

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение высшего образования
«Новгородский государственный университет имени Ярослава Мудрого»
Институт электронных и информационных систем
Кафедра физики твердого тела и микроэлектроники



УТВЕРЖДАЮ
Директор ИЭИС

С.И.Эминов

«25» января 2020 г.

РАБОЧАЯ ПРОГРАММА
учебной дисциплины

ОБРАБОТКА И АНАЛИЗ ДАННЫХ

по направлению подготовки

11.03.04 Электроника и нанoeлектроника

направленность (профиль) Микроэлектроника и твердотельная электроника

СОГЛАСОВАНО

Зам. директора ИЭИС

Ариас Е.А.Ариас
«24» 12 2020 г.

Разработал

Доцент КФТТМ

И.С.Телина
«25» 11 2020 г.

Принято на заседании кафедры ФТТМ

Протокол № 4 от «26» 11 2020 г.

Заведующий кафедрой ФТТМ

Б.И.Селезнев
«26» 11 2020 г.

1 Цели и задачи освоения учебной дисциплины

Цель освоения учебной дисциплины: формирование компетентности обучающихся в области обработки и анализа данных, способствующей подготовке обучающихся к научно-исследовательской деятельности, а также к решению задач профессиональной деятельности.

Задачи:

- а) формирование представления у обучающихся о месте и роли анализа данных в процессах моделирования и исследования современных технических систем;
- б) формирование системы базовых знаний о методах математической статистики;
- в) знакомство с возможностями языка Python для анализа и обработки данных;
- г) формирование умений решать прикладные задачи по обработке, анализу и визуализации данных с использованием Python.

2 Место учебной дисциплины в структуре ОПОП

Учебная дисциплина относится к обязательной части учебного плана основной профессиональной образовательной программы направления подготовки 11.03.04 Электроника и нанoeлектроника и направленности (профилю) Микроэлектроника и твердотельная электроника (далее – ОПОП). Для освоения дисциплины обучающимся необходимы знания информационных технологий и прикладной математики. Полученные в процессе изучения знания могут использоваться при прохождении учебной практики, написании курсовых и выпускных квалификационных работ.

3 Требования к результатам освоения учебной дисциплины

Перечень компетенций, которые формируются в процессе освоения учебной дисциплины:

Общепрофессиональные компетенции:

ОПК-3 Способен применять методы поиска, хранения, обработки, анализа и представления в требуемом формате информации из различных источников и баз данных, соблюдая при этом основные требования информационной безопасности;

ОПК-4 Способен применять современные компьютерные технологии для подготовки текстовой и конструкторско-технологической документации с учетом требований нормативной документации.

Результаты освоения учебной дисциплины представлены в таблице 1.

Таблица 1 - Результаты освоения учебной дисциплины

Код и наименование компетенции	Результаты освоения учебной дисциплины (индикаторы достижения компетенций)		
ОПК-3 Способен применять методы поиска, хранения, обработки, анализа и представления в требуемом формате информации из различных источников и баз данных, соблюдая при этом основные требования информационной безопасности	Знать - как использовать информационно-коммуникационные технологии при поиске необходимой информации - современные принципы поиска, хранения, обработки, анализа и представления в требуемом формате информации	Уметь - решать задачи обработки данных с помощью современных средств автоматизации	Владеть - навыками обеспечения информационной безопасности

Код и наименование компетенции	Результаты освоения учебной дисциплины (индикаторы достижения компетенций)		
ОПК-4 Способен применять современные компьютерные технологии для подготовки текстовой и конструкторско-технологической документации с учетом требований нормативной документации	Знать - как использовать информационно-коммуникационные технологии при поиске необходимой информации - современные интерактивные программные комплексы для выполнения и редактирования текстов, изображений и чертежей	Уметь - проектировать решение конкретной задачи проекта, выбирая оптимальный способ ее решения, исходя из действующих правовых норм и имеющихся ресурсов и ограничений - использовать современные средства автоматизации разработки и выполнения конструкторской документации	Владеть - современными программными средствами подготовки конструкторско-технологической документации

4 Структура и содержание учебной дисциплины

4.1 Трудоемкость учебной дисциплины

Таблица 2 - Трудоемкость учебной дисциплины для очной формы обучения

Части учебной дисциплины (модуля)	Всего	Распределение по семестрам
		3 семестр
1. Трудоемкость учебной дисциплины (модуля) в зачетных единицах (ЗЕТ)	3	3
2. Контактная аудиторная работа в академических часах (АЧ)	42	42
3. Курсовая работа/курсовой проект (АЧ) <i>(при наличии)</i>	-	-
4. Внеаудиторная СРС в академических часах (АЧ)	66	66
5. Промежуточная аттестация (зачет; дифференцированный зачет; экзамен) (АЧ)	ДЗ	ДЗ

4.2 Содержание учебной дисциплины

Раздел № 1 Основы статистических методов обработки и анализа данных

- 1.1 Основы статистического анализа данных. Генеральная совокупность и выборка. Источники данных. Формат данных (бинарные файлы, текстовые файлы, базы данных). Удаление выбросов, заполнение пропусков, фильтрация, объединение и сортировка данных, изменение типов данных. Расчёт описательных статистик выборки. Анализ закона распределения одномерных случайных величин. Графический анализ данных.
- 1.2 Исследование статистических взаимосвязей. Проверка статистических гипотез. Критерии согласия. Поиск взаимосвязей в данных и оценка их статистической значимости. Сравнение групп данных. Одновыборочные и двухвыборочные критерии. Сравнение нескольких выборок.

- Раздел № 2 Программные системы для решения задач обработки и анализа данных
- 2.1 Особенности языка Python. Основы синтаксиса. Среда разработки (PyCharm и блокноты Jupyter Notebook). Простые типы данных. Переменные. Списки. Строки. Регулярные выражения. Словари. Условия. Циклы. Функции. Классы. Файлы и исключения. Модули Python.
- 2.2 Анализ данных с использованием Python. Работа с библиотеками Numpy и SciPy. Анализ данных с помощью библиотеки Pandas. Работа с объектами Series и DataFrame. Группировка и агрегирование данных. Сводные таблицы. Визуальный анализ данных. Библиотека Matplotlib. Методы библиотек Seaborn и Plotly.

4.3 Трудоемкость разделов учебной дисциплины и контактной работы

Таблица 3 - Трудоемкость разделов учебной дисциплины для очной формы обучения

№	Наименование разделов (тем) учебной дисциплины (модуля), УЭМ, наличие КП/КР	Контактная работа (в АЧ)			Внеауд. СРС (в АЧ)	Формы текущего контроля	
		Аудиторная					
		ЛЕК	ПЗ	ЛР			
		В	т.ч.	СРС			
1.	Раздел № 1 Основы статистических методов обработки и анализа данных	7	14	0	3	33	
2.	1.1 Основы статистического анализа данных	3	7			15	задание
3.	1.2 Исследование статистических взаимосвязей	4	7		3	18	задание
4.	Раздел № 2 Программные системы для решения задач обработки и анализа данных	0	21	0	3	33	
5.	2.1 Особенности языка Python		10			20	задание
6.	2.2 Анализ данных с использованием Python		11		3	13	задание
7.	Промежуточная аттестация						ДЗ
	ИТОГО	7	35	0	6	66	

4.4 Лабораторные работы и курсовые работы/курсовые проекты

4.5.1 Перечень тем лабораторных работ:

Лабораторные работы не предусмотрены учебным планом.

4.5.2 Примерные темы курсовых работ/курсовых проектов:

Курсовые работы/курсовые проекты не предусмотрены учебным планом.

5 Методические рекомендации по организации освоения учебной дисциплины

Таблица 4 - Методические рекомендации по организации лекций

№	Темы лекционных занятий (форма проведения)	Трудоемкость в АЧ
1.	Основы статистического анализа данных. Генеральная совокупность и выборка. Источники данных. Формат данных (бинарные файлы, текстовые файлы, базы данных). Удаление выбросов, заполнение пропусков, фильтрация, объединение и сортировка данных, изменение типов данных. Расчёт описательных статистик выборки. Анализ закона распределения одномерных случайных величин. Графический анализ данных (<i>информационная лекция</i>)	3
2.	Исследование статистических взаимосвязей. Проверка статистических гипотез. Критерии согласия. Поиск взаимосвязей в данных и оценка их статистической значимости. Сравнение групп данных. Одновыборочные и двухвыборочные критерии. Сравнение нескольких выборок (<i>информационная лекция</i>)	4
	ИТОГО	7

Таблица 5 - Методические рекомендации по организации практических занятий

№	Темы практических занятий (форма проведения)	Трудоем- кость в АЧ
1.	Предобработка данных. Удаление выбросов, заполнение пропусков, фильтрация, объединение и сортировка данных, изменение типов данных. Расчёт описательных статистик выборки. Анализ закона распределения одномерных случайных величин. Графический анализ данных (задание)	7
2.	Исследование статистических взаимосвязей. Проверка статистических гипотез. Критерии согласия. Поиск взаимосвязей в данных и оценка их статистической значимости. Сравнение групп данных. Одновыборочные и двухвыборочные критерии. Сравнение нескольких выборок (задание)	7
3.	Особенности языка Python. Основы синтаксиса. Среды разработки (PyCharm и блокноты Jupyter Notebook). Простые типы данных. Переменные. Списки. Строки. Регулярные выражения. Словари. Условия. Циклы. Функции. Классы. Файлы и исключения. Модули Python (задание)	10
4.	Анализ данных с использованием Python. Работа с библиотеками Numpy и SciPy. Анализ данных с помощью библиотеки Pandas. Работа с объектами Series и DataFrame. Группировка и агрегирование данных. Сводные таблицы. Визуальный анализ данных. Библиотека Matplotlib. Методы библиотек Seaborn и Plotly (задание)	11
ИТОГО		35

6 Фонд оценочных средств учебной дисциплины

Фонд оценочных средств представлен в Приложении А.

7 Условия освоения учебной дисциплины

7.1 Учебно-методическое обеспечение

Учебно-методическое обеспечение учебной дисциплины (модуля) представлено в Приложении Б.

7.2 Материально-техническое обеспечение

Таблица 6 - Материально-техническое обеспечение учебной дисциплины

№	Требование к материально-техническому обеспечению согласно ФГОС ВО	Наличие материально-технического оборудования	
1.	Учебные аудитории для проведения учебных занятий	аудитория для проведения лекционных: учебная мебель (столы, стулья, доска)	
		компьютерный класс с выходом в Интернет, в том числе для проведения практических занятий	
		помещения для самостоятельной работы (наличие компьютера, выход в Интернет)	
2.	Мультимедийное оборудование	проектор, экран, выход в интернет	
3.	Программное обеспечение		
Наименование программного продукта		Обоснование для использования (лицензия, договор, счёт, акт или иное)	Дата выдачи
Microsoft Windows 7 Professional		Dreamspark (Imagine) № 370aef61-476a-4b9f-bd7c-84bb13374212	30.04.2015
Microsoft Windows 10 for Educational Use		Dreamspark (Imagine) № 370aef61-476a-4b9f-bd7c-84bb13374212	30.04.2015
Microsoft Office 2013 Standard		Open License № 62018256	31.07.2016
Microsoft Imagine (Microsoft Azure Dev Tools for Teaching) Standard		Договор №243/ю, 370aef61-476a-4b9f-bd7c-84bb13374212	19.12.2018

ABBYY FineReader PDF 15 Business. Версия для скачивания (годовая лицензия с академической скидкой)*	Договор №191/Ю	16.11.2020
Kaspersky Endpoint Security для бизнеса – Стандартный Russian Edition. 500-999. Node 1 year Educational Renewal License *	Договор №148/ЕП(У)20-ВБ, 1С1С-200914-092322-497-674	11.09.2020
Антиплагиат. Вуз.*	Договор №3341/12/ЕП(У)21-ВБ	29.01.2021
Подписка Microsoft Office 365	свободно распространяемое для вузов	-
Adobe Acrobat	свободно распространяемое	-
Teams	свободно распространяемое	-
Skype	свободно распространяемое	-
Zoom	свободно распространяемое	-
IDE PyCharm Community Edition, Anaconda (Python 3.7 или выше, Jupyter Notebook и библиотеки: numpy, pandas, scipy, matplotlib, seaborn и др.)	свободно распространяемое	-

* отечественное производство

Приложение А
(обязательное)
Фонд оценочных средств
учебной дисциплины «Обработка и анализ данных»

1 Структура фонда оценочных средств

Фонд оценочных средств состоит из двух частей:

а) открытая часть - общая информация об оценочных средствах (название оценочных средств, проверяемые компетенции, баллы, количество вариантов заданий, методические рекомендации для применения оценочных средств и пр.), которая представлена в данном документе, а также те вопросы и задания, которые могут быть доступны для обучающегося;

б) закрытая часть - фонд вопросов и заданий, которая не может быть заранее доступна для обучающихся (экзаменационные билеты, вопросы к контрольной работе и пр.) и которая хранится на кафедре.

2 Перечень оценочных средств текущего контроля и форм промежуточной аттестации

№	Оценочные средства для текущего контроля	Разделы (темы) учебного модуля	Баллы	Проверяемые компетенции
1.	Задание	Темы: 1.1, 1.2, 2.1, 2.2	30х5	ОПК-3, ОПК-4
<i>Промежуточная аттестация</i>				
	Дифференцированный зачет		-	ОПК-3, ОПК-4
	ИТОГО		150	

3 Рекомендации к использованию оценочных средств

1) Задание

Критерии оценки	Количество вариантов заданий
1 Правильность выполнения задания	10
2 Корректность кода	
3 Полнота ответов на уточняющие вопросы	

Все задания выполняются с использованием языка Python.

Примерные задания:

- Загрузите данные из табличного документа (на выбор преподавателя). Выполните предобработку данных:
 - удалите дубликаты строк и столбцов, если есть;
 - измените типы данных, если это необходимо;
 - найдите пропуски и удалите/заполните их;
 - найдите аномальные значения и удалите/замените их;
 - сохраните данные в файл.
- Загрузите данные из табличного документа (на выбор преподавателя). Выполните следующее:
 - постройте парные графики признаков. Сделайте вывод о взаимном влиянии переменных;

- рассчитайте корреляционную матрицу. Визуализируйте результат. Сделайте вывод;
- сравните распределения двух переменных (на выбор), используя подходящий статистический критерий;
- оцените согласованность частот в двух группах (на выбор), используя подходящий статистический критерий.

Все материалы для проведения промежуточного контроля хранятся на кафедре.

Приложение Б
(обязательное)
Карта учебно-методического обеспечения
учебной дисциплины «Обработка и анализ данных»

1. Основная литература

Библиографическое описание издания (автор, наименование, вид, место и год издания, кол. стр.)	Кол. экз. в библ. НовГУ	Наличие в ЭБС
Печатные источники		
1 Миркин Б. Г. Введение в анализ данных : учебник и практикум для вузов / Б. Г. Миркин ; Нац. исслед. ун-т - Высш. шк. экономики. - Москва : Юрайт, 2014. - 173, [2] с. : ил. - (Авторский учебник). - Библиогр.: с. 173-174. - Прил.: с. 158-170. - Указ.: с. 171-172. - Кн. доступна в ЭБС biblio-online.ru. - ISBN 978-5-9916-4120-3	2	
2 Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP : учебное пособие для вузов. - 2-е изд., перераб. и доп. - Санкт-Петербург : БХВ-Петербург, 2007. - 375с. : ил.+ CD-ROM. - (Учебное пособие). - Библиогр.: с. 368-371. - Указ.: с. 372-375. - На обл.: Хранилища данных - ISBN 5-94157-991-8	5	
3 Чашкин Ю. Р. Математическая статистика. Анализ и обработка данных : учебное пособие для вузов / Ю. Р. Чашкин. - 2-е изд., перераб. и доп. - Ростов-на-Дону : Феникс, 2010. - 236, [2] с. : ил. - (Высшее образование). - Библиогр.: с. 232-234. - Прил.: с. 210-231. - ISBN 978-5-222-16474-7	11	
Электронные ресурсы		
1 Официальная документация по библиотеке pandas. – Pandas: http://pandas.pydata.org/pandas-docs/stable/		
2 Официальная документация по библиотеке matplotlib. – Matplotlib: https://matplotlib.org/contents.html		
3 Официальный tutorial по библиотеке seaborn. – Seaborn: https://seaborn.pydata.org/tutorial.html		
4 Официальная документация по библиотеке scikit-learn. – Scikit-learn: https://scikit-learn.org/0.20/documentation.html		
5 Официальная документация по Python - https://docs.python.org/3/ https://pythoner.name/documentation		

2. Дополнительная литература

Библиографическое описание издания (автор, наименование, вид, место и год издания, кол. стр.)	Кол. экз. в библ. НовГУ	Наличие в ЭБС
Печатные источники		
1 Кацко И.А. Практикум по анализу данных на компьютере : учебно-практическое пособие для вузов / Под ред. Г.В. Гореловой. - Москва : КолосС, 2009. - 276, [3] с. : ил. - (Учебники и учебные пособия для студентов высших учебных заведений). - Библиогр.: с. 273-274. - ISBN 978-5-9532-0624-2	2	
2 Сузи Р.А. Язык программирования Python. - 2-е изд. - Санкт-Петербург : БХВ-Петербург, 2006. - 648с. : ил. - (Основы информационных технологий). - ISBN 5-9556-0058-2. - ISBN 5-94774-442-2	1	
Электронные ресурсы		
1 Automate the Boring Stuff with Python [Электронный ресурс]. – Режим доступа: WWW.URL: https://automatetheboringstuff.com/		

Новгородский государственный
университет им. Ярослава Мудрого
Научная библиотека
Сектор учета

3. Информационное обеспечение

Наименование ресурса	Договор	Срок договора
Профессиональные базы данных		
База данных электронной библиотечной системы вуза «Электронный читальный зал-БиблиоТех» https://www.novsu.ru/dept/1114/bibliotech/	Договор № БТ-46/11 от 17.12.2014	бессрочный
Электронный каталог научной библиотеки http://mars.novsu.ac.ru/MarcWeb/	База собственной генерации	бессрочный
База данных «Аналитика» (картотека статей) http://mars.novsu.ac.ru/MarcWeb/	База собственной генерации	бессрочный
База данных Научной электронной библиотеки eLIBRARY.RU https://elibrary.ru/	в открытом доступе	-
Национальная подписка в рамках проекта Министерства образования и науки РФ (Госзадание № 4/2017 г.) к наукометрическим БД Scopus и Web of Science https://www.webofscience.com/wos/woscc/basic-search https://www.scopus.com/search/form.uri?display=basic#basic	регистрация (территория вуза)	2022
База данных электронно-библиотечной системы «Национальная электронная библиотека» https://нэб.пф	в открытом доступе	-
Информационные справочные системы		
Университетская информационная система «РОССИЯ» https://uisrussia.msu.ru	в открытом доступе	-
Национальный портал онлайн обучения «Открытое образование» https://openedu.ru	в открытом доступе	-
Официальный сайт Федерального агентства по техническому регулированию и метрологии http://protect.gost.ru/	в открытом доступе	-

Проверено НБ НовГУ

Зав. кафедрой ФТТМ Б.И.Селезнев
«26» 11 2020 г.



Приложение В

(обязательное)

Лист актуализации рабочей программы учебной дисциплины «Обработка и анализ данных»

Рабочая программа актуализирована на 20__/20__ учебный год.

Протокол № ____ заседания кафедры от « ____ » _____ 20__ г.

Разработчик: _____

Зав. кафедрой _____

Рабочая программа актуализирована на 20__/20__ учебный год.

Протокол № ____ заседания кафедры от « ____ » _____ 20__ г.

Разработчик: _____

Зав. кафедрой _____

Рабочая программа актуализирована на 20__/20__ учебный год.

Протокол № ____ заседания кафедры от « ____ » _____ 20__ г.

Разработчик: _____

Зав. кафедрой _____

Перечень изменений, внесенных в рабочую программу:

Номер изменения	№ и дата протокола заседания кафедры	Содержание изменений	Зав. кафедрой	Подпись
1	Протокол заседания КФТТМ № 10а от 08.06.2021 г.	В рабочую программу добавлена компетенция ОПК-5. Компетенция ОПК-4 дана в новой формулировке.	Селезнев Б.И.	

1 Содержание изменений:

1.1 Добавлена компетенция ОПК-5

Код и наименование общепрофессиональной компетенции	Код и наименование индикатора достижения общепрофессиональной компетенции
ОПК-5. Способен разрабатывать алгоритмы и компьютерные программы, пригодные для практического применения	ОПК-5.1 Знать методы алгоритмизации, языки и технологии программирования, пригодные для практического применения; ОПК-5.2 Уметь применять методы алгоритмизации, языки и технологии программирования при решении профессиональных задач; ОПК-5.3 Владеть навыками программирования, отладки и тестирования прототипов программного обеспечения

1.2 Изменена формулировка компетенции ОПК-4

Код и наименование общепрофессиональной компетенции	Код и наименование индикатора достижения общепрофессиональной компетенции
ОПК-4. Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности	ОПК-4.1 Знать принципы работы современных информационных технологий; ОПК-4.2 Уметь использовать современные информационные технологии для решения задач профессиональной деятельности; ОПК-4.5 Владеть навыками использования современных информационных технологий для решения задач профессиональной деятельности.