

УТВЕРЖДАЮ

Первый проректор, проректор по
УР

_____ А.Е. Рудин

«21» ___ 02 ___ 2023 года

Рабочая программа дисциплины

Б1.В.ДВ.03.02 Обработка данных

Учебный план: 2023-2024 42.04.03 ВШПМ Изд страт в медиапростр ОО №2-1-127.plx

Кафедра: **21** Информационных и управляющих систем

Направление подготовки:
(специальность) 42.04.03 Издательское дело

Профиль подготовки: Издательские стратегии в медиапространстве
(специализация)

Уровень образования: магистратура

Форма обучения: очная

План учебного процесса

Семестр (курс для ЗАО)		Контактная работа обучающихс я	Сам. работа	Контроль, час.	Трудоё мкость, ЗЕТ	Форма промежуточной аттестации
		Лаб. занятия				
4	УП	36	71,75	0,25	3	Зачет
	РПД	36	71,75	0,25	3	
Итого	УП	36	71,75	0,25	3	
	РПД	36	71,75	0,25	3	

Рабочая программа дисциплины составлена в соответствии с федеральным государственным образовательным стандартом высшего образования по направлению подготовки 42.04.03 Издательское дело, утверждённым приказом Минобрнауки России от 08.06.2017 г. № 513

Составитель (и):

кандидат технических наук, Доцент

Шефер Е.А.

От кафедры составителя:

Заведующий кафедрой информационных и управляющих систем

Горина
Владимировна

Елена

От выпускающей кафедры:

Заведующий кафедрой

Лезунова Наталья
Борисовна

Методический отдел:

1 ВВЕДЕНИЕ К РАБОЧЕЙ ПРОГРАММЕ ДИСЦИПЛИНЫ

1.1 Цель дисциплины: Сформировать компетенции обучающегося в области сбора, обработки, поиска информации согласно поставленной задачи, подготовки исходных данных для анализа в ручном и автоматическом режимах, анализа полученных результатов.

1.2 Задачи дисциплины:

- научить осуществлять поиск исходных данных для решения поставленной профессиональной задачи;
- научить проводить первоначальную подготовку данных к анализу;
- познакомить с прикладным математическим инструментарием статистики;
- научить работать в составе многопрофильных команд, занимающихся созданием сложных анализом данных.

1.3 Требования к предварительной подготовке обучающегося:

Предварительная подготовка предполагает создание основы для формирования компетенций, указанных в п. 2, при изучении дисциплин:

- Цифровые технологии в издательском деле
- Стратегия и тактика управления издательством

2 КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ

ПК-4: Способен организовывать и поддерживать обратную связь с аудиторией и контрагентами в профессиональной среде
Знать: Термины, категории, методы и алгоритмы, относящиеся к восприятию, анализу, преобразованию, обобщению информации с использованием программных средств
Уметь: Осуществлять анализ, обобщение и оформление результатов проведенных исследований
Владеть: Навыками построения информационных моделей объектов, процессов, ситуаций с целью совершенствования маркетинговой политики издательства

3 РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ

Наименование и содержание разделов, тем и учебных занятий	Семестр (курс для ЗАО)	Контактная работа	СР (часы)	Инновац. формы занятий	Форма текущего контроля
		Лаб. (часы)			
Раздел 1. Интеллектуальный анализ данных – основные понятия.	4				С
Тема 1. Data Mining. Задачи Data Mining. Классификация задач Data Mining. Лабораторная работа 1. Решение задач классификации и регрессии. Лабораторная работа 2. Решение задач поиска ассоциативных правил и задач кластеризации.		8	10	АС	
Тема 2. Практическое применение Data Mining. Интернет-технологии. Торговля. Телекоммуникации. Промышленное производство. Медицина. Банковское дело. Страховой бизнес. Другие области применения. Лабораторная работа 3. Применение Data Mining в анализе статистических данных.		4	10		

Раздел 2. Виды анализа данных.				
Тема 3. Визуальный анализ данных — Visual Mining. Выполнение визуального анализа данных. Характеристики средств визуализации. Методы визуализации. Методы геометрических преобразований. Отображение иконок. Методы, ориентированные на пикселы. Иерархические образы. Лабораторная работа 4. Использование различных методов визуального анализа данных в решении задач обработки данных.	4	10		С
Тема 4. Анализ текстовой информации — Text Mining. Извлечение ключевых понятий из текста. Классификация текстовых документов. Методы кластеризации текстовых документов. Средства анализа текстовой информации. Лабораторная работа 5. Решение задач анализа текстов. Лабораторная работа 6. Решение задач аннотирования текстов.	8	10	АС	
Тема 5. Распределенный анализ данных. Системы мобильных агентов. Использование мобильных агентов для анализа данных. Лабораторная работа 7. Решение задач анализа распределенных данных.	4	10		
Раздел 3. Инструменты Data Mining.				
Тема 6. Идея Data Mining в реальном времени. Рекомендательные машины. Инструменты Data Mining в реальном времени. Лабораторная работа 8. Моделирование анализа данных в реальном времени.	4	10	АС	С

Тема 7. Автоматизация выполнения бизнес -процессов. Анализ процессов. Методы Process Mining. Библиотека алгоритмов Process Mining — ProM. Лабораторная работа 9. Решение задач анализа бизнес-процессов.	4	11,75		
Итого в семестре (на курсе для ЗАО)	36	71,75		
Консультации и промежуточная аттестация (Зачет)	0,25			
Всего контактная работа и СР по дисциплине	36,25	71,75		

4 КУРСОВОЕ ПРОЕКТИРОВАНИЕ

Курсовое проектирование учебным планом не предусмотрено

5. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

5.1 Описание показателей, критериев и системы оценивания результатов обучения

5.1.1 Показатели оценивания

Код компетенции	Показатели оценивания результатов обучения	Наименование оценочного средства
ПК-4	<p>Называет основные формы и методы сбора, обработки информации.</p> <p>Использует основные компьютерные программы, решающие задачи статистической обработки данных; осуществляет выбор и применяет методы статистической обработки при проведении в доступных формах научных исследований в сфере профессиональной деятельности.</p> <p>Использует методы и алгоритмы анализа и оценки процессов в профессиональной сфере.</p>	<p>Вопросы для устного собеседования.</p> <p>Практико-ориентированное задание.</p>

5.1.2 Система и критерии оценивания

Шкала оценивания	Критерии оценивания сформированности компетенций	
	Устное собеседование	Письменная работа
Зачтено	<p>Обучающийся твердо знает программный материал, грамотно и по существу излагает его, не допускает существенных неточностей в ответе на вопросы, способен правильно применить основные методы и инструменты при решении практических задач, владеет необходимыми навыками и приемами их выполнения.</p> <p>Учитываются баллы, накопленные в течение семестра.</p>	
Не зачтено	<p>Обучающийся не может изложить значительной части программного материала, допускает существенные ошибки, допускает неточности в формулировках и доказательствах, нарушения в последовательности изложения программного материала; неуверенно, с большими затруднениями выполняет практические задания.</p> <p>Не учитываются баллы, накопленные в течение семестра.</p>	

5.2 Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности

5.2.1 Перечень контрольных вопросов

№ п/п	Формулировки вопросов
Семестр 4	
1	Приведите наиболее распространенные формы представления исходных данных для многомерного статистического анализа. Приведите примеры использования различных форм.
2	Особенности работы с многомерным признаковым пространством.
3	Виды зависимостей, исследуемых в многомерном статистическом анализе.
4	Основные этапы многомерного статистического анализа.
5	Числовые характеристики закона распределения многомерной случайной величины.
6	Основные показатели тесноты связи между категоризированными переменными. Коэффициент Крамера.
7	Основные показатели тесноты связи между категоризированными переменными. Информационная мера связи.
8	Проверка взаимной независимости переменных. Критерий Мантеля-Хензеля.
9	Описать суть агломеративных методов многомерного анализа. Привести примеры сферы их применения.
10	Приведите обобщенную характеристику методов кластерного анализа.
11	Охарактеризуйте задачи, решаемые с помощью метода главных компонент.
12	Как интерпретируются результаты компонентного анализа?

13	Опишите основные методики факторного анализа.
14	Опишите алгоритм выделения матрицы факторных нагрузок.
15	Выполнение анализа данных в реальном времени.

5.2.2 Типовые тестовые задания

Не предусмотрены.

5.2.3 Типовые практико-ориентированные задания (задачи, кейсы)

1. Исследовать линейную зависимость между количественными переменными с помощью автоматической пошаговой регрессии в пакете STATGRAPHICS.

- Выбрать окончательную модель методом селекции вперед.
- Выполнить дисперсионный анализ выбранной модели регрессии.
- Записать найденное уравнение регрессии.
- Сделать вывод о его адекватности и информационной способности.

2. Построить дерево решений и продукционные правила для прогнозирования результатов в Deductor. Указать вклад признаков в разделение множеств объектов.

С помощью визуализатора «Правила» представить список правил для всех листьев конечного дерева. С помощью визуализатора «Значимость атрибутов» оценить значимость признаков. С помощью визуализатора «Таблица сопряженности» оценить качество классификации данных.

3. Выполнить поиск ассоциативных правил на примере данных о продажах кондитерских товаров с целью применения результатов для стимулирования продаж. Использовать визуализаторы «Популярные наборы», «Правила», «Дерево правил», «Что если». Произвести анализ предпочтений.

5.3 Методические материалы, определяющие процедуры оценивания знаний, умений, владений (навыков и (или) практического опыта деятельности)

5.3.1 Условия допуска обучающегося к промежуточной аттестации и порядок ликвидации академической задолженности

Проведение промежуточной аттестации регламентировано локальным нормативным актом СПбГУПТД «Положение о проведении текущего контроля успеваемости и промежуточной аттестации обучающихся»

5.3.2 Форма проведения промежуточной аттестации по дисциплине

Устная Письменная Компьютерное тестирование Иная

5.3.3 Особенности проведения промежуточной аттестации по дисциплине

При проведении зачета время, отводимое на подготовку к ответу, составляет не более 40 мин. Для выполнения практического задания обучающему предоставляется необходимая справочная информация, файлы исходных текстов и рисунков.

Сообщение результатов обучающемуся производится сразу после проверки выполнения практической задачи и принятия устного ответа.

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

6.1 Учебная литература

Автор	Заглавие	Издательство	Год издания	Ссылка
6.1.1 Основная учебная литература				
Мельниченко, А. С.	Математическая статистика и анализ данных	Москва: Издательский Дом МИСиС	2018	http://www.iprbookshop.ru/78563.html
Шахова, О. А.	Статистическая обработка результатов исследований	Тюмень: Издательство «Титул»	2022	https://www.iprbookshop.ru/119099.html
Замятин, А. В.	Интеллектуальный анализ данных	Томск: Издательский Дом Томского государственного университета	2020	https://www.iprbookshop.ru/116889.html
6.1.2 Дополнительная учебная литература				
Пальмов, С. В.	Интеллектуальный анализ данных	Самара: Поволжский государственный университет телекоммуникаций и информатики	2017	http://www.iprbookshop.ru/75376.html
Пименов В. И., Пименов И. В.	Интеллектуальный анализ данных	СПб.: СПбГУПТД	2017	http://publish.sutd.ru/tp_ext_inf_publish.php?id=201748

Истомина, А. П.	Анализ качественных исследований данных	Ставрополь: Северо-Кавказский федеральный университет	2018	http://www.iprbookshop.ru/92674.html
Жуковский, О. И.	Информационные технологии и анализ данных	Томск: Томский государственный университет систем управления и радиоэлектроники, Эль Контент	2014	http://www.iprbookshop.ru/72106.html

6.2 Перечень профессиональных баз данных и информационно-справочных систем

Портал Росстандарта по стандартизации [Электронный ресурс]. URL: <http://standard.gost.ru/wps/portal/>
 Электронно-библиотечная система IPRbooks [Электронный ресурс]. URL: <http://www.iprbookshop.ru/>

6.3 Перечень лицензионного и свободно распространяемого программного обеспечения

MicrosoftOfficeProfessional
 Microsoft Windows
 Deductor Academic

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

Аудитория	Оснащение
Лекционная аудитория	Мультимедийное оборудование, специализированная мебель, доска
Компьютерный класс	Мультимедийное оборудование, компьютерная техника с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду