
МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение высшего
профессионального образования

«ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ УПРАВЛЕНИЯ И
РАДИОЭЛЕКТРОНИКИ» (ТУСУР)

УТВЕРЖДАЮ

Заведующий кафедрой ЭМИС

_____ И. Г. Боровской

«__» _____ 2012 г.

Боровской И.Г. Колесникова С.И.

Научно-исследовательская работа магистра

*Методические указания по выполнению самостоятельных работ
для направления 23010.68 «Информатика и вычислительная техника»*

2012

Составители: Боровской И.Г., Колесникова С.И., каф.ЭМИС

АННОТАЦИЯ

Цели настоящих методических указаний: 1) освоение основных принципов и методов научного исследования (анализа данных); 2) приобретение практических навыков в построении постановок прикладных задач и разработки алгоритмов их анализа, поиска закономерностей и распознавания характерных черт в поведении сложных объектов, анализа качества алгоритмов. В четырех частях указаний приведены примеры организации научных семинаров, где обсуждаются постановки задач и методы их решения (анализа методов их возможного решения) на следующие темы:

1. Математические основы формализации и исследования предметных областей и современных прикладных задач. Логическая схема научного исследования. Приоритетные направления исследований.
2. Модели и методы анализа временных рядов и прогнозирования состояний сложных объектов по стохастическим временным рядам.
3. Модели и методы интеллектуального анализа данных и их использование для управления сложными системами. Алгебраический подход к решению слабоформализованных прикладных задач.
4. Методы корректной обработки статистических данных и форм представления результатов при осуществлении экспериментальных исследований (при оформлении магистерской диссертации).

Ссылки на теоретический материал приведены *только те и в том объеме*, которые были актуальны на момент составления пособия и могут дополняться другими источниками, однако, приоритет отдан классическим источникам, в которых зачастую уже есть решение поставленной задачи в более формализованном виде, «расшифровать» которое не всегда удается. Задания к семинарам являются постановками задач, востребованными на мировом ранке, предназначенными не только для усвоения принципов и основ современных методов исследования и анализа данных. Предполагается, что студенты знают все дисциплины учебного плана данного направления.

Методические указания предназначены для студентов экономического факультета.

СОДЕРЖАНИЕ МЕТОДИЧЕСКИХ УКАЗАНИЙ К САМОСТОЯТЕЛЬНОЙ РАБОТЕ СТУДЕНТОВ

по дисциплине «Анализ данных»

для студентов направления 230100.62 – Информатика и вычислительная техника.

Профиль - Автоматизированное управление бизнес-процессами и финансами

Краткое содержание тем и планируемых результатов их освоения.....	3
УКАЗАНИЯ К САМОСТОЯТЕЛЬНОЙ РАБОТЕ СТУДЕНТОВ.....	5
ПОДГОТОВКА К научным проблемно-ориентированным интерактивным семинарам....	5
1. УКАЗАНИЯ К САМОСТОЯТЕЛЬНОЙ РАБОТЕ СТУДЕНТОВ по разделу 1.....	5
Примеры обсуждаемых проблем и составляющих их задач	6
2. УКАЗАНИЯ К САМОСТОЯТЕЛЬНОЙ РАБОТЕ СТУДЕНТОВ по разделу 2.....	11
Примеры обсуждаемых задач и соответствующих алгоритмов	12
3. УКАЗАНИЯ К САМОСТОЯТЕЛЬНОЙ РАБОТЕ СТУДЕНТОВ по разделу 3.....	15
Примеры обсуждаемых задач и соответствующих алгоритмов	15
4. УКАЗАНИЯ К САМОСТОЯТЕЛЬНОЙ РАБОТЕ СТУДЕНТОВ по разделу 4.....	17
Примеры обсуждаемых задач (магистерских работ по каф. ЭМИС прошлых лет).17	
Примеры решения практических задач по обработке данных.....	19
Использованная литература	20

Обозначения: ИДЗ - индивидуальные домашние задания

СРС - самостоятельная работа студентов

НПС - научные проблемно-ориентированные интерактивные семинары

З-Эл – знания элементарные (определения, понятия, умение приводить иллюстрирующие примеры);

З-Пр – знания продуктивные (умение применить знания элементарные для решения учебных задач);

У-Эл – «умения» элементарные (уметь пользоваться готовыми частными алгоритмами для решения типовых задач), умение решать задачи по шаблону (копировать);

У-Пр – «умения» продуктивные (применять положения и известные частные алгоритмы дисциплины для решения практических задач);

В-Эл – элементарное владение методами дисциплины и уверенное осуществление (построение) основных операций для решения типовых задач;

В-Пр – продуктивно распознавать проблемы, алгоритмизировать их анализ и применять методы дисциплины для решения практических задач;

Краткое содержание тем и планируемых результатов их освоения

Таблица 0.1

Тема практических занятий	Деятельность студента. Решая задачи, студент:	Отрабатываемые компетенции и/ожидаемый уровень освоения
1. Математические основы формализации и исследования предметных областей и	<ul style="list-style-type: none"> • <i>использует</i> определения теории вероятностей и математической статистики, анализа данных, теории управления, теории распознавания образов, теории массового обслуживания, теории алгоритмов и программирования; • <i>выбирает</i> способ построения модели 	ОК-4, ОК-5, ОК-6, ОК-7 / Уровни: З-Эл, У-Эл, В-Эл ПК-3; ПК-4; ПК-

Составители: Боровской И.Г., Колесникова С.И., каф.ЭМИС

<p>современных прикладных задач. Логическая схема научного исследования. Приоритетные направления исследований.</p>	<p>анализируемого объекта;</p> <ul style="list-style-type: none"> • <i>устанавливает</i> роль агента для анализа объекта; • <i>использует</i> знания, полученные ранее и самостоятельно решает задачи выбора методики исследования; • <i>учится</i> применять статистические критерии; • совместно с преподавателем <i>разрабатывает</i> методику решения сложных (плохоформализуемых) задач. • <i>использует</i> понятия интеллектуального анализа данных и Data mining; 	<p>6/ Уровни: 3-Пр, У-Пр, В-Пр</p>
<p>2 Модели и методы анализа временных рядов и прогнозирования состояний сложных объектов по стохастическим временным рядам.</p>	<ul style="list-style-type: none"> • <i>знакомится</i> с методами анализа временных рядов; • <i>изучает</i> свойства хаотических временных рядов; • <i>решает</i> задачи прогнозирования стохастических временных рядов на базе известных методов и алгоритмов; • <i>изучает</i> модели оценивания параметров модели. • <i>использует</i> методы инвариантных многообразий для управления сложными системами. 	<p>ОК-4, ОК-5, ОК-6, ОК-7 / Уровни: 3-Эл, У-Эл, В-Эл</p> <p>ПК-3; ПК-4; ПК-6/ Уровни: 3-Пр, У-Пр, В-Пр</p>
<p>3. Модели и методы интеллектуального анализа данных и их использование для управления сложными системами. Алгебраический подход к решению слабоформализованных прикладных задач.</p>	<ul style="list-style-type: none"> • <i>учится</i> корректно ставить задачи классификации и прогнозирования; • <i>определяет</i> корректные методы исследования выборок; • <i>определяет</i> коллективы мер близости выборок; • <i>изучает</i> алгоритмические композиции анализа и обработки выборок; • <i>применяет</i> адаптивные методы сглаживания, выделения тренда, оценки дисперсии временных рядов и других выборок; • <i>анализирует</i> системы анализа распределённых данных и принципы их построения; • <i>знакомится</i> с существующими стандартами Data mining; • <i>использует</i> методы инвариантных многообразий для управления сложными системами. 	<p>ОК-4, ОК-5, ОК-6, ОК-7 / Уровни: 3-Эл, У-Эл, В-Эл</p> <p>ПК-3; ПК-4; ПК-6/ Уровни: 3-Пр, У-Пр, В-Пр</p>
<p>4 Методы корректной обработки статистических данных и форм представления результатов при осуществлении экспериментальных исследований (при оформлении магистерской диссертации).</p>	<ul style="list-style-type: none"> • <i>анализирует</i> системы анализа распределённых данных и принципы их построения; • <i>знакомится</i> с существующими стандартами автоматизированной обработки данных; • <i>изучает</i> документацию по написанию научных статей, диссертаций. 	<p>ОК-4, ОК-5, ОК-6, ОК-7 / Уровни: 3-Эл, У-Эл, В-Эл</p> <p>ПК-3; ПК-4; ПК-6/ Уровни: 3-Пр, У-Пр, В-Пр</p>

УКАЗАНИЯ К САМОСТОЯТЕЛЬНОЙ РАБОТЕ СТУДЕНТОВ.

Подготовка к научным проблемно-ориентированным интерактивным семинарам

1. Ознакомиться со справочными интернет-сведениями (СРС)
2. Ознакомиться с указанной темой в основной и дополнительной литературе.

Основная литература

1. З. Брандт. Анализ данных. Статистические и вычислительные методы для научных работников и инженеров : Пер. с англ. : Учебное пособие / З. Брандт ; пер. : О. И. Волкова; ред. пер. : Е. В. Чепурин. - М. : Мир, 2003 ; М. : АСТ, 2003. – 686 с.
2. Лапко А.В. Непараметрические системы обработки информации : Учебное пособие для вузов / А. В. Лапко, С. В. Ченцов; Российская Академия наук. Сибирское отделение, Институт вычислительного моделирования. - М. : Наука, 2000. - 349 с.
3. Загоруйко Н.Г. Прикладные методы анализа данных и знаний. - Новосибирск: ИМ СО РАН, 1999, стр. 270 http://www.sernam.ru/book_zg.php

Дополнительная литература

1. Горелик А. Л., Скрипкин В. А. Методы распознавания: Учебное пособие для вузов. - 4-е изд., испр. - М.: Высшая школа, 2004. – 260 с.
 2. Воронцов К.В. Лекции по методам оценивания и выбора моделей. 2007. Режим доступа: www.ccas.ru/voron/download/Modeling.pdf.
 3. Воронцов К.В. Обзор современных исследований по проблеме качества обучения алгоритмов. Таврический вестник информатики и математики. – 2004. – № 1. – С. 5 – 24. <http://www.ccas.ru/frc/papers/voron04twim.pdf>
 4. Дюк В. Обработка данных на ПК в примерах / Вячеслав Дюк. - СПб.:Питер, 1997. - 240с.
 5. Прикладная статистика. Классификация и снижение размерности : справочное издание / С. А. Айвазян [и др.] ; ред. С. А. Айвазян. - М. : Финансы и статистика, 1989. - 608 с.
 6. Тюрин, Ю.Н.. Анализ данных на компьютере : учебное пособие для вузов / Ю. Н. Тюрин, А. А. Макаров. - 4-е изд., перераб. - М. : Форум, 2008. - 366,
 7. Кокс, Д. Р. Анализ данных типа времени жизни : Пер. с англ. / Д. Р. Кокс, Д. Оукс; Пер. О. В. Селезнева, Ред. Ю. К. Беляев, Предисл. Ю. К. Беляев. - М. : Финансы и статистика, 1988. – 189с.
http://www.sernam.ru/book_zg.php
<http://www.statlab.kubsu.ru/node/4>
<http://www.intuit.ru/department/database/dataanalysis/>
<http://itteach.ru/predstavlenie-znaniy/metodi-i-sredstva-intellektualnogo-analiza-dannich>
<http://www.bsu.ru/content/hec/biometria/modules/stdatmin.html>
- Другие источники, согласно теме предстоящего доклада.

3. Ознакомиться с принципом постановок прикладных задач.
4. Рекомендуется привести примеры постановок прикладных задач (в рамках изучения обзоров научной литературы известных ученых мировой), «близких» к рассматриваемой.
5. Ознакомиться с планом проведения интерактивных занятий-семинаров в случае их проведения, прилагающегося к каждому разделу, и принципом подготовки к нему.
6. Составить и предоставить преподавателю вариант доклада о работе, если он входит в форму отчетности по данному разделу знаний. Обсудить детали доклада.
7. Представить доклад-презентацию о проведенной исследовательской работе на семинаре.

1. УКАЗАНИЯ К САМОСТОЯТЕЛЬНОЙ РАБОТЕ СТУДЕНТОВ по разделу 1

Тема. Формализации постановок прикладных задач. Основы теории анализа данных (данные, знания, гипотеза, закономерность, признак). Логическая схема научного исследования. Анализ вариантов постановок прикладных задач в магистерских диссертациях прошлых лет. Известные постановки нерешенных прикладных задач. Обзорные статьи ведущих ученых мира.

ЦЕЛЬ РАБОТЫ

Выбор направлений исследований, определение проблем и вытекающих из нее частных целей и задач. Определяется цель исследования, обосновывается предмет и объект исследования. Подготовка к исследованию. Определение логики исследования.

Знакомство с основными математическими дисциплинами, изучающими закономерности массовых случайных явлений и основными этапами исследования любой сложной (в том числе, плохоформализуемой) проблемы:

- 1) определение проблемы;
разработка подхода (подходов в случае коллективного принятия решения) к решению проблемы;
- 2) разработка плана исследования;
- 3) сбор (определение) данных (обучающих и контрольных)
- 4) предобработка данных;
- 5) исследование и решение задачи;
- 6) подготовка отчета и презентации-доклада.
- 7) уточнить примеры типовых заданий на СРС.

Форма текущего контроля освоения компетенций ОК-4, ОК-5, ОК-6, ОК-7, уровни З-Эл, У-Эл, В-Эл; ПК-3; ПК-4; ПК-6, уровни З-Пр, У-Пр, В-Пр (см. табл.1): *отчет* – в форме доклада на семинарах.

Примеры обсуждаемых проблем и составляющих их задач

Проблема 1.1. *Исследование автоматизированных системы обработки данных.* Назначение. Эффективность (параметры оценки). Примеры удачных и «неудачных» разработок АИС в различных областях.

Постановки задач: АИС в медицинских обследованиях; АИС в геологии и геофизики.

Проблема 1.2. *Исследование возможностей (недостатков) численных методов для решения расчетных физических задач.*

Постановки задач:

Задача теплопроводности.

Задача расчета гидродинамических сопротивлений.

Задача диффузии с подвижной границей.

Проблема 1.3. *Исследование современного состояния методов разработки Web-приложений.*

Постановки задач:

Принципы разработки интернет-магазина.

Web-приложение для работы с базой данных "Геофизические сигналы" (заказ лаборатории ТПУ).

Системы интернет-тестирования: анализ достоинств и недостатков. Уровни защиты системы тестирования

Указания 1.3.

Дополнительная литература.

1. Кларк Дж. Системология. Автоматизация решения системных задач / Пер. с англ. А.И. Горлин. – М.: Радио и связь, 1990. – 538 с.
2. Поиск электромагнитных предвестников землетрясений / Под ред. М.Б. Гохберга. - М.: ИФЗ АН СССР, 1988. - 167 с.
3. Степанов М. В. Изучение краткосрочных предвестников землетрясений методом электромагнитной эмиссии: Дис. канд. физ.-мат. наук. М., 2004.
4. Семенов Н.А. Методы автоматизированного проектирования системы прогнозирования землетрясений: Дис. канд. техн. наук. Санкт-Петербург, 2007. 136с.

Проблема 1.4. *Исследование современного состояния криптографии.* Анализ недостатков.

Постановки задач:

Применение хаотических систем для кодирования информации.

Система AES шифрования/дешифрования на базе хаотической динамики.

Особенности применения нелинейных систем в качестве хаотического генератора.

Указания 1.4.

Дополнительная литература.

1. L. Kosarev. Chaos-based cryptography: a brief overview //Circuits and systems. - 2001- №. 3. - С. 6–21.
2. Птицын Н. Приложение теории детерминированного хаоса в криптографии. - Изд.: МГТУ им. Н.Э. Баумана. – 2002.- 81 с.
3. Болтенков В.А., Никольский Е.С. Анализ алгоритмов хаотического шифрования изображений // Цифровые технологии. – 2010. – № 7. – С. 61-66.

Проблема 1.5. *Исследование современного состояния результативности методов математической статистики.* Анализ недостатков. Ложные регрессии и последствия их использования. Основы непараметрической статистики. Цели применения непараметрических методов. Работа с малыми выборками. Непараметрические критерии: критерии тест Манна-Уитни, W-тест Уилкоксона и др. Условия применимости.

Постановки задач:

Апробировать критерии на короткой выборке.

Апробировать критерии на выборке неустойчивой статистически.

Исследование стохастических временных рядов.

Указания 1.5.

Дополнительная литература.

1. Сошникова Л.А., Тамашевич В.Н. и др. Многомерный статистический анализ в экономике: Учеб. Пособие для вузов/Под ред. проф. Тамашевича. – М.: ЮНИТИ-ДАНА, 1999. –598с.
2. http://ru.wikipedia.org/wiki/%D1%F2%E0%F2%E8%F1%F2%E8%F7%E5%F1%EA%E8%E9_%EA%F0%E8%F2%E5%F0%E8%E9
3. Полигон алгоритмов [Электронный ресурс]: система для массового тестирования алгоритмов классификации на реальных задачах. – Электрон.дан. –М., [2011]. – Режим доступа: <http://poligon.machinelearning.ru/>, свободный. – Загл. с экрана.
4. Гольденберг Л.М., Матюшкин Б.Д., Поляк М.Н. Цифровая обработка сигналов. Справочник – М.: Радио и связь, 1985. 312 с.
5. Афанасьев В.Н., Юзбашев М.М. Анализ временных рядов и прогнозирование. – М.:

Финансы и статистика, 2001. – 228 с

6. Безручко Б.П., Смирнов Д.А. Математическое моделирование и хаотические временные ряды. – Саратов: ГосУНЦ «Колледж», 2005. – 320 с.

Ряды данных: результаты пулковских и международных наблюдений международной службы вращения Земли (<http://hpiers.obspm.fr/>), в дальнейшем C01 и C02;

Проблема 1.6. *Исследование современного состояния теории управления. Анализ недостатков по обзорным статьям академика А.А.Красовского.*

Постановки задач:

Аналитические методы управления. Условия применения. Достигнутые результаты.

Методы управления сложными нелинейными системами.

Метод инвариантных многообразий для управления сложными системами (Халил, А.А.Колесников).

Указания 1.6. Разобрать работы А.А.Колесникова. Разобрать работы Халила.

Дополнительная литература

1. Колесников А.А. Основы синергетической теории управления. – М.: ИСПО-Сервис, 2000. – 260 с.

2. Колесников А.А. Синергетические методы управления сложными системами: теория системного синтеза. – М.: Едиториал УРСС, 2005. – 230 с.

Проблема 1.7. *Исследование современных методов интеллектуального анализа данных (ИАД). Анализ недостатков по обзорным статьям ВЦ РАН (К.В.Воронцов).*

Постановки задач:

Обзор систем анализа распределённых данных и принципы их построения.

Существующие стандарты Data mining и OLAP.

Понятие «мобильных агентов» и системы мобильных агентов, используемых для анализа данных.

Указания 1.7. Разобрать методы Data Mining:

- комплексный системный анализ;
- краткосрочный и долгосрочный прогноз развития ситуаций;
- выработка вариантов оптимизационных решений;
- прогнозирование ситуаций в зависимости от некоторых параметров технологического процесса;
- обнаружение скрытых закономерностей развития процессов;
- прогнозирование закономерностей развития процессов;
- обнаружение скрытых факторов влияния;
- обнаружение и идентификация ранее неизвестных взаимосвязей между параметрами и факторами влияния;
- анализ среды взаимодействия процессов и прогнозирование изменения ее характеристик;
- выработка оптимизационных рекомендаций по управлению процессами;
- визуализация результатов анализа, подготовку предварительных отчетов и проектов допустимых решений с оценками достоверности и эффективности возможных реализаций.

Проблема 1.8. *Классические основы теории временных рядов и системы их прогнозирования. Моделирование рядов динамики.*

Постановки задач:

Методы нелинейной фильтрации стохастических временных рядов.

Методы модовой декомпозиции EMD (Empirical Mode Decomposition) и преобразование Гильберта-Хуанга ННТ (Hilbert-Huang Transform).
Анализ свойств временных рядов, порождаемых хаотическими моделями. Роль хаотических моделей в описании реальных ситуаций.
Свойства одномерной хаотической модели Хатчинсона.
Свойства одномерной хаотической модели Фейгенбаума

Указания 1.8.

- Классифицировать методы анализа стационарных стохастических временных рядов.
- Классифицировать методы анализа нестационарных стохастических временных рядов.
- Классифицировать типы нестационарности и методы обнаружения характера нестационарности.
- Суть нелинейной фильтрации сигнала с аддитивным шумом.
- Виды ядерных функций. Непараметрические оценки плотности вероятности типа Розенблатта – Парзена.
- Динамический хаос. Одномерные хаотические модели.

Дополнительная литература

1. Давыдов В.А., Давыдов А.В. Очистка геофизических данных от шумов с использованием преобразования Гильберта-Хуанга.// Электронное научное издание "Актуальные инновационные исследования: наука и практика", 2010, № 1. <http://www.actualresearch.ru>.
 2. Афанасьев В.Н., Юзбашев М.М. Анализ временных рядов и прогнозирование. – М.: Финансы и статистика, 2001. – 228 с.
 3. Бассвиль. М., Банвениста А. Обнаружение изменения свойств сигналов и динамических систем. Пер. с англ.- М.: Мир, 1989. - 278 с.
- Безручко Б.П., Смирнов Д.А. Математическое моделирование и хаотические временные ряды. – Саратов: ГосУНЦ «Колледж», 2005. – 320 с.

Проблема 1.9. *Возможности нейро-алгоритмов и их реализация в реальном времени.*

Постановки задач:

Классификационные правила и их реализация в нейросети.

Классификация на базе конструирования нейросети в специализированном нейропакете».

Примеры успешного решения нелинейных проблем на базе нейро-алгоритмов.

Нелинейное управление на базе нейросети. Примеры.

Указания 1.9.

Дополнительная литература для подготовки к занятию:

- 1) <http://www.alleng.ru/d/econ/econ292.htm> Юдин С.В. Математика в экономике. Тула: РГТЭУ, 2009. — 228 с.
- 2) http://www.aup.ru/books/m155/2_12.htm Орлов А.И. Математика случая. Вероятность и статистика – основные факты. Учебное пособие. М.: МЗ-Пресс, 2004.

Ознакомиться со следующим материалом (по указанным источникам):

- 1) Нейросеть и реализация в специализированном компьютерном нейропакете.
- 2) Классификационные правила и их реализация в нейросети;
- 3) Возможности нейро-алгоритмов.
- 4) Примеры успешного решения нелинейных проблем на базе нейро-алгоритмов.

Проблема 1.10. Алгебраический подход к решению слабоформализованных прикладных задач.

Постановки задач:

Анализ прогностических свойств алгебраических композиций базовых алгоритмов. Сравнительный анализ методов прогнозирования финансовых рядов на основе отдельных (базовых) алгоритмов.

Построение алгебраических композиций для прогнозирования по временным рядам на базе трех методов: авторегрессии, метода ARADS и генетического алгоритма.

Указания 1.10.

- Классифицировать базовые алгоритмы анализа стохастических временных рядов.
- Классифицировать методы фильтрации нестационарных стохастических временных рядов.
- Классифицировать методы выделения тренда в стохастических временных рядах.
- Классифицировать методы анализа динамического хаоса.

Дополнительная литература

1. Рудаков К.В., Чехович Ю.В. Алгебраический подход к проблеме синтеза обучаемых алгоритмов выделения трендов // Доклады РАН. - 2003. -Т. 388. - № 1.- С. 33-36.
2. Журавлев Ю.И., Рязанов В.В., Сенько О.В. Распознавание. Математические методы. Программная система. Практические применения. – Москва: Фазис. – 2006. – 176 с.
3. Журбенко И.Г., Кожевникова И.А. Стохастическое моделирование процессов. – М.: МГУ, 1990. – 147 с.
4. Загоруйко Н.Г. Прикладные методы анализа данных и знаний. – Изд. ИМ СО РАН, Новосибирск, 1999 г. – 273 с.
5. Загоруйко Н.Г. Интеллектуальный анализ данных, основанный на функции конкурентного сходства // Автометрия. – 2008. – Т. 44. – №3. – С. 31-40.

Подготовка занятий по теме №1. Выбор ведущих студентов, ответственных за выбор и подачу необходимой информации и обсуждение с ним алгоритма занятий.

Таблица 1.1

№	№ задач и	Вид (совмещение нескольких видов) интерактивной работы	Трудоемкость (час)	Отрабатываемые компетенции/ожидаемый уровень освоения	Оценка личностных качеств	Контроль выполнения работы (участие в полемике, доклад, реферат)
1	1.1-1.10	Работа в команде: распределение задач. Решение задач: этапность исследования.	4×4	ОК-4, ОК-5, ОК-6, ОК-7	Качество работы; своевременность докладов по выбранным постановкам проблем	Критерии оценивания поведения на занятии: активность, инициативность, грамотность, обоснованность защищаемой позиции.
Всего			16			

Вступление. Сообщение темы и обоснование ее актуальности через вышеуказанные проблемы.

Основная часть:

- I. Сообщение в виде доклада-презентации ответственными студентами за проведение занятий по теме 1, в котором излагается суть обсуждаемых положений:
 - 1) Схема исследования задач;
 - 2) Степень завершенности решения задач.
 - 3) Место конкретных задач в «шахматке» задач по указанной проблеме.
 - II. Выяснение позиций участников с зафиксированными точками зрения на решение вышеизложенных задач.
- Итог II-го этапа: формирование целевых групп по общности позиций каждой из групп.
- III. Организация коммуникации между группами: 1) выяснение позиции-варианта решения выявленных групп и защита занятой позиции; 2) формирование нового набора вариантов решений на основании общего обсуждения; 3) перспективность задачи и актуальность ее решения.
 - IV. Выбор задачи как основы для магистерской диссертации.

Выводы: реализован самостоятельный поиск учащимися путей и вариантов решения поставленной задачи-проблемы (выбор одного из предложенных вариантов или нахождение собственного варианта и обоснование решения на базе коллективной интерактивной работы).

Итог занятий по теме №1: Оценивание компетенций (ОК-4, ОК-5, ОК-6, ОК-7, уровни 3-Эл, У-Эл, В-Эл; ПК-26 уровни 3-Пр, У-Пр, В-Пр) по результатам работы на занятиях (активность, инициативность, грамотность, обоснованность защищаемой позиции) и своевременности реакции при подготовке к семинарам.

2. УКАЗАНИЯ К САМОСТОЯТЕЛЬНОЙ РАБОТЕ СТУДЕНТОВ по разделу 2

Тема. Модели и методы анализа временных рядов и прогнозирования состояний сложных объектов по временным рядам.

ЦЕЛЬ РАБОТЫ

Выбор направлений исследований, определение проблем и вытекающих из нее частных целей и задач. Определяется цель исследования, обосновывается предмет и объект исследования. Подготовка к исследованию. Определение логики исследования. Знакомство с основными математическими методами, изучающими закономерности и их выявление в отдельных реализациях случайного процесса – временных рядах и основными этапами исследования стохастического временного ряда.

- 1) определение проблемы;
- 2) разработка плана исследования;
- 3) сбор (определение) данных (обучающих и контрольных)
- 4) предобработка данных;
- 5) разработка подхода (подходов в случае коллективного принятия решения) к решению проблемы;
- 6) исследование особенностей выбранного метода (методов) для решения задачи; проверка условий его применимости.
- 7) подготовка отчета и презентации-доклада.
- 8) уточнить примеры типовых заданий на СРС.

Форма текущего контроля освоения компетенций ОК-4, ОК-5, ОК-6, ОК-7, уровни 3-Эл, У-Эл, В-Эл; ПК-3; ПК-4; ПК-6, уровни 3-Пр, У-Пр, В-Пр (см. табл.1): *отчет* – в форме доклада на семинарах.

Примеры обсуждаемых задач и указания к соответствующим алгоритмам их решения

Задача 2.1. *Исследование автоматизированных систем обработки временных рядов.* Назначение. Эффективность (параметры оценки). Примеры удачных и «неудачных» применений компьютерных пакетов в различных областях.

Постановки частных задач:

АИС в медицинских обследованиях.

АИС в геологии и геофизике.

Ложные регрессии в экономических моделях.

Задача 2.2. *Методы интеллектуального анализа данных (ИАД) в исследовании временных рядов.* Назначение. Эффективность (параметры оценки). Примеры удачных и «неудачных» применений компьютерных пакетов в различных областях.

Постановки частных задач:

ИАД в медицинских обследованиях.

ИАД в геологии и геофизике.

Эталонные модели в разметке состояний по временным рядам.

ИАД в оценке прогнозов по финансовым рядам данных.

Регрессионные модели. Достоинства и недостатки.

Задача 2.3. *Исследование эффективности алгоритмов выделения тренда в стохастических временных рядах.* Назначение. Эффективность (параметры оценки). Примеры удачных и «неудачных» применений фильтрации в различных областях.

Постановки частных задач:

Исследование эффективности алгоритма выделения тренда в стохастических временных рядах на основе эталонной разметки.

Применение метода модовой декомпозиции для исследования стохастических временных рядов.

Метод Хуанга-Гильберта.

Указания 2.3. *Дополнительная литература для подготовки к занятию:*

1. Давыдов В.А., Давыдов А.В. Очистка геофизических данных от шумов с использованием преобразования Гильберта-Хуанга // Электронное научное издание "Актуальные инновационные исследования: наука и практика", 2010, № 1. <http://www.actualresearch.ru>.
2. Дюк В. Обработка данных на ПК в примерах / Вячеслав Дюк. - СПб.: Питер, 1997. - 240с.
3. Прикладная статистика. Классификация и снижение размерности : справочное издание / С. А. Айвазян [и др.] ; ред. С. А. Айвазян. - М. : Финансы и статистика, 1989. - 608 с.

Задача 2.4. *Исследование эффективности алгоритмов анализа временных финансовых рядов.* Назначение. Эффективность (параметры оценки). Примеры алгоритмов.

Постановки частных задач:

Алгоритм идентификации разворотных точек на примере финансовых рядов.

Анализ прогностических свойств адаптационной полиномиальной аппроксимации.

Сравнительный анализ методов прогнозирования финансовых рядов на основе карт Кохонена авторегрессионного моделирования.

Модификация квантильной регрессии для анализа временных рядов.

Модификация фильтра Ходрика-Прескотта для исследования финансовых рядов.

Задача 2.5. *Алгоритмы анализа временных геофизических рядов.* Назначение. Эффективность (параметры оценки). Примеры алгоритмов.

Постановки частных задач:

АИС "Геофизические сигналы" для мониторинга и прогнозирования. Методы обработки данных.

Прогноз предразрушений горных пород с помощью генетического алгоритма.

Прогнозирование геофизических временных рядов с помощью нейронных сетей.

Указания 2.5. *Дополнительная литература для подготовки к занятию:*

- 1) <http://www.alleng.ru/d/econ/econ292.htm> Юдин С.В. Математика в экономике. Тула: РГТЭУ, 2009. — 228 с.
- 2) http://www.aup.ru/books/m155/2_12.htm Орлов А.И. Математика случая. Вероятность и статистика – основные факты. Учебное пособие. М.: МЗ-Пресс, 2004.
- 3) Самостоятельный интернет-поиск.

Рассмотреть пример решения задачи «Выдавать ли кредит клиенту» в аналитическом специализированном нейроките (например, Deductor, BaseGroup). В качестве обучающего набора данных выступает база данных, содержащая информацию о клиентах, в частности: Сумма кредита, Срок кредита, Цель кредитования, Возраст, Пол, Образование, Частная собственность, Квартира, Площадь квартиры. На основе этих данных необходимо построить модель, которая сможет дать ответ, входит ли клиент, желающий получить кредит, в группу риска невозврата кредита, т.е. пользователь должен получить ответ на вопрос «Выдавать ли кредит клиенту?». Задача относится к группе задач классификации, т.е. обучения с учителем.

Ознакомиться со следующим материалом (по указанным источникам):

- 5) Перцептрон Розенблатта.
- 6) Классификационные правила и их реализация в нейросети;
- 7) Возможности нейро-алгоритмов.
- 8) Примеры успешного решения нелинейных проблем на базе нейро-алгоритмов.

Задача 2.6. *Авторегрессионное моделирование временных рядов. Связь с регрессионными моделями.* Эффективность (параметры оценки) моделирования. Примеры алгоритмов.

Постановки частных задач:

Исследование эффективности метода Бока для реконструкции временных рядов.

Метод ARADS как обобщение регрессионных моделей. Сравнение с методом Бока.

Указания 2.6. Типичной в сфере маркетинга является задача прогнозирования рынков. В результате решения данной задачи оцениваются перспективы развития конъюнктуры определенного рынка, изменения рыночных условий на будущие периоды, определяются тенденции рынка (структурные изменения, потребности покупателей, изменения цен).

Обычно в этой области решаются следующие практические задачи:

- прогноз продаж товаров (например, с целью определения нормы товарного запаса);
- прогнозирование продаж товаров, оказывающих влияние друг на друга;
- прогноза продаж в зависимости от внешних факторов.

Для заданной *базы данных* (student\Колесникова\ Анализ данных) торговых временных рядов провести исследование по всем вышеперечисленным задачам на базе методов «Гусеница» и модовой декомпозиции. Сравнить результаты.

Ознакомиться со следующим материалом (по указанным источникам):

- 1) Методы главных компонент и «Гусеница»;
- 2) Метод модовой декомпозиции и преобразование Гильберта-Хуанга;
- 3) Метод Бока;
- 4) Карты Кохонена
- 5) Авторегрессионное моделирование.
- 6) Квантильная регрессия для анализа временных рядов.
- 7) Фильтр Ходрика-Прескотта для исследования финансовых рядов.

Подготовка занятий по теме №2. Выбор ведущих студентов, ответственных за выбор и подачу необходимой информации и обсуждение с ним алгоритма занятий.

Таблица 2.1

№	№ задач и	Вид (совмещение нескольких видов) интерактивной работы	Трудоемкость (час)	Отрабатываемые компетенции/ожидаемый уровень освоения	Оценка личностных качеств	Контроль выполнения работы (участие в полемике, доклад, реферат)
1	2.1-2.6	Работа в команде: распределение задач. Решение задач: этапность исследования.	8×4	ОК-4, ОК-5, ОК-6, ОК-7, ПК-3; ПК-4; ПК-6	Качество работы; своевременность докладов по выбранным постановкам проблем	Критерии оценивания поведения на занятии: активность, инициативность, грамотность, обоснованность защищаемой позиции.
Всего			32			

Вступление. Сообщение темы и обоснование ее актуальности через вышеуказанные проблемы.

Основная часть:

- I. Сообщение в виде доклада-презентации ответственными студентами за проведение занятий по теме 2, в котором излагается суть обсуждаемых положений:
 - 1) Схема исследования задач;
 - 2) Степень завершенности решения задач.
 - 3) Место конкретных задач в «шахматке» задач по указанной проблеме.
 - II. Выяснение позиций участников с зафиксированными точками зрения на решение вышеизложенных задач.
- Итог II-го этапа: формирование целевых групп по общности позиций каждой из групп.
- III. Организация коммуникации между группами: 1) выяснение позиции-варианта решения выявленных групп и защита занятой позиции; 2) формирование нового набора вариантов решений на основании общего обсуждения; 3) перспективность задачи и актуальность ее решения.
 - IV. Постановки задач (варианты постановок) как основы для магистерской диссертации.

Выводы: реализован самостоятельный поиск учащимися путей и вариантов решения поставленной задачи-проблемы (выбор одного из предложенных вариантов или нахождение собственного варианта и обоснование решения на базе коллективной интерактивной работы).

Итог занятий по теме №2: Оценивание компетенций (ОК-4, ОК-5, ОК-6, ОК-7, уровни 3-Эл, У-Эл, В-Эл; ПК-26 уровни 3-Пр, У-Пр, В-Пр) по результатам работы на занятиях (активность, инициативность, грамотность, обоснованность защищаемой позиции) и своевременности реакции при подготовке к семинарам.

3. УКАЗАНИЯ К САМОСТОЯТЕЛЬНОЙ РАБОТЕ СТУДЕНТОВ по разделу 3

Тема. Модели и методы безмодельного управления на базе ансамбля реализаций (временных рядов) сложными объектами (много связными, нелинейными).

ЦЕЛЬ РАБОТЫ

Выбор направлений исследований, определение проблем и вытекающих из нее частных целей и задач. Определяется цель исследования, обосновывается предмет и объект исследования. Подготовка к исследованию. Определение логики исследования. Знакомство с основными математическими методами, изучающими управление нелинейными объектами.

- 1) определение проблемы;
- 2) разработка плана исследования;
- 3) сбор (определение) данных (обучающих и контрольных)
- 4) предобработка данных;
- 5) разработка подхода (подходов в случае коллективного принятия решения) к решению проблемы;
- 6) исследование особенностей выбранного метода (методов) для решения задачи; проверка условий его применимости.
- 7) подготовка отчета и презентации-доклада.
- 8) уточнить примеры типовых заданий на СРС.

Форма текущего контроля освоения компетенций ОК-4, ОК-5, ОК-6, ОК-7, уровни 3-Эл, У-Эл, В-Эл; ПК-3; ПК-4; ПК-6, уровни 3-Пр, У-Пр, В-Пр (см. табл.1): *отчет* – в форме доклада на семинарах.

Примеры обсуждаемых задач и соответствующих алгоритмов

Задача 3.1. *Исследование автоматизированных систем обработки временных рядов.* Назначение. Эффективность (параметры оценки). Примеры удачных и «неудачных» применений компьютерных пакетов в различных областях.

Постановки частных задач:

Исследование свойств нелинейного управления на многообразиях в хаотических системах.

Синергетическое управление асинхронным двигателем с нелинейным наблюдателем состояний.

Анализ свойств нелинейных регуляторов на примере электро-механического объекта. Нелинейное управление в модели Фейгенбаума.

Задача 3.2. *Исследование методов построения «наблюдателей» неизмеряемых координат сложного объекта.* Назначение. Эффективность (параметры оценки).

Постановки частных задач:

Примеры применений «наблюдателей» в различных технических объектах.

Исследование свойств линейных «наблюдателей» состояний сложного объекта.

Исследование свойств нелинейного «наблюдателя» состояний при нелинейном управлении на многообразиях в хаотических системах.

Задача 3.3. Исследование «безмодельных» методов управления сложными объектами. Назначение. Эффективность (параметры оценки).

Постановки частных задач:

Синтез регуляторов температуры в термосетях.

Ситуационный центр управления энергоэффективностью.

Подготовка занятий по теме №3. Выбор ведущих студентов, ответственных за выбор и подачу необходимой информации и обсуждение с ним алгоритма занятий.

Таблица 3.1

№	№ задачи и	Вид (совмещение нескольких видов) интерактивной работы	Трудоемкость (час)	Отрабатываемые компетенции/ожидаемый уровень освоения	Оценка личностных качеств	Контроль выполнения работы (участие в полемике, доклад, реферат)
1	3.1-3.3	Работа в команде: распределение задач. Решение задач: этапность исследования.	8×4	ОК-4, ОК-5, ОК-6, ОК-7, ПК-3; ПК-4; ПК-6	Качество работы; своевременность докладов по выбранным постановкам проблем	Критерии оценивания поведения на занятии: активность, инициативность, грамотность, обоснованность защищаемой позиции.
Всего			32			

Вступление. Сообщение темы и обоснование ее актуальности через вышеуказанные проблемы.

Основная часть:

- I. Сообщение в виде доклада-презентации ответственными студентами за проведение занятий по теме 3, в котором излагается суть обсуждаемых положений:
 - 1) Схема исследования задач;
 - 2) Степень завершенности решения задач.
 - 3) Место конкретных задач в «шахматке» задач по указанной проблеме.
 - II. Выяснение позиций участников с зафиксированными точками зрения на решение вышеизложенных задач.
- Итог II-го этапа: формирование целевых групп по общности позиций каждой из групп.
- III. Организация коммуникации между группами: 1) выяснение позиции-варианта решения выявленных групп и защита занятой позиции; 2) формирование нового набора вариантов решений на основании общего обсуждения; 3) перспективность задачи и актуальность ее решения.
 - IV. Постановки задач (варианты постановок) как основы для магистерской диссертации.

Выводы: реализован самостоятельный поиск учащимися путей и вариантов решения поставленной задачи-проблемы (выбор одного из предложенных вариантов или нахождение собственного варианта и обоснование решения на базе коллективной интерактивной работы).

Итог занятий по теме №3: Оценивание компетенций (ОК-4, ОК-5, ОК-6, ОК-7, уровни 3-Эл, У-Эл, В-Эл; ПК-26 уровни 3-Пр, У-Пр, В-Пр) по результатам работы на занятиях (активность, инициативность, грамотность, обоснованность защищаемой позиции) и своевременности реакции при подготовке к семинарам.

4. УКАЗАНИЯ К САМОСТОЯТЕЛЬНОЙ РАБОТЕ СТУДЕНТОВ по разделу 4

Тема. Обсуждение результатов исследований и работ при подготовке к печати по темам НИР (тезисы докладов, статьи). Методы корректной обработки статистических данных и форм представления результатов при осуществлении экспериментальных исследований (при оформлении магистерской диссертации). Направления дальнейших исследований.

ЦЕЛЬ РАБОТЫ

Выбор направлений исследований, определение проблем и вытекающих из нее частных целей и задач. Определяется цель исследования, обосновывается предмет и объект исследования. Подготовка к исследованию. Определение логики исследования. Знакомство с основными математическими методами корректной обработки выборок (длинных, коротких). Знакомство с основными компьютерными пакетами обработки данных и формами представления результатов. Когнитивные средства интерпретации данных.

Программа работы студента-магистранта выбирается индивидуально в соответствии с выбранной темой диссертации и включает естественные этапы:

Этап 1. Выбор направления и подготовка к исследованию. Определяется цель исследования, обосновывается предмет и объект исследования. Подготовка к исследованию.

Этап 2. Библиографический поиск, составление научного обзора, проектирование «шахматки» работ в выбранном направлении и выявление места конкретной задачи в этой «шахматке». Осуществление сбора, обработки, анализа, сопоставления и систематизации сведений по теме диссертации. Проведение патентного поиска. Обоснование актуальности задачи и необходимости выполнения данного исследования, формирование рабочей гипотезы и локализация задачи исследования.

Этап 3. Разработка плана и общего алгоритма исследования. Планирование, подготовка и проведение экспериментов. Подготовка отчета.

Этап 4. Моделирование и численные (экспериментальные) исследования. Разработка алгоритмического и программного обеспечения.

Этап 5. Обсуждение полученных результатов, формулирование выводов. Проведение системного анализа решаемой задачи, формулирование научных выводов и перспектив дальнейшего развития решения данной задачи.

Этап 6. Оформление пояснительной записки (текста диссертации). В тексте должна быть поставлена задача, кратко изложена теоретическая и/или практическая часть, план экспериментов, результаты и их обсуждение.

Форма текущего контроля освоения компетенций ОК-4, ОК-5, ОК-6, ОК-7, уровни 3-Эл, У-Эл, В-Эл; ПК-3; ПК-4; ПК-6, уровни 3-Пр, У-Пр, В-Пр (см. табл.0.1): *отчет* – в форме доклада на семинарах.

Примеры обсуждаемых задач (магистерских работ по каф. ЭМИС прошлых лет)

- Задача 4.1.** Синтез регуляторов температуры в термосетях.
- Задача 4.2.** Автоматизация обмена данными с ЦЗУ через USB порт.
- Задача 4.3.** Программный модуль обработки видеопотока от DV камер.
- Задача 4.4.** Модификация метода сетевого рендеринга.
- Задача 4.5.** Реализация автоматизированной системы обработки коротких сообщений.
- Задача 4.6.** Реализация web - сервиса по технологии ASP.NET с использованием платформы Enfity Frame work для доступа к управляемым данным.
- Задача 4.7.** Свойства системы шифрования/дешифрования на базе хаотической динамики.
- Задача 4.8.** Анализ прогностических свойств метода японских свечей.
- Задача 4.9.** Анализ прогностических свойств адаптационной полиномиальной аппроксимации.
- Задача 4.10.** Исследование эффективности алгоритма выделения тренда в стохастических временных рядах на основе эталонной разметки.
- Задача 4.11.** Применение метода модовой декомпозиции для исследования стохастических временных рядов.
- Задача 4.12.** Алгоритм идентификации разворотных точек на примере финансовых рядов.
- Задача 4.13.** Прогнозирование временных рядов на основе коллективных правил.
- Задача 4.14.** Исследование свойств корректной модификации метода анализа иерархий.
- Задача 4.15.** Модификация квантильной регрессии для анализа временных рядов.
- Задача 4.16.** Модификация фильтра Ходрика-Прескотта для исследования финансовых рядов.
- Задача 4.17.** Исследование свойств нелинейного управления на многообразиях в хаотических системах.
- Задача 4.18.** Синергетическое управление асинхронным двигателем с нелинейным наблюдателем состояний.
- Задача 4.19.** Анализ свойств нелинейных регуляторов на примере электро-механического объекта.
- Задача 4.20.** Анализ свойств нелинейных регуляторов на примере электро-механического объекта.
- Задача 4.21.** Нелинейное управление в модели Фейгенбаума.

Подготовка занятий по теме №4. Выбор ведущих студентов, ответственных за выбор и подачу необходимой информации и обсуждение с ним алгоритма занятий.

Таблица 4.1

№	№ задачи	Вид (совмещение нескольких видов) интерактивной работы	Трудоемкость (час)	Отрабатываемые компетенции/ожидаемый уровень освоения	Оценка личностных качеств	Контроль выполнения работы (участие в полемике, доклад, реферат)
1	4.1-4.21	Работа в команде: распределение задач. Решение задач: этапность исследования.	16×4	ОК-4, ОК-5, ОК-6, ОК-7, ПК-3; ПК-4; ПК-6	Качество работы; своевременность докладов по выбранным постановкам проблем	Критерии оценивания поведения на занятии: активность, инициативность, грамотность, обоснованность защищаемой позиции.
Всего			64			

Вступление. Сообщение темы и обоснование ее актуальности через вышеуказанные проблемы.

Основная часть:

- I. Сообщение в виде доклада-презентации ответственными студентами за проведение занятий по теме 3, в котором излагается суть обсуждаемых положений:
 - 1) Схема исследования задач;
 - 2) Степень завершенности решения задач.
 - 3) Место конкретных задач в «шахматке» задач по указанной проблеме.
 - II. Выяснение позиций участников с зафиксированными точками зрения на решение вышеизложенных задач.
- Итог II-го этапа: формирование целевых групп по общности позиций каждой из групп.
- III. Организация коммуникации между группами: 1) выяснение позиции-варианта решения выявленных групп и защита занятой позиции; 2) формирование нового набора вариантов решений на основании общего обсуждения; 3) перспективность задачи и актуальность ее решения.
 - IV. Постановки задач как основы для магистерской диссертации.
 - V. Доклады-отчеты о проделанной работе над диссертацией.

Выводы: реализован самостоятельный поиск учащимися путей и вариантов решения поставленной задачи-проблемы (выбор одного из предложенных вариантов или нахождение собственного варианта и обоснование решения на базе коллективной интерактивной работы).

Итог занятий по теме №4: Оценивание компетенций (ОК-4, ОК-5, ОК-6, ОК-7, уровни 3-Эл, У-Эл, В-Эл; ПК-26 уровни 3-Пр, У-Пр, В-Пр) по результатам работы на занятиях (активность, инициативность, грамотность, обоснованность защищаемой позиции) и своевременности реакции при подготовке к семинарам.

Указания к решению практических задач по обработке данных

Основные модели управления данными, многомерный анализ данных

Практические занятия П4.1. Комплексный подход к внедрению Data Mining, OLAP и хранилищ данных в СППР.

- 1) Осуществить классификацию систем поддержки принятия решений (СППР), учитывая одну из существующих по типу ориентации: на данные, на модели, на знания, на документы, на коммуникации, интер- и интра-СППР, на специфику функций, на базу Web.
- 2) Процесс-исследование Data Mining: основные этапы. Охарактеризовать этапы.
- 3) Исследовать чувствительность Data Mining к выбросам. Методы очистки данных.

Указания 4.1. Разобрать суть примеров в применении Data Mining в промышленном производстве:

- комплексный системный анализ производственных ситуаций;
- краткосрочный и долгосрочный прогноз развития производственных ситуаций;
- выработка вариантов оптимизационных решений;
- прогнозирование качества изделия в зависимости от некоторых параметров технологического процесса;
- обнаружение скрытых тенденций и закономерностей развития производственных процессов;
- прогнозирование закономерностей развития производственных процессов;
- обнаружение скрытых факторов влияния;

- обнаружение и идентификация ранее неизвестных взаимосвязей между производственными параметрами и факторами влияния;
- анализ среды взаимодействия производственных процессов и прогнозирование изменения ее характеристик;
- выработка оптимизационных рекомендаций по управлению производственными процессами;
- визуализация результатов анализа, подготовку предварительных отчетов и проектов допустимых решений с оценками достоверности и эффективности возможных реализаций.

Практические занятия П4.2. Подходы к определению знаний и данных, информации.

- 1) Задать понятия: знания, данные, информация своими свойствами. Привести разные определения указанным понятиям.
- 2) Модели представления данных и знаний (по Дюку В.Д., Янковской А.Е. и др.).
- 3) OLAP: оперативная аналитическая обработка данных (On-Line Analytical Processing).
- 4) Условия корректной статистической обработки данных.

Указания 4.2. Рекомендуемые темы:

- Классификация с помощью деревьев решений;
- Байесовская классификация;
- Классификация при помощи искусственных нейронных сетей;
- Классификация методом опорных векторов;
- Статистические методы, в частности, линейная регрессия;
- Классификация при помощи метода ближайшего соседа;
- Классификация сbr-методом (case based reasoning, метод рассуждения по аналогии);
- Классификация при помощи генетических алгоритмов.

Практические занятия П4.3. Системы анализа распределённых данных. Анализ способов¹ анализа распределённых данных. Системы анализа распределённых данных. Разработка Web-сайта на основе базы данных.

Указания 4.3. Темы для предварительного изучения. Настройка Internet Information Server, знакомство с HTML, ASP, настройка DSN по эл. источнику: Сороковиков В. Н. Разработка Web-сайта на основе базы данных. – Томск. – 2002.

Опубликовать спроектированную БД согласно указанному алгоритму.

Ознакомиться со следующим материалом (по указанным источникам):

- 1) Обзор систем анализа распределённых данных и принципы их построения;
- 2) Существующие стандарты Data mining и OLAP;
- 3) Понятие «мобильных агентов» и системы мобильных агентов, используемых для анализа данных;

ИСПОЛЬЗОВАННАЯ ЛИТЕРАТУРА

1. Васильев В.А. Непараметрическое оценивание функционалов от распределений стационарных последовательностей / В.А. Васильев, А.В. Добровидов, Г.М. Кошкин. – М.: Наука, 2004. – 508 с
2. Воронцов К.В. Лекции по методам оценивания и выбора моделей. 2007. Режим доступа: www.ccas.ru/voron/download/Modeling.pdf.

¹ <http://www.rarus.nn.ru/products/olap/30.htm>

3. Воронцов К.В. Обзор современных исследований по проблеме качества обучения алгоритмов. Таврический вестник информатики и математики. – 2004. – № 1. – С. 5 – 24. <http://www.ccas.ru/frc/papers/voron04twim.pdf>.
4. Горелик А. Л., Скрипкин В. А. Методы распознавания: Учебное пособие для вузов. - 4-е изд., испр. - М.: Высшая школа, 2004. – 260 с.
5. Дюк В. Обработка данных на ПК в примерах / Вячеслав Дюк. - СПб.:Питер, 1997. - 240с.
6. З. Брандт. Анализ данных. Статистические и вычислительные методы для научных работников и инженеров: Пер. с англ.: Учебное пособие/ З. Брандт; пер.: О.И.Волкова; ред. пер.: Е. В. Чепурин. - М. :Мир, 2003 ; М.: АСТ, 2003. – 686 с.
7. Загоруйко Н.Г. Прикладные методы анализа данных и знаний. - Новосибирск: ИМ СО РАН, 1999, стр. 270 http://www.sernam.ru/book_zg.php
8. Кокс, Д. Р. Анализ данных типа времени жизни : Пер. с англ. / Д. Р. Кокс, Д. Оукс; Пер. О. В. Селезнева, Ред. Ю. К. Беляев, Предисл. Ю. К. Беляев. - М. : Финансы и статистика, 1988. – 189с.
9. Лапко А.В. Непараметрические системы обработки информации : Учебное пособие для вузов / А. В. Лапко, С. В. Ченцов; Российская Академия наук. Сибирское отделение, Институт вычислительного моделирования. - М. : Наука, 2000. - 349 с.
10. Прикладная статистика. Классификация и снижение размерности : справочное издание / С. А.Айвазян [и др.] ; ред. С. А. Айвазян. - М. : Финансы и статистика, 1989. - 608 с.
11. Тюрин, Ю.Н.. Анализ данных на компьютере : учебное пособие для вузов / Ю. Н. Тюрин, А. А. Макаров. - 4-е изд., перераб. - М. : Форум, 2008. – 366.
12. <http://itteach.ru/predstavlenie-znaniy/metodi-i-sredstva-intellektualnogo-analiza-dannich><http://www.bsu.ru/content/hec/biometria/modules/stdatmin.html>
13. <http://www.arshinov74.ru/files/files/10.pdf>
14. <http://www.intuit.ru/department/database/dataanalysis/>
15. <http://www.olap.ru/basic/dm2.asp>
16. http://www.sernam.ru/book_zg.php
17. <http://www.statlab.kubsu.ru/node/4>