Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Выборнова Любовь Алексеевна
Должность: ректор
Дата подписания: 29.10.2023 10:52:15
Уникальный программный ключ:
c3b3b9c625f6c113afa2a2c42baff9e05a38b76e

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

федеральное государственное бюджетное образовательное учреждение высшего образования

«Поволжский государственный университет сервиса» (ФГБОУ ВО «ПВГУС»)

Высшая школа интеллектуальных систем и кибертехнологий

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Б1.В.ДВ.01.5 «Интеллектуальный анализ данных»

Направление подготовки:
09.03.02 «Информационные системы и технологии»

Направленность (профиль):
«Информационные системы и технологии»

Квалификация выпускника: **бакалавр**

Тольятти 2023

Рабочая программа дисциплины *«Интеллектуальный анализ данных»* разработана в соответствии с федеральным государственным образовательным стандартом высшего образования - *бакалавриат* по направлению подготовки *09.03.02 «Информационные системы и технологии»*, утверждённым приказом Министерства образования и науки Российской Федерации от 19.09.2017 г. № 926.

Составители:

д.т.н., профессор	В.И. Воловач
(ученая степень, ученое звание)	(ФИО)

РПД обсуждена на заседании высшей школы интеллектуальных систем и кибертехнологий 15.12.2023 г., протокол № 4

Директор высшей школы интеллектуальных систем и кибертехнологий

к. э. н., доцент	/О.А. Филиппова
(уч.степень, уч.звание)	(ФИО)

1. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫХ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

1.1. Цель освоения дисциплины

Целью освоения дисциплины является:

- формирование у обучающихся общепрофессиональных и профессиональных компетенций, направленных на решение задач профессиональной деятельности;
- развитие навыков профессиональной деятельности.

1.2. Перечень планируемых результатов обучения по дисциплине

Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине	Основание (ПС) *для профессиональных компетенций
ОПК-2. Способен понимать принципы работы современных информационных технологий и программных средств, в том числе отечественного производства, и использовать их при решении задач профессиональной деятельности	ИОПК-2.4 Умеет анализировать массивы больших данных с использованием современных программных средств	**Знает:** требования к оформлению научных публикаций по теме исследования в сфере информационной безопасности **Умеет:** разрабатывать планы и программы проведения научных исследований в сфере информационной безопасности **Владеет:** навыками проектирования технических разработок защищенных информационных, информационно-аналитических и экспертных систем	
ПК-3 Способен к выполнению работ по управлению, настройке, разработке баз данных ИС	ИПК-3.1. Осуществляет разработку структуры баз данных ИС в соответствии с архитектурной спецификацией	**Знает:** принципы проведения предпроектного исследования, принципы и методы анализа предметной области исследования; принципы функционирования информационно-аналитических и экспертных систем; преимущества интеллектуальных технологий и систем **Умеет:** обосновывать выбор технологий и компонент создаваемых интеллектуальных и информационно-аналитических систем и формировать базу знаний и базу правил в области выбора конфигураций и состава информационно-аналитических и экспертных систем **Владеет:** навыками извлечения и применения знаний проектирования интеллектуальных и информационно-аналитических систем	06.015 Специалист по информационным системам

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина относится к части, формируемой участниками образовательных отношений, Блока 1. Дисциплины (модули) образовательной программы (Специализация 1. Искусственный интеллект и его применение Б1.В.ДВ.01).

3. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

3.1. Объем и структура дисциплины

Общая трудоѐмкость дисциплины составляет **5 з.е. (180 час.)**, их распределение по видами работ и семестрам представлено в таблице.

Виды учебных занятий и работы обучающихся	Трудоемкость, час
Общая трудоемкость дисциплины, час	**180**
Контактная работа обучающихся с преподавателем по видам учебных занятий (всего), в т.ч.:	**48 / 16**
занятия лекционного типа (лекции)	18 / 6
занятия семинарского типа (семинары, практические занятия, практикумы, коллоквиумы и иные аналогичные занятия)	18 / 6
лабораторные работы	12 / 4
Самостоятельная работа всего, в т.ч.:	**105 / 155**
Самоподготовка по темам (разделам) дисциплины	69 / 119
Выполнение курсового проекта /курсовой работы	36 / 36
Контроль (часы на экзамен, зачет)	**27 / 9**
Промежуточная аттестация	**Экзамен, КП**

Примечание: *-/- объем часов соответственно для очной и очно-заочной форм обучения*

Дисциплина реализуется посредством проведения контактной работы с обучающимися (включая проведение текущего контроля успеваемости), самостоятельной работы обучающихся и промежуточной аттестации.

В процессе освоения дисциплины может применяться электронное обучение и дистанционные образовательные технологии.

В процессе освоения дисциплины обучающиеся обеспечены доступом к электронной информационно-образовательной среде и электронно-библиотечным системам.

3.2. Содержание дисциплины, структурированное по темам

Планируемые результаты освоения: код формируемой компетенции и индикаторы достижения компетенций	Наименование разделов, тем	Виды учебной работы				Формы текущего контроля (наименование оценочного средства)
		Контактная работа			Самостоятельная работа, час	
		Лекции, час	Лабораторные работы, час	Практические занятия, час		
ПК-3 ИПК-3.1	**Тема 1. Технологии и средства интеллектуального анализа данных** Теоретические основы технологий интеллектуального анализа и экспертных систем для оценки безопасности ИС	6/2				Доклад/ сообщение
	Практическое занятие № 1. Практическое применение технологии Data Mining и экспертных систем			9/3		Выполнение практических заданий
	Самостоятельная работа .Подготовка к практическому занятию.				23/40	Самостоятельное изучение учебных материалов
ОПК-2 ИОПК-2.4	**Тема 2. Интеллектуальный анализ деятельности предприятия** 2.1 Построение аналитических предложений 2.2 Оперативный анализ интеллектуальных данных 2.3 Поиск и построение ассоциативных правил 2.4 Построение дерева целей и дерева решений	6/2				Доклад/ сообщение

Планируемые результаты освоения: код формируемой компетенции и индикаторы достижения компетенций	Наименование разделов, тем	Виды учебной работы				Формы текущего контроля (наименование оценочного средства)
		Контактная работа			Самостоятельная работа, час	
		Лекции, час	Лабораторные работы, час	Практические занятия, час		
	2.5 Самоорганизующие карты Кохонена.					
	Практическое занятие № 2. Построение самоорганизующей карты Кохонена.			9/3		Выполнение практических заданий
	Самостоятельная работа .Подготовка к практическому занятию.				23/40	Самостоятельное изучение учебных материалов
ПК-3 ИПК-3.1	**Тема 3. Применение систем извлечения знаний методами Data Mining и интеллектуального анализа данных для комплексной оценки безопасности автоматизированных систем управления и разработки систем проактивной защиты информации на основе анализа событий в информационной системе**	6/2				Доклад/ сообщение
	Лабораторная работа. Предиктивная аналитика		12/4			Выполнение практических заданий
	Самостоятельная работа .Подготовка к практическому занятию.				23/39	Самостоятельное изучение учебных материалов
	Выполнение курсового проекта /курсовой работы				36/36	Самостоятельное выполнение с консультацией преподавателя
	ИТОГО	**18/6**	**12/4**	**18/6**	**105/155**	

Примечание: *-/- объем часов соответственно для очной и очно-заочной форм обучения*

4. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮДИСЦИПЛИНЫ

4.1. Общие методические рекомендации по освоению дисциплины, образовательные технологии

Дисциплина реализуется посредством проведения контактной работы с обучающимися(включая проведение текущего контроля успеваемости), самостоятельной работы обучающихся и промежуточной аттестации.

При проведении учебных занятий по дисциплине обеспечивается развитие у обучающихся навыков командной работы, межличностной коммуникации, принятия решений, лидерских качеств (включая проведение интерактивных лекций, групповых дискуссий, ролевых игр, тренингов, анализ ситуаций и имитационных моделей, преподавание дисциплины в форме курса, составленного на основе результатов научных исследований, проводимых университетом, в том числе с учетом региональных особенностей профессиональной деятельности выпускников и потребностей работодателей).

Преподавание дисциплины ведется с применением следующих видов **образовательных технологий:**

- *балльно-рейтинговая технология оценивания;*
- *электронное обучение;*
- *проблемное обучение;*
- *разбор конкретных ситуаций;*
- *информационные технологии: Miro, Google-документы, Zoom.*

Для оценки знаний, умений, навыков и уровня сформированности компетенции по дисциплине применяется балльно-рейтинговая система контроля и оценки успеваемости студентов. В основу балльно-рейтинговой системы положены принципы, в соответствии с которыми формирование рейтинга студента осуществляется в ходе текущего контроля успеваемости. Максимальное количество баллов в семестре – 100.

4.2. Методические указания для обучающихся по освоению дисциплины на занятиях лекционного типа

Лекционный курс предполагает систематизированное изложение основных вопросов тематического плана. В ходе лекционных занятий раскрываются базовые вопросы в рамках каждой темы дисциплины. Обозначаются ключевые аспекты тем, а также делаются акценты на наиболее сложные и важные положения изучаемого материала.

Лекционные занятия проводятся в поточной аудитории с применением мультимедийного проектора в виде учебной презентации или в ЭИОС университета.

Отдельные темы предлагаются для самостоятельного изучения (конспектируются).

Материалы лекций являются опорной основой для подготовки обучающихся к практическим занятиям и выполнения заданий самостоятельной работы, а также к мероприятиям текущего контроля успеваемости и промежуточной аттестации по дисциплине.

4.3. Методические указания для обучающихся по освоению дисциплины на лабораторных работах

Подготовку к каждой лабораторной работе студент должен начать с ознакомления с планом занятия, который отражает содержание предложенной темы. Каждая выполненная работа с оформленным отчетом по ней подлежит защите преподавателю.

При оценивании лабораторных работ учитывается следующее:

–качество выполнения экспериментально-практической части работы и степень соответствия результатов работы заданным требованиям;

–качество оформления отчета по работе;

–качество устных ответов на контрольные вопросы при защите работы.

Лабораторные работы организуются, в том числе в форме практической подготовки, которая предусматривает участие обучающихся в выполнении отдельных элементов работ, связанных с будущей профессиональной деятельностью.

4.4. Методические указания для обучающихся по освоению дисциплины на занятиях семинарского типа/ на практических занятиях

Практические (семинарские) занятия представляют собой детализацию лекционного теоретического материала, проводятся в целях закрепления курса и охватывают все основные разделы. Основной формой проведения семинаров и практических занятий является обсуждение наиболее проблемных и сложных вопросов по отдельным темам, а также решение задач и разбор примеров и ситуаций в аудиторных условиях.

Практические (семинарские) занятия обучающихся обеспечивают:

– проверку и уточнение знаний, полученных на лекциях;

– получение умений и навыков составления докладов и сообщений, обсуждения вопросов

– по учебному материалу дисциплины;

– подведение итогов занятий по рейтинговой системе, согласно технологической карте дисциплины.

Практические занятия организуются, в том числе в форме практической подготовки, которая предусматривает участие обучающихся в выполнении отдельных элементов работ, связанных с будущей профессиональной деятельностью.

Практическая подготовка предусматривает: выполнение практических заданий – темы 1, 2, 3.

4.5. Методические указания по самостоятельной работе обучающихся

Самостоятельная работа обеспечивает подготовку обучающегося к аудиторным занятиями мероприятиям текущего контроля и промежуточной аттестации по изучаемой дисциплине. Результаты этой подготовки проявляются в активности обучающегося на занятиях и в качестве выполненных практических заданий и других форм текущего контроля.

Самостоятельная работа студентов включает:

– изучение учебной литературы по курсу;
– решение практических ситуаций и задач;
– работу с ресурсами Интернет;
– решение практических ситуаций в виде кейсов;
– подготовку к тестированию по темам курса;
– подготовку к промежуточной аттестации по курсу и др.

При выполнении заданий для самостоятельной работы рекомендуется проработка материалов лекций по каждой пройденной теме, а также изучение рекомендуемой литературы.

Для обучающихся по очно-заочной форме обучения самостоятельная работа является основным видом учебной деятельности.

Для обеспечения самостоятельной работы обучающихся используется электронный учебный курс, созданный в ЭИОС университета http://sdo.tolgas.ru/

4.6. Методические указания для выполнения курсового проекта / работы

Выполнение курсового проекта/ работы способствует лучшему освоению обучающимися учебного материала, формирует практический опыт и умения по изучаемой дисциплине, способствует формированию у обучающихся готовности к самостоятельной профессиональной деятельности, является этапом к выполнению выпускной квалификационной работы.

Примерная тематика курсовых работ (проектов)

1. Исходные данные для Data Mining.
2. Типы закономерностей, выявляемых методами Data Mining.
3. Методы исследования данных в Data Mining.
4. Постановка задач обучения по прецедентам.
5. Объекты и признаки.
6. Типы шкал: бинарные, номинальные, порядковые, количественные.

7. Типы задач: классификация, регрессия, прогнозирование, кластеризация, ассоциация.
8. Основные понятия: модель алгоритмов, метод обучения, функция потерь и функционал качества, принцип минимизации эмпирического риска, обобщающая способность, скользящий контроль.
9. Нейрон и межнейронное взаимодействие.
10. Модель технического нейрона.
11. Искусственные нейронные сети и их архитектура.
12. Структура многослойной нейронной сети.
13. Функции активации.
14. Проблема полноты. Задача исключающего или.
15. Полнота двухслойных сетей в пространстве булевских функций. Алгоритм обратного распространения ошибокБольшие данные (Big Data): современные подходы к обработке и хранению.
16. Проблема множественного сравнения данных.
17. Модель вычислений MapReduce.
18. Реализации алгоритмов на MapReduce.
19. Обзор задач и методов машинного обучения и интеллектуального анализа больших данных.
20. Аналитическая обработка сложноструктурированных больших данных.
21. Построение аналитических предложений
22. Оперативный анализ интеллектуальных данных
23. Поиск и построение ассоциативных правил
24. Построение дерева целей и дерева решений
25. Самоорганизующие карты Кохонена.
26. Понятие информационного хранилища.
27. Принципы построения информационных хранилищ.
28. Задачи и средства администрирования ИАС.
29. Требования к качеству данных и способы его обеспечения при загрузке в информационное хранилище.
30. Проблемы, разрешаемые при приведении данных единой структуре информационного хранилища.
31. Концепции построения структур хранилищ данных.
32. Назначение, состав и выполняемые функции базы метаданных – репозитория хранилищ данных.
33. Принципы создания репозитория хранилищ данных.
34. Элементы моделей данных хранилища (факт-таблица, таблицы измерений, консольная таблица).
35. Схемы представления – модели многомерных данных.

5. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

5.1. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

Вся литература, включенная в данный перечень, представлена в виде электронных ресурсов в электронной библиотеке университета (ЭБС). Литература, используемая в печатном виде, представлена в научной библиотеке университета в объеме не менее 0,25 экземпляров на одного обучающегося.

Основная литература

1. Алексеев, Д. С. Технологии интеллектуального анализа данных : учебник / Д. С. Алексеев, О. В. Щекочихин. - Изд. 2-е, стер. - Документ Reader. - Санкт-Петербург [и др.] : Лань, 2024. - 176 с. - URL: https://reader.lanbook.com/book/362915 (дата обращения: 09.12.2021). - Режим доступа: для авториз. пользователей. - ISBN 978-5-507-48763-9. - Текст : электронный. URL: https://reader.lanbook.com/book/362915

2. Кулаичев, А. П. Методы и средства комплексного статистического анализа данных : учеб. пособие для вузов по дисциплинам "Мат. статистика" и "Информатика" / А. П. Кулаичев. - 5-е изд., перераб. и доп. - Документ read. - Москва : ИНФРА-М, 2022. - 484 с. : ил. - Прил. - Предм. указ. - URL: https://znanium.com/read?id=379982 (дата обращения: 21.02.2022). - Режим доступа: для авториз. пользователей. - ISBN 978-5-16-012834-4. - 978-5-16-103357-9. - Текст : электронный.URL: https://znanium.com/read?id=379982

3. Макшанов, А. В. Технологии интеллектуального анализа данных : учеб. пособие / А. В. Макшанов, А. Е. Журавлев. - Изд. 2-е, стер. - Документ Reader. - Санкт-Петербург [и др.] : Лань, 2022. - 209 с. - (Учебники для вузов. Специальная литература). - URL: https://reader.lanbook.com/book/206711 (дата обращения: 17.10.2022). - Режим доступа: для авториз. пользователей. - ISBN 978-5-8114-4493-9. - Текст : электронный. URL: https://reader.lanbook.com/book/206711A

4. Цехановский, В. В. Технология интеллектуального анализа данных в процессах и системах : учебник / В. В. Цехановский, В. Д. Чертовской. - Документ read. - Санкт-Петербург [и др.] : Лань, 2021. - 165 с. - Прил. - URL: https://reader.lanbook.com/book/302753 (дата обращения: 09.02.2023). - Режим доступа: для авториз. пользователей. - ISBN 978-5-507-45404-4. - Текст : электронный. URL: https://reader.lanbook.com/book/302753

Дополнительная литература

5. Миркин, Б. Г. Введение в анализ данных : учеб. и практикум для бакалавриата и магистратуры по инж.-техн., естеств.-науч. и экон. направлениям и специальностям / Б. Г. Миркин ; Высш. шк. экономики, Нац. исслед. ун-т. - Москва : ЮРАЙТ, 2016. - 174 с. : ил., табл. - (Авторский учебник). - Предм. указ. - Прил. - ISBN 978-5-9916-5009-0 : 394-90. - Текст : непосредственный.

6. Петрунин, Ю. Ю. Информационные технологии анализа данных. Data Analysis : учеб. пособие для вузов по дисциплине "Информатика" по упр. и экон. специальностям и направлениям / Ю. Ю. Петрунин ; МГУ им. М. В. Ломоносова, Фак. гос. упр. - 2-е изд. - Москва : Университет, 2010. - 292 с. : ил., табл. - ISBN 978-5-98227-701-5 : 309-87. - Текст : непосредственный.

5.2. Профессиональные базы данных, информационно-справочные системы, интернет-ресурсы

1. eLIBRARY.RU : научная электронная библиотека : сайт. – Москва, 2000 - . - URL: https://elibrary.ru (дата обращения: 03.12.2021). – Режим доступа: для зарегистрир. пользователей. – Текст: электронный.

2. КонсультантПлюс : справочная правовая система : сайт / ЗАО «КонсультантПлюс». – Москва, 1992 - . - URL: http://www.consultant.ru (дата обращения 03.12.2021). - Текст : электронный.

3. Электронная библиотечная система Поволжского государственного университета сервиса : сайт / ФГБОУ ВО «ПВГУС». – Тольятти, 2010 - . - URL. : http://elib.tolgas.ru(дата обращения 03.12.2021). - Режим доступа: для авториз. пользователей. - Текст : электронный.

4. Электронно-библиотечная система Znanium.com: сайт / ООО "ЗНАНИУМ". – Москва, 2011 - . - URL: https://znanium.com/ (дата обращения 03.12.2021). - Режим доступа: для авториз. пользователей. - Текст : электронный.

5. Электронно-библиотечная система Лань : сайт / ООО "ЭБС ЛАНЬ". - Москва, 2011 - . - URL: https://e.lanbook.com/ (дата обращения 03.12.2021). - Режим доступа: для авториз. пользователей. - Текст : электронный.

5.3. Программное обеспечение

Информационное обеспечение учебного процесса по дисциплине осуществляется с использованием следующего программного обеспечения (лицензионного и свободно распространяемого), в том числе отечественного производства:

№п/п	Наименование	Условия доступа
1	Microsoft Windows	из внутренней сети университета (лицензионный договор)
2	MicrosoftOffice	из внутренней сети университета (лицензионный договор)
3	КонсультантПлюс	из внутренней сети университета (лицензионный договор)
4	СДО MOODLE	из любой точки, в которой имеется доступ к сети Интернет (лицензионный договор)

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

Помещения представляют собой учебные аудитории для проведения учебных занятий, предусмотренных учебным планом и рабочей программой дисциплины, оснащенные оборудованием и техническими средствами обучения.

Занятия лекционного типа. Учебные аудитории для занятий лекционного типа укомплектованы мебелью и техническими средствами обучения, служащими для представления учебной информации (стационарные или переносные наборы демонстрационного оборудования (проектор, экран, компьютер/ноутбук), учебно-наглядные пособия (презентации по темам лекций), обеспечивающие тематические иллюстрации, соответствующие данной программе дисциплины.

Занятия семинарского типа. Учебные аудитории для занятий семинарского типа укомплектованы мебелью и техническими средствами обучения, служащими для представления учебной информации (стационарные или переносные наборы демонстрационного оборудования (проектор, экран, компьютер/ноутбук).

Промежуточная аттестация. Для проведения промежуточной аттестации по дисциплине используются компьютерные классы, оснащенные компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду университета и/или учебные аудитории, укомплектованные мебелью и техническими средствами обучения.

Самостоятельная работа. Помещения для самостоятельной работы оснащены компьютерной техникой с возможностью подключения к сети «Интеренет» и доступом к электронной информационно-образовательной среде университета. Для организации самостоятельной работы обучающихся используются:

– компьютерные классы университета;

– библиотека (медиазал), имеющая места для обучающихся, оснащенные компьютерами с доступом к базам данных и сети «Интернет».

Электронная информационно-образовательная среда университета (ЭИОС).

Каждыйобучающийсявтечениевсегопериодаобученияобеспечениндивидуальнымнеограниченным доступом к электронной информационно-образовательной среде университета (ЭИОС) http://sdo.tolgas.ru/ из любой точки, в которой имеется доступ к информационно-телекоммуникационной сети «Интернет», как на территории университета, так и вне ее.

ЭИОС университета обеспечивает:

– доступ к учебным планам, рабочим программам дисциплин (модулей), программам практик, электронным учебным изданиям и электронным образовательным ресурсам, указанным в рабочих программах дисциплин (модулей), программах практик;

– формирование электронного портфолио обучающегося, в том числе сохранение его работ и оценок за эти работы.

В случае реализации образовательной программы с применением электронного обучения, дистанционных образовательных технологий ЭИОС дополнительно обеспечивает:

– фиксацию хода образовательного процесса, результатов промежуточной аттестации и результатов освоения образовательной программы;

– проведение учебных занятий, процедур оценки результатов обучения, реализация которых предусмотрена с применением электронного обучения, дистанционных образовательных технологий;

– взаимодействие между участниками образовательного процесса, в том числе синхронное и (или) асинхронное взаимодействия посредством сети «Интернет».

7. ОСОБЕННОСТИ ОРГАНИЗАЦИИ ОБУЧЕНИЯ ДЛЯ ЛИЦ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ И ИНВАЛИДОВ

При необходимости рабочая программа дисциплины может быть адаптирована для обеспечения образовательного процесса инвалидов и лиц с ограниченными возможностями здоровья, в том числе для дистанционного обучения. Для этого требуется заявление студента (его законного представителя) и заключение психолого-медико-педагогической комиссии (ПМПК).

В случае необходимости, обучающимся из числа лиц с ограниченными возможностями здоровья (по заявлению обучающегося) а для инвалидов также в соответствии с индивидуальной программой реабилитации инвалида, могут предлагаться следующие варианты восприятия учебной информации с учетом их индивидуальных психофизических особенностей, в том числе с применением электронного обучения и дистанционных технологий:

– для лиц с нарушениями зрения: в печатной форме увеличенным шрифтом; в форме электронного документа; в форме аудиофайла (перевод учебных материалов в аудиоформат); в печатной форме на языке Брайля; индивидуальные консультации с привлечением тифлосурдопереводчика; индивидуальные задания и консультации;

– для лиц с нарушениями слуха: в печатной форме; в форме электронного документа; видеоматериалы с субтитрами; индивидуальные консультации с привлечением сурдопереводчика; индивидуальные задания и консультации;

– для лиц с нарушениями опорно-двигательного аппарата: в печатной форме; в форме электронного документа; в форме аудиофайла; индивидуальные задания и консультации.

8. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ (ФОНД ОЦЕНОЧНЫХ СРЕДСТВ) ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

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

Для оценки знаний, умений, навыков и уровня сформированности компетенции по дисциплине применяется балльно-рейтинговая система контроля и оценки успеваемости студентов. В основу балльно-рейтинговой системы положены принципы, в соответствии с которыми формирование рейтинга студента осуществляется в ходе текущего контроля успеваемости. Максимальное количество баллов в семестре – 100.

Шкала оценки результатов освоения дисциплины, сформированности результатов обучения

Форма проведения промежуточной аттестации	Шкалы оценки уровня сформированности результатов обучения		Шкала оценки уровня освоения дисциплины	
	Уровневая шкала оценки компетенций	100 бальная шкала, %	100 бальная шкала, %	5-балльная шкала, дифференцированная оценка/балл
Экзамен КП	допороговый	ниже 61	ниже 61	«неудовлетворительно» / 2
	пороговый	61-85,9	61-69,9	«удовлетворительно» / 3
			70-85,9	«хорошо» / 4
	повышенный	86-100	86-100	«отлично» / 5

По итогам текущей успеваемости студенту может быть выставлена оценка по промежуточной аттестации в соответствии за набранными за семестр баллами (по накопительному рейтингу).Студентам, набравшим в ходе текущего контроля успеваемости по дисциплине от 61 до 100 баллов и выполнившим все обязательные виды запланированных учебных занятий, по решению преподавателя без прохождения промежуточной аттестации выставляется оценка в соответствии со шкалой оценки результатов освоения дисциплины.

Результат обучения считается сформированным (повышенный уровень), если теоретическое содержание курса освоено полностью; при устных собеседованиях студент исчерпывающе, последовательно, четко и логически стройно излагает учебный материал; свободно справляется с задачами, вопросами и другими видами заданий, требующих применения знаний, использует в ответе дополнительный материал; все предусмотренные рабочей учебной программой задания выполнены в соответствии с установленными требованиями, студент способен анализировать полученные результаты; проявляет самостоятельность при выполнении заданий, качество их выполнения оценено числом баллов от 86 до 100, что соответствует повышенному уровню сформированности результатов обучения.

Результат обучения считается сформированным (пороговый уровень), если теоретическое содержание курса освоено полностью; при устных собеседованиях студент последовательно, четко и логически стройно излагает учебный материал; справляется с задачами, вопросами и другими видами заданий, требующих применения знаний; все предусмотренные рабочей учебной программой задания выполнены в соответствии с установленными требованиями, студент способен анализировать полученные результаты; проявляет самостоятельность при выполнении заданий, качество их выполнения оценено числом баллов от 61 до 85,9, что соответствует пороговому уровню сформированности результатов обучения.

Результат обучения считается несформированным, если студент при выполнении заданий не демонстрирует знаний учебного материала, допускает ошибки, неуверенно, с большими затруднениями выполняет задания, не демонстрирует необходимых умений, качество выполненных заданий не соответствует установленным требованиям, качество их выполнения оценено числом баллов ниже 61, что соответствует допороговому уровню.

Формы текущего контроля успеваемости

Формы текущего контроля	Количество контрольных точек	Количество баллов за 1 контр. точку	Макс. возм. кол-во баллов
Доклад/сообщение	3	10	30
Отчет по практическимработам	3	20	60
Творческий рейтинг (участие в научных, научно-практических конференциях, грантах и т.д.)	1	10	10
		Итого	**100 баллов**

Система оценивания представлена в электронном учебном курсе по дисциплине http://sdo.tolgas.ru/.

8.2. Типовые контрольные задания или иные материалы для ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ

8.2.1. Типовые задания к практическим (семинарским) занятиям (темы докладов/сообщений)

Практическое занятие № 1. Практическое применение технологии Data Mining и экспертных систем

Цель: формирование профессиональной компетенции сфере применения цифровых инструментов сбора, обработки и анализа больших данных.

Теоретическая часть

Big Data — это серия подходов, инструментов и методов, используемых для обработки структурированных и неструктурированных данных огромных объемов и значительного разнообразия для получения результатов, воспринимаемых людьми, которые доказывают свою эффективность в условиях непрерывного роста. Большие данные служат альтернативой традиционным системам управления базами данных и решениям в рамках Business Intelligence. Таким образом, большие данные не относятся к определенному количеству данных или даже к самим данным. Вместо этого термин означает методы обработки данных, которые позволяют распределенную обработку информации. Эти методы могут применяться как к огромным наборам данных (например, содержание всех страниц в Интернете), так и к небольшим (например, к содержанию этой статьи). Большие данные имеют важное значение для глобального бизнеса, поскольку большее количество данных приводит к более точному анализу, который, в свою очередь, обеспечивает более эффективное принятие решений, повышение операционной эффективности и сокращению затрат. Apache HTTP-сервер — свободный веб-сервер. Apache является кроссплатформенным ПО, поддерживает операционные системы Linux, BSD, Mac OS, Microsoft Windows, Novell NetWare, BeOS. Основными достоинствами Apache считаются надёжность и гибкость конфигурации

Вопросы и задания для обсуждения:

1. Языки программирования для анализа больших данных
2. Платные и бесплатные инструменты для обработки и анализа больших данных
3. Задачи классификации Data Mining
4. Задачи кластеризации Data Mining
5. Методы прогнозирования в Data Mining
6. Характеристика основных методов обнаружения вторжений
7. Особенности технологий Data Mining для обнаружения атак
8. Классификация задач Data Mining
9. Цель технологии Data Mining
10. Предсказательные модели
11. Описательные модели
12. Базовые методы Data Mining

13. Нечеткая логика
14. Генетические алгоритмы (ГА)
15. Нейронные сети
16. Нейро-нечеткая система для обнаружения вторжений в сеть

Перечень докладов по теме 1 «Технологии и средства интеллектуального анализа данных»

36. Исходные данные для Data Mining.
37. Типы закономерностей, выявляемых методами Data Mining.
38. Методы исследования данных в Data Mining.
39. Постановка задач обучения по прецедентам.
40. Объекты и признаки.
41. Типы шкал: бинарные, номинальные, порядковые, количественные.
42. Типы задач: классификация, регрессия, прогнозирование, кластеризация, ассоциация.
43. Основные понятия: модель алгоритмов, метод обучения, функция потерь и функционал качества, принцип минимизации эмпирического риска, обобщающая способность, скользящий контроль.
44. Нейрон и межнейронное взаимодействие.
45. Модель технического нейрона.
46. Искусственные нейронные сети и их архитектура.
47. Структура многослойной нейронной сети.
48. Функции активации.
49. Проблема полноты. Задача исключающего или.
50. Полнота двухслойных сетей в пространстве булевских функций. Алгоритм обратного распространения ошибокБольшие данные (Big Data): современные подходы к обработке и хранению.
51. Проблема множественного сравнения данных.
52. Модель вычислений MapReduce.
53. Реализации алгоритмов на MapReduce.
54. Обзор задач и методов машинного обучения и интеллектуального анализа больших данных.
55. Аналитическая обработка сложноструктурированных больших данных.

Практическое занятие № 2. Построение самоорганизующей карты Кохонена

Цель: Научиться осуществлять сбор, обработку и интеллектуальный анализ научно-технической информации по теме исследования на примере построения самоорганизующей карты Кохонена

Задание

1. Показать структуру кластеров карты Кохонена.
2. Составить таблица, полученная при содержательной интерпретации кластеров.

Порядок выполнения работы:

1.Ознакомиться с заданием.
2.Выполнить задания в соответствии с вариантом.
3.Ответьте на контрольные вопросы.

Содержание отчёта:

1. Название и цель работы.
2.Указать задание.
3.Представить решение задач.
4.Ответы на контрольные вопросы.

Контрольные вопросы:
1. Какова структура сети Кохонена?
2.В чем заключается процесс обучения без учителя (самоорганизация)?
3. В чем отличие процесса обучения с учителем, от обучения без учителя?
4. Какой нейрон называется нейроном-победителем в конкурентном обучении сети Кохонена?
5. Как оценить результаты кластеризации по раскраске карты Кохонена?
6. В чем заключается нарушение топологического подобия при визуализации результатов кластеризации?
7. Опишите структуру карты Кохонена и принцип выбора числа ячеек в ней.

Задание: Пусть имеется файл с данными о состоянии районов РТ: доля оплаченных счетов, доля долгов, недопоставки района и субсидии району (Районы.xls). Необходимо разбить районы Самарской области на группы (кластеры) в зависимости от четырех указанных показателей. Для этого выполните этапы, представленные ниже.

Ход работы

1. Для работы *выделите область данных* на листе книги MSExcel. По определению данные на листе располагаются следующим образом: входы и выходы – столбцы, а строки – это обучающие примеры. Выделяемая область может включать в первой строке названия входов как на русском, так и на английском языке. Кроме того, в первой колонке могут находиться данные, используемые не для обучения сети Кохонена, а для идентификации примеров при последующем анализе результатов.



Рис. 2. 1.

2. *Кликните мышкой* по кнопке  с цветной картой Кохонена на панели инструментов «NeuralAnalysis». В ответ появится диалоговое окно «Selectdatasource», предлагающее уточнить параметры области данных для работы.

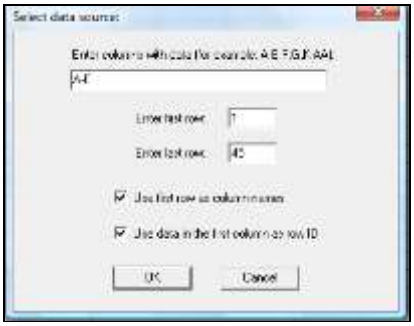


Рис. 2. 2.

В нашем примере предлагается использовать первую строку как названия столбцов, а первый столбец как названия строк, т.е. кластеризуемых данных.

3. Если параметры ввода Вас удовлетворяют, *нажмите* кнопку «ОК».

4. Перед Вами откроется основное окно программы Kohonen Map, которое содержит два листа «Project» и «Results».

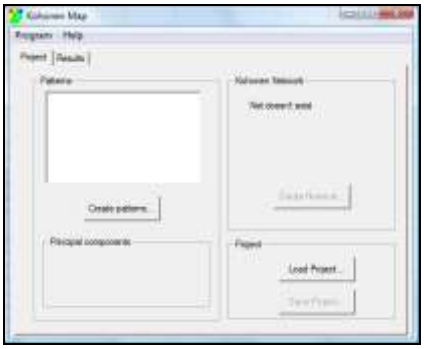


Рис. 2. 3.

5. Открывшийся перед Вами лист «Project» позволяет определить и предобработать данные для последующего использования. Кроме того, с этого листа можно сохранить обученную нейросеть («SaveProject…») или загрузить уже сохраненный в прошлом проект («LoadProject…»). Первое, что вам надо сделать, - это определить входы, для чего нажмите на кнопку «Createpatterns…».

6. В открывшемся диалоговом окне «Selectrelevantcolumns» выберете в окне левого списка «Allcolumns» необходимые входы и с помощью кнопок «>» или «>>» переведите в окно списка «Selected». Корректировку выбранных входов можно провести используя кнопки «<» или «<<». Отметим, что все данные являются входными и участвуют в обучении.

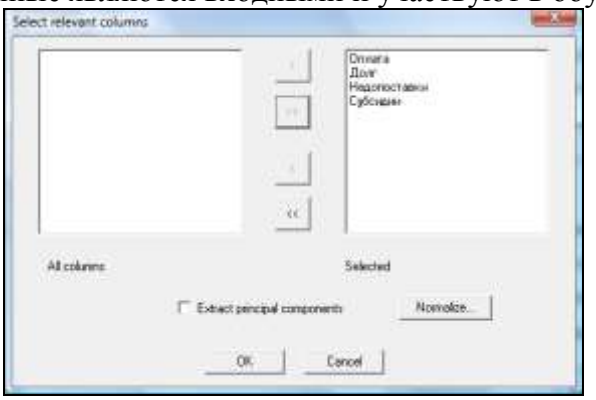


Рис. 2. 4.

7. Поскольку конкретные значения входов могут быть любого диапазона, то рекомендуется провести их нормировку. Для большинства случаев подходит нормировка входных значений «Mean/Variance». Нажмите кнопку «Normalize…» и выберите в открывшемся окне «Inputsnormalization» соответствующую позицию переключателя. Подтвердите Ваш выбор нажатием кнопки «Ок» и вернитесьв окно «Selectrelevantcolumns».

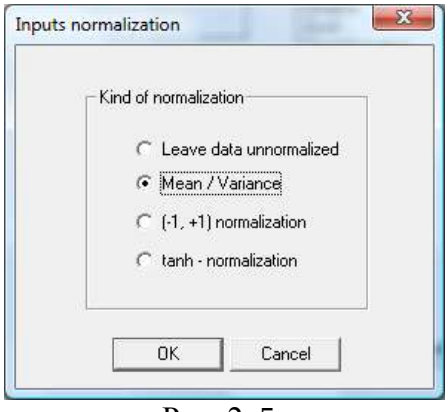


Рис. 2. 5.

8. Следующий этап – создание сети – карты Кохонена. Нажатием кнопки «CreateNetwork…» перейдите в диалоговое окно «Dialog» и задайте параметры сети: число ячеек по горизонтали и вертикали.

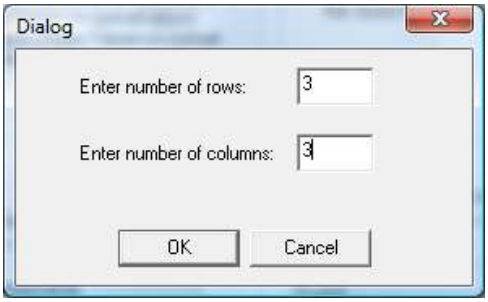


Рис. 2. 6.

В нашем примере мы выбрали размеры карты 3x3, т.е. все районы РТ будут разбиты на 9 кластеров.

9. Подтвердите выбранную Вами конфигурацию нейросети нажатием кнопки «Ок». Далее автоматически стартует процесс обучения.

10. По завершении процесса обучения для анализа результатов перейдите на закладку «Results».

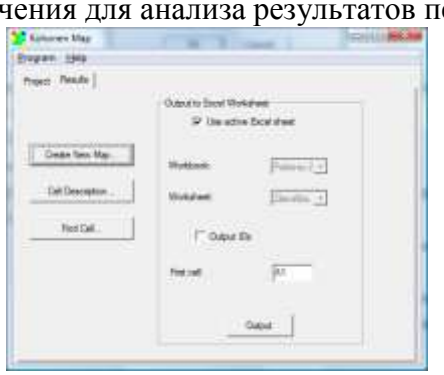


Рис. 2. 7.

Здесь доступны следующие функции.

Кнопка «CreateNewMap…» - цветовая раскраска карты Кохонена по любому параметру с выбранной степенью градации.

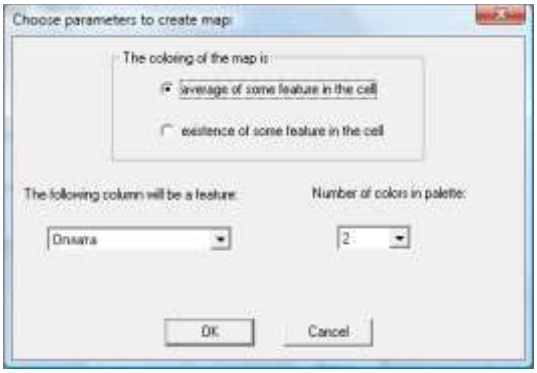


Рис. 2. 8.

Выбор раскраски карты по оплате с пятью степенями градации цвета.



Рис. 2. 9.Результат раскраски по оплате.

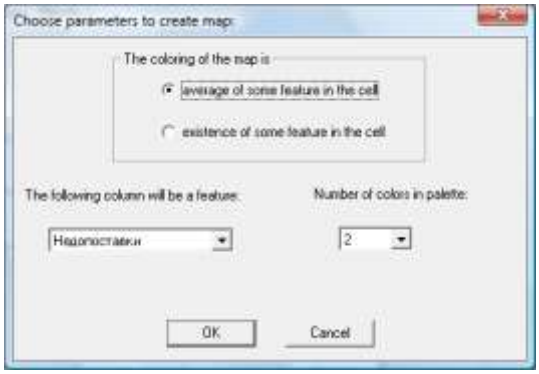


Рис. 2. 10.Выбор раскраски карты по недопоставкам с четырьмя степенями градации цвета.



Рис. 2. 11.Результат раскраски по недопоставкам.

Размеры каждого квадрата-кластера пропорциональны числу примеров, принадлежащих данному кластеру после обучения.

Кнопка «CellDescription…» - определение усредненных значений входных параметров для данного кластера (ячейки) и принадлежащих ему примеров.

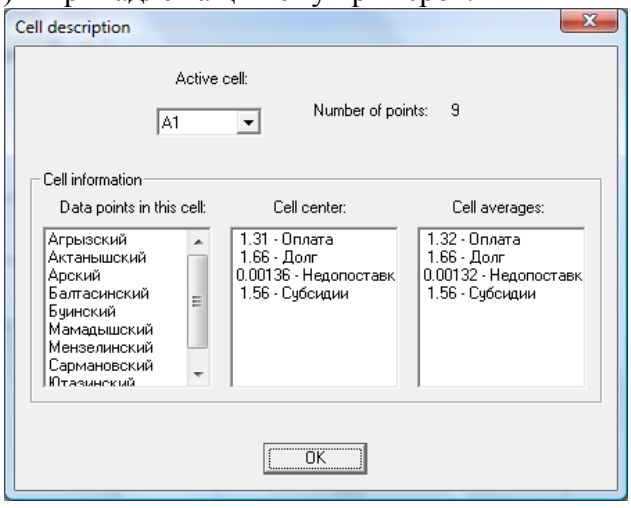


Рис. 2. 12.Описание кластера A1.

Кнопка «FindCell…» - поиск кластера, которому принадлежит данный пример.

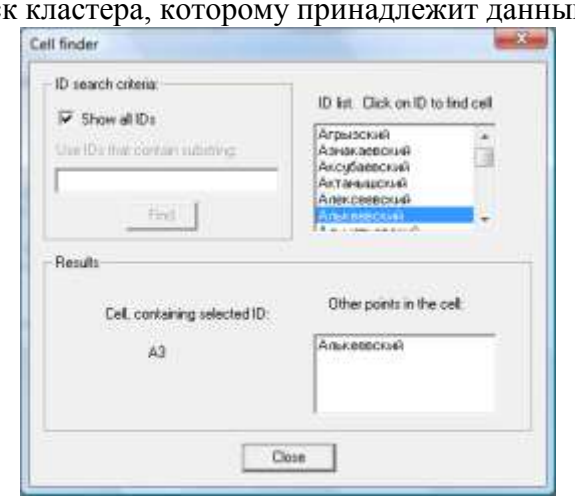


Рис. 2. 13.Определение кластера, к которому принадлежит Алькеевский район (результат – кластер А3).

Кнопка «Output» - сохранение результатов в книге MSExcel.

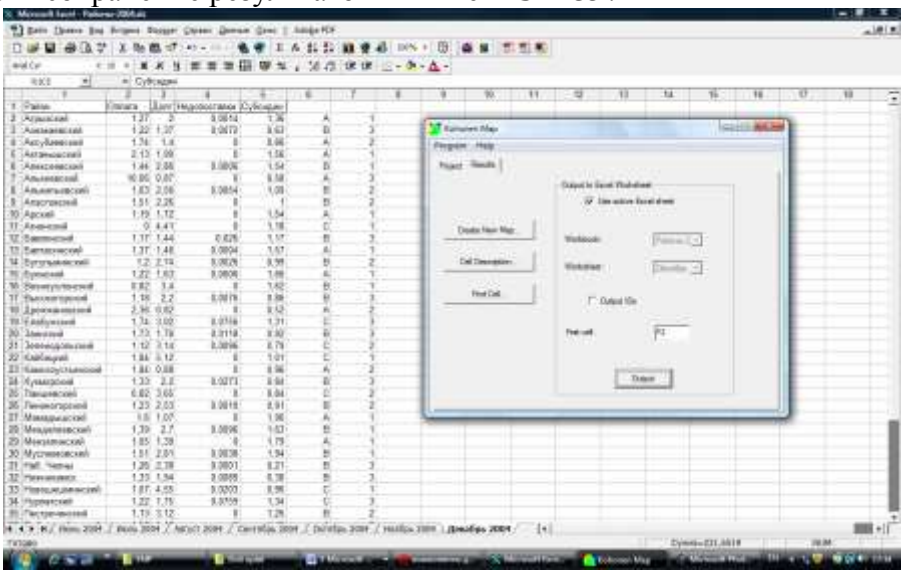


Рис. 2. 14.Выгрузка данных для каждого района в таблицу Excel.

В нашем примере данные выгружаются в ячейки, начиная с F2. Для каждого района указывается кластер, к которому он принадлежит. Кластеры обозначаются двойной маркировкой БУКВА_ЦИФРА.

11. Программа позволяет также управлять параметрами процесса обучения и изменения цветовой палитры раскраски карты Кохонена. Для этого выберите пункт меню «Program», пункт «Preferences» и далее «SetCustom…».

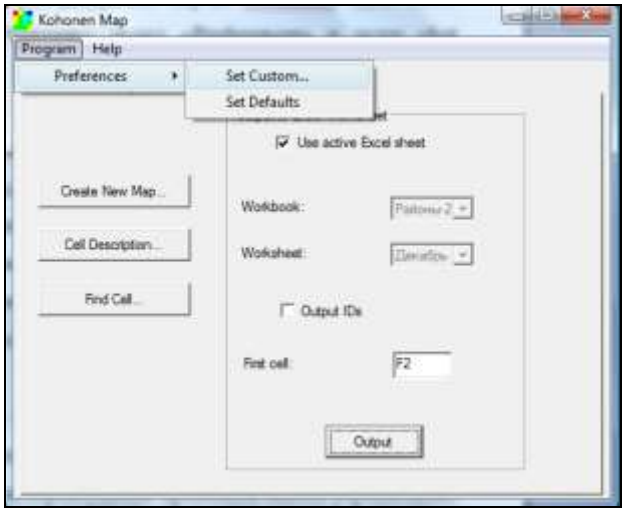


Рис. 2. 15.

В открывшемся окне «Programpreferences» на закладке «Colors» можно установить другие цвета градационной раскраски карты,

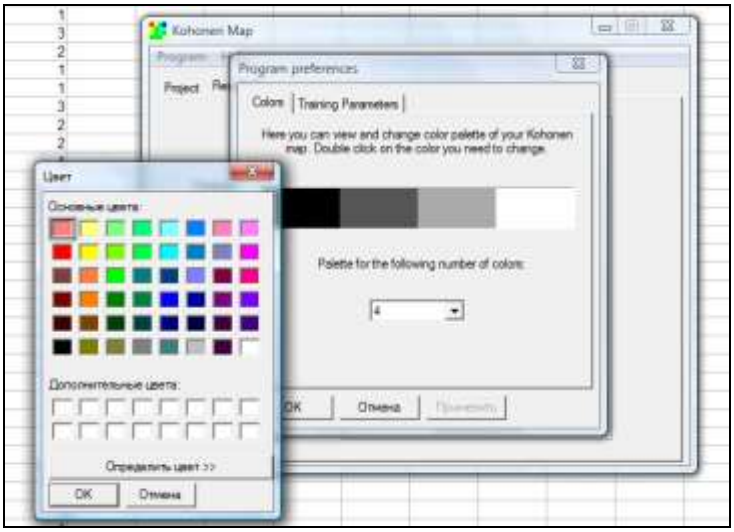


Рис. 2. 16.

а на странице «Trainingparameters» - параметры обучения сети.

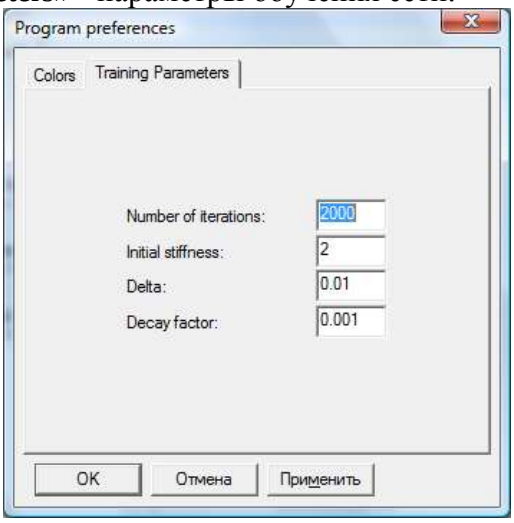


Рис. 2. 17.

Восстановить параметры программы по умолчанию можно, выбрав там же подпункт меню «Set Defaults».

12. Создаваемая при нажатии кнопки «CreateNewMap…» карта является активной: при двойном щелчке мыши на какой-либо ячейке открывается окно «Celldescription», в котором удобно проводить анализ усредненных значений параметров.

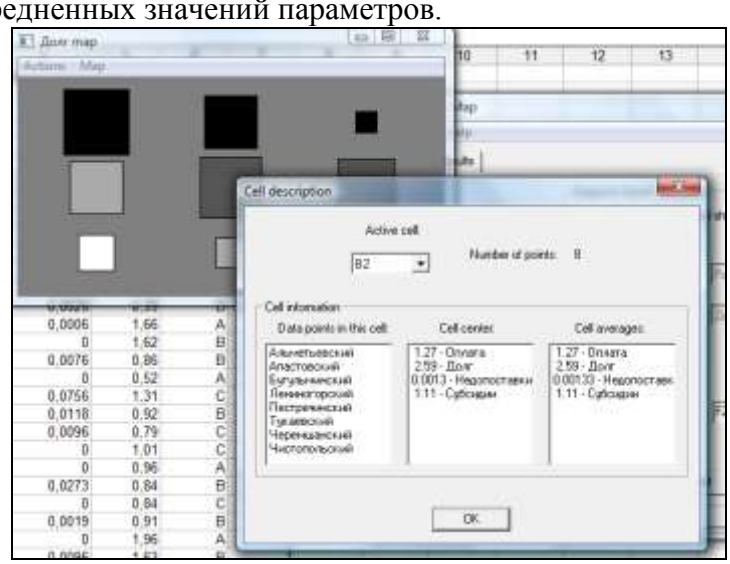


Рис. 2. 18.

Дополнительно предусмотрена возможность сохранения изображения карты раскраски в формате *.bmp для последующего экспорта через клипборд в любые документы MSOffice. Для этого в окне карты выберите пункт меню «Actions» и далее «CopytoBitmap».

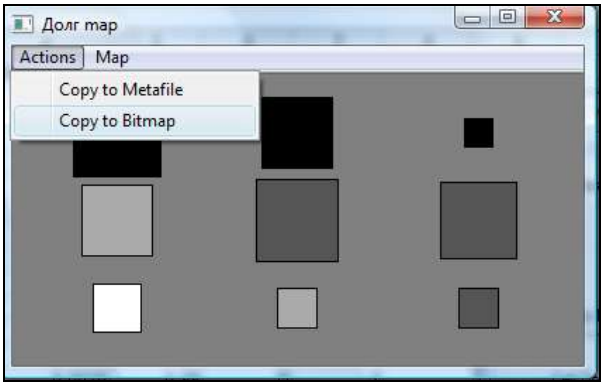


Рис. 2. 19.

Далее вернитесь в документ MSOffice и произведите вставку рисунка командой «Paste» меню «Edit».

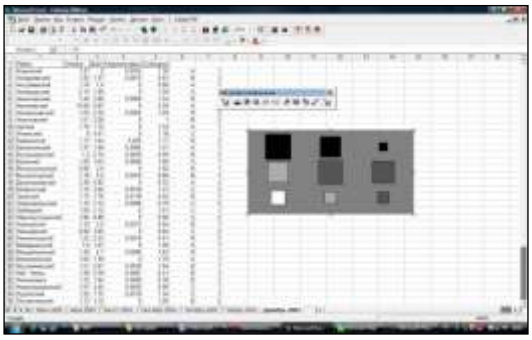


Рис. 2. 20.

13. Теперь осталось сохранить результаты работы. В программе предусмотрены функции сохранения проекта (кнопка «SaveProject…») и экспорта результатов назад в книгу MSExcel. Для экспорта результатов перейдите на закладку «Project», задайте необходимые параметры и сохраните результаты нажатием кнопки «Ок».

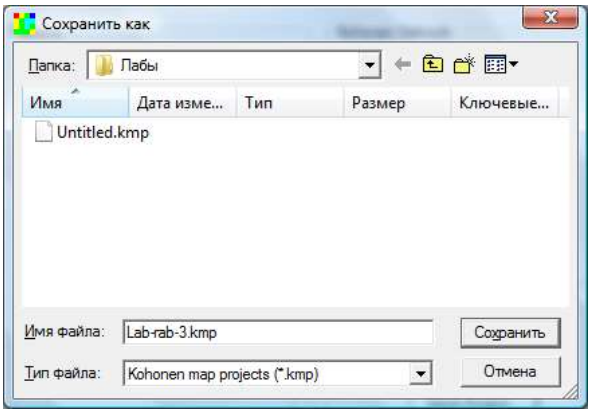


Рис. 2. 21.

Перечень докладов по теме 2 «Интеллектуальный анализ деятельности предприятия»

1. Построение аналитических предложений
2. Оперативный анализ интеллектуальных данных
3. Поиск и построение ассоциативных правил
4. Построение дерева целей и дерева решений
5. Самоорганизующие карты Кохонена.

Лабораторная работа. Предиктивная аналитика

Цель: формирование профессиональной компетенции сфере применения цифровых инструментов сбора, обработки и анализа больших данных.

Варианты и задания

Необходимо выполнить задания по поиску ассоциативных правил в соответствии с предложенным вариантом.

Варианты 1.

Тестовые варианты.

Вариант 1. Сформировать набор из 30 транзакций на основе чеков покупок в продуктовом магазине.

Вариант 2. Сформировать набор из 30 транзакций на основе чеков покупок в аптеке.

Вариант 3. Сформировать набор из 30 транзакций на основе чеков покупок в магазине компьютерной техники.

Вариант 4. Сформировать набор из 30 транзакций на основе чеков покупок в магазине одежды.

Вариант 5. Сформировать набор из 30 транзакций на основе продуктов, употребляемых на завтрак, обед и ужин.

Вариант 6. Сформировать набор из 30 транзакций на основе чеков покупок в магазине канцтоваров.

Вариант 7. Сформировать набор из 30 транзакций на основе чеков покупок в магазине бытовой техники.

Вариант 8. Сформировать набор из 30 транзакций на основе чеков покупок в кондитерском отделе.

Вариант 9. Сформировать набор из 30 транзакций на основе чеков покупок в магазине косметики.

Вариант 10. Сформировать набор из 30 транзакций на основе чеков покупок в магазине спорттоваров.

2. Варианты из репозиториев.

Вариант 1. Набор данных: http://archive.ics.uci.edu/ml/datasets/Online+Retail

Вариант 2. Набор данных: https://github.com/viktree/curly-octo-chainsaw/blob/master/BreadBasket_DMS.csv P.S. Можно предложить собственные «большие» наборы данных.

Задания 1. Применить для тестовых вариантов и вариантов из репозиториев различные алгоритмы поиска ассоциативных правил при одинаковых начальных условиях (при одинаковых пороговых значениях для поддержки и достоверности) и сравнить полученные результаты. Для тестовых вариантов выполнить ручные расчеты (например, с применением MS Excel) и расчеты с применением программных библиотек на языке Python.

Для вариантов из репозиториев выполнить расчеты с применением программных библиотек на языке Python. В качестве алгоритмов поиска ассоциативных правил использовать алгоритмы:

Apriori (https://pypi.org/project/apriori-python/);

Efficient Apriori (https://pypi.org/project/efficient-apriori/);

FPGrowth (https://pypi.org/project/fpgrowth-py/).

Задания 2. Сформировать базы ассоциативных правил с уровнем минимальной достоверности 60% и 80%. Вычислить для ассоциативных правил поддержку, достоверность, значимость.

Задания 3. Оценить время формирования искомых ассоциативных правил с применением различных алгоритмов и построить диаграммы, позволяющие выполнить сравнительный анализ.

Задания 4. Выполнить визуализацию ассоциативных правил (https://pypi.org/project/pyarmviz/)

Перечень докладов по теме 3. «Применение систем извлечения знаний методами Data Mining и интеллектуального анализа данных для комплексной оценки безопасности автоматизированных систем управления и разработки систем проактивной защиты информации на основе анализа событий в информационной системе»

1. Понятие информационного хранилища.
2. Принципы построения информационных хранилищ.
3. Задачи и средства администрирования ИАС.
4. Требования к качеству данных и способы его обеспечения при загрузке в информационное хранилище.
5. Проблемы, разрешаемые при приведении данных единой структуре информационного хранилища.
6. Концепции построения структур хранилищ данных.
7. Назначение, состав и выполняемые функции базы метаданных – репозитория хранилищ данных.
8. Принципы создания репозитория хранилищ данных.
9. Элементы моделей данных хранилища (факт-таблица, таблицы измерений, консольная таблица).
10. Схемы представления – модели многомерных данных.

8.3. Типовые контрольные задания или иные материалы для проведения ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

Форма проведения промежуточной аттестации по дисциплине: экзамен/ защита курсового проекта (курсовой работы) (по результатам накопительного рейтинга или в форме компьютерного тестирования

Устно-письменная форма по вопросам к экзамену предполагается, как правило, для сдачи академической задолженности.

Защита курсового проекта/ работы. *Результаты защиты курсового проекта/ работы выставляются по пятибалльной системе оценивания ("отлично", "хорошо", "удовлетворительно", "неудовлетворительно") с обязательным проставлением количества баллов, набранных в соответствии с балльно-рейтинговой системой (по стобалльной шкале).*

Перечень вопросов к защите курсового проекта

1. Языки программирования для анализа больших данных
2. Платные и бесплатные инструменты для обработки и анализа больших данных
3. Методы прогнозирования в Data Mining
4. Характеристика основных методов обнаружения вторжений
5. Особенности технологий Data Mining для обнаружения атак
6. Классификация задач Data Mining
7. Цель технологии Data Mining
8. 1. Какова структура сети Кохонена?
9. 2.В чем заключается процесс обучения без учителя (самоорганизация)?
10. 3. В чем отличие процесса обучения с учителем, от обучения без учителя?
11. 4. Какой нейрон называется нейроном-победителем в конкурентном обучении сети Кохонена?

Примерный перечень вопросов и заданий для подготовки к экзамену

ОПК-2. Способен понимать принципы работы современных информационных технологий и программных средств, в том числе отечественного производства, и использовать их при решении задач профессиональной деятельности

1. Перечислите задачи интеллектуального анализа данных

2. Что понимается под кластером объектов данных, в чем его преимущества перед классификацией
3. Перечислите требования к кластеризации в интеллектуальном анализе данных
4. Назовите категории, на которые можно разделить методы кластеризации
5. Приведите примеры методов кластеризации
6. Перечислите характеристики кластера и дайте их краткое описание:
7. Перечислите основные недостатки методов кластеризации
8. Назовите основные проблемы кластерного анализа
9. Название расстояния, используемое в алгоритмах кластеризации, вычисляемое по данной формуле

$$d_2(x_i, x_j) = \sqrt{\sum_{t=1}^{m} (x_{it} - x_{jt})^2} .$$

10. Название расстояния, используемое в алгоритмах кластеризации, вычисляемое по данной формуле

$$d_H(x_i, x_j) = \sum_{t=1}^{m} | x_{it} - x_{jt} | .$$

11. Название расстояния, используемое в алгоритмах кластеризации, вычисляемое по данной формуле

$$d_\infty(x_i, x_j) = \max_{1 \le t \le m} | x_{it} - x_{jt} |$$

12. Название расстояния, используемое в алгоритмах кластеризации, вычисляемое по данной формуле

$$d_M(x_i, x_j) = (x_i - x_j) S^{-1} (x_i - x_j)^t .$$

13. Раскройте определение Дерево решений (DecisionTree) —
14. По какому алгоритму осуществляется построение дерева
15. Классификация как метод интеллектуального анализа данных
16. Дайте характеристику простой классификации
17. Основными этапами процесса классификации являются:
18. Укажите цель процесса классификации
19. Назовите достоинства нейронной сети Кохонена
20. Назовите недостатки нейронной сети Кохонена

ПК-3 Способен к выполнению работ по управлению, настройке, разработке баз данных ИС

1. Что понимается под Data Mining? В каком году и кем было дано данное определение?
2. Какова классификация задач Data Mining?
3. Дайте краткое описание задаче анализа данных - классификация
4. Дайте краткое описание задаче анализа данных - регрессия
5. Дайте краткое описание задаче анализа данных - кластеризация
6. Дайте краткое описание задаче анализа данных - поиск ассоциативных правил
7. В чем заключается цель технологии Data Mining. Какие существуют виды моделей Data Mining?
8. Какие подходы и методы относят к базовым методам Data Mining? Укажите их недостатки
9. Дайте краткую характеристику генетическим алгоритмам
10. Дайте краткую характеристику нейронным сетям
11. Перечислите этапы интеллектуального анализа данных
12. Каково назначение стандарта CWM (Common WarehouseMetamodel) Data Mining?
13. Перечислите основные стандарты в области DataMining.
14. На какие концептуальные области разделен Пакет Data Mining стандарта CWM?
15. Какова структура стандарта CWM-DM?