

Утверждено приказом ГОУ ВПО ДонГУУ от 23.08.2016г. №675

ДОНЕЦКАЯ НАРОДНАЯ РЕСПУБЛИКА
МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«ДОНЕЦКАЯ АКАДЕМИЯ УПРАВЛЕНИЯ И ГОСУДАРСТВЕННОЙ СЛУЖБЫ
ПРИ ГЛАВЕ ДОНЕЦКОЙ НАРОДНОЙ РЕСПУБЛИКИ»

ФАКУЛЬТЕТ ГОУДАРСТВЕННОЙ СЛУЖБЫ И УПРАВЛЕНИЯ
КАФЕДРА ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

УТВЕРЖДАЮ

Проректор по учебной работе

Л.Н.Костина


20.08.2016 г.

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

«Современные технологии анализа информации»

Направление подготовки

09.04.03 «Прикладная информатика»

Донецк

2017

Рабочая программа учебной дисциплины «Современные технологии анализе информации» для студентов 1 курса образовательного уровня «магистр» направления подготовки 09.04.03 «Прикладная информатика» очной формы обучения

Автор(ы),
разработчик(и): доцент, к. ф.-м. н., доцент., А.Н. Верзилов

Программа рассмотрена на
заседании ПМК кафедры

«Прикладная информатика»

Протокол заседания ПМК от

08.06.2017

№ 10

дата

Председатель ПМК



А.Н.Верзилов

Программа рассмотрена на
заседании кафедры

Информационных технологий

Протокол заседания кафедры от

09.06.2017

№ 13

дата

Заведующий кафедрой



Н.В.Брадуд

1. Цель освоения дисциплины и планируемые результаты обучения по дисциплине (соотнесенные с планируемыми результатами освоения образовательной программы)

Цель изучения дисциплины – дополнение и совершенствование знаний в области методов многомерного статистического анализа социально-экономических процессов, изучение современных проблем, направлений и методов на базе современных технологий и систем статистического анализа с использованием ПЭВМ.

Задачи учебной дисциплины:

- ознакомить студентов с методикой и современными методами анализа данных;
- сформировать у слушателей базу знаний о современных направлениях и возможностях методов многомерной статистики, тенденциях их развития, современных вычислительных технологиях и инструментах многомерного экономико-статистического анализа в социальной среде и экономической жизни общества;
- научить обобщать и анализировать информацию;
- дать навыки применения различных методов анализа данных.

Компетенции обучающегося, формируемые в результате освоения дисциплины:

Код соответствующей компетенции по ФГОС	Наименование компетенций	Результат освоения (знать, уметь, владеть)
ОК -1	- способность к абстрактному мышлению, анализу, синтезу	Знать: - способы приобретения и использования в практической деятельности новых знаний и умений Уметь: - приобретать и использовать в практической деятельности новые знания и умения, стремиться к саморазвитию Владеть: - методами приобретения и использования в практической деятельности новых знаний и умений
ОПК - 4	- способность исследовать закономерности становления и развития информационного общества в конкретной прикладной области	Знать: - теоретические основы исследования проблем развития информационного общества в области управления социально-экономическими процессами Уметь: - анализировать и применять традиционные и новые методы исследования информационного потенциала общества при решении конкретных прикладных задач Владеть: - навыками подбора, анализа и обработки информации, необходимой для выполнения конкретных расчетов, подготовки исходных данных, расчета и анализа специфических для сфер деятельности показателей на основе типовых методов и методик
ОПК - 5	- способность на практике применять новые научные принципы и методы исследований	Знать: - современные методы исследования и информационно – коммуникационные технологии; - основные технологии анализа данных Уметь:

		- применять современные математико-статистические методы и инструментальные средства для исследования объектов профессиональной деятельности Владеть: - методами статистического анализа и прогнозирования
ПК - 8	- способность анализировать данные и оценивать требуемые знания для решения нестандартных задач с использованием математических методов и методов компьютерного моделирования	Знать: - современные методы исследования и информационно – коммуникационные технологии; - основные технологии анализа данных Уметь: - использовать математические, инструментальные средства для обработки, анализа и систематизации информации по теме исследования Владеть: - навыками проведения факторного и кластерного анализа

2. Место дисциплины в структуре основной образовательной программы

Дисциплина относится к базовой части (М.2.Б) профессионального цикла (М.2)

2.1. Требования к предварительной подготовке обучающегося

Дисциплины, которые необходимо изучить, прежде чем приступить к освоению настоящей дисциплины: линейная алгебра, теория вероятности и математическая статистика, эконометрика, экономика предприятия, исследование операций и методы оптимизации.

2.2. Дисциплины и/или практики, для которых освоение данной дисциплины необходимо как предшествующее:

Дисциплины, для которых данная дисциплина является предшествующей: математическое и компьютерное моделирование, маркетинговые исследования в области информационных систем, информационные хранилища, технологии облачных вычислений, научно – исследовательская практика

3. Объем дисциплины в кредитах (зачетных единицах) с указанием количества академических часов, выделенных на аудиторную (по видам учебных занятий) и самостоятельную работу студента

	Зачетные единицы (кредиты ECTS)	Всего часов		Форма обучения (вносятся данные по реализуемым формам)	
		О	З	Очная	Заочная
				Семестр №1	Семестр №2
Общая трудоемкость	4,0	144	X	Количество часов на вид работы:	
Виды учебной работы, из них:					
Аудиторные занятия (всего)				72	X
В том числе:					
Лекции				36	X
Практические занятия				36	X
Самостоятельная работа (всего)				72	X
Промежуточная аттестация					
В том числе:					
зачет /экзамен				экзамен	

4. Содержание дисциплины, структурированное по разделам (темам) с указанием отведенного на них количества академических часов и видов учебных занятий

4.1. Разделы (темы) дисциплины с указанием отведенного на них количества академических часов и видов учебных занятий

Наименование раздела, темы дисциплины	Виды учебной работы (бюджет времени) (вносятся данные по реализуемым формам)									
	Очная форма обучения					Заочная форма обучения				
	Лекции	Практические занятия	Семинарские занятия	Самостоятельная работа	Всего	Лекции	Практические занятия	Семинарские занятия	Самостоятельная работа	Всего
1	2	3	4	5	6	7	8	9	10	11
Раздел 1. Технологии многомерного статистического анализа информации										
Тема 1.1. Введение в предмет. Задачи СППР	2	2		4	8					
Тема 1.2. Организация хранилища данных	2	2		4	8					
Тема 1.3. Многомерная модель данных	2	2		8	12					
Тема 1.4. Архитектура OLAP - систем	2	2		4	8					
Итого по разделу:	8	8		20	36					
Раздел 2. Современные технологии интеллектуального анализа данных										
Тема 2.1. Задачи и модели Data Mining	4	4		6	14					
Тема 2.2. Методы и средства Data Mining	4	4		6	14					
Тема 2.3. Визуальный анализ данных Visual Mining	2	2		4	8					
Итого по разделу:	10	10		16	36					
Раздел 3. Решение основных задач анализа информации										
Тема 3.1. Классификация и регрессия	8	8		16	32					
Тема 3.2. Поиск ассоциативных правил	4	4		8	16					
Тема 3.3. Кластеризация	6	6		12	24					

Наименование раздела, темы дисциплины	Виды учебной работы (бюджет времени) (вносятся данные по реализуемым формам)									
	Очная форма обучения					Заочная форма обучения				
	Лекции	Практические занятия	Семинарские занятия	Самостоятельная работа	Всего	Лекции	Практические занятия	Семинарские занятия	Самостоятельная работа	Всего
1	2	3	4	5	6	7	8	9	10	11
Итого по разделу:	18	18		36	72					
Всего за семестр:	36	36		72	144					

4.2. Содержание разделов дисциплины:

Наименование раздела, темы дисциплины	Содержание разделов дисциплины	Содержание семинарских/практических занятий		
			Кол-во часов	
			0	3
1	2	3	4	5
Раздел 1. Технологии многомерного статистического анализа информации				
Тема 1.1. Введение в предмет. Задачи СППР.	Задачи систем поддержки принятия решений. Основная задача СППР. Классификация задач анализа информации. Использование OLTP-систем и их неэффективность	Практическое занятие № 1. Версии системы «Statistica», общая технология статистического анализа в системе «Statistica»	2	
Тема 1.2. Организация хранилища данных	Концепция хранилища данных. Организация хранилища данных. Проблемы очистки данных. Методы очистки данных	Практическое занятие № 1. Направления современного многомерного статистического анализа данных в системе «Statistica»	2	
Тема 1.3. Многомерная модель данных	Определение OLAP-систем. Концептуальное многомерное представление. Правила Кодда. Тест FASMI	Практическое занятие № 1. Современный многомерный линейный регрессионный анализ в системе «Statistica».	2	
Тема 1.4. Архитектура OLAP - систем	Основные компоненты OLAP-системы. Преимущества использования многомерных БД. Реляционный OLAP. Гибридный OLAP/	Практическое занятие № 1. Современный многомерный нелинейный регрессионный анализ в системе «Statistica».	2	
Раздел 2. Современные технологии интеллектуального анализа данных				
Тема 2.1. Задачи и модели Data Mining	Классификация задач Data Mining. Практическое применение Data Mining.	Практическое занятие № 1. Архитектура хранилища данных в	2	

Наименование раздела, темы дисциплины	Содержание разделов дисциплины	Содержание семинарских/практических занятий		
			Кол-во часов	
			0	3
1	2	3	4	5
	Предсказательные модели. Описательные модели	Deductor Warehouse 2. Создание хранилища данных в Deductor	2	
Тема 2.2. Методы и средства Data Mining	Базовые методы Data Mining. Нечеткая логика. Генетические алгоритмы. Нейронные сети. Процесс обнаружения знаний. Управление знаниями. Программные реализации методов Data Mining	Практическое занятие № 1.Data Mining в задачах прогнозирования	2	
		2.Создание сценария прогнозирования объема продаж	2	
Тема 2.3. Визуальный анализ данных Visual Mining	Выполнение визуального анализа данных. Характеристики средств визуализации данных. Метод геометрических преобразований. Иерархические образы	Практическое занятие № 1.Трансформация данных в Deductor Studio	2	
Раздел 3. Решение основных задач анализа информации				
Тема 3.1. Классификация и регрессия	Постановка задачи. Способы представления результатов. Алгоритм построения 1-правил. Метод Naïve Bayes. Методика «разделяй и властвуй». Алгоритм ID3. Алгоритм C4.5. Алгоритм покрытия	Практическое занятие № 1. Задачи сегментации и классификации	2	
		2. Классификация с помощью деревьев решений	2	
		3. Логистическая регрессия	2	
		4. Прогнозирование	2	
Тема 3.2. Поиск ассоциативных правил	Постановка задачи. Способы представления результатов. Алгоритм Apriori. Разновидности алгоритма Apriori.	Практическое занятие № 1. Генерация и интерпретация ассоциативных правил	2	
		2. Визуализатор «Что-если» в ассоциативных правилах.	2	
Тема 3.3. Кластеризация	Постановка задачи и базовые алгоритмы кластеризации. Адаптивные методы кластеризации. Выбор наилучшего решения. Использование формальных критериев. Примеры адаптивной кластеризации	Практическое занятие № 1. Исследование методов кластерного анализа при интеллектуальной обработке данных	2	
		2. Классификация объектов по параметрам в Deductor Warehouse	2	
		3. Кластеризация с помощью самоорганизующейся карты Кохонена	2	

5. Перечень учебной литературы, необходимой для освоения дисциплины

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

Элементы учебно-методического комплекса дисциплины утверждены на заседании кафедры информационных технологий (протокол №1 от 29.08.2017).

Контрольные вопросы для самоподготовки

1. Системы обработки статистических данных. Основные характеристики и возможности.
2. Типы документов в системе «Statistica». Основные направления и терминология статистического анализа данных в системе «Statistica»
3. Показатели многомерного регрессионного и корреляционного анализа в системе «Statistica»
4. Особенности и виды многомерных статистических графиков в в системе «Statistica»
5. Показатели робастности первичных данных в системе «Statistica»
6. Основные процедуры факторного и компонентного анализа в системе «Statistica»
7. Процедуры выбора общностей в факторном анализе в системе «Statistica». Анализ на главных осях.
8. Назначение современного анализа канонических корреляций. Возможности анализа канонических корреляций в системе «Statistica».
9. Теоретические основы метода канонических корреляций. Основные статистические характеристики метода в системе «Statistica».
10. Использование аппарата канонических корреляций для исследования множественных регрессий.
11. Теоретические основы и реализация метода многомерного метрического шкалирования в системе «Statistica».
12. Оценка близости объектов, экспертная оценка и построение матрицы расстояний методом метрического шкалирования в системе «Statistica».
13. Реализация метода многомерного шкалирования в системе «Statistica».
14. Интерпретация результатов процедуры многомерного метрического шкалирования совокупности объектов в системе «Statistica».
15. Определение базовой и вероятностной нейронной сети в системе «Statistica».
16. Построение обобщенной регрессионной нейронной сети в системе «Statistica».
17. Примеры применения нейросетевого моделирования в системе «Statistica».
18. Методы, альтернативные нейросетевому моделированию в системе «Statistica».
19. Многомерный дисперсионный анализ регрессий и корреляций в системе «Statistica».
20. Моделирование многомерных распределений в системе Statistica.
21. Понятие искусственного интеллекта.
22. Предыстория искусственного интеллекта. Появление термина искусственного интеллекта и его определение.
23. Направления искусственного интеллекта, сущность и основные идеи.
24. Транспьютерная технология, где и когда она зародилась и применялась.
25. Предыстория искусственного интеллекта. Перечислить 3 подхода к созданию нейросетей.
26. Понятие искусственного интеллекта. Когда зародился искусственный интеллект в России. Принцип направления кибернетики «черного ящика» и его основные идеи.
27. Появление термина искусственного интеллекта и его определение. Когда зародился искусственный интеллект в России.
28. Понятие модели лабиринтного поиска, ее появление, примеры.
29. Предыстория искусственного интеллекта. Методы математической логики, перечислите их и привести примеры.

30. Нейрокибернетика и ее основная идея. Устройство персептрон и его применение.
31. Когда появились первые нейросети. В каком направлении искусственного интеллекта они используются.
32. Фреймы и формализованные модели фреймов. Структура фрейма в двух форматах. Способы получения слотом значений во фрейме-экземпляре.
33. Свойство теории фреймов. Формальные логические модели.
34. Основные функции и правила срабатывания в машине вывода. Правило *modus ponens*. Правила и функции в управляющем компоненте машины вывода.
35. Схема функционирования интерпретатора. Системы и методы стратегии управления выводом. Понятие лингвистической переменной
36. Классификаторы по критерию минимума расстояния. Привести примеры.
37. Меры сходства и критерии кластеризации.

5.2. Перечень основной учебной литературы

1. Кричевский М.Л. Интеллектуальный анализ данных в менеджменте: Учебное пособие. - СПб.: ГУАП, 2005. - 208 с. Электронный ресурс: <http://window.edu.ru/resource/889/44889>
2. Филимонов В.А. Интеллектуальные системы и экспертный анализ: Учебное пособие.- Омск: Изд-во ОмГУ, 2002.- 32 с. Электронный ресурс: <http://window.edu.ru/resource/814/80814>
3. Технологии извлечения, обработки и анализа информации: Учебно-методический комплекс. - СПб.: Изд-во РГПУ им. А.И. Герцена, 2008. - 246 с. Электронный ресурс: <http://window.edu.ru/resource/268/64268>
4. Анализ данных и процессов: учеб. пособие / А. А. Барсегян, М. С. Куприянов, И. И. Холод, М. Д. Тесс, С. И. Елизаров. — 3-е изд., перераб. и доп. — СПб.: БХВ-Петербург, 2009. — 512 с.: ил.
5. Боровиков, Владимир Павлович. Прогнозирование в системе Statistica в среде Windows : Основы теории и интенсивная практика на компьютере : учеб. пособие для высш. учеб. заведений по специальности "Прикладная математика" / В. П. Боровиков, Г. И. Ивченко .- 2-е изд., перераб. и доп. - М. : Финансы и статистика, 2006 .- 368 с.

5.3. Перечень дополнительной литературы

1. Романов В.П. Интеллектуальные информационные системы в экономике: Учебное пособие. -. «Экзамен», 2003. – 496 с.
2. Гаврилова Т.А., Хорошевский, С.В. Базы знаний интеллектуальных систем: учебное пособие. – СПб.: Питер,2006. -382 с.
3. Гаскаров Д.В. Интеллектуальные информационные системы: Учебник для вузов. - М.: ВШ, 2005. – 432
4. Паклин Н.Б., Орешков, В.И. Бизнес аналитика: от данных к знаниям: Учеб. пособие .2-е изд. – СПб.: Питер, 2010. – 704 с. , прил. CD
5. Тельнов Ю.Ф. Интеллектуальные информационные системы в экономике. Учебное пособие. -М.: СИНТЕГ. 2001. – 316 с.

6. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

1. Elibrary.ru [Электронный ресурс]: электронная библиотечная система: база данных

- содержит сведения об отечественных книгах и периодических изданиях по науке, технологии, медицине и образованию / Рос. информ. портал. – Москва, 2000– . – Режим доступа: <http://elibrary.ru>.
2. Издательство "Лань" [Электронный ресурс]: электронная библиотечная система: содержит электронные версии книг издательства «Лань» и других ведущих издательств учебной литературы, так и электронные версии периодических изданий по естественным, техническим и гуманитарным наукам. – Москва, 2010– . Режим доступа: <http://e.lanbook.com>.
 3. Электронный архив УГЛТУ [Электронный ресурс]: содержит электронные версии научных, учебных и учебно-методических разработок авторов - ученых УГЛТУ. Режим доступа: <http://elar.usfeu.ru>.
 4. Znanium.com [Электронный ресурс]: электронная библиотечная система : содержит электронные версии книг издательства Инфра-М и других ведущих издательств учебной литературы, так и электронные версии периодических изданий по естественным, техническим и гуманитарным наукам. Режим доступа: <http://znanium.com>.
 5. Российская государственная библиотека [Электронный ресурс]: содержит электронные версии книг, учебников, монографий, сборников научных трудов как отечественных, так и зарубежных авторов, периодических изданий. Режим доступа: [http:// www.rbc.ru](http://www.rbc.ru).
 6. Портал искусственного интеллекта. <http://www.aiportal.ru/>

7. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень программного обеспечения и информационных справочных систем

7.1. Перечень информационных технологий (при необходимости)

При выполнении индивидуальных работ, в ходе практических занятий студентами используются:

- Автоматизированные системы поиска информации;
- Технологии интеллектуального анализа информации Data Mining; Технология распределенной обработки данных.

7.2. Перечень программного обеспечения (при необходимости)

- Операционная система MS Windows;
- Офисный пакет MS Office 2003 или выше;
- Система STATISTICA v10;
- Аналитическая платформа Deductor Academic v 5.3.

7.3. Перечень информационных справочных систем (при необходимости)

Автоматизированная информационно-библиотечная система, предназначенная для осуществления доступа читателей к электронному каталогу Научной библиотеки ДонГУУ поиска, просмотра, заказа и бронирования документов.

Зарубежные электронные научные информационные ресурсы: European Library. Свободный доступ к ресурсам 47 Национальных библиотек Европы, Австралии. Национальная библиотека.

8. Фонд оценочных средств для контроля уровня сформированности компетенций

8.1. Виды промежуточной аттестации

Текущий контроль успеваемости позволяет оценить уровень сформированности элементов компетенций (знаний и умений), компетенций с последующим объединением оценок и проводится в форме устного опроса (фронтальный, индивидуальный, комплексный), письменной проверки (ответы на вопросы, тестовые задания), включая задания для самостоятельной работы.

Промежуточная аттестация в форме экзамена позволяет оценить уровень сформированности компетенций в целом по дисциплине и может осуществляться как в письменной так и в устной форме.

8.2. Показатели и критерии оценки результатов освоения дисциплины

Средним баллом за дисциплину является средний балл за текущую учебную деятельность.

Механизм конвертации результатов изучения студентом дисциплины в оценки по государственной шкале и шкале ECTS представлен в таблице.

Средний балл по дисциплине (текущая успеваемость)	Отношение полученного студентом среднего балла по дисциплине к максимально возможной величине этого показателя	Оценка по государственной шкале	Оценка по шкале ECTS	Определение
4,5 – 5,0	90% – 100%	5	A	отлично – отличное выполнение с незначительным количеством неточностей (до 10%)
4,0 – 4,49	80% – 89%	4	B	хорошо – в целом правильно выполненная работа с незначительным количеством ошибок (до 20%)
3,75 – 3,99	75% – 79%	4	C	хорошо – в целом правильно выполненная работа с незначительным количеством ошибок (до 25%)
3,25 – 3,74	65% – 74%	3	D	удовлетворительно – неплохо, но со значительным количеством недостатков (до 35%)
3,0 – 3,24	60% – 64%	3	E	достаточно – выполнение удовлетворяет минимальные критерии, но со

				значительным количеством недостатков (до 40%)
до 3,0	35% – 59%	2	FX	неудовлетворительно с возможностью повторной сдачи (ошибок свыше 40%)
	0 – 34%	2	F	неудовлетворительно – надо поработать над тем, как получить положительную оценку (ошибок свыше 65%)

8.3. Критерии оценки работы студента

При усвоении каждой темы за текущую учебную деятельность студента выставляются оценки по 5-балльной (государственной) шкале. Оценка за каждое задание в процессе текущей учебной деятельности определяется на основе процентного отношения операций, правильно выполненных студентом во время выполнения задания:

- 90-100% – «5»,
- 75-89% – «4»,
- 60-74% – «3»,
- менее 60% – «2».

Если на занятии студент выполняет несколько заданий, оценка за каждое задание выставляется отдельно.

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

Индивидуальные задания

Раздел 1. Технологии многомерного статистического анализа информации

Индивидуальное задание 1.

По 20 предприятиям отрасли изучается зависимость выработки продукции на 1 работника (y), тыс.руб. - "ВЫРАБОТКА" от ввода в действие новых основных фондов в % от стоимости фондов на конец года (x_1) - "ФОНДЫ" и от удельного веса рабочих высокой квалификации в общей численности рабочих (x_2), % - "РАБОЧИЕ". Данные записаны в файле пакета STATISTICA и представлены на рис.1.

	1 ВЫРАБОТКА	2 ФОНДЫ	3 РАБОЧИЕ
1	7	3,9	10
2	7	3,9	14
3	7	3,7	15
4	7	4,0	16
5	7	3,8	17
6	7	4,8	19
7	8	5,4	19
8	8	4,4	20
9	8	5,3	20
10	10	6,8	20
11	9	6,0	21
12	11	6,4	22
13	9	6,8	22
14	11	7,2	25
15	12	8,0	28
16	12	8,2	29
17	12	8,1	30
18	12	8,5	31
19	14	9,6	32
20	14	9,0	36

Рисунок 1

ЗАДАНИЕ

1. Получить дескриптивные статистики по каждому признаку. Оценить показатели вариации каждого признака и сделать вывод о возможностях применения метода наименьших квадратов для их изучения.
2. Составить уравнение множественной регрессии, оценить его параметры, пояснить их экономический смысл.
3. Рассчитать частные коэффициенты эластичности и дать на их основе сравнительную оценку силы влияния факторов на результат.
4. Проанализировать линейные коэффициенты парной и частной корреляции.
5. Оценить значения скорректированного и нескорректированного линейных коэффициентов множественной корреляции.
6. С помощью F-критерия Фишера оценить статистическую надежность уравнения регрессии в целом.

Раздел 3. Решение основных задач анализа информации

Индивидуальная работа 2. Сегментация клиентов телекоммуникационной компании

Описание бизнес-задачи

В такой высокотехнологической отрасли, как телекоммуникации, методы и подходы Data Mining получили широкое применение. Решаемые задачи прежде всего связаны с программами лояльности и удержанием существующей клиентской базы, а также с привлечением новых потребителей услуг.

В биллинговых системах телекоммуникационных компаний накапливаются большие объемы данных. В первую очередь это информация об абонентах и статистика использования услуг. Анализ такой информации ручными и полуручными методами малоэффективен.

Постановка задачи.

Руководство филиала региональной телекоммуникационной компании, предоставляющей услуги мобильной связи, поставило задачу сегментации абонентской базы. Ее целями являются:

- построение профилей абонентов путем выявления их схожего поведения в плане частоты, длительности и времени звонков, а также ежемесячных расходов;
- оценка наиболее и наименее доходных сегментов.

Эта информация может в дальнейшем использоваться для:

- разработки маркетинговых акций, направленных на определённые группы клиентов;
- разработки новых тарифных планов;
- оптимизации расходов на адресную SMS-рассылку о новых услугах и тарифах;
- предотвращения оттока клиентов в другие компании.

Данные за последние несколько месяцев, взятые из биллинговой системы представляют собой таблицу со следующими полями (табл. 1)

Таблица 1. Данные по абонентам из биллинговой системы

№	Поле	Описание	Тип
1	Возраст	Возраст клиента	Целый
2	Среднемесячный расход	Сколько в среднем в месяц тратит абонент на мобильную связь	Вещественный
3	Средняя продолжительность разговора	Сколько в среднем минут на исходящие звонки тратит абонент за месяц	Вещественный
4	Звонков днем за месяц	Количество исходящих звонков в утреннее и дневное время	Целый
5	Звонков вечером за месяц	Количество исходящих звонков в вечернее время	Целый
6	Звонков ночью за месяц	Количество исходящих звонков в ночное время	Целый
7	Звонки в другие города	Количество исходящих звонков в другие города	Целый
8	Звонки в другие страны	Число исходящих международных звонков	Целый
9	Доля звонков на стационарные телефоны	-	Вещественный
10	Количество SMS	Число исходящих SMS-сообщений в месяц	Целый

Исходные данные.

Были отобраны только активные абоненты, которые регулярно пользовались услугами сотовой связи в течение последних нескольких месяцев. Данные находятся в файле mobile.txt, который можно взять в системе MOODLE либо у преподавателя.

Контрольные вопросы для проведения текущего контроля и промежуточной аттестации:

1. Прикладная статистика. Основные понятия.
2. Формы записи исходных статистических данных
3. Центральные проблемы прикладной статистики и этапы статистического анализа
4. Системы обработки статистических данных. Основные характеристики и возможности.

5. Типы документов в системе «Statistica». Основные направления и терминология статистического анализа данных в системе «Statistica»
6. Показатели многомерного регрессионного и корреляционного анализа в системе «Statistica»
7. Особенности и виды многомерных статистических графиков в в системе «Statistica»
8. 5.Показатели робастности первичных данных в системе «Statistica»
9. Основные процедуры факторного и компонентного анализа в системе «Statistica»
10. Процедуры выбора общностей в факторном анализе в системе «Statistica». Анализ на главных осях.
11. Назначение современного анализа канонических корреляций. Возможности анализа канонических корреляций в системе «Statistica».
12. Теоретические основы метода канонических корреляций. Основные статистические характеристики метода в системе «Statistica».
13. Использование аппарата канонических корреляций для исследования множественных регрессий.
14. Теоретические основы и реализация метода многомерного метрического шкалирования в системе «Statistica».
15. Оценка близости объектов, экспертная оценка и построение матрицы расстояний методом метрического шкалирования в системе «Statistica».
16. Реализация метода многомерного шкалирования в системе «Statistica».
17. Интерпретация результатов процедуры многомерного метрического шкалирования совокупности объектов в системе «Statistica».
18. Определение базовой и вероятностной нейронной сети в системе «Statistica».
19. Построение обобщенной регрессионной нейронной сети в системе «Statistica».
20. Примеры применения нейросетевого моделирования в системе «Statistica».
21. Методы, альтернативные нейросетевому моделированию в системе «Statistica».
22. Многомерный дисперсионный анализ регрессий и корреляций в системе «Statistica».
23. Моделирование многомерных распределений в системе Statistica.
24. Понятие искусственного интеллекта.
25. Предыстория искусственного интеллекта. Появление термина искусственного интеллекта и его определение.
26. Направления искусственного интеллекта, сущность и основные идеи.
27. Транспьютерная технология, где и когда она зародилась и применялась.
28. Предыстория искусственного интеллекта. Перечислить 3 подхода к созданию нейросетей.
29. Понятие искусственного интеллекта. Когда зародился искусственный интеллект в России. Принцип направления кибернетики «черного ящика» и его основные идеи.
30. Появление термина искусственного интеллекта и его определение. Когда зародился искусственный интеллект в России.
31. Понятие модели лабиринтного поиска, ее появление, примеры.
32. Предыстория искусственного интеллекта. Методы математической логики, перечислите их и привести примеры.
33. Нейрокибернетика и ее основная идея. Устройство персептрон и его применение.
34. Когда появились первые нейросети. В каком направлении искусственного интеллекта они используются.
35. Фреймы и формализованные модели фреймов. Структура фрейма в двух форматах. Способы получения слотом значений во фрейме-экземпляре.
36. Свойство теории фреймов. Формальные логические модели.
37. Основные функции и правила срабатывания в машине вывода. Правило *modus ponens*. Правила и функции в управляющем компоненте машины вывода.
38. Схема функционирования интерпретатора. Системы и методы стратегии управления выводом. Понятие лингвистической переменной

39. Классификаторы по критерию минимума расстояния. Привести примеры.
40. Меры сходства и критерии кластеризации.
41. Алгоритм максиминного расстояния.
42. Алгоритм К - внутригрупповых средних (метод К - средних).
43. Алгоритм К - внутригрупповых средних с адаптацией метрикой.
44. Классификация объектов как задача теории статистических решений.
45. Классификация с минимальным уровнем ошибки.
46. Разделяющие функции для случая нормальной плотности.
47. Оценка параметров и обучение с учителем.
48. Алгоритмы обучения классификаторов, основанные на аппроксимации плотностей распределения.
49. Алгоритмы обучения классификаторов, основанные на аппроксимации апостериорных вероятностей.
50. Сущность проблемы снижения размерности. Принципы организации методов снижения размерности.
51. Метод главных компонент.
52. Многомерное шкалирование.
53. Задачи и методы выбора информативной системы признаков.

Критерии оценивания компетенций (результатов) по уровням освоения учебного материала:

1 – репродуктивный (освоение знаний, выполнение деятельности по образцу, инструкции или под руководством), если самостоятельно (или с помощью преподавателя) выполнены все пункты работы;

2 – продуктивный (планирование и самостоятельное выполнение деятельности, решение проблемных задач; применение умений в новых условиях), если выполнены все пункты работы самостоятельно и улучшена точность результата;

3 – творческий (самостоятельное проектирование экспериментальной деятельности; оценка и самооценка инновационной деятельности), если предложен более рациональный алгоритм решения задачи.

8.3.2. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и/или опыта деятельности

– оценивание проводится преподавателем в течении всего учебного процесса на основе выполнения текущих контрольных и индивидуальных заданий, самостоятельной работы за компьютером;

– результаты выполнения практических работ предъявляются в виде отчетов оформленных в тетради;

– оценивание практических работ осуществляет преподаватель, который проводит семинарские занятия.

– экзамен принимает преподаватель, который проводит лекционные занятия.

9. Методические рекомендации (указания) для обучающихся по освоению дисциплины

Рекомендации, позволяющие обучающимся оптимальным образом организовать процесс изучения как теоретического учебного материала дисциплины, так и подготовки к практическим занятиям: изучение лекций, коллективное обсуждение тем на практических занятиях, индивидуальная работа за компьютером, самостоятельная работа над текущими темами, самостоятельная работа над индивидуальными заданиями.

По работе студент должен:

1. разобрать метод решения поставленной задачи и имеющиеся указания к её выполнению;
2. построить алгоритм для реализации задачи на ЭВМ;

3. написать программу по этому алгоритму;
4. отладить программу;
5. просчитать тестовый пример;
6. получить численные результаты для поставленной задачи;
7. убедиться в достоверности полученных результатов;
8. отчитаться перед преподавателем по теоретической и практической части индивидуальной работы.

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

Реализация данной учебной дисциплины осуществляется с использованием материально-технической базы, обеспечивающей проведение всех видов учебных занятий и научно-исследовательской работы обучающихся, предусмотренных программой учебной дисциплины и соответствующей действующим санитарным и противопожарным правилам, и нормам, в соответствующим образом оборудованных аудиториях. Лекционные аудитории должны быть оснащены: доской; интерактивной доской + ноутбук (по возможности).

11. Иные сведения и (или) материалы: (включаются на основании решения кафедры)

Оформление сведений о дополнении и изменении рабочей программы учебной дисциплины

Рабочие программы учебных дисциплин ежегодно обсуждаются, актуализируются на заседаниях ПМК, рассматриваются на заседаниях кафедр и утверждаются проректором по учебной работе, информация об изменениях отражается в листе сведений о дополнении и изменении рабочей программы учебной дисциплины. В случае существенных изменений программа полностью переоформляется. Обновленный электронный вариант программы размещается на сервере ГОУ ВПО «ДонАУиГС».

Изменения в РПУД могут вноситься в следующих случаях:

- изменение государственных образовательных стандартов или других нормативных документов, в том числе локальных нормативных актов;
- изменение требований работодателей к выпускникам;
- разработка новых методик преподавания и контроля знаний студентов.

Ответственность за актуализацию РПУД несут преподаватели, реализующие дисциплину.

СВЕДЕНИЯ О ДОПОЛНЕНИИ И ИЗМЕНЕНИИ РАБОЧЕЙ ПРОГРАММЫ УЧЕБНОЙ ДИСЦИПЛИНЫ НА 20___/20___ УЧЕБНЫЙ ГОД

«Название дисциплины»

Направление подготовки
(профиль/магистерская программа)

ДОПОЛНЕНО (с указанием раздела РПУД)

ИЗМЕНЕНО (с указанием раздела РПУД)

УДАЛЕНО (с указанием раздела РПУД)

Реквизиты протокола заседания кафедры от _____ № _____ дата
