

Рабочая программа составлена на основании требований:

- Федерального государственного образовательного стандарта высшего образования – специалитет по специальности 10.05.03 Информационная безопасность автоматизированных систем, утвержденного приказом Минобрнауки России от 26.11.2020 №1457
- учебного плана, утвержденного ученым советом БГТУ им. В.Г. Шухова в 2021 году.

Составитель: к.т.н.  (Гвоздевский И.Н.)
(ученая степень и звание, подпись) (инициалы, фамилия)

Рабочая программа обсуждена на заседании кафедры

« 14 » 05 2021 г., протокол № 8

Заведующий кафедрой: к.т.н., доцент  (Поляков В.М.)
(ученая степень и звание, подпись) (инициалы, фамилия)


Рабочая программа согласована с выпускающей кафедрой программного обеспечения вычислительной техники и автоматизированных систем
(наименование кафедры/кафедр)

Заведующий кафедрой: к.т.н., доцент  (Поляков В.М.)
(ученая степень и звание, подпись) (инициалы, фамилия)

« 14 » 05 2021 г.

Рабочая программа одобрена методической комиссией института

« 20 » 05 2021 г., протокол № 9

Председатель к.т.н., доцент  (Семернин А.Н.)
(ученая степень и звание, подпись) (инициалы, фамилия)

1. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ

Категория (группа) компетенций	Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Наименование показателя оценивания результата обучения по дисциплине
Профессиональные компетенции	ПК-1 Способен разрабатывать системы защиты информации, содержащие элементы современных интеллектуальных технологий (проектный)	ПК-1.4 Использует и разрабатывает методы интеллектуального анализа данных	<p>Знания</p> <ul style="list-style-type: none"> - методов интеллектуального анализа данных; - современных средств интеллектуального анализа данных; - методов работы с распределенными хранилищами данных <p>Умения</p> <ul style="list-style-type: none"> - применять технологии распределенных вычислений для анализа больших данных. <p>Навыки</p> <ul style="list-style-type: none"> - использования системы Hadoop

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

1. Компетенция ПК-1 Способен разрабатывать системы защиты информации, содержащие элементы современных интеллектуальных технологий

Данная компетенция формируется следующими дисциплинами и практиками.

Стадия	Наименования дисциплины
1.	Программирование микроконтроллеров
2.	Основы искусственного интеллекта
3.	Теория принятия решений
4.	Интеллектуальный анализ больших данных
5.	Интеллектуальные системы информационной безопасности
6.	Анализ рисков информационной безопасности
7.	Системы и среды программирования
8.	Производственная преддипломная практика
9.	Подготовка к процедуре защиты и защита выпускной квалификационной работы

3. ОБЪЕМ ДИСЦИПЛИНЫ

Общая трудоемкость дисциплины составляет 4 зач. единицы, 144 часа.

Форма промежуточной аттестации __ экзамен

(экзамен, дифференцированный зачет, зачет)

Вид учебной работы	Всего часов	Семестр № 9
Общая трудоемкость дисциплины, час	144	144
Контактная работа (аудиторные занятия), в т.ч.:	73	73
лекции	34	34
лабораторные	17	17
практические	17	17
групповые консультации в период теоретического обучения и промежуточной аттестации	5	5
Самостоятельная работа студентов, включая индивидуальные и групповые консультации, в том числе:	71	71
Курсовой проект		
Курсовая работа		
Расчетно-графическое задание		
Индивидуальное домашнее задание		
Самостоятельная работа на подготовку к аудиторным занятиям (лекции, лабораторные занятия)	37	37
Экзамен (9 семестр)	34	34

4. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

4.1 Наименование тем, их содержание и объем Курс 5 Семестр 1

№ п/п	Наименование раздела (краткое содержание)	Объем на тематический раздел по видам учебной нагрузки, час			
		Лекции	Практические занятия	Лабораторные занятия	Самостоятельная работа на подготовку к аудиторным
1. Данные в открытых системах. Большие данные					
	Представления данных. Организация данных. Унификация форматов. Параметризация и трансформации данных. Поток данных. Организация хранения и доступа к данным.	10	4	-	10
2. Технологии распределенных вычислений. MapReduce					
	Трансформация данных в пары «ключ-значение». Сущность технологии MapReduce. Примеры использования: вычисление тренда, поточная кластеризация.	10	6	8	12
3. Распределенная файловая система Hadoop и интеллектуальные методы анализа данных в ней					
	Распределенная файловая система. Избыточное хранение. Системы управления вычислениями. Вычислительный кластер. Алгоритмы вычислений на кластере. Машинное обучение на кластере.	14	7	9	15
	ВСЕГО	34	17	17	37

4.2. Содержание практических (семинарских) занятий

№ п/п	Наименование раздела дисциплины	Тема практического занятия	К-во часов	Самостоятельная работа на подготовку к практическим занятиям
семестр № 9				
1	Данные в открытых системах. Большие данные	Представления больших данных. Преобразования в json	4	2
5	Технологии распределенных вычислений. MapReduce	Приведение к требуемой размерности	2	1
6		Организация работы по технологии MapReduce	2	1
7		Оценка эффективности вычислений	2	1
9	Распределенная файловая система	Архитектура вычислительного кластера	2	1
10	Hadoop и интеллектуальные методы анализа	Алгоритмизация вычислений трендов на потоке данных в кластере	3	2
11		Алгоритмизация поточной	2	1

	данных в ней	кластеризации и выявления аномалий		
			ИТОГО:	17
			ВСЕГО:	9
				26

4.3. Содержание лабораторных занятий

№ п/п	Наименование раздела дисциплины	Тема лабораторного занятия	К-во часов	Самостоятельная работа на подготовку к лабораторным занятиям
семестр № 9				
1	Технологии распределенных вычислений. MapReduce	Реализация mapper на узлах кластера	4	3
2		Реализация reducer на узлах кластера	4	3
5	Распределенная файловая система Hadoop и интеллектуальные методы анализа данных в ней	Установка и настройка Hadoop	2	1
6		Проведение вычислений в однонодовом режиме	2	1
7		Проведение вычислений в многонодовом режиме	5	3
			ИТОГО:	17
			ВСЕГО:	11
				28

4.4. Содержание курсового проекта/работы

Учебным планом не предусмотрено

4.5. Содержание расчетно-графического задания, индивидуальных домашних заданий

Учебным планом не предусмотрены

5. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕГО КОНТРОЛЯ, ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

5.1. Реализация компетенций

1 Компетенция ПК-1 Способен разрабатывать системы защиты информации, содержащие элементы современных интеллектуальных технологий

Наименование индикатора достижения компетенции	Используемые средства оценивания
ПК-1.4 Использует и разрабатывает методы интеллектуального анализа данных	Собеседование, защита лабораторных работ, экзамен

5.2. Типовые контрольные задания для промежуточной аттестации

5.2.1. Перечень контрольных вопросов (типовых заданий) для экзамена

№ п/п	Наименование раздела дисциплины	Содержание вопросов (типовых заданий)
1	Данные в открытых системах. Большие данные	<ol style="list-style-type: none">1. Как обычно выглядят данные на узлах?2. Каким образом можно организовать сбор табличных данных?3. Опишите способ журналирования определенных событий на узлах.4. Что такое большие данные?5. Как выполнить унификацию данных разных типов?6. Как настроить поток данных на сервер с узлов?7. Что такое технология master/slave? В чем отличие от технологии клиент/сервер?8. В чем основная трудность анализа больших данных?9. Что такое пакетный обмен данными и онлайн обмен данными?10. Как учитывается время на передачу данных при синхронизации потоков данных на сервере?
2	Технологии распределенных вычислений. MapReduce	<ol style="list-style-type: none">1. Что такое распределенные вычисления?2. На основе каких протоколов передачи данных можно построить вычислительный кластер?3. В чем основная идея технологии MapReduce?4. Как организовать выполнение на узлах стадии Map чтобы мастер получал однородную информацию со всех узлов?5. Как выполняется стадия Shuffle?6. Стадия Reduce: что в итоге получит мастер-узел?7. Алгоритм вычисления значений функциональных трендов (авторегрессии).8. Распределенный алгоритм кластеризации.9. Алгоритм распознавания многовекторной активности.10. Алгоритм распознавания DDoS атаки.
3	Распределенная файловая система Hadoop и интеллектуальные методы анализа данных в ней	<ol style="list-style-type: none">1. Какой язык является нативным для Hadoop? К каким последствиям это приводит?2. Опишите работу распределенной файловой системы Hadoop.3. В чем особенности установки Hadoop? Какие ее шаги зависят от конфигурации узла?4. Что такое yarn?5. Как запустить mapper и reducer в вычислительном кластере под управлением Hadoop, если они написаны на Python?6. В чем особенности алгоритма распределенной кластеризации при выполнении в многонодовом режиме Hadoop?7. Как реализовать алгоритм диагностики аномалий в кластере?8. Предложите способ пакетной обработки потоковых данных в кластере.

		9. Можно ли организовать работу искусственной нейронной сети в вычислительном кластере? Если да, то на основе какого алгоритма? 10. Как решить задачу классификации в кластере?
--	--	--

5.2.2. Перечень контрольных материалов для защиты курсового проекта/ курсовой работы

Учебным планом не предусмотрены

5.3. Типовые контрольные задания (материалы) для текущего контроля в семестре

Наименование раздела дисциплины	Типовые контрольные задания
Данные в открытых системах. Большие данные. ПК-1.4	Написать детектор аномальных состояний компьютерной системы. Написать DLP-систему.
Технологии распределенных вычислений. MapReduce. ПК-1.4	Разработать интеллектуальный honeypot с данными в форме персональных данных, а также детектор проникновений в него.
Распределенная файловая система Hadoop и интеллектуальные методы анализа данных в ней. ПК-1.4	Предложите способ быстрого распознавания DDoS атаки на узлы кластера.

5.4. Описание критериев оценивания компетенций и шкалы оценивания

Промежуточная аттестация предусмотрена в форме экзамена, используется следующая шкала оценивания: 2 – неудовлетворительно, 3 – удовлетворительно, 4 – хорошо, 5 – отлично.

Критериями оценивания достижений показателей являются:

Наименование показателя оценивания результата обучения по дисциплине	Критерий оценивания
Знания	Знание терминологии
	Знание основных принципов и методов работы с большими данными
	Знания о применимости методов анализа больших данных в сфере профессиональной деятельности
	Объем и полнота освоенного материала
Умения	Способность уместно применять терминологию
	Способность сопоставить метод решения задаче анализа больших данных
	Способность аргументировать предложение методов интеллектуального анализа больших данных
	Способность уверенно работать в распределенном кластере Hadoop
	Способность проводить распределенный анализ больших данных
Навыки	Установки и настройки Hadoop
	Подготовки данных на узлах кластера

Оценка преподавателем выставляется интегрально с учётом всех показателей и критериев оценивания.

Оценка сформированности компетенций по показателю **знания**.

Критерий	Уровень освоения и оценка			
	2	3	4	5
Знание терминологии	Не знает терминов и определений	Знает термины и определения, но допускает неточности формулировок	Знает термины и определения	Знает термины и определения, может корректно сформулировать их самостоятельно
Знание основных принципов и методов работы с большими данными	Не знает принципов и методов работы с большими данными	Имеет неполные знания о принципах и методах работы с большими данными	Проявляет достаточно полные знания о принципах и методах работы с большими данными	Знает принципы и методы работы с большими данными, может их свободно интерпретировать и применить
Знания о применимости методов анализа больших данных в сфере профессиональной деятельности	Не имеет представлений о применении методов анализа больших данных в области информационной безопасности	Знает примеры применения методов анализа больших данных в области информационной безопасности	Демонстрирует понимание применения методов анализа больших данных в области информационной безопасности	Показывает полное понимание работы методов анализа больших данных в профессиональной сфере, демонстрирует творческий подход
Объем и полнота освоенного материала	Не знает значительной части материала дисциплины	Знает только основной материал дисциплины, не усвоил его деталей	Знает материал дисциплины в достаточном объеме	Обладает твердым и полным знанием материала дисциплины, владеет дополнительными знаниями

Оценка сформированности компетенций по показателю **умения**.

Критерий	Уровень освоения и оценка			
	2	3	4	5
Способность уместно применять терминологию	Не применяет терминологию или применяет ее неуместно	Применяет терминологию, но допускает и бытовое описание	Применяет терминологию везде, где это необходимо	Уверенно применяет терминологию и умеет, при необходимости, интерпретировать ее
Способность сопоставить метод решения задаче анализа больших данных	Не может корректно подобрать метод для решения задачи	Предлагает наиболее уместные методы из известных ему, но знает только основные методы	Находит нужный метод для любой предложенной задачи	Выбирает несколько методов решения для предложенной задачи и может рассмотреть их сильные и слабые стороны
Способность аргументировать предложение	Не имеет своей точки зрения или не способен ее защитить	Может аргументировать предложение интеллектуально	Показывает способность выработать и обосновать	Имеет обоснованную и устойчивую точку зрения на

методов интеллектуальног о анализа больших данных		х методов анализа больших данных	применение метода интеллектуальног о анализа больших данных по проблеме, конструктивно обсуждать ее	соответствующий проблеме метод интеллектуального анализа больших данных
Способность уверенно ориентироваться в программных кодах на языке Python	Не знает языка Python	Применяет только основные конструкции и типы данных языка	Применяет все, упомянутые при изучении дисциплины конструкции и типы данных, уверенно владеет синтаксисом	Применяет современные паттерны программирования , осуществляет поиск новых решений при разработке программ
Способность провести распределенный анализ больших данных	Не способен проводить распределенный анализ данных	Способен проводить простую подготовку и распределенный анализ данных	Умеет работать с данными в распределенном кластере, подбирает интеллектуальные средства анализа	Умеет организовать свою работу аналитика больших данных в распределенном кластере, подготовить обоснованный аналитический отчет

Оценка сформированности компетенций по показателю *навыки*.

Критерий	Уровень освоения и оценка			
	2	3	4	5
Установка и настройка Hadoop	Не работает в Hadoop	Способен настроить и запустить обработку в кластере	Уверенно программирует приложения для работы в распределенном кластере	Уверенно программирует в том числе сложные распределенные алгоритмы
Подготовка данных на узлах кластера	Не может корректно подобрать метод подготовки данных	Предлагает наиболее уместные методы из известных ему, но знает только основные методы	Находит нужный метод для любой предложенной конфигурации данных	Выбирает несколько методов решения для предложенной конфигурации и задачи, может рассмотреть их сильные и слабые стороны

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ И УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ

6.1. Материально-техническое обеспечение

№	Наименование специальных помещений и помещений для самостоятельной работы	Оснащенность специальных помещений и помещений для самостоятельной работы
1	Аудитория 425 ГУК	Компьютерный класс с выходом в интернет
2	Аудитория 430 ГУК	Компьютерный класс с моделью автоматизированной системы

6.2. Лицензионное и свободно распространяемое программное обеспечение

№	Перечень лицензионного программного обеспечения.	Реквизиты подтверждающего документа
1	Операционная система Ubuntu 18	Свободно распространяемое ПО
2	Операционная система Linux Mint 20	Свободно распространяемое ПО

6.3. Перечень учебных изданий и учебно-методических материалов

1. Радченко И.А, Николаев И.Н. Технологии и инфраструктура Big Data. – СПб: Университет ИТМО, 2018. – 52 с.
2. Лесковец, Ю. Анализ больших наборов данных / Ю. Лесковец, А. Раджараман. — М.: ДМК, 2016. — 498 с.
3. Гитис Л.Х. Кластерный анализ в задачах классификации, оптимизации и прогнозирования. / Л.Х. Гитис. – М: Изд-во МГГУ, 2001. – 104 с.
4. Большие данные : учеб. пособие / И. Б. Тесленко [и др.] ; Владим. гос. ун-т им. А.Г. и Н.Г. Столетовых. – Владимир : Изд-во ВлГУ, 2021. – 123 с.

6.4. Перечень интернет ресурсов, профессиональных баз данных, информационно-справочных систем

1. Ресурс для аналитиков данных <https://kaggle.com/>
2. Система elibrary: <https://elibrary.ru/>
3. Система Google Scholar: <https://scholar.google.com/>
4. Научная библиотека БГТУ им. В.Г. Шухова: <http://ntb.bstu.ru/jirbis2/>

7. УТВЕРЖДЕНИЕ РАБОЧЕЙ ПРОГРАММЫ

Рабочая программа утверждена на 20____ /20____ учебный год
без изменений / с изменениями, дополнениями

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

Заведующий кафедрой _____
подпись, ФИО

Директор института _____
подпись, ФИО