализа данных о демографической ситуации в России. Для каждого варианта приведены таблицы с указанием вида исходных данных, которые будут анализироваться средствами бизнес-аналитики

5. Оценочные материалы промежуточной аттестации по дисциплине

5.1 Зачет с оценкой проводится с применением следующих методов (средств): в виде теста. На зачет с оценкой выносится материал в объеме, предусмотренном рабочей программой учебной дисциплины.

При реализации промежуточной аттестации в ЭО/ДОТ могут быть использованы следующие формы: – при проведении занятий лекционного типа: устный опрос; при проведении занятий семинарского типа: устный опрос, тестирование.

Промежуточная аттестация проводится в форме собеседования по вопросам.
Тестирование в СДО с прокторингом.

Фонд тестовых заданий по дисциплине находится на кафедре экономики.

5.2. Оценочные материалы промежуточной аттестации

Компонент компетенции	Промежуточный / ключевой индикатор оценивания	Критерий оценивания
УК-2.1 Способен оценивать ресурсное обеспечение проекта с учетом последовательности этапов его жизненного цикла	Оценивает ресурсное обеспечение проекта с учетом последовательности этапов его жизненного цикла	Определено ресурсное обеспечение проекта с учетом последовательности этапов его жизненного цикла
ПКс-8.1 Способен определять критерии социально-экономической эффективности и разрабатывает альтернативные варианты управленческих решений	Определяет критерии социально-экономической эффективности и разрабатывает альтернативные варианты управленческих решений	Определены критерии социально-экономической эффективности и разрабатывает альтернативные варианты управленческих решений
ПКс-8.2 Способен обосновывать выбор оптимального управленческого решения на основе критериев социально-экономической эффективности	Обосновывает выбор оптимального управленческого решения на основе критериев социально-экономической эффективности	Выбрано оптимальное управленческое решение на основе критериев социально-экономической эффективности

Типовые оценочные материалы промежуточной аттестации

Примеры задач (допуск к зачёту с оценкой)

1. Создать файл QlikView с данными о Российских банках. Создать списки, содержащие данные о городе, банке, бюджете, количестве рабочих и активах. Построить гистограмму количества рабочих в банке.

2. Найти описательную статистику и оценить характеристики вариационного ряда. Построить гистограмму распределения для данных на листе Описательная статистика в файле Excel с заданиями на зачет.
3. Создать файл QlikView с данными о Российских банках. Создать списки, содержащие данные о городе, банке, бюджете, количестве рабочих и активах. Создать объект «диаграмма» с измерением «город», содержащий данные о минимальном, максимальном и среднем числе рабочих в банке.
4. Создать файл QlikView с данными о Российских банках. Создать списки, содержащие данные о городе, банке, бюджете, количестве рабочих и активах. Создать объект «диаграмма» с измерением «банк», содержащий данные об активах банков. Отсортировать данные по убыванию. Выбрать 10 банков с максимальными активами. Найти статистические характеристики
5. Выполнить предобработку данных и построить диаграмму «ящик с усами», если в качестве границ ящика использовать – первый и третий квартили. Для определения аномальных значений использовать $1,5 \times [\text{межквартильное расстояние}]$. Для использования крайних значений использовать $2 \times [\text{межквартильное расстояние}]$. Усы строить после удаления аномалий.
6. Создать файл QlikView с данными о нагрузке. Построить гистограмму суммарной нагрузки в группах, используя лист Дисциплины, группы, преподаватели. Построить сводную таблицу с размерностями Группа, преподаватели, фактом – суммарное число часов. Отобразить схему данных с помощью обозревателя таблиц.

Перечень вопросов для зачета с оценкой

1. Системы поддержки принятия решений. Хранилища данных.
2. Размерностные модели. OLAP-куб. Таблица размерностей. Таблица фактов.
3. Сравнительный анализ OLAP и OLTP-систем.
4. Понятие бизнес-аналитики. Классификация средств «бизнес-аналитики».
5. Этапы анализа данных. KDD.
6. Data Mining. Средства обработки Data Mining
7. Элементы математической статистики. Описательная статистика. Операции агрегирования данных.
8. Графические средства анализа. Диаграмма рассеяния. Гистограмма.
9. Начальные этапы KDD. ETL. Средства очистки и трансформации данных.
10. Классификация методов предобработки и очистки данных.
11. Методы борьбы с аномалиями. Ящечная диаграмма.
12. Общая характеристика задач кластерного анализа.
13. Метрики кластерного анализа.
14. Методы определения близости между кластерами.
15. Иерархическая кластеризация. Дендограмма.
16. Метод k-средних.
17. Ассоциативные правила. Свойства антимонотонности.
18. Метрики построения ассоциативных правил.
19. Алгоритм построения ассоциативных правил a’priori.
20. Общая характеристика деревьев решений.
21. Алгоритмы построения деревьев решений.
22. Оценка качества классификации. Задачи классификации. ROC-кривая. Таблица сопряженности.
23. Определение регрессионной модели. Логистическая регрессионная модель. Использование логистической модели для классификации.
24. Нейронные сети. Перцептрон. Радиальные базисные сети.
25. Использование карты Кохоннена для решения задач классификации.

26. Общая характеристика QlikView.
27. Общая характеристика DeductorAcademic.
28. Общая характеристика средств интеллектуального анализа SQLServer.

Шкала оценивания.

На оценку «**Отлично**» студент должен продемонстрировать:

Успешное и систематическое применение навыков анализа нормативных актов, регулирующих отношения в экономической сфере деятельности.

Сформированные систематические представления о правовых нормах действующего законодательства, регулирующие отношения в экономической деятельности.

Находить необходимую правовую информацию для решения проблем в экономической деятельности хозяйствующих субъектов.

А также излагает выводы и предложения, правильно отвечает на все дополнительные вопросы, ответ должен быть логичным и последовательным.

На оценку «**Хорошо**» студент должен продемонстрировать:

В целом успешное, но содержащее отдельные пробелы применение навыков анализа нормативных актов, регулирующих отношения в экономической деятельности.

Сформированные, но содержащие отдельные пробелы представления о правовых нормах действующего законодательства, регулирующие отношения в экономической деятельности.

Находить необходимую правовую информацию для решения проблем в экономической деятельности хозяйствующих субъектов.

А также излагает выводы и предложения. Вместе с тем, не вполне правильно отвечает на все дополнительные вопросы. Ответ не отличается логичностью и последовательностью.

На оценку «**Удовлетворительно**» студент должен продемонстрировать:

В целом успешное, но не систематическое применение навыков анализа нормативных актов, регулирующих отношения в экономической деятельности.

Частично находить необходимую правовую информацию для решения проблем в экономической деятельности хозяйствующих субъектов.

Неполные представления о правовых нормах действующего законодательства, регулирующие отношения в экономической деятельности.

В тоже время не может изложить выводы и предложить рекомендации. Не вполне правильно отвечает на все дополнительные вопросы. Ответ не отличается логичностью и последовательностью.

На оценку «**Неудовлетворительно**» студент не может продемонстрировать:

Фрагментарное применение навыков анализа нормативных актов, регулирующих отношения в экономической деятельности.

Фрагментарные представления о правовых нормах действующего законодательства, регулирующие отношения в экономической деятельности.

Не может найти необходимую правовую информацию для решения проблем в экономической деятельности хозяйствующих субъектов.

Не может изложить выводы и предложить рекомендации. Неправильно отвечает на все дополнительные вопросы. Ответ не отличается логичностью и последовательностью.

6. Методические материалы для освоения дисциплины

Экзамены организуются в период сессии в соответствии с текущим графиком учебного процесса, утвержденным в соответствии с установленным в СЗИУ порядком. Продолжительность экзамена для каждого студента не может превышать четырех академических часов. Экзамен не может начинаться ранее 9.00 часов и заканчиваться позднее 21.00 часа. Экзамен проводится в аудитории, в которую запускаются одновременно

не более 5 человек. Время на подготовку ответов по билету каждому обучающемуся отводится 30-40 минут. При явке на экзамен с оценкой обучающийся должен иметь при себе зачетную книжку. Во время экзамена обучающиеся по решению преподавателя могут пользоваться учебной программой дисциплины и справочной литературой.

Рабочей программой дисциплины предусмотрены следующие виды аудиторных занятий: лекции, практические занятия. На лекциях рассматривается наиболее сложный материал дисциплины. Лекция сопровождается презентациями, компьютерными текстами лекции, что позволяет магистранту самостоятельно работать над повторением и закреплением лекционного материала. Для этого магистранту должно быть предоставлено право самостоятельно работать в компьютерных классах в сети Интернет.

Для работы с печатными и электронными ресурсами СЗИУ имеется возможность доступа к электронным ресурсам. Организация работы студентов с электронной библиотекой указана на сайте института (странице сайта – «Научная библиотека»).

№ п/п	Наименование темы или раздела дисциплины (модуля)	Список рекомендуемой литературы		Вопросы для самопроверки
		Основная	Дополнительная	
1	Основы анализа данных. Системы поддержки принятия решений	№№ 1, 2	№ 1,2	<ol style="list-style-type: none"> 1. Дайте сравнительный анализ OLAP и OLTP систем. Сферы их применения. 2. В чем отличие информационного хранилища от баз данных? 3. Принципы построения информационных хранилищ. Классификация информационных хранилищ. 4. Модели информационных хранилищ. Многомерная модель данных. Нормальная форма. Денормализация моделей данных. 5. Правила Кодда. Зачем применяется денормализация моделей? 6. Размерностные модели. В чем отличие таблицы фактов от размерностной таблицы? 7. Дайте характеристику стандартам DataMining. 8. Что такое «большие данные»? 9. Назовите средства бизнес-аналитики
2	Предобработка и очистка данных	№№ 1,2	№№ 1,2	<ol style="list-style-type: none"> 1. Дайте характеристикуэтапа ETL (Extracting Transforming and Loading). 2. Какие задачи решаются DataMining? 3. Каково предназначение и средства разведочный анализ данных? Дайте характеристику диаграммы «ящик с усами» 4. Назовите какие операции выполняются при агрегировании данных. 5. Приведите примеры использования статистических пакетов для разведочного анализа.

				<p>6. Назовите и выполните сравнительный анализ графических средств анализа. Дайте характеристику биржевых диаграмм.</p> <p>7. Для чего используются диаграммы рассеяния?</p>
	Классификационный анализ без обучения. Кластерный анализ	№ 1,2	№№ 1,2	<p>1. Что понимается под кластером? Назовите характеристики кластера. Что такое «центроид» кластера?</p> <p>2. Дайте классификацию методов кластерного анализа. Приведите примеры их применения в практической жизни.</p> <p>3. Зачем используются меры близости? Назовите методы определения близости между кластерами.</p> <p>4. Когда применяется метод ближнего соседа, дальнего соседа? Сравните их.</p> <p>5. Дайте характеристику метрик кластерного анализа.</p> <p>6. Поясните содержание «дендограммы» и организацию ее применения.</p> <p>7. Что понимается под профилем кластера.</p> <p>8. Использование статистических пакетов для решения задач кластерного анализа.</p> <p>9. Дайте характеристику метода k-средних.</p>
3	Анализ взаимосвязей между переменными, ассоциативные правила	№№ 1,2	№№ 1,2	<p>1. Зачем используются ассоциативные правила? Приведите примеры задач использования ассоциативных правил.</p> <p>2. Дайте определение ассоциативного правила. Зачем используются обобщенные правила? Что такое транзакция. Приведите примеры транзакций.</p> <p>3. Какие показатели используются для построения правила?</p> <p>4. Алгоритмы построения ассоциативных правил. Алгоритм apriori.</p> <p>5. Общая характеристика пакета Deductor.</p> <p>6. Использование пакета Deductor для решения задач интеллектуального анализа данных.</p>
4	Классификационный анализ с обучением	№№ 1, 2	№№ 1,2	<p>1. Дайте определение задачи классификации. Какие методы решения задачи классификации Вы знаете?</p> <p>2. Особенности решения задач классификации с обучением.</p> <p>3. Деревья классификации и их свойства.</p> <p>4. Приведите примеры алгоритмы</p>

				<p>построения деревьев.</p> <p>5. Как определяется правило остановки построения дерева?</p> <p>6. Алгоритм CART? Приведите пример его использования.</p> <p>7. Дайте определение искусственного нейрона. Из чего состоит нейрон?</p> <p>8. Дайте классификацию решающих функций нейрона. Выполните их сравнительный анализ.</p> <p>9. Искусственные нейронные сети. Что понимается под архитектурой нейронных сетей?</p>
--	--	--	--	--

7. Учебная литература и ресурсы информационно-телекоммуникационной сети "Интернет"

7.1. Основная литература

1. Миркин Б.Г. Введение в анализ данных [Электронный ресурс]: учебник и практикум / Б. Г. Миркин; Нац. исслед. ун-т Высш. шк. экономики. - Электрон. дан. - М. : Юрайт, 2016. - 174 с.
2. Паклин Н.Б. Бизнес-аналитика: от данных к знаниям : [хранилища данных и OLAP, очистка и предобработка данных, основные алгоритмы DataMining, сравнение и ансамбли моделей, решение бизнес задач на аналитической платформе Deductor] : учеб. пособие / Н. Паклин, В. Орешков. - 2-е изд., испр. - СПб.[и др.]: Питер, 2013. - 701

7.2. Дополнительная литература

1. Наумов В.Н. Средства бизнес-аналитики. – СПб.: СЗИУ, 2016.
2. Наумов В.Н. Проектирование информационных систем: Учебное пособие. - СПб.: СЗИУ, 2018.

7.3. Нормативные правовые документы и иная правовая информация

1. Конституция Российской Федерации.- М., 1993.
2. Бюджетный кодекс Российской Федерации от 31 июля 1998г. № 145-ФЗ (с изм. и доп. от 20 августа 2007г.). – М., 2007.
3. Гражданский кодекс Российской Федерации. Часть I. ФЗ РФ № 51-ФЗ от 30 ноября 1994г. Часть II. ФЗ РФ № 314-ФЗ от 26 января 1996г. Часть III. ФЗ РФ от 26 ноября 2001г. Часть IV. ФЗ РФ № 18 декабря 2006г. – М., 2008.
4. Жилищный кодекс Российской Федерации. – М., 2005.
5. Налоговый кодекс Российской Федерации. – М., 2008.
6. Трудовой кодекс Российской Федерации. – М., 2002.

7.4. Интернет-ресурсы

Русскоязычные ресурсы

1. <http://nwapa.spb.ru> – официальный сайт научной библиотеки СЗИУ
2. Единый центр аудита [Электронный ресурс]. – Режим доступа: [http:// www.eaudit.ru](http://www.eaudit.ru).
3. МедиаПилотline [Электронный ресурс]. – Режим доступа: [http:// www.mpilot.ru](http://www.mpilot.ru).
4. Национальное рейтинговое агентство [Электронный ресурс]. – Режим доступа: [http:// www.ra-national.ru](http://www.ra-national.ru).

5. Официальный сайт Счетной палаты Российской Федерации [Электронный ресурс]. – Режим доступа: [http:// www.ach.cov.ru](http://www.ach.cov.ru).
6. Оценка интеллектуальной собственности [Электронный ресурс]. – Режим доступа: [http:// www.woku.ru/oczenka](http://www.woku.ru/oczenka).
7. Официальный сайт Центрального Банка Российской Федерации [Электронный ресурс]. – Режим доступа: [http:// www.gbr.ru](http://www.gbr.ru).
8. Правовая система «Гарант-Интернет» [Электронный ресурс]. – Режим доступа: [http:// www.garweb.ru](http://www.garweb.ru).
9. Правовая система «КонсультантПлюс» [Электронный ресурс]. – Режим доступа: [http:// www.consultant.ru](http://www.consultant.ru).
10. Рейтинговое агентство ЗАО «Рус-Рейтинг» [Электронный ресурс]. – Режим доступа: [http:// www.rusrating.ru](http://www.rusrating.ru).
11. РосБизнесКонсалтинг [Электронный ресурс]. – Режим доступа: [http:// www.rbc.ru](http://www.rbc.ru).
12. Центр экономического анализа и экспертизы [Электронный ресурс]. – Режим доступа: [http:// www.ceae.ru](http://www.ceae.ru).
13. Электронные учебники электронно-библиотечной системы (ЭБС) «Айбукс» http://www.nwapa.spb.ru/index.php?page_id=76
14. Научно-практические статьи по экономике и финансам Электронной библиотеки ИД «Гребенников» http://www.nwapa.spb.ru/index.php?page_id=76
15. Статьи из журналов и статистических изданий Ист Вью http://www.nwapa.spb.ru/index.php?page_id=76
16. Сайт рейтинговой компании Standard&Poors [Электронный ресурс]. – Режим доступа: [http:// www.standartpoors.ru](http://www.standartpoors.ru)
17. Сайт рейтингового агентства «Эксперт РА» [Электронный ресурс]. – Режим доступа: [http:// www.raexpert.ru](http://www.raexpert.ru)
18. <http://vopreco.ru> - официальный сайт журнала «Вопросы экономики»
19. <http://www.bankdelo.ru> – официальный сайт журнала «Банковское дело»
20. <http://www.cbr.ru> – официальный сайт Центрального Банка РФ
21. <http://www.kommersant.ru> – официальный сайт издательского дома Коммерсант
22. <http://finans.rusba.ru> – официальный сайт журнала «Финансы и экономика»
23. <http://www.expert.ru> официальный сайт медиа-холдинга Эксперт
24. <http://www.cbr.ru> - Официальный сайт Центрального банка России
25. <http://http://www.rbc.ru/> - официальный сайт РосБизнесКонсалтинг

7.5. Иные источники

Не используются.

8. Материально-техническая база, информационные технологии, программное обеспечение и информационные справочные системы

Курс включает использование программного обеспечения Microsoft Excel, Microsoft Word, Microsoft Power Point для подготовки текстового и табличного материала, графических иллюстраций.

Методы обучения предполагают использование информационных технологий (компьютерное тестирование, демонстрация мультимедийных материалов).

Задействованы Интернет-сервисы и электронные ресурсы (поисковые системы, электронная почта, профессиональные тематические чаты и форумы, системы аудио и видео конференций, онлайн энциклопедии, справочники, библиотеки, электронные учебные и учебно-методические материалы).

Допускается применение системы дистанционного обучения.