

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«КАЛИНИНГРАДСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ»

А.В. Снытников

**ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ИССЛЕДОВАТЕЛЬСКОЙ
ДЕЯТЕЛЬНОСТИ**

Учебно-методическое пособие по изучению дисциплины для студентов
магистратуры по направлению подготовки
09.04.01 – «Информатика и вычислительная техника»

Калининград
Издательство ФГБОУ ВО «КГТУ»
2023

УДК 519.6

Рецензент:

кандидат педагогических наук, доцент кафедры прикладной информатики
ФГБОУ ВО «Калининградский государственный технический университет»
Заболотнова Елена Юрьевна

Снытников, А.В.

Информационные технологии в исследовательской деятельности: учебно-методическое пособие по изучению дисциплины для студентов магистратуры по направлению подготовки 09.04.01- «Информатика и вычислительная техника»/
А.В. Снытников – Калининград: Изд-во ФГБОУ ВО «КГТУ», 2023. – 17 с.

Учебно-методическое пособие является руководством по изучению дисциплины «Информационные технологии в исследовательской деятельности» для студентов магистратуры по направлению подготовки 09.04.01 – «Информатика и вычислительная техника». Содержит характеристику дисциплины (цель и планируемые результаты изучения дисциплины, место дисциплины в структуре основной профессиональной образовательной программы), тематический план с описанием для каждой темы формы проведения занятия, вопросы для изучения, методические материалы к занятиям.

Учебно-методическое пособие по изучению дисциплины рекомендовано к использованию в учебном процессе в качестве локального электронного методического материала методической комиссией ИЦТ от 05.07.2023 г., протокол № 08.

© Федеральное государственное бюджетное
образовательное учреждение высшего образования
"Калининградский государственный технический
университет", 2023 г.
© Снытников А.В., 2023 г.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	4
1 ТЕМАТИЧЕСКИЙ ПЛАН.....	5
2. СОДЕРЖАНИЕ И МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ИЗУЧЕНИЮ ДИСЦИПЛИНЫ.....	5
Тема 1. Редакционно-издательская система LaTeX.	6
Тема 2. Символьные вычисления в языке Python.	7
Тема 3. Парная линейная регрессия.	8
Тема 4. Множественная регрессия	9
Тема 5. Прогнозирование во временных рядах на основе рекуррентных нейронных сетей.	10
Тема 6. Прогнозирование во временных рядах на основе модели ARIMA.....	10
Тема 7. Факторный анализ.	11
4 ТРЕБОВАНИЯ К АТТЕСТАЦИИ ПО ДИСЦИПЛИНЕ	12
Библиографический список	15

ВВЕДЕНИЕ

Дисциплина «Информационные технологии в исследовательской деятельности» является частью образовательной программы магистратуры по направлению подготовки 09.04.01 - «Информатика и вычислительная техника».

Дисциплина «Информационные технологии в исследовательской деятельности» относится к модулю «Наукоемкие информационные технологии».

Общая трудоемкость дисциплины составляет 3 зачетных единицы (з.е.), т.е. 108 академических часов контактной и самостоятельной учебной работы студента; работой, связанной с текущей и промежуточной (заключительной) аттестацией по дисциплине.

Целью освоения дисциплины «Информационные технологии в исследовательской деятельности» является формирование у студентов способности осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий, выбирать современные технологии и инструментальные средства для задач сбора и систематизации информации по проблеме с последующей её оценкой адекватности и достоверности.

В результате освоения дисциплины обучающийся должен:

Знать:

- теоретические и практические основы современных информационных технологий.

Уметь:

- применять методы математического моделирования для решения технических и исследовательских задач;

- использовать современные прикладные программные средства общего и специального назначения.

Владеть:

- навыками применения современного инструментария для решения технических задач;

- методикой построения и анализа математических моделей для оценки состояния и прогноза развития технических явлений и процессов.

В данном учебно-методическом пособии представлены методические материалы по изучению дисциплины, включающие тематический план занятий с перечнем ключевых вопросов для каждой лекции, рекомендуемой литературой, методическими указаниями и вопросами для самоконтроля. Изложены методические рекомендации по подготовке к лабораторным работам.

1 ТЕМАТИЧЕСКИЙ ПЛАН

Тематический план для студентов очной формы обучения приведен в табл.1.

Таблица 1 - Трудоёмкость освоения дисциплины в третьем семестре по очной форме обучения

№ п/п	Раздел (модуль) дисциплины	Контактная работа с преподавателем					СРС	Подготовка и аттестация в период сессии	
		ЛК	ЛР	ПР	РЭ	КА			
1	Редакционно-издательская система LaTeX								
2	Символьные вычисления в языке Python								
3	Парная линейная регрессия								
4	Множественная регрессия								
5	Прогнозирование во временных рядах на основе рекуррентных нейронных сетей								
6	Прогнозирование во временных рядах на основе модели ARIMA								
7	Факторный анализ								
ИТОГО:		16	30		14	0,15	47,85		
Всего		108							

2. СОДЕРЖАНИЕ И МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ИЗУЧЕНИЮ ДИСЦИПЛИНЫ

Приступая к изучению данного курса, рекомендуется придерживаться последовательности, соответствующей тематическому плану, приведенному в данном учебно-методическом пособии.

Приходя на очередную лекцию, необходимо обязательно предварительно просмотреть конспект предыдущей лекции. В случае вынужденного пропуска лекции нужно изучить ее самостоятельно, используя список рекомендованных литературных источников, приведенный в данном учебно-методическом пособии. Ответы на возникшие при этом вопросы можно решить с помощью рекомендованной литературы или на консультации у преподавателя.

Ниже приведен тематический план лекционных занятий.

Тематический план лекционных занятий

Таблица 2 – План лекционных занятий

Тема 1	Редакционно-издательская система LaTeX
Тема 2	Символьные вычисления в языке Python
Тема 3	Парная линейная регрессия
Тема 4	Множественная регрессия
Тема 5	Прогнозирование во временных рядах на основе рекуррентных нейронных сетей
Тема 6	Прогнозирование во временных рядах на основе модели ARIMA
Тема 7	Факторный анализ

Тема 1. Редакционно-издательская система LaTeX.

Ключевые вопросы темы

1. Команды системы LaTeX.
2. Формат документа LaTeX.
3. Математические выражения.
4. Библиография в LaTeX.
5. Оформление презентаций в LaTeX
6. Использование рисунков в документе LaTeX.
7. Шаблон для написания ВКР в LaTeX..

Источники

[1, с. 65-71], [2, с. 7-35].

Методические рекомендации

Начать изучение работы в редакционно-издательской системе LaTeX целесообразно с простого текста, содержащего математические формулы с греческими буквами, знаками интеграла, дробями, сложными скобками и матрицами. Студенты должны убедиться в относительной простоте набора математических текстов в LaTeX по сравнению с MS Word.

Для того, чтобы исключить при первом знакомстве с LaTeX вопросы конфигурирования и установки, лучше начинать работу в одном из онлайн-редакторов LaTeX (overleaf.com, papeeria.com).

Следующим вопросом должно быть построение списка литературы в LaTeX на основе генератора библиографий BIBTEX. Студенты должны научиться составлять список нужной им литературы автоматически, без ручного набора библиографических данных с оформлением списка литературы по требуемому стандарту (например, по ГОСТу).

Далее студенты должны научиться вставлять в документ рисунки и делать ссылки на них в тексте. Итогом изучения данной темы является презентация каждого студента по своей ВКР бакалавра, набранная в системе LaTeX.

Вопросы для самоконтроля

1. Как оформляются индексы в LaTeX?
2. Как набрать знак интеграла с пределами в LaTeX?
3. Чем отличаются стили документа article и book?
4. Какие форматы рисунков можно использовать?
5. Какой стиль документа используется для оформления презентаций?
6. Как набрать таблицу в LaTeX?
7. Как вставить содержание документа (распределение заголовков по страницам)?

Тема 2. Символьные вычисления в языке Python.

Ключевые вопросы темы

1. Библиотека SymPy. Введение
2. Символьные переменные.
3. Упрощение выражений, раскрытие скобок.
4. Символьное дифференцирование.
5. Символьное интегрирование, определенные и неопределенные интегралы.
6. Преобразование символьных выражений в формат LaTeX.
7. Подстановка значений символьных переменных.

Источники

[1, с. 71-91], [2, с. 46-64], [3, с. 8,13,28,32-34].

Методические рекомендации

При изучении данной темы необходимо начать с ключевого понятия – символьных переменных, которые обеспечивают возможность преобразования выражений, дифференцирования и интегрирования, студентам следует четко уяснить отличие символьных переменных от обычных переменных языка Python.

Далее рассматривается преобразование выражений, раскрытие скобок и приведение подобных. Затем - символьное дифференцирование и интегрирование, с обязательным рассмотрением неопределенных интегралов, а также интегралов по бесконечному интервалу.

Заканчивать тему следует вопросом о вычислении символьных выражений путем подстановки числовых значений символьных переменных.

Вопросы для самоконтроля

1. В чем отличие символьных переменных от обычных?
2. Какая функция библиотеки SymPy производит раскрытие скобок?
3. Как выполнить дифференцирование до 4-го порядка?
4. Можно ли вычислить интеграл по бесконечному интервалу средствами SymPy?
5. Как преобразовать символьное выражение в формат LaTeX?
6. Какая функция библиотеки SymPy производит подстановку значения символьной переменной?

Тема 3. Парная линейная регрессия.

Ключевые вопросы темы

1. Понятие регрессии. Корреляция
2. Условие применимости линейной регрессии.
3. Библиотека Scikitlearn.
4. Построение парной линейной регрессии средствами языка Python.
5. Погрешность парной линейной регрессии.

Источники

[3, с. 94], [2, с. 91-115], [3, с. 45,49].

Методические рекомендации

Перед началом изучения парной линейной регрессии необходимо ознакомиться с самим понятием регрессии, ее задачами и основными методами их решения.

Следующим целесообразно рассмотреть вопросом применимости парной линейной регрессии, и основанный на коэффициенте корреляции Пирсона критерий.

Далее приводится краткое описание библиотеки Scikitlearn и ее основных возможностей, и в частности, класс LinearRegression, содержащий функции для вычисления линейной регрессии.

Завершает изучение этой темы вопрос об ошибке аппроксимации и характеристиках правильно подобранной модели.

Вопросы для самоконтроля

1. Назовите необходимое и достаточное условие применимости линейной регрессии.
2. Что входит в понятие построения линейной регрессии (что выполняет функция fit)?
3. Как вывести коэффициент и свободный член линейной регрессии для построенной модели?

4. В каком интервале значений независимой переменной применимы предсказания, полученные на основе линейной регрессии?
5. Как вычисляется ошибка аппроксимации?
6. Каковы характеристики хорошо подобранной модели

Тема 4. Множественная регрессия

Ключевые вопросы темы

1. Общее описание множественной регрессии. Метод наименьших квадратов.
2. Условие применимости множественной линейной регрессии. Матрица корреляции.
3. Библиотека Scikitlearn. Функции для работы с большими задачами
4. Построение множественной линейной регрессии средствами языка Python.
5. Погрешность множественной линейной регрессии.

Источники

[1, с. 5-21], [2, с. 142-183], [3, с. 64-67].

Методические рекомендации

Начинать рассмотрение данной темы следует с метода, которым находится решение задачи о множественной регрессии в библиотеке Scikitlearn, т.е. с метода наименьших квадратов, далее необходимо рассмотреть вопрос о применимости процедуры множественной регрессии на основе матрицы корреляции.

Затем следует подробно рассмотреть процедуры и функции библиотеки Scikitlearn, позволяющие решать многомерные задачи, и в особенности класс LinearRegression и его применение для решения задачи множественной регрессии.

Завершить рассмотрение данной темы следует решением вопроса о нахождении наиболее значимой среди независимых переменных путем сортировки массива

Вопросы для самоконтроля

1. Каковы условия применимости множественной регрессии?
2. Как измерить ошибки предсказаний, выполненных с помощью множественной регрессии?
3. Как вывести коэффициенты и свободный член множественной регрессии?
4. Могут ли для множественной регрессии использоваться нечисловые переменные?
5. Как найти наиболее значимую из независимых переменных (переменную, которая оказывает наибольшее влияние на результат)?

Тема 5. Прогнозирование во временных рядах на основе рекуррентных нейронных сетей.

Ключевые вопросы темы

1. Построение обучающего множества на основе временного ряда.
2. Вид функции потерь для задачи прогнозирования во временных рядах.

Роль функции потерь в задачах машинного обучения.

3. Общее описание рекуррентных нейронных сетей. Сети LSTM и GRU.
4. Функции активации.
5. Многослойные нейронные сети. Особенности глубокого обучения

Источники

[1, с. 21-26], [2, с. 184-203], [3, с. 67,68,72]

Методические рекомендации

В рамках данной темы необходимо ознакомиться с основами функционирования и обучения нейронных сетей.

Начать рекомендуется с изучения формирования обучающего множества для задачи прогнозирования во временных рядах.

Далее следует рассмотреть вид функции потерь для данной конкретной задачи и роль функции потерь в задачах машинного обучения и нейронных сетях в целом.

Затем перейти к основному для данной темы вопросу об архитектуре рекуррентных нейронных сетей, и подробно рассмотреть два основных на данный момент вида рекуррентных нейронных сетей – LSTM и GRU.

Завершить изучение данной темы рассмотрением влияния большого количества слоев нейронной сети на качество решения задачи прогнозирования.

Вопросы для самоконтроля

1. Для чего нужна функция потерь?
2. В чем отличие обучающего и проверочного множества?
3. Какие существуют функции активации?
4. Как выполняется построение обучающего множества на основе временного ряда?
5. В чем преимущество многослойных нейронных сетей?

Тема 6. Прогнозирование во временных рядах на основе модели ARIMA.

Ключевые вопросы темы

1. Определение авторегрессионных моделей.
2. Модели скользящего среднего.
3. Задача о приближении заданной кривой рядом на основе линейных функций.
4. Понятие стационарности ряда.

5. Минимизация ошибок. Статистика Акаике.
6. Параметры модели ARIMA. Метод подбора оптимальных значений параметров
7. Критерии Акаике, Байеса и Ханнана-Куина.

Источники

[1, с. 26-37], [2, с. 204-119, 234-256], [3, с. 72,73,79,80]

Методические рекомендации

Рассмотрение данной темы необходимо начинать с определения авторегрессионных моделей, а также с условия их применимости, а именно автокорреляции различных интервалов временного ряда.

В рамках данной темы необходимо разобраться как определяются три основных параметра модели ARIMA: параметр авторегрессии, параметр дифференцирования и параметр скользящего среднего. Важно научиться находить значения этих параметров, оптимальной с точки зрения результатов прогнозирования.

Далее рассматривается формирование ошибки аппроксимации и используемый при этом аппарат математической статистики, в частности, статистика Х.Акаике.

В завершение темы рассматриваются три информационных критерия, позволяющие оценивать качество построенной модели: критерий Акаике, Байеса и Ханнана-Куина

Вопросы для самоконтроля

1. Расшифруйте название модели — ARIMA
2. Объясните смысл трех основных параметров модели — $model = ARIMA(series, order=(5,1,0))$.
3. В чем значение графика автокорреляции
4. В чем значение графика плотности остаточных ошибок
5. Что означают величины AIC, BIC и HQIC?

Тема 7. Факторный анализ.

Ключевые вопросы темы

1. Фактор.
2. Факторная нагрузка.
3. Коммунальность.
4. Собственное число.
5. Факторное пространство.

Источники

[1, с. 52-57], [2, с. 275-303], [3, с. 104,109-111].

Методические рекомендации

Изучение данной темы необходимо начинать с изучения понятий и определений, которые входят в его теоретический аппарат факторного анализа. Наиболее важными из них являются понятия «фактор», «факторная нагрузка», «коммунальность», «собственное число», «факторный пространство».

Далее студенты осваивают основной принцип работы факторного анализа, который заключается в том, что анализируется матрица данных, где каждый столбец представляет собой переменную, а каждая строка – наблюдение.

Затем рассматривается основной принцип работы факторного анализа, а именно, группировка переменных по свойствам и их факторизация. При этом необходимо научиться отслеживать разделение на основу общих свойств, которые присутствуют в их корреляционной структуре. При этом студенты должны начинать видеть зависимости между переменными и определять, какие из них наиболее значимы, а какие не влияют на общую картину.

Вопросы для самоконтроля

1. Что такое фактор?
2. Как определяются и какую роль играют собственные числа?
3. Что такое факторная нагрузка?
4. Что такое коммунальность?
5. Где может применяться факторный анализ?
6. Что такое факторное пространство?

4 ТРЕБОВАНИЯ К АТТЕСТАЦИИ ПО ДИСЦИПЛИНЕ

Текущий контроль предназначен для проверки хода и качества усвоения студентами учебного материала и стимулирования их учебной работы. Он может осуществляться в ходе всех видов занятий в форме, избранной преподавателем или предусмотренной рабочей программой дисциплины.

Текущий контроль предполагает постоянный контроль преподавателем качества усвоения учебного материала, активизацию учебной деятельности студентов на занятиях, побуждение их к самостоятельной систематической работе. Он необходим обучающимся для самоконтроля на разных этапах обучения.

Промежуточная аттестация по дисциплине проводится в форме зачета на основе результатов текущей успеваемости по дисциплине.

Таблица 2 – Система оценок и критерии выставления оценки

Система оценок Критерий	2	3	4	5
	0-40%	41-60%	61-80 %	81-100 %
	«неудовлетворительно»	«удовлетворительно»	«хорошо»	«отлично»
	«не зачтено»	«зачтено»		
1 Системность и полнота знаний в отношении изучаемых объектов	Обладает частичными и разрозненными знаниями, которые не может научно-корректно связывать между собой (только некоторые из которых может связывать между собой)	Обладает минимальным набором знаний, необходимым для системного взгляда на изучаемый объект	Обладает набором знаний, достаточным для системного взгляда на изучаемый объект	Обладает полнотой знаний и системным взглядом на изучаемый объект
2 Работа с информацией	Не в состоянии находить необходимую информацию, либо в состоянии находить отдельные фрагменты информации в рамках поставленной задачи	Может найти необходимую информацию в рамках поставленной задачи	Может найти, интерпретировать и систематизировать необходимую информацию в рамках поставленной задачи	Может найти, систематизировать необходимую информацию, а также выявить новые, дополнительные источники информации в рамках поставленной задачи
3. Научное осмысление изучаемого явления, процесса, объекта	Не может делать научно корректных выводов из имеющихся у него сведений, в состоянии проанализировать только некоторые из имеющихся у него сведений	В состоянии осуществлять научно корректный анализ предоставленной информации	В состоянии осуществлять систематический и научно корректный анализ предоставленной информации, вовлекает в исследование новые релевантные задаче данные	В состоянии осуществлять систематический и научно-корректный анализ предоставленной информации, вовлекает в исследование новые релевантные поставленной задаче данные, предлагает новые ракурсы поставленной задачи

Система оценок Критерий	2	3	4	5
	0-40%	41-60%	61-80 %	81-100 %
	«неудовлетворительно»	«удовлетворительно»	«хорошо»	«отлично»
	«не зачтено»	«зачтено»		
4. Освоение стандартных алгоритмов решения профессиональных задач	В состоянии решать только фрагменты поставленной задачи в соответствии с заданным алгоритмом, не освоил предложенный алгоритм, допускает ошибки	В состоянии решать поставленные задачи в соответствии с заданным алгоритмом	В состоянии решать поставленные задачи в соответствии с заданным алгоритмом, понимает основы предложенного алгоритма	Не только владеет алгоритмом и понимает его основы, но и предлагает новые решения в рамках поставленной задачи

Библиографический список

Основная литература:

1. Рочев, К. В. Информационные технологии. Анализ и проектирование информационных систем: учебное пособие для вузов / К. В. Рочев. — 3-е изд., стер. — Санкт-Петербург: Лань, 2022. — 128 с. — Режим доступа: для авториз. пользователей. — Лань: электронно-библиотечная система. — URL: <https://e.lanbook.com/book/223442> (дата обращения: 23.02.2023). — ISBN 978- 5-507-44339-0. — Текст: электронный.
2. Яхонтова, И. М. Информационные технологии в науке, производстве и образовании: учебное пособие / И. М. Яхонтова, Т. А. Крамаренко. — Краснодар: КубГАУ, 2020. — 122 с. — Режим доступа: для авториз. пользователей. — Лань: электронно-библиотечная система. — URL: <https://e.lanbook.com/book/254285> (дата обращения: 24.02.2023). — ISBN 978-5-907346-88-8. — Текст: электронный.

Дополнительная литература:

1. Информационные технологии: учебник / Ю. Ю. Громов, И. В. Дидрих, О. Г. Иванова, и др.; Тамбовский государственный технический университет. — Тамбов: Тамбовский государственный технический университет (ТГТУ), 2015. — 260 с. — Режим доступа: по подписке. — URL: <https://biblioclub.ru/index.php?page=book&id=444641> (дата обращения: 23.02.2023). — ISBN 978-5-8265-1428-3. — Текст: электронный.
2. Советов, Б. Я. Информационные технологии: теоретические основы: учебное пособие / Б. Я. Советов, В. В. Цехановский. — 2-е изд., стер. — СанктПетербург: Лань, 2022. — 444 с. — Режим доступа: для авториз. пользователей. — Лань: электронно-библиотечная система. — URL: <https://e.lanbook.com/book/209876> (дата обращения: 24.02.2023). — ISBN 978- 5-8114-1912-8. — Текст: электронный.
3. Топоркова, О. М. Информационные технологии: учеб. пособие для студентов высш. учеб. заведений, обучающихся в бакалавриате по направлению подгот.: "Информатика и вычисл. техника" ; "Прикладная информатика" / О. М. Топоркова ; Калинингр. гос. техн. ун-т. - Калининград: КГТУ, 2015. - 103, [1] с. - Текст: непосредственный.
4. Исакова, А. И. Информационные технологии: учебное пособие / А. И. Исакова; Томский Государственный университет систем управления и радиоэлектроники (ТУСУР), Кафедра автоматизированных систем управления (АСУ). — Томск: ТУСУР, 2013. — 207 с. — Режим доступа: по подписке. — URL: <https://biblioclub.ru/index.php?page=book&id=480610> (дата обращения: 23.02.2023). — Текст: электронный.
5. Чикунова, Н. Ф. Проектирование баз данных и организация их защиты в СУБД ACCESS: учеб. пособие по дисциплине "Безопасность систем баз данных" для студентов специальности 10.05.03 "Информ. безопасность автоматизир. систем" очной формы обучения / Н. Ф. Чикунова; Федер.

агентство по рыболовству [и др.]. - Калининград: БГАРФ, 2019 -. - Текст: непосредственный. Ч. 1. - 2019. - 106 с.

6. Чикунова, Н. Ф. Проектирование баз данных и организация их защиты в СУБД MySQL: учеб. пособие по дисциплине "Безопасность систем баз данных" для студентов специальности 10.05.03 "Информ. безопасность автоматизир. систем" / Н. Ф. Чикунова ; Федер. агентство по рыболовству [и др.]. - Калининград: БГАРФ, 2020 -. - Текст: непосредственный. Ч. 2. - 2020. - 92 с.

Локальный электронный методический материал

Снытников Алексей Владимирович

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ИССЛЕДОВАТЕЛЬСКОЙ
ДЕЯТЕЛЬНОСТИ

Редактор

Локальное электронное издание

Уч.-изд. л. . Печ. л.

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Калининградский государственный технический университет».
236022, Калининград, Советский проспект, 1