

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«Оренбургский государственный университет»

Кафедра информационных систем и технологий

УТВЕРЖДАЮ
Декан факультета информационных технологий
Пищухин А.М.
(подпись, расшифровка подписи)



24 " апреля 2015 г

РАБОЧАЯ ПРОГРАММА

ДИСЦИПЛИНЫ

«Б.1.В.ДВ.4.2 Технология интеллектуального анализа данных»

Уровень высшего образования

БАКАЛАВРИАТ

Направление подготовки

09.03.02 Информационные системы и технологии
(код и наименование направления подготовки)

Общий профиль

(наименование направленности (профиля) образовательной программы)

Тип образовательной программы

Программа академического бакалавриата

Квалификация

Бакалавр

Форма обучения

Очная

Рабочая программа дисциплины «Б.1.В.ДВ.4.2 Технология интеллектуального анализа данных» /сост. М.М. Пирязев - Оренбург: ОГУ, 2014

Рабочая программа предназначена студентам очной формы обучения по направлению подготовки 09.03.02 Информационные системы и технологии

Содержание

1 Цели и задачи освоения дисциплины	4
2 Место дисциплины в структуре образовательной программы	4
3 Требования к результатам обучения по дисциплине	5
4 Структура и содержание дисциплины	5
4.1 Структура дисциплины	5
4.2 Содержание разделов дисциплины	6
4.3 Лабораторные работы	7
5 Учебно-методическое обеспечение дисциплины	7
5.1 Основная литература	7
5.2 Дополнительная литература	8
5.3 Периодические издания	8
5.4 Интернет-ресурсы	8
5.5 Методические указания к лабораторным занятиям	9
5.6 Программное обеспечение, профессиональные базы данных и информационные справочные системы современных информационных технологий	9
6 Материально-техническое обеспечение дисциплины	9
Лист согласования рабочей программы дисциплины	10
Дополнения и изменения в рабочей программе дисциплины	11
Приложения:	
Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине	
Методические указания для обучающихся по освоению дисциплины	

1 Цели и задачи освоения дисциплины

Цели освоения дисциплины: формирование у студентов теоретических знаний о современных методах интеллектуального анализа данных (Data mining), а также получение практических навыков программной реализации алгоритмов извлечения знаний из больших массивов данных.

Задачи:

- получить представление о состоянии и перспективе развития технологий анализа данных как направления теории искусственного интеллекта;
- изучить основные методы и алгоритмы интеллектуального анализа данных при решении задач обработки больших массивов данных;
- овладеть навыками программной реализации алгоритмов интеллектуального анализа данных с использованием библиотеки Xelopes.

2 Место дисциплины в структуре образовательной программы

Дисциплина относится к дисциплинам (модулям) по выбору вариативной части блока 1 «Дисциплины (модули)»

Пререквизиты дисциплины: *Б.1.Б.14 Технологии обработки информации, Б.1.В.ОД.2 Технологии программирования, Б.2.В.П.2 Научно-исследовательская работа (распределенная)*

Требования к входным результатам обучения, необходимым для освоения дисциплины

Предварительные результаты обучения, которые должны быть сформированы у обучающегося до начала изучения дисциплины	Компетенции
Знать: базовые концепции технологий программирования, основные этапы и принципы создания программного продукта, методологию построения алгоритмов, конструктивные компоненты и структуру компьютерных программ, основные структуры данных, способы их представления и обработки, преимущества использования объектно-ориентированного подхода при создании сложных программных продуктов, методы обработки исключений, методы и средства отладки и тестирования программ. Уметь: разрабатывать алгоритмы; создавать программы на языке программирования высокого уровня; использовать готовые библиотеки классов; применять современные системные и инструментальные программные средства для реализации алгоритмических и программных информационных технологий; выполнять отладку и тестирование программ; оформлять программную документацию. Владеть: навыками объектно-ориентированного программирования задач обработки данных в различных предметных областях; опытом отладки, тестирования и документирования программ.	ОПК-1 владением широкой общей подготовкой (базовыми знаниями) для решения практических задач в области информационных систем и технологий ПК-12 способностью разрабатывать средства реализации информационных технологий (методические, информационные, математические, алгоритмические, технические и программные)
Знать: основные методы математической статистики. Уметь: использовать методы математической статистики для обработки информации и проводить оценку адекватности результатов. Владеть: инструментальными средствами обработки информации методами математической статистики.	ПК-24 способностью обосновывать правильность выбранной модели, сопоставляя результаты экспериментальных данных и полученных решений
Знать: основные виды и процедуры обработки информации, модели и методы решения задач обработки информации. Уметь: осуществлять математическую и информационную постановку задач по обработке информации, использовать алгоритмы обработки информации для различных приложений.	ПК-25 способностью использовать математические методы обработки, анализа и синтеза результатов профессиональных

Предварительные результаты обучения, которые должны быть сформированы у обучающегося до начала изучения дисциплины	Компетенции
Владеть: инструментальными средствами обработки информации.	исследований

Постреквизиты дисциплины: *Б.2.В.П.1 Преддипломная практика*

3 Требования к результатам обучения по дисциплине

Процесс изучения дисциплины направлен на формирование следующих результатов обучения

Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций	Формируемые компетенции
Знать: отечественный и зарубежный опыт интеллектуального анализа данных; основные понятия и определения интеллектуального анализа данных; методы построения моделей и анализа зависимостей в больших массивах данных; современные программные средства для проектирования и разработки систем интеллектуального анализа данных. Уметь: проводить сбор, анализ научно-технической информации, отечественного и зарубежного опыта интеллектуального анализа данных; обосновывать выбор методов и средств реализации систем интеллектуального анализа данных при решении соответствующих практических задач; использовать современные программные средства для проектирования и исследования систем интеллектуального анализа данных; создавать законченные прикладные решения в области анализа данных на основе библиотеки Xelopes. Владеть: методами и инструментальными средствами интеллектуального анализа данных при реализации современных информационных технологий.	ПК-12 способностью разрабатывать средства реализации информационных технологий (методические, информационные, математические, алгоритмические, технические и программные) ПК-22 способностью проводить сбор, анализ научно-технической информации, отечественного и зарубежного опыта по тематике исследования ПК-25 способностью использовать математические методы обработки, анализа и синтеза результатов профессиональных исследований

4 Структура и содержание дисциплины

4.1 Структура дисциплины

Общая трудоемкость дисциплины составляет 4 зачетных единиц (144 академических часов).

Вид работы	Трудоемкость, академических часов	
	8 семестр	всего
Общая трудоёмкость	144	144
Контактная работа:	54,25	54,25
Лекции (Л)	28	28
Лабораторные работы (ЛР)	26	26
Промежуточная аттестация (зачет, экзамен)	0,25	0,25
Самостоятельная работа: - самоподготовка (проработка и повторение лекционного материала и материала учебников и учебных пособий; - подготовка к лабораторным занятиям; - подготовка к коллоквиумам; - подготовка к рубежному контролю и т.п.)	89,75	89,75

Вид работы	Трудоемкость, академических часов	
	8 семестр	всего
Вид итогового контроля (зачет, экзамен, дифференцированный зачет)	диф. зач.	

Разделы дисциплины, изучаемые в 8 семестре

№ раздела	Наименование разделов	Количество часов				
		всего	аудиторная работа			внеауд. работа
			Л	ПЗ	ЛР	
1	Хранилища данных и OLAP-системы	8	2	-	-	6
2	Data Mining	37	2	-	20	15
3	Классификация и регрессия	12	2	-	4	6
4	Ассоциативные правила	12	2	-	4	6
5	Кластеризация	12	2	-	4	6
6	Визуальный анализ данных	8	2	-	-	6
7	Анализ текстовой информации	8	2	-	-	6
8	Распределенный анализ данных и Data Mining в реальном времени	12	2	-	4	6
9	Средства анализа процессов	8	2	-	-	6
	Итого:	144	28		26	90
	Всего:	144	28		26	90

4.2 Содержание разделов дисциплины

1 Хранилища данных и OLAP-системы. Хранилища данных (ХД). Концепция хранилища данных. Организация ХД. Очистка данных. Концепция хранилища данных и анализ.

OLAP-системы. Многомерная модель данных. Определение OLAP-систем. Концептуальное многомерное представление. Архитектура OLAP-систем

2 Data Mining. Добыча данных - Data Mining. Задачи Data Mining. Практическое применение Data Mining. Модели Data Mining. Методы Data Mining. Процесс обнаружения знаний. Управление знаниями (Knowledge Management). Средства Data Mining.

Стандарты Data Mining. Стандарт CWM. Стандарт CRISP. Стандарт PMML. Другие стандарты Data Mining.

Библиотека Xelopes

3 Классификация и регрессия. Классификация и регрессия. Постановка задачи. Представление результатов. Методы построения правил классификации. Методы построения деревьев решений. Методы построения математических функций. Прогнозирование временных рядов

4 Ассоциативные правила. Поиск ассоциативных правил. Постановка задачи. Представление результатов. Алгоритмы

5 Кластеризация. Кластеризация. Постановка задачи кластеризации. Представление результатов. Базовые алгоритмы кластеризации. Адаптивные методы кластеризации

6 Визуальный анализ данных. Визуальный анализ данных - Visual Mining. Выполнение визуального анализа данных. Характеристики средств визуализации данных. Методы визуализации

7 Анализ текстовой информации. Анализ текстовой информации - Text Mining. Задача анализа текстов. Извлечение ключевых понятий из текста. Классификация текстовых документов. Методы кластеризации текстовых документов. Задача аннотирования текстов. Средства анализа текстовой информации

8 Распределенный анализ данных и Data Mining в реальном времени. Распределенный анализ данных. Системы мобильных агентов. Использование мобильных агентов для анализа данных. Система анализа распределенных данных.

Data Mining в реальном времени (Real-Time Data Mining). Идея Data Mining в реальном времени. Рекомендательные машины. Инструменты Data Mining в реальном времени

9 Средства анализа процессов. Средства анализа процессов - Process Mining. Автоматизация выполнения бизнес-процессов. Анализ процессов. Методы Process Mining. Библиотека алгоритмов Process Mining – ProM

4.3 Лабораторные работы

№ ЛР	№ раздела	Наименование лабораторных работ	Кол-во часов
1	2	Изучение библиотеки Xelopes. Диаграмма Model	2
2	2	Изучение библиотеки Xelopes. Диаграмма Settings	2
3	2	Изучение библиотеки Xelopes. Диаграмма Attribute	2
4	2	Изучение библиотеки Xelopes. Диаграмма Algorithms	2
5	2	Изучение библиотеки Xelopes. Диаграмма DataAccess	2
6	2	Изучение библиотеки Xelopes. Диаграмма Transformation	2
7	2	Знакомство с GUI-интерфейсом библиотеки алгоритмов Data Mining	2
8	2	Использование библиотеки Xelopes. Реализация алгоритмов построения моделей unsupervised	3
9	2	Использование библиотеки Xelopes. Реализация алгоритмов построения моделей supervised	3
10	3	Использование библиотеки Xelopes. Решение задачи поиска ассоциативных правил	4
11	4	Использование библиотеки Xelopes. Решение задачи кластеризации	4
12	5	Использование библиотеки Xelopes. Решение задачи классификации	4
13	8	Использование библиотеки Xelopes. Разработка агента интеллектуального анализа данных	4
		Итого:	26

5 Учебно-методическое обеспечение дисциплины

5.1 Основная литература

1 Барсегян, А. А. Анализ данных и процессов [Электронный ресурс]: учеб. пособие / А. А. Барсегян, М. С. Куприянов, И. И. Холод, М. Д. Тесс, С. И. Елизаров. — 3-е изд., перераб. и доп. . – Электрон. текстовые дан. - СПб.: БХВ-Петербург, 2009. - 512 с. - ISBN 978-5-9775-0368-6. – Режим доступа: <http://znanium.com/catalog.php?bookinfo=350638>.

2 Чубукова, И.А. Data Mining : учеб. пособие / И.А. Чубукова. - 2-е изд., испр. - М. : Интернет-Университет Информационных Технологий, 2008. - 383 с. - (Основы информационных технологий). - ISBN 978-5-94774-819-2 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=233055>.

3 Барсегян, А. А. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP [Электронный ресурс] : учебное пособие / А. А. Барсегян, М. С. Куприянов, В. В. Степаненко. - 2-е изд., перераб. и доп. – Электрон. текстовые дан. - СПб.: БХВ-Петербург, 2007. - 384 с. - ISBN 5-94157-991-8. – Режим доступа: <http://znanium.com/catalog.php?bookinfo=489445>.

4 Чубукова, И. А. Data Mining [Текст] : учеб. пособие / И. А. Чубукова. - М. : ИНТУИТ. РУ : БИНОМ. Лаборатория знаний, 2006. - 382 с. : ил. + табл. - (Основы информационных технологий). - Библиогр.: с. 375-382. - ISBN 5-9556-0064-7. - ISBN 5-94774-522-4.

5.2 Дополнительная литература

- 1 Дайитбегов, Д. М. Компьютерные технологии анализа данных в эконометрике [Текст] : монография / Д. М. Дайитбегов.- 3-е изд., доп. - Москва : ИНФРА-М : Вузовский учебник, 2013. - 587 с. - (Научная книга) - ISBN 978-5-9558-0275-6. - ISBN 978-5-16-006145-0.
- 2 Дайитбегов, Д. М. Компьютерные технологии анализа данных в эконометрике [Электронный ресурс]: монография / Д. М. Дайитбегов.- 3-е изд., доп. - Электрон. текстовые дан. - Москва : ИНФРА-М : Вузовский учебник, 2013. - 587 с. - (Научная книга) - ISBN 978-5-9558-0275-6. - ISBN 978-5-16-006145-0. - Режим доступа: <http://znanium.com/catalog.php?bookinfo=365692>.
- 3 Информационные аналитические системы [Электронный ресурс] : учебник / Т. В. Алексеева, Ю. В. Амириди, В. В. Дик и др.; под ред. В. В. Дика.- Электрон. текстовые дан. - М.: МФПУ Синергия, 2013. - 384 с. - (Университетская серия). - ISBN 978-5-4257-0092-6. - Режим доступа: <http://znanium.com/catalog.php?bookinfo=451186>.
- 4 Низаметдинов, Ш.У. Анализ данных : учебное пособие / Ш.У. Низаметдинов, В.П. Румянцев. - М. : МИФИ, 2012. - 286 с. - ISBN 978-5-7262-1687-4 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=231829>.
- 5 Симчера, В. М. Методы многомерного анализа статистических данных [Текст] : учеб. пособие для вузов / В. М. Симчера. - М. : Финансы и статистика, 2008. - 400 с. - Библиогр. в конце гл. - Слов. терминов: с. 372-391. - ISBN 978-5-279-03184-9.
- 6 Симчера, В.М. Методы многомерного анализа статистических данных : учебное пособие / В.М. Симчера. - М. : Финансы и статистика, 2008. - 398 с. - ISBN 978-5-279-03184-9 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=59559>.
- 7 Microsoft SQL Server 2005 Analysis Services. OLAP и многомерный анализ данных [Текст] : наиболее полное руководство / А. Б. Бергер [и др.]; общ. ред. А. Бергера, И. Горбач. - СПб. : БВХ-Петербург, 2007. - 928 с. : ил. - Предм. указ.: с. 901-905. - ISBN 978-5-94157-158-1.
- 8 Брандт, З. Анализ данных = Data Analysis. Statistical and Computational Methods for Scientists and Engineers [Текст] : стат. и вычисл. методы для науч. работников и инженеров / З. Брандт; пер. с англ. О. И. Волковой. - М. : Мир : АСТ, 2003. - 686 с. : ил. - Парал. тит. л. англ. - Библиогр.: с. 651-654. - Предм. указ.: с. 668-674. - ISBN 5-03-003478-1. - ISBN 5-17-019778-0.

5.3 Периодические издания

- 1 Информационные технологии : журнал // Информационные технологии с ежемесячным приложением. - М. : Агенство "Роспечать".
- 2 Прикладная информатика : журнал. - М. : АПР.
- 3 Компьютерпресс : журнал. - М. : АРЗИ.
- 4 Программные продукты и системы : журнал. - М. : Агенство "Роспечать".
- 5 Открытые системы. СУБД : журнал. - М. : Агенство "Роспечать".

5.4 Интернет-ресурсы

- 1 Интеллектуальный анализ данных [Электронный ресурс]. – Режим доступа: http://www.machinelearning.ru/wiki/index.php?title=Data_Mining.
- 2 Чубукова, И.А. Data Mining: учебный курс [Электронный ресурс]. – Режим доступа: <http://www.intuit.ru/departament/database/datamining/>.
- 3 Зиновьев, А. Ю. Визуализация многомерных данных [Электронный ресурс]. – Режим доступа: <http://pca.narod.ru/ZINANN.htm>.

5.5 Методические указания к лабораторным занятиям

Для выполнения лабораторных работ используются разработанные в электронном виде методические указания к лабораторным занятиям, содержащие условия задач, примеры и указания к их решению.

5.6 Программное обеспечение, профессиональные базы данных и информационные справочные системы современных информационных технологий

Для обеспечения лабораторных занятий необходимо программное обеспечение:

- операционная система Microsoft Windows;
- интегрированная среда разработки Microsoft Visual Studio;
- система управления базами данных Microsoft SQL Server.

6 Материально-техническое обеспечение дисциплины

Лабораторные занятия проводятся в компьютерных классах. Все компьютеры оснащены лицензионным программным обеспечением, необходимым для реализации данной дисциплины. Для чтения мультимедийных лекций компьютерные классы оснащены проекторами и экранами.

ЛИСТ

согласования рабочей программы

Направление подготовки: 09.03.02 Информационные системы и технологии

код и наименование

Профиль: Общий профиль

Дисциплина: Б.1.В.ДВ.4.2 Технология интеллектуального анализа данных

Форма обучения: очная

(очная, очно-заочная, заочная)

Год набора 2014

РЕКОМЕНДОВАНА заседанием кафедры
информационных систем и технологий

наименование кафедры

протокол № 8 от "15" апреля 2015 г.

Ответственный исполнитель, заведующий кафедрой
информационных систем и технологий

наименование кафедры

подпись

Чепасов В.И.

расшифровка подписи

дата

Исполнители:

доцент

должность

подпись

Пирязев М.М.

расшифровка подписи

дата

должность

подпись

расшифровка подписи

дата

СОГЛАСОВАНО:

Председатель методической комиссии по направлению подготовки

09.03.02 Информационные системы и технологии

код наименование

личная подпись

Чепасов В.И.

расшифровка подписи

дата

Заведующий отделом комплектования научной библиотеки

личная подпись

Истомина Т.В.

расшифровка подписи

дата

Начальник отдела информационных образовательных технологий ЦИТ

личная подпись

Дырдина Е.В.

расшифровка подписи

дата

Внесенные изменения на 2015 год набора

" 24 " апреля 2015 г.

data