МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ

Федеральное государственное бюджетное образовательное учреждение высшего образования

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

УТВЕРЖДАЮ
Проректор по учебной работе
_____ П. Е. Троян
«___» _____ 20__ г.

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

Анализ данных

Уровень образования: **высшее образование - бакалавриат**
Направление подготовки (специальность): **09.03.02 Информационные системы и технологии**
Направленность (профиль): **Информационные системы и технологии**
Форма обучения: **очная**
Факультет: **ФВС, Факультет вычислительных систем**
Кафедра: **ЭМИС, Кафедра экономической математики, информатики и статистики**
Курс: **2**
Семестр: **4**
Учебный план набора 2014 года

Распределение рабочего времени

№	Виды учебной деятельности	4 семестр	Всего	Единицы
1	Лекции	34	34	часов
2	Практические занятия	34	34	часов
3	Всего аудиторных занятий	68	68	часов
4	Самостоятельная работа	112	112	часов
5	Всего (без экзамена)	180	180	часов
6	Подготовка и сдача экзамена	36	36	часов
7	Общая трудоемкость	216	216	часов
		6.0	6.0	З.Е

Экзамен: 4 семестр

Томск 2017

Согласована на портале № 13379

Рассмотрена и одобрена на заседании кафедры
протокол № 5 от «19» 1 2017 г.

ЛИСТ СОГЛАСОВАНИЙ

Рабочая программа составлена с учетом требований Федерального Государственного образовательного стандарта высшего образования (ФГОС ВО) по направлению подготовки (специальности) 09.03.02 Информационные системы и технологии, утвержденного 2015-03-12 года, рассмотрена и утверждена на заседании кафедры «___» _____ 20__ года, протокол №_____.

Разработчики:

профессор каф. ЭМИС	_____	Колесникова С. И.
Заведующий обеспечивающей каф. ЭМИС	_____	Боровской И. Г.

Рабочая программа согласована с факультетом, профилирующей и выпускающей кафедрами направления подготовки (специальности).

Декан ФВС	_____	Козлова Л. А.
Заведующий выпускающей каф. ЭМИС	_____	Боровской И. Г.

Эксперты:

доцент кафедры экономики Томский государственный университет систем управления и радиоэлектроники	_____	1111111111 2. 3.

Согласована на портале № 13379

1. Цели и задачи дисциплины

1.1. Цели дисциплины

- знакомство с основными достижениями о процессе построения и анализа математических моделей реальных процессов и явлений, учитывающих случайные факторы со статистически устойчивыми свойствами;
- усвоение фундаментальных понятий теории вероятностей и математической статистики;
- овладение студентами основными методами постановки и решения задач обработки данных на основе методов математической статистики.

1.2. Задачи дисциплины

– 1) Освоить основной понятийный аппарат теории вероятностей и математической статистики.

– 2) Усвоить условия корректного применения методов теории вероятностей и математической статистики.

– 3) Овладеть способами решения простых вероятностных задач.

– 4) Усвоить основные модели и соответствующие программные средства обработки статистического материала.

– 5) Овладеть основными методами математической статистики для решения профессиональных задач.

–

2. Место дисциплины в структуре ОПОП

Дисциплина «Анализ данных» (Б1.В.ДВ.12.2) относится к блоку 1 (вариативная часть).

Предшествующими дисциплинами, формирующими начальные знания, являются следующие дисциплины: Математика.

Последующими дисциплинами являются: .

3. Требования к результатам освоения дисциплины

Процесс изучения дисциплины направлен на формирование следующих компетенций:

– ПК-25 способностью использовать математические методы обработки, анализа и синтеза результатов профессиональных исследований;

В результате изучения дисциплины студент должен:

– **знать** о основные понятия теории вероятности: аксиоматика теории вероятности, случайные события и основные теоремы теории вероятности; о методы описания и определения одно- и многомерных случайных величин; о предельные теоремы теории вероятности.

– **уметь** о вычислять вероятности случайных событий; о находить числовые характеристики случайных величин; о решать прикладные задачи на основе методов теории вероятности и математической статистики.

– **владеть** о методами решения вероятностных задач; о основными операциями над событиями и комбинаторными методами вычисления вероятности событий; о методами определения вероятностей сложных событий; о методами определения числовых характеристик случайных величин и функций от случайных величин; о методами точечного и интервального оценивания; о методами мышления: логическим, комбинаторно-вероятностным.

4. Объем дисциплины и виды учебной работы

Общая трудоемкость дисциплины составляет 6.0 зачетных единицы и представлена в таблице 4.1.

Таблица 4.1 – Трудоемкость дисциплины

Виды учебной деятельности	Всего часов	Семестры
		4 семестр
Аудиторные занятия (всего)	68	68
Лекции	34	34

Согласована на портале № 13379

Практические занятия	34	34
Самостоятельная работа (всего)	112	112
Подготовка к практическим занятиям, семинарам	112	112
Всего (без экзамена)	180	180
Подготовка и сдача экзамена	36	36
Общая трудоемкость ч	216	216
Зачетные Единицы	6.0	6.0

5. Содержание дисциплины

5.1. Разделы дисциплины и виды занятий

Разделы дисциплины и виды занятий приведены в таблице 5.1.

Таблица 5.1 – Разделы дисциплины и виды занятий

Названия разделов дисциплины	Лекции	Практические занятия	Самостоятельная работа	Всего часов (без экзамена)	Формируемые компетенции
4 семестр					
1 Методы и средства анализа данных	8	8	20	36	ПК-25
2 Методы классификации и кластеризации	8	8	20	36	ПК-25
3 Методы построения математических моделей и прогнозирования временных рядов	8	10	33	51	ПК-25
4 Основные понятия математической статистики	10	8	39	57	ПК-25
Итого за семестр	34	34	112	180	
Итого	34	34	112	180	

5.2. Содержание разделов дисциплины (по лекциям)

Содержание разделов дисциплин (по лекциям) приведено в таблице 5.2.

Таблица 5.2 - Содержание разделов дисциплин (по лекциям)

Названия разделов	Содержание разделов дисциплины по лекциям	Трудоемкость, ч	Формируемые компетенции
4 семестр			
1 Методы и средства анализа данных	Постановка задачи анализа данных. Сущность и условия применимости методов анализа данных.Модели Data Mining. Классификация методов анализа данных.Системы распознавания образов на базе Data	8	ПК-25

Согласована на портале № 13379

	Mining и принципы их построения.Примеры практического применения Data Mining.		
	Итого	8	
2 Методы классификации и кластеризации	Постановка задачи классификации и классификационные решающие правила.Методы построения правил классификации.Постановка задачи кластеризации, меры близости.Неиерархические алгоритмы кластеризации.Адаптивные методы кластеризации.	8	ПК-25
	Итого	8	
3 Методы построения математических моделей и прогнозирования временных рядов	Обзор методов анализа временных рядов.Существующие стандарты и системы анализа временных рядов.Обзор методов прогнозирования стационарных временных рядов. Подходы к анализу нестационарных временных рядов. Анализ свойств одномерных хаотических моделей.Метод разметки стохастических временных рядов и решаемые на его основе задачи.	8	ПК-25
	Итого	8	
4 Основные понятия математической статистики	Статистическая структура. Статистические решения. Выборка. Выборочные моменты, их асимптотические свойства. Порядковые статистики. Эмпирическая функция распределения. Выборочная медиана.Статистическое оценивание. Методы оценивания плотности распределения. Гистограмма. Статистические методы обработки экспериментальных данных.Точечные оценки, несмещенность, состоятельность, оптимальность оценок. Функция правдоподобия. Неравенство Рао - Крамера. Эффективные оценки. Достаточные статистики. Критерий факторизации. Оценки максимального правдоподобия и их свойства. Метод моментов. Свойства оценок, полученных по методу моментов.Интервальное оценивание. Построение доверительных интервалов с помощью центральной случайной величины и распределения точечной оценки. Проверка статистических гипотез. Распределения, связанные с	10	ПК-25

Согласована на портале № 13379

	нормальным: распределения хи-квадрат, Стьюдента. Статистические выводы о параметрах нормального распределения. Критерии согласия хи-квадрат и Колмогорова. Линейная регрессионная модель.		
	Итого	10	
Итого за семестр		34	

5.3. Разделы дисциплины и междисциплинарные связи с обеспечивающими (предыдущими) и обеспечиваемыми (последующими) дисциплинами

Разделы дисциплины и междисциплинарные связи с обеспечивающими (предыдущими) и обеспечиваемыми (последующими) дисциплинами представлены в таблице 5.3.

Таблица 5.3 - Разделы дисциплины и междисциплинарные связи

Наименование дисциплин	№ разделов данной дисциплины, для которых необходимо изучение обеспечивающих и обеспечиваемых дисциплин			
	1	2	3	4
Предшествующие дисциплины				
1 Математика	+	+	+	+

5.4. Соответствие компетенций, формируемых при изучении дисциплины, и видов занятий

Соответствие компетенций, формируемых при изучении дисциплины, и видов занятий представлено в таблице 5.4

Таблица 5.4 – Соответствие компетенций и видов занятий, формируемых при изучении дисциплины

Компетенции	Виды занятий			Формы контроля
	Лекции	Практические занятия	Самостоятельная работа	
ПК-25	+	+	+	Домашнее задание, Конспект самоподготовки, Опрос на занятиях, Дифференцированный зачет

6. Интерактивные методы и формы организации обучения

Не предусмотрено РУП

7. Лабораторные работы

Не предусмотрено РУП

8. Практические занятия (семинары)

Наименование практических занятий (семинаров) приведено в таблице 8.1.

Согласована на портале № 13379

Таблица 8. 1 – Наименование практических занятий (семинаров)

Названия разделов	Наименование практических занятий (семинаров)	Трудоемкость, ч	Формируемые компетенции
4 семестр			
1 Методы и средства анализа данных	Статистический анализ данных. Трендовый анализ. Регрессионный анализ.Основы непараметрической статистики. Анализ реальной проблемы и выбор методов и средства анализа данных.	8	ПК-25
	Итого	8	
2 Методы классификации и кластеризации	Задачи классификации и классификационные решающие правила.Типы данных. Виды метрик.Использование генетических и нейро-алгоритмов для классификации и кластеризации.Классификация на базе конструирования нейросети в специализированном нейропакете.	8	ПК-25
	Итого	8	
3 Методы построения математических моделей и прогнозирования временных рядов	Методы нелинейной фильтрации стохастических временных рядов.Анализ свойств временных рядов, порождаемых хаотическими моделями. Методы анализа нестационарных стохастических временных рядов. Метод модовой декомпозиции и Хуанга-Гильберта.Методы главных компонент и «Гусеница».	10	ПК-25
	Итого	10	
4 Основные понятия математической статистики	Выборка. Выборочные моменты, их асимптотические свойства.Порядковые статистики. Эмпирическая функция распределения. Выборочная медиана.Гистограмма. Статистические методы обработки экспериментальных данных.Точечные оценки, несмещенность, состоятельность, оптимальность оценок. Функция правдоподобия. Построение доверительных интервалов с помощью центральной случайной величины и распределения точечной оценки.	8	ПК-25
	Итого	8	
Итого за семестр		34	

Согласована на портале № 13379

9. Самостоятельная работа

Виды самостоятельной работы, трудоемкость и формируемые компетенции представлены в таблице 9.1.

Таблица 9.1 - Виды самостоятельной работы, трудоемкость и формируемые компетенции

Названия разделов	Виды самостоятельной работы	Трудоемкость, ч	Формируемые компетенции	Формы контроля
4 семестр				
1 Методы и средства анализа данных	Подготовка к практическим занятиям, семинарам	20	ПК-25	Домашнее задание, Опрос на занятиях
	Итого	20		
2 Методы классификации и кластеризации	Подготовка к практическим занятиям, семинарам	20	ПК-25	Домашнее задание, Опрос на занятиях
	Итого	20		
3 Методы построения математических моделей и прогнозирования временных рядов	Подготовка к практическим занятиям, семинарам	33	ПК-25	Домашнее задание, Конспект самоподготовки, Опрос на занятиях
	Итого	33		
4 Основные понятия математической статистики	Подготовка к практическим занятиям, семинарам	39	ПК-25	Домашнее задание, Опрос на занятиях
	Итого	39		
Итого за семестр		112		
	Подготовка и сдача экзамена	36		Экзамен
Итого		148		

9.1. Вопросы для подготовки к практическим занятиям, семинарам

1. Биномиальное и геометрическое распределения. Теорема Пуассона, оценка отклонения биномиальных вероятностей от пуассоновских.
2. Непрерывные распределения: нормальное, показательное, равномерное.
3. Гистограмма. Статистические методы обработки экспериментальных данных.
4. Точечные оценки, несмещенность, состоятельность, оптимальность оценок. Функция правдоподобия. Построение доверительных интервалов с помощью центральной случайной величины и распределения точечной оценки.
5. Подготовка к проверочной индивидуальной работе и ее выполнение на темы:
6. Вероятность, аксиомы вероятности (по Колмогорову). Элементы комбинаторики.
7. Основные теоремы теории вероятностей. Условные вероятности. Независимость событий.
8. Формулы полной вероятности и Байеса.
9. Схема независимых испытаний. Формулы Бернулли. Биномиальное распределение.
10. Понятие случайного процесса. Пуассоновский процесс. Случайные потоки.
11. Цепи Маркова и их использование в моделировании социально-экономических процессов.
12. Теория массового обслуживания: основные модели.

Согласована на портале № 13379

10. Курсовая работа (проект)

Не предусмотрено РУП

11. Рейтинговая система для оценки успеваемости студентов

11.1. Балльные оценки для элементов контроля

Таблица 11.1 – Балльные оценки для элементов контроля

Элементы учебной деятельности	Максимальный балл на 1-ую КТ с начала семестра	Максимальный балл за период между 1КТ и 2КТ	Максимальный балл за период между 2КТ и на конец семестра	Всего за семестр
4 семестр				
Домашнее задание	5	5		10
Конспект самоподготовки	5	5		10
Опрос на занятиях	10	20	20	50
Итого максимум за период	20	30	20	70
Экзамен				30
Нарастающим итогом	20	50	70	100

11.2. Пересчет баллов в оценки за контрольные точки

Пересчет баллов в оценки за контрольные точки представлен в таблице 11.2.

Таблица 11. 2 – Пересчет баллов в оценки за контрольные точки

Баллы на дату контрольной точки	Оценка
≥ 90% от максимальной суммы баллов на дату КТ	5
От 70% до 89% от максимальной суммы баллов на дату КТ	4
От 60% до 69% от максимальной суммы баллов на дату КТ	3
< 60% от максимальной суммы баллов на дату КТ	2

11.3. Пересчет суммы баллов в традиционную и международную оценку

Пересчет суммы баллов в традиционную и международную оценку представлен в таблице 11.3.

Таблица 11. 3 – Пересчет суммы баллов в традиционную и международную оценку

Оценка (ГОС)	Итоговая сумма баллов, учитывает успешно сданный экзамен	Оценка (ECTS)
5 (отлично) (зачтено)	90 - 100	A (отлично)
4 (хорошо) (зачтено)	85 - 89	B (очень хорошо)
	75 - 84	C (хорошо)
	70 - 74	D (удовлетворительно)
3 (удовлетворительно) (зачтено)	65 - 69	
	60 - 64	E (посредственно)
2 (неудовлетворительно) (не зачтено)	Ниже 60 баллов	F (неудовлетворительно)

Согласована на портале № 13379

12. Учебно-методическое и информационное обеспечение дисциплины

12.1. Основная литература

1. Горлач, Б.А. Теория вероятностей и математическая статистика. [Электронный ресурс] — Электрон. дан. — СПб. : Лань, 2013. — 320 с. [Электронный ресурс]. - https://e.lanbook.com/book/4864#book_name

12.2. Дополнительная литература

1. Вентцель Е.С.,Овчаров Л.А. Задачи и упражнения по теории вероятностей : Учебное пособие для втузов / Е. С. Вентцель, Л. А. Овчаров. - 6-е изд., стереотип. - М. : Академия, 2005. (наличие в библиотеке ТУСУР - 99 экз.)
2. П. Е.Данко. Высшая математика в упражнениях и задачах [с решениями]: учебное пособие: в 2 ч. / П. Е.Данко [и др.]. - Ч. 2. - 7-е изд., испр. . - М. : ОНИКС, 2009. - 448с (наличие в библиотеке ТУСУР - 6 экз.)
3. В. Е. Гмурман Теория вероятностей и математическая статистика : Учебное пособие для вузов / В. Е. Гмурман ; Министерство образования и науки Российской Федерации. - 12-е изд., перераб. - М. : Высшее образование, 2006. – 478с. (наличие в библиотеке ТУСУР - 21 экз.)

12.3 Учебно-методические пособия

12.3.1. Обязательные учебно-методические пособия

1. Колесникова С.И. Высшая математика III. Основы теории вероятностей. Элементы математической статистики. Методическое пособие. Томск: ТУСУР, 2007.–106 с. (наличие в библиотеке ТУСУР - 54 экз.)
2. Теория вероятностей и математическая статистика: Методические указания по выполнению практических работ / Колесникова С. И. - 2012. 28 с. [Электронный ресурс] - Режим доступа: http://edu.tusur.ru/publications/880, дата обращения: 09.02.2017.
3. Теория вероятностей и математическая статистика: Методические указания к самостоятельной работе студентов / Колесникова С. И. - 2012. 16 с. [Электронный ресурс] - Режим доступа: http://edu.tusur.ru/publications/881, дата обращения: 09.02.2017.
4. Анализ данных: Методические указания по самостоятельной работе / Колесникова С. И. - 2012. 18 с. [Электронный ресурс] - Режим доступа: http://edu.tusur.ru/publications/3053, дата обращения: 09.02.2017.

12.3.2 Учебно-методические пособия для лиц с ограниченными возможностями здоровья

Учебно-методические материалы для самостоятельной и аудиторной работы обучающихся из числа инвалидов предоставляются в формах, адаптированных к ограничениям их здоровья и восприятия информации.

Для лиц с нарушениями зрения:

- в форме электронного документа;
- в печатной форме увеличенным шрифтом.

Для лиц с нарушениями слуха:

- в форме электронного документа;
- в печатной форме.

Для лиц с нарушениями опорно-двигательного аппарата:

- в форме электронного документа;
- в печатной форме.

12.4. Базы данных, информационно-справочные, поисковые системы и требуемое программное обеспечение

1. http://www.intuit.ru/department/mathematics/ptams/lit.html
2. http://www.exponenta.ru/Теории вероятностей
3. http://www.math-portal.ru/ycebnikiteorver
4. http://www.MatBuro.ru/Учебник по теории вероятностей
5. http://www.teorver-online.narod.ru

Согласована на портале № 13379

13. Материально-техническое обеспечение дисциплины

13.1. Общие требования к материально-техническому обеспечению дисциплины

13.1.1. Материально-техническое обеспечение для лекционных занятий

Лекционные и практические занятия: о лекционные аудитории, в том числе оснащенные презентационной техникой с выходом в Интернет; о аудитории для проведения интерактивных лекций: видеопроектор, экран настенный, др. оборудование о электронный учебно-методический комплекс курса, размещенный на сервере ЭФ по адресу: student\Колесникова\ ТВиМС; о методические указания по практическим работам: Самостоятельная работа студентов. о рабочие места преподавателя и студентов с выходом в Интернет, предназначенные для работы в электронной образовательной среде; о электронный учебно-методический комплекс курса, размещенный на сервере ЭФ по адресу: student\Колесникова\ ТВиМС; о методические указания по самостоятельной работе студентов.

13.1.2. Материально-техническое обеспечение для практических занятий

Для проведения практических (семинарских) занятий используется учебная аудитория, расположенная по адресу 634034, Томская область, г. Томск, Вершинина улица, д. 74, 4 этаж, ауд. 425. Состав оборудования: Учебная мебель; Доска магнитно-маркерная -1шт.; Коммутатор D-Link Switch 24 port - 1шт.; Компьютеры класса не ниже ПЭВМ INTEL Celeron D336 2.8ГГц. -14 шт. Используется лицензионное программное обеспечение, пакеты версией не ниже: Microsoft Windows XP Professional with SP3/Microsoft Windows 7 Professional with SP1; Microsoft Windows Server 2008 R2; Visual Studio 2008 EE with SP1; Microsoft Office Visio 2010; VirtualBox 6.2. Имеется помещения для хранения и профилактического обслуживания учебного оборудования.

13.1.3. Материально-техническое обеспечение для самостоятельной работы

Для проведения практических (семинарских) занятий используется учебная аудитория, расположенная по адресу 634034, Томская область, г. Томск, Вершинина улица, д. 74, 4 этаж, ауд. 425. Состав оборудования: Учебная мебель; Доска магнитно-маркерная -1шт.; Коммутатор D-Link Switch 24 port - 1шт.; Компьютеры класса не ниже ПЭВМ INTEL Celeron D336 2.8ГГц. -14 шт. Используется лицензионное программное обеспечение, пакеты версией не ниже: Microsoft Windows XP Professional with SP3/Microsoft Windows 7 Professional with SP1; Microsoft Windows Server 2008 R2; Visual Studio 2008 EE with SP1; Microsoft Office Visio 2010; VirtualBox 6.2. Имеется помещения для хранения и профилактического обслуживания учебного оборудования.

13.2. Материально-техническое обеспечение дисциплины для лиц с ограниченными возможностями здоровья

Освоение дисциплины лицами с ОВЗ осуществляется с использованием средств обучения общего и специального назначения.

При обучении студентов **с нарушениями слуха** предусмотрено использование звукоусиливающей аппаратуры, мультимедийных средств и других технических средств приема/передачи учебной информации в доступных формах для студентов с нарушениями слуха, мобильной системы обучения для студентов с инвалидностью, портативной индукционной системы. Учебная аудитория, в которой обучаются студенты с нарушением слуха, оборудована компьютерной техникой, аудиотехникой, видеотехникой, электронной доской, мультимедийной системой.

При обучении студентов **с нарушениями зрениями** предусмотрено использование в лекционных и учебных аудиториях возможности просмотра удаленных объектов (например, текста на доске или слайда на экране) при помощи видеоувеличителей для удаленного просмотра.

При обучении студентов **с нарушениями опорно-двигательного аппарата** используются альтернативные устройства ввода информации и другие технические средства приема/передачи учебной информации в доступных формах для студентов с нарушениями опорно-двигательного аппарата, мобильной системы обучения для людей с инвалидностью.

14. Фонд оценочных средств

14.1. Основные требования к фонду оценочных средств и методические рекомендации

Фонд оценочных средств и типовые контрольные задания, используемые для оценки

Согласована на портале № 13379

сформированности и освоения закрепленных за дисциплиной компетенций при проведении текущей, промежуточной аттестации по дисциплине приведен в приложении к рабочей программе.

14.2 Требования к фонду оценочных средств для лиц с ограниченными возможностями здоровья

Для студентов с инвалидностью предусмотрены дополнительные оценочные средства, перечень которых указан в таблице.

Таблица 14 – Дополнительные средства оценивания для студентов с инвалидностью

Категории студентов	Виды дополнительных оценочных средств	Формы контроля и оценки результатов обучения
С нарушениями слуха	Тесты, письменные самостоятельные работы, вопросы к зачету, контрольные работы	Преимущественно письменная проверка
С нарушениями зрения	Собеседование по вопросам к зачету, опрос по терминам	Преимущественно устная проверка (индивидуально)
С нарушениями опорно-двигательного аппарата	Решение дистанционных тестов, контрольные работы, письменные самостоятельные работы, вопросы к зачету	Преимущественно дистанционными методами
С ограничениями по общемедицинским показаниям	Тесты, письменные самостоятельные работы, вопросы к зачету, контрольные работы, устные ответы	Преимущественно проверка методами, исходя из состояния обучающегося на момент проверки

14.3 Методические рекомендации по оценочным средствам для лиц с ограниченными возможностями здоровья

Для студентов с ОВЗ предусматривается доступная форма предоставления заданий оценочных средств, а именно:

- в печатной форме;
- в печатной форме с увеличенным шрифтом;
- в форме электронного документа;
- методом чтения ассистентом задания вслух;
- предоставление задания с использованием сурдоперевода.

Студентам с инвалидностью увеличивается время на подготовку ответов на контрольные вопросы. Для таких студентов предусматривается доступная форма предоставления ответов на задания, а именно:

- письменно на бумаге;
- набор ответов на компьютере;
- набор ответов с использованием услуг ассистента;
- представление ответов устно.

Процедура оценивания результатов обучения инвалидов по дисциплине предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в форме электронного документа;
- в печатной форме увеличенным шрифтом.

Для лиц с нарушениями слуха:

- в форме электронного документа;
- в печатной форме.

Для лиц с нарушениями опорно-двигательного аппарата:

- в форме электронного документа;
- в печатной форме.

При необходимости для обучающихся с инвалидностью процедура оценивания результатов

Согласована на портале № 13379

обучения может проводиться в несколько этапов.

Согласована на портале № 13379

Приложение 1

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ

Федеральное государственное бюджетное образовательное учреждение высшего образования

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

УТВЕРЖДАЮ
Проректор по учебной работе
_____ П. Е. Троян
«___» _____ 20__ г.

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ПО УЧЕБНОЙ ДИСЦИПЛИНЕ

Анализ данных

Уровень образования: **высшее образование - бакалавриат**
Направление подготовки (специальность): **09.03.02 Информационные системы и технологии**
Направленность (профиль): **Информационные системы и технологии**
Форма обучения: **очная**
Факультет: **ФВС, Факультет вычислительных систем**
Кафедра: **ЭМИС, Кафедра экономической математики, информатики и статистики**
Курс: **2**
Семестр: **4**

Учебный план набора 2014 года

Разработчики:

- профессор каф. ЭМИС Колесникова С. И.

Экзамен: 4 семестр

Томск 2017

Согласована на портале № 13379

1. Введение

Фонд оценочных средств (ФОС) является приложением к рабочей программе дисциплины (практики) и представляет собой совокупность контрольно-измерительных материалов (типовые задачи (задания), контрольные работы, тесты и др.) и методов их использования, предназначенных для измерения уровня достижения студентом установленных результатов обучения.

ФОС по дисциплине (практике) используется при проведении текущего контроля успеваемости и промежуточной аттестации студентов.

Перечень закрепленных за дисциплиной (практикой) компетенций приведен в таблице 1.

Таблица 1 – Перечень закрепленных за дисциплиной компетенций

Код	Формулировка компетенции	Этапы формирования компетенций
ПК-25	способностью использовать математические методы обработки, анализа и синтеза результатов профессиональных исследований	Должен знать о основные понятия теории вероятности: аксиоматика теории вероятности, случайные события и основные теоремы теории вероятности; о методы описания и определения одно- и многомерных случайных величин; о предельные теоремы теории вероятности. ; Должен уметь о вычислять вероятности случайных событий; о находить числовые характеристики случайных величин; о решать прикладные задачи на основе методов теории вероятности и математической статистики. ; Должен владеть о методами решения вероятностных задач; о основными операциями над событиями и комбинаторными методами вычисления вероятности событий; о методами определения вероятностей сложных событий; о методами определения числовых характеристик случайных величин и функций от случайных величин; о методами точечного и интервального оценивания; о методами мышления: логическим, комбинаторно-вероятностным. ;

Общие характеристики показателей и критериев оценивания компетенций на всех этапах приведены в таблице 2.

Таблица 2 – Общие характеристики показателей и критериев оценивания компетенций по этапам

Показатели и критерии	Знать	Уметь	Владеть
Отлично (высокий уровень)	Обладает фактическими и теоретическими знаниями в пределах изучаемой области с пониманием границ применимости	Обладает диапазоном практических умений, требуемых для развития творческих решений, абстрагирования проблем	Контролирует работу, проводит оценку, совершенствует действия работы
Хорошо (базовый уровень)	Знает факты, принципы, процессы, общие понятия в пределах изучаемой области	Обладает диапазоном практических умений, требуемых для решения определенных проблем в	Берет ответственность за завершение задач в исследовании, приспосабливает свое

Согласована на портале № 13379

		области исследования	поведение к обстоятельствам в решении проблем
Удовлетворительно (пороговый уровень)	Обладает базовыми общими знаниями	Обладает основными умениями, требуемыми для выполнения простых задач	Работает при прямом наблюдении

2 Реализация компетенций

2.1 Компетенция ПК-25

ПК-25: способностью использовать математические методы обработки, анализа и синтеза результатов профессиональных исследований.

Для формирования компетенции необходимо осуществить ряд этапов. Этапы формирования компетенции, применяемые для этого виды занятий и используемые средства оценивания представлены в таблице 3.

Таблица 3 – Этапы формирования компетенции и используемые средства оценивания

Состав	Знать	Уметь	Владеть
Содержание этапов	математические методы обработки, анализа и синтеза результатов профессиональных исследований на основе основных принципов и понятий теории и методов анализа данных..	применять математические методы обработки, анализа многомерных данных	математическими методами и способами синтеза результатов профессиональных исследований для решения типовых задач машинного обучения.
Виды занятий	• Практические занятия; • Лекции; • Самостоятельная работа;	• Практические занятия; • Лекции; • Самостоятельная работа;	• Самостоятельная работа;
Используемые средства оценивания	• Домашнее задание; • Опрос на занятиях; • Конспект самоподготовки; • Дифференцированный зачет; • Экзамен;	• Домашнее задание; • Опрос на занятиях; • Конспект самоподготовки; • Дифференцированный зачет; • Экзамен;	• Домашнее задание; • Дифференцированный зачет; • Экзамен;

Формулировка показателей и критериев оценивания данной компетенции приведена в таблице 4.

Таблица 4 – Показатели и критерии оценивания компетенции на этапах

Состав	Знать	Уметь	Владеть
Отлично (высокий уровень)	• основные программные средства для создания математических моделей для эмпирических процессов и явлений; условия применимости стандартных пакетов	• свободно обосновывать и применять методы проведения численного моделирования и вычислительных экспериментов при обработке выборочных данных; применять	• несколькими программными средствами при проведении вычислительных экспериментов; способностью руководить междисциплинарной

Согласована на портале № 13379

	прикладных про-грамм для проведения численного моделирования; методы проведения численного моделирования. ;	основные программные средства для получения статистик на базе выборочных данных различного прикладного назначения. ;	командой для решения комплексных задач; статистическими методиками получения численных моделей процессов и явлений. ;
Хорошо (базовый уровень)	• аргументацию выбора программных средств для проведения численного моделирования; некоторые программные средства для создания математических моделей процессов и явлений; графические пакеты для иллюстрации решения задач. ;	• применять методы проведения численного моделирования при создании математических моделей эмпирических данных; применять некоторые программные средства для проведения статистической об-работки эмпирических данных ;	• способностью работать в междисциплинарной команде; некоторыми программными средствами при проведении численного моделирования; методикой проведения статистической обработки эмпирических данных разного прикладного на-значения. ;
Удовлетворительно (пороговый уровень)	• основные понятия и правила обработки данных; по крайней мере одну из прикладных программ для проведения статистической обработки эмпирических данных ;	• проводить статистическую обработку эмпирических данных процессов и объектов на базе готовых алгоритмов; представлять результаты своей работы. ;	• терминологией дисциплины и понятиями дисциплины ; основными методами проведения статистической обработки эмпирических данных с применением стандартных программных средств (по готовому алгоритму). ;

3 Типовые контрольные задания

Для реализации вышеперечисленных задач обучения используются типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы, в следующем составе.

3.1 Вопросы на самоподготовку

- Внедрение Data Mining, OLAP и хранилищ данных в СППР.
- Подходы к определению знаний и данных, информации.
- Концепция OLAP: оперативная аналитическая обработка данных.
- Системы анализа распределённых данных.
- Основные модели управления данными и распределённый анализ данных.

3.2 Темы домашних заданий

- Методы и принципы машинного обучения.
- Методы классификации и кластеризации.FRIS-функция и методики относительного оценивания расстояний.
- Методы построения математических моделей и прогнозирования временных рядов.
- Основные модели управления данными, многомерный анализ данных.

Согласована на портале № 13379

3.3 Темы опросов на занятиях

– Постановка задачи анализа данных. Сущность и условия применимости методов анализа данных.

– Модели Data Mining.

– Классификация методов анализа данных.

– Системы распознавания образов на базе Data Mining и принципы их построения.

– Примеры практического применения Data Mining.

3.4 Экзаменационные вопросы

– Технология Data Mining 1) это «раскопка данных», процесс получения из “сырых” данных новой, потенциально полезной информации о предметной области (скрытых знаний). В основу этих технологий положена концепция шаблонов (паттернов) и зависимостей, отражающих многоаспектные взаимоотношения в данных. Поиск паттернов производится автоматическими методами, не ограниченными рамками априорных предположений о структуре выборки и виде распределений значений анализируемых показателей; 2) это технология вычисления оценок (голосования); 3) основана на исчислении высказываний, в частности на аппарате алгебры логики; 4) это технология, в которой принадлежность объектов к кластерам определяется таким образом, чтобы: а) минимизировать изменчивость (различия) объектов внутри кластеров; б) максимизировать изменчивость объектов между кластерами.

– Кластерный анализ (самообучение, обучение без учителя, таксономия) применяется 1) при автоматическом формировании перечня образов по обучающей выборке; 2) в методе потенциальных функций; 3) при структурном (лингвистическом) подходе.

– Качество решающих правил оценивается 1) по репрезентативной выборке достаточно полно представляет генеральную совокупность (гипотетическое множество всех возможных объектов каждого образа); 2) по обучающей выборке - множеству объектов, заданных значениями признаков и принадлежность которых к тому или иному классу достоверно известна "учителю" и сообщается учителем "обучаемой" системе; 3) по контрольной (экзаменационной) выборке, в которую входят объекты, заданные значениями признаков, и принадлежность которых тому или иному образу известна только учителю.

– Имеется 20 стандартных ящиков однородных деталей. Вероятность того, что в одном взятом наудачу ящике Кластеризация – 1) это операция автоматической классификации, в ходе которой объекты объединяются в группы (кластеры) таким образом, что внутри групп различия между объектами минимальны, а между группами – максимальны. При этом в ходе кластеризации не только определяется состав кластеров, но и сам их набор и границы. 2) это операция, основанная на трудно формализуемых знаниях и интуиции исследователя. При этом исследователь сам определяет, какую информацию и каким образом система должна использовать для достижения требуемого эффекта классификации. 3) процесс формирования обобщенных образов классов, на основе обучающей выборки, содержащей характеристики конкретных объектов как в описательных, так и в классификационных шкалах и градациях.

– Обучение без учителя (самообучение) 1) процесс формирования обобщенных образов классов, на основе обучающей выборки, содержащей характеристики конкретных объектов как в описательных, так и в классификационных шкалах и градациях. 2) процесс формирования обобщенных образов классов, на основе обучающей выборки, содержащей характеристики конкретных объектов, причем только в описательных шкалах и градациях. 3) этот процесс называют кластерным анализом (таксономией).

– Интеллектуальный анализ данных (ИАД или data mining) – 1) Объекты описываются не множеством числовых значений признаков , а структурой объекта. Иерархия предполагает описание сложных объектов с помощью более простых подобъектов. Те, в свою очередь, могут быть описаны с помощью подобъектов следующего уровня и т.д. Этот подход основан на аналогии между структурой объектов и синтаксисом языков. Распознавание объекта состоит в распознавании непроизводных его элементов и синтаксическом анализе "предложения", описывающего данный объект. 2) это процесс обнаружения в "сырых" данных ранее неизвестных, нетривиальных, практически полезных и доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности; 3) это совокупность

Согласована на портале № 13379

математических моделей, численных методов, программных средств и информационных технологий, обеспечивающих обнаружение в эмпирических данных доступной для интерпретации информации и синтез на основе этой информации ранее неизвестных, нетривиальных и практически полезных для достижения определенных целей знаний.

3.5 Вопросы дифференцированного зачета

– 1. Основные шаги при осуществлении OLAP-анализа

– 2. Охарактеризуйте методы построения деревьев решений

– 3. Управляемые динамические стохастические временные ряды

– 4. Хаотические системы и их роль в описании реальных процессов. Изложить суть метода инвариантных многообразий

– 5. Условные вероятности. Формула полной вероятности. Формула Байеса.

– 6. Независимые испытания. Схема Бернулли. Формула Бернулли. Полиномиальная схема.

– 7. Предельные теоремы в схеме Бернулли. Локальная предельная теорема Муавра-Лапласа. Интегральная предельная теорема Муавра-Лапласа. Предельные теоремы в схеме Бернулли. Теорема Пуассона.

– 8. Что понимается под стандартами Data mining. Назовите стандартные типы закономерностей, которые позволяют выявлять методы Data Mining.

– 9. Закон распределения случайной величины, функция распределения вероятностей случайной величины. Ее свойства. Плотность распределения вероятностей случайной величины. Ее свойства.

– 10. Основные виды распределений дискретных и непрерывных случайных величин, их числовые характеристики. Подсчет вероятностей.

– 11. Принципиальные отличия временного ряда от простой последовательности наблюдений. Вопросы, на которые необходимо ответить для решения задачи прогнозирования.

– 12. Эмпирический риск. Оценивание качества алгоритма.

– 13. Классификационные правила: приведите примеры. Дайте определение опорным векторам

– 14. Дать краткую характеристику основным статистическим методам: корреляционный анализ, дисперсионный анализ, факторный анализ, кластерный анализ, дискриминантный анализ, регрессионный анализ, многомерное шкалирование

– 15. Типы данных. Виды метрик. Влияние метрики на качество классификации и кластеризации.

– 16. Центральная предельная теорема.

– 17. Основные задачи математической статистики. Описательная статистика. Эмпирическая функция распределения. Гистограмма, полигон частот. Выборочные характеристики.

– 18. Задачи теории оценивания. Точечное оценивание. Свойства точечных оценок. Оценки математического ожидания и дисперсии, их свойства.

– 19. Методы получения оценок: метод моментов. Пример. Методы получения оценок: метод максимального правдоподобия. Свойства МП-оценок.

– 20. Интервальное (доверительное) оценивание. Доверительный интервал для математического ожидания генеральной совокупности выборки.

– 21. Доверительное оценивание параметров нормального распределения.

– 22. Основы теории проверки статистических гипотез. Критерий согласия. Уровень значимости и мощность критерия. Критерий согласия Колмогорова о виде распределения (для простой гипотезы). Критерий согласия Хи-квадрат о виде распределения (для простой гипотезы). Проверка параметрических гипотез. Проверка гипотез, связанных с параметрами нормального распределения. Простая линейная регрессионная модель, оценки параметров по методу наименьших квадратов (МНК), значимость модели, адекватность модели.

4 Методические материалы

Для обеспечения процесса обучения и решения задач обучения используются следующие

Согласована на портале № 13379

материалы:

– методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы фор-мирования компетенций, согласно п. 12 рабочей программы.

4.1. Основная литература

1. Горлач, Б.А. Теория вероятностей и математическая статистика. [Электронный ресурс] — Электрон. дан. — СПб. : Лань, 2013. — 320 с. [Электронный ресурс]. - https://e.lanbook.com/book/4864#book_name

4.2. Дополнительная литература

1. Вентцель Е.С.,Овчаров Л.А. Задачи и упражнения по теории вероятностей : Учебное пособие для втузов / Е. С. Вентцель, Л. А. Овчаров. - 6-е изд., стереотип. - М. : Академия, 2005. (наличие в библиотеке ТУСУР - 99 экз.)

2. П. Е.Данко. Высшая математика в упражнениях и задачах [с решениями]: учебное пособие: в 2 ч. / П. Е.Данко [и др.]. - Ч. 2. - 7-е изд., испр. . - М. : ОНИКС, 2009. - 448с (наличие в библиотеке ТУСУР - 6 экз.)

3. В. Е. Гмурман Теория вероятностей и математическая статистика : Учебное пособие для вузов / В. Е. Гмурман ; Министерство образования и науки Российской Федерации. - 12-е изд., перераб. - М. : Высшее образование, 2006. – 478с. (наличие в библиотеке ТУСУР - 21 экз.)

4.3. Обязательные учебно-методические пособия

1. Колесникова С.И. Высшая математика III. Основы теории вероятностей. Элементы математической статистики. Методическое пособие. Томск: ТУСУР, 2007.–106 с. (наличие в библиотеке ТУСУР - 54 экз.)

2. Теория вероятностей и математическая статистика: Методические указания по выполнению практических работ / Колесникова С. И. - 2012. 28 с. [Электронный ресурс] - Режим доступа: http://edu.tusur.ru/publications/880, свободный.

3. Теория вероятностей и математическая статистика: Методические указания к самостоятельной работе студентов / Колесникова С. И. - 2012. 16 с. [Электронный ресурс] - Режим доступа: http://edu.tusur.ru/publications/881, свободный.

4. Анализ данных: Методические указания по самостоятельной работе / Колесникова С. И. - 2012. 18 с. [Электронный ресурс] - Режим доступа: http://edu.tusur.ru/publications/3053, свободный.

4.4. Базы данных, информационно справочные и поисковые системы

1. http://www.intuit.ru/department/mathematics/ptams/lit.html
2. http://www.exponenta.ru/Теории вероятностей
3. http://www.math-portal.ru/ycebnikiteorver
4. http://www.MatBuro.ru/Учебник по теории вероятностей
5. http://www.teorver-online.narod.ru

Согласована на портале № 13379