

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ
федеральное государственное автономное образовательное учреждение высшего
образования
"САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
АЭРОКОСМИЧЕСКОГО ПРИБОРОСТРОЕНИЯ"

Кафедра № 14

УТВЕРЖДАЮ
Руководитель направления
доц., к.т.н., доц.
(должность, уч. степень, звание)
А.В. Шагомиров
(инициалы, фамилия)
(подпись)
«28» февраля 2022г

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

«Обработка и управление большими данными»
(Наименование дисциплины)

Код направления подготовки/ специальности	09.05.01
Наименование направления подготовки/ специальности	Применение и эксплуатация автоматизированных систем специального назначения
Наименование направленности	Автоматизированные системы обработки информации и управления
Форма обучения	очная

Лист согласования рабочей программы дисциплины

Программу составил (а)

д.т.н., проф.
(должность, уч. степень, звание)

(подпись, дата)

Ф.А. Таубин
(инициалы, фамилия)

Программа одобрена на заседании кафедры № 14

«28» февраля 2022г, протокол №8

Заведующий кафедрой № 14

к.т.н., доц.
(уч. степень, звание)

(подпись, дата)

В.Л. Оленев
(инициалы, фамилия)

Ответственный за ОП ВО 09.05.01(02)

доц., к.т.н., доц.
(должность, уч. степень,
звание)

(подпись, дата)

А.В. Шагомиров
(инициалы, фамилия)

Заместитель директора института №1 по методической работе

ст. преп.
(должность, уч. степень,
звание)

(подпись, дата)

В.Е. Таратун
(инициалы, фамилия)

Аннотация

Дисциплина «Обработка и управление большими данными» входит в вариативную часть образовательной программы подготовки обучающихся по специальности «09.05.01 «Применение и эксплуатация автоматизированных систем специального назначения» направленность «Автоматизированные системы обработки информации и управления». Дисциплина реализуется кафедрой №14.

Дисциплина нацелена на формирование у выпускника

общефессиональных компетенций:

ОПК-9 «способность использовать достижения отечественной и зарубежной науки, техники и технологии»;

профессиональных компетенций:

ПК-14 «способность осуществлять контроль качества аппаратного, программного и информационного обеспечений в автоматизированных системах специального назначения».

Содержание дисциплины охватывает круг вопросов, связанных с обработкой и управлением большими данными.

Преподавание дисциплины предусматривает следующие формы организации учебного процесса: лекции, лабораторные работы, самостоятельная работа обучающегося.

Программой дисциплины предусмотрены следующие виды контроля: текущий контроль успеваемости, промежуточная аттестация в форме экзамена.

Общая трудоемкость освоения дисциплины составляет 4 зачетных единицы, 144 часа.

Язык обучения по дисциплине «русский».

1. Перечень планируемых результатов обучения по дисциплине

1.1. Цели преподавания дисциплины

Цель – изучение и приобретение студентами магистерской программы подготовки навыков обработки и управления большими данными.

1.2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения ОП

В результате освоения дисциплины обучающийся должен обладать следующими компетенциями:

ОПК-9 «способность использовать достижения отечественной и зарубежной науки, техники и технологии»:

знать – основные методы решения задач обработки данных.

уметь – применять основные методы решения задач обработки данных.

владеть – основными методами решения задач обработки данных

ПК-14 «способность осуществлять контроль качества аппаратного, программного и информационного обеспечений в автоматизированных системах специального назначения»:

знать – способы сбора и обработки данных

уметь - выбирать инструментальные средства для обработки данных в соответствии с поставленной задачей.

владеть – навыками самостоятельного построения алгоритма и его анализа.

2. Место дисциплины в структуре ОП

Дисциплина базируется на знаниях, ранее приобретенных обучающимися при изучении следующих дисциплин:

– Базы данных

Знания, полученные при изучении материала данной дисциплины, имеют самостоятельное значение

3. Объем дисциплины в ЗЕ/академ. час

Данные об общем объеме дисциплины, трудоемкости отдельных видов учебной работы по дисциплине (и распределение этой трудоемкости по семестрам) представлены в таблице 1

Таблица 1 – Объем и трудоемкость дисциплины

Вид учебной работы	Всего	Трудоемкость по семестрам
		№9
1	2	3
Общая трудоемкость дисциплины, ЗЕ/(час)	4/ 144	4/ 144
<i>Аудиторные занятия, всего час., В том числе</i>	34	34
лекции (Л), (час)	17	17
лабораторные работы (ЛР), (час)	17	17

Экзамен, (час)	36	36
Самостоятельная работа , всего	74	74
Вид промежуточного контроля: зачет, дифф. зачет, экзамен (Зачет, Дифф. зач, Экз.)	Экз.	Экз.

4. Содержание дисциплины

4.1. Распределение трудоемкости дисциплины по разделам и видам занятий

Разделы и темы дисциплины и их трудоемкость приведены в таблице 2.

Таблица 2. – Разделы, темы дисциплины и их трудоемкость

Разделы, темы дисциплины	Лекции (час)	ПЗ (СЗ) (час)	ЛР (час)	КП (час)	СРС (час)
Семестр 9					
Раздел 1. Введение.	4		2		14
Раздел 2. Технологии обработки больших данных.	4		5		20
Раздел 3. Технологии и программные средства для обработки данных.	6		5		20
Раздел 4. Системы Data Mining.	3		5		20
Итого в семестре:	17		17		74
Итого:	17	0	17	0	74

4.2. Содержание разделов и тем лекционных занятий

Содержание разделов и тем лекционных занятий приведено в таблице 3.

Таблица 3 - Содержание разделов и тем лекционных занятий

Номер раздела	Название и содержание разделов и тем лекционных занятий
1	Раздел 1. Введение. Тема 1.1. Определение больших данных и базовая модель. Тема 1.2. Применение больших данных. Тема 1.3. Большие данные как феномен современных информационных технологий.
2	Раздел 2. Технологии обработки больших данных. Тема 2.1. Консолидация данных. Тема 2.2. Визуализация данных. Тема 2.3. Классификация и кластеризация данных. Тема 2.4. Регрессионный анализ данных. Тема 2.5. Анализ ассоциативных правил. Тема 2.6. Нейронные сети.
3	Раздел 3. Технологии и программные средства для обработки данных. Тема 3.1. Работа с NoSQL СУБД. Тема 3.2. Язык программирования R для анализа больших данных. Тема 3.3. Технология MapReduce. Тема 3.4. Apache Hadoop. Тема 3.5. Google Bigtable.

	Тема 3.6. Система потоковой обработки Storm.
4	Раздел 4. Системы Data Mining Тема 4.1. Понятие о технологии Data Mining. Тема 4.2. Реализация в пакетах прикладных программ. Тема 4.3. Сетевые технологии Data Mining.

4.3. Практические (семинарские) занятия

Темы практических занятий и их трудоемкость приведены в таблице 4.

Таблица 4 – Практические занятия и их трудоемкость

№ п/п	Темы практических занятий	Формы практических занятий	Трудоемкость, (час)	№ раздела дисциплины
Учебным планом не предусмотрено				

4.4. Лабораторные занятия

Темы лабораторных занятий и их трудоемкость приведены в таблице 5.

Таблица 5 – Лабораторные занятия и их трудоемкость

№ п/п	Наименование лабораторных работ	Трудоемкость, (час)	№ раздела дисциплины
Семестр 9			
1	Визуализация данных	2	2
2	Кластерный анализ	4	2
3	Проектирование реляционной базы данных (предметная область выбирается студентом самостоятельно).	4	3
4	Реализация спроектированной базы данных в СУБД	4	3
5	Реализация технологии Data Mining в пакетах прикладных программ.	3	4
Всего:		17	

4.5. Курсовое проектирование (работа)

Учебным планом не предусмотрено

4.6. Самостоятельная работа обучающихся

Виды самостоятельной работы и ее трудоемкость приведены в таблице 6.

Таблица 6 Виды самостоятельной работы и ее трудоемкость

Вид самостоятельной работы	Всего, час	Семестр 9, час
1	2	3
Самостоятельная работа, всего	74	74

изучение теоретического материала дисциплины (ТО)	70	70
Подготовка к текущему контролю (ТК)	4	4

5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)

Учебно-методические материалы для самостоятельной работы обучающихся указаны в п.п. 8-10.

6. Перечень основной и дополнительной литературы

6.1. Основная литература

Перечень основной литературы приведен в таблице 7.

Таблица 7 – Перечень основной литературы

Шифр	Библиографическая ссылка / URL адрес	Количество экземпляров в библиотеке (кроме электронных экземпляров)
	Г.С. Лбов. Анализ данных и знаний. Новосибирск, Новосибирский государственный университет, 2010.	

6.2. Дополнительная литература

Перечень дополнительной литературы приведен в таблице 8.

Таблица 8 – Перечень дополнительной литературы

Шифр	Библиографическая ссылка / URL адрес	Количество экземпляров в библиотеке (кроме электронных экземпляров)
	В. Майер-Шенбергер, К. Кукьер «Большие данные» Манн, Иванов и Фербер, 2013	
	Д. Силен, А. Мейсман, М. Али. Основы Data Science и Big Data. Python и наука о данных. Питер, 2017.	

7. Перечень ресурсов информационно-телекоммуникационной сети ИНТЕРНЕТ, необходимых для освоения дисциплины

Перечень ресурсов информационно-телекоммуникационной сети ИНТЕРНЕТ, необходимых для освоения дисциплины приведен в таблице 9.

Таблица 9 – Перечень ресурсов информационно-телекоммуникационной сети ИНТЕРНЕТ, необходимых для освоения дисциплины

URL адрес	Наименование
e.lanbook.com/books/	Электронно-библиотечная система
www.gpntb.ru/	Государственная публичная научно-техническая библиотека.
www.nlr.ru/	Российская национальная библиотека.
www.nns.ru/	Национальная электронная библиотека.
www.rsl.ru/	Российская государственная библиотека

www.microinform.ru/	Учебный центр компьютерных технологий «Микроинформ»
www.intuit.ru/	Образовательный сайт
www.window.edu.ru/	Библиотека учебной и методической литературы
www.osp.ru/	Журнал «Открытые системы»
www.ihtika.lib.ru/	Библиотека учебной и методической литературы

8. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине

8.1. Перечень программного обеспечения

Перечень используемого программного обеспечения представлен в таблице 10.

Таблица 10 – Перечень программного обеспечения

№ п/п	Наименование
	Не предусмотрено

8.2. Перечень информационно-справочных систем

Перечень используемых информационно-справочных систем представлен в таблице 11.

Таблица 11 – Перечень информационно-справочных систем

№ п/п	Наименование
	Не предусмотрено

9. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине

Состав материально-технической базы представлен в таблице 12.

Таблица 12 – Состав материально-технической базы

№ п/п	Наименование составной части материально-технической базы	Номер аудитории (при необходимости)
1	Мультимедийная лекционная аудитория	
2	Компьютерный класс	

10. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине

10.1. Состав фонда оценочных средств приведен в таблице 13

Таблица 13 - Состав фонда оценочных средств для промежуточной аттестации

Вид промежуточной аттестации	Примерный перечень оценочных средств
Экзамен	Список вопросов к экзамену

10.2. Перечень компетенций, относящихся к дисциплине, и этапы их формирования в процессе освоения образовательной программы приведены в таблице 14.

Таблица 14 – Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы

Номер семестра	Этапы формирования компетенций по дисциплинам/практикам в процессе освоения ОП
ОПК-9	«способность использовать достижения отечественной и зарубежной науки, техники

и технологии»	
6	ЭВМ и периферийные устройства
9	Обработка и управление большими данными
ПК-14 «способность осуществлять контроль качества аппаратного, программного и информационного обеспечений в автоматизированных системах специального назначения»	
4	Учебная (технологическая) практика
7	Спецификация и верификация систем и сетей
8	Разработка и стандартизация программных комплексов
8	Производственная (технологическая) практика
8	Надежность автоматизированных систем
9	Системы на кристалле
9	Обработка и управление большими данными

10.3. В качестве критериев оценки уровня сформированности (освоения) у обучающихся компетенций применяется шкала модульно–рейтинговой системы университета. В таблице 15 представлена 100–балльная и 4–балльная шкалы для оценки сформированности компетенций.

Таблица 15 –Критерии оценки уровня сформированности компетенций

Оценка компетенции		Характеристика сформированных компетенций
100-балльная шкала	4-балльная шкала	
$85 \leq K \leq 100$	«отлично» «зачтено»	- обучающийся глубоко и всесторонне усвоил программный материал; - уверенно, логично, последовательно и грамотно его излагает; - опираясь на знания основной и дополнительной литературы, тесно привязывает усвоенные научные положения с практической деятельностью направления; - умело обосновывает и аргументирует выдвигаемые им идеи; - делает выводы и обобщения; - свободно владеет системой специализированных понятий.
$70 \leq K \leq 84$	«хорошо» «зачтено»	- обучающийся твердо усвоил программный материал, грамотно и по существу излагает его, опираясь на знания основной литературы; - не допускает существенных неточностей; - увязывает усвоенные знания с практической деятельностью направления; - аргументирует научные положения; - делает выводы и обобщения; - владеет системой специализированных понятий.
$55 \leq K \leq 69$	«удовлетворительно» «зачтено»	- обучающийся усвоил только основной программный материал, по существу излагает его, опираясь на знания только основной литературы; - допускает несущественные ошибки и неточности; - испытывает затруднения в практическом применении знаний направления; - слабо аргументирует научные положения; - затрудняется в формулировании выводов и обобщений; - частично владеет системой специализированных понятий.
$K \leq 54$	«неудовлетворительно» «не зачтено»	- обучающийся не усвоил значительной части программного материала; - допускает существенные ошибки и неточности при рассмотрении проблем в конкретном направлении; - испытывает трудности в практическом применении знаний;

		- не может аргументировать научные положения; - не формулирует выводов и обобщений.
--	--	--

10.4. Типовые контрольные задания или иные материалы:

1. Вопросы (задачи) для экзамена (таблица 16)

Таблица 16 – Вопросы (задачи) для экзамена

№ п/п	Перечень вопросов (задач) для экзамена
1	Определение больших данных и базовая модель. Применение больших данных.
2	Консолидация данных.
3	Визуализация данных.
4	Классификация и кластеризация данных.
5	Кластерный анализ.
6	Нейронные сети.
7	Программные средства для обработки данных: NoSQL СУБД.
8	Язык программирования R для анализа больших данных.
9	Программные средства для обработки данных: Технология MapReduce..
10	Программные средства для обработки данных: Apache Hadoop.
11	Программные средства для обработки данных: система потоковой обработки Storm.
12	Программные средства для обработки данных: Google Bigtable.
13	Системы Data Mining: Понятие о технологии Data Mining.
14	Системы Data Mining: Реализация в пакетах прикладных программ.
15	Системы Data Mining: Сетевые технологии Data Mining.

2. Вопросы (задачи) для зачета / дифференцированного зачета (таблица 17)

Таблица 17 – Вопросы (задачи) для зачета / дифф. зачета

№ п/п	Перечень вопросов (задач) для зачета / дифференцированного зачета
	Учебным планом не предусмотрено

3. Темы и задание для выполнения курсовой работы / выполнения курсового проекта (таблица 18)

Таблица 18 – Примерный перечень тем для выполнения курсовой работы / выполнения курсового проекта

№ п/п	Примерный перечень тем для выполнения курсовой работы / выполнения курсового проекта
	Учебным планом не предусмотрено

4. Вопросы для проведения промежуточной аттестации при тестировании (таблица 19)

Таблица 19 – Примерный перечень вопросов для тестов

№ п/п	Примерный перечень вопросов для тестов
	Учебным планом не предусмотрено

5. Контрольные и практические задачи / задания по дисциплине (таблица 20)

Таблица 20 – Примерный перечень контрольных и практических задач / заданий

№ п/п	Примерный перечень контрольных и практических задач / заданий
	Учебным планом не предусмотрено

10.5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и / или опыта деятельности, характеризующих этапы формирования компетенций, содержатся в Положениях «О текущем контроле успеваемости и промежуточной аттестации студентов ГУАП, обучающихся по программам высшего образования» и «О модульно-рейтинговой системе оценки качества учебной работы студентов в ГУАП».

11. Методические указания для обучающихся по освоению дисциплины

Целью дисциплины является – получение студентами необходимых знаний, умений и навыков в области обработки и управления большими данными.

Методические указания для обучающихся по освоению лекционного материала

Основное назначение лекционного материала – логически стройное, системное, глубокое и ясное изложение учебного материала. Назначение современной лекции в рамках дисциплины не в том, чтобы получить всю информацию по теме, а в освоении фундаментальных проблем дисциплины, методов научного познания, новейших достижений научной мысли. В учебном процессе лекция выполняет методологическую, организационную и информационную функции. Лекция раскрывает понятийный аппарат конкретной области знания, её проблемы, дает цельное представление о дисциплине, показывает взаимосвязь с другими дисциплинами.

Планируемые результаты при освоении обучающимся лекционного материала:

- получение современных, целостных, взаимосвязанных знаний, уровень которых определяется целевой установкой к каждой конкретной теме;
- получение опыта творческой работы совместно с преподавателем;
- развитие профессионально-деловых качеств, любви к предмету и самостоятельного творческого мышления.
- появление необходимого интереса, необходимого для самостоятельной работы;
- получение знаний о современном уровне развития науки и техники и о прогнозе их развития на ближайшие годы;
- научиться методически обрабатывать материал (выделять главные мысли и положения, приходить к конкретным выводам, повторять их в различных формулировках);
- получение точного понимания всех необходимых терминов и понятий.

Лекционный материал может сопровождаться демонстрацией слайдов и использованием раздаточного материала при проведении коротких дискуссий об особенностях применения отдельных тематик по дисциплине.

Методические указания для обучающихся по прохождению лабораторных работ

В ходе выполнения лабораторных работ обучающийся должен углубить и закрепить знания, практические навыки, овладеть современной методикой и техникой эксперимента в

соответствии с квалификационной характеристикой обучающегося. Выполнение лабораторных работ состоит из экспериментально-практической, расчетно-аналитической частей и контрольных мероприятий.

Выполнение лабораторных работ обучающимся является неотъемлемой частью изучения дисциплины, определяемой учебным планом, и относится к средствам, обеспечивающим решение следующих основных задач у обучающегося:

- приобретение навыков исследования процессов, явлений и объектов, изучаемых в рамках данной дисциплины;
- закрепление, развитие и детализация теоретических знаний, полученных на лекциях;
- получение новой информации по изучаемой дисциплине;
- приобретение навыков самостоятельной работы с лабораторным оборудованием и приборами.

Структура и форма отчета о лабораторной работе

Титульный лист.

Содержание.

Задание. Номер варианта.

Постановка задачи.

Теоретические выкладки.

Расчеты.

Листинг программы(если требуется).

Выводы.

Список литературы.

Требования к оформлению отчета о лабораторной работе

Изложение текста и оформление работ следует выполнять в соответствии с требованиями ГОСТ 7.32 – 2001.

http://guap.ru/guap/standart/prav_main.shtml

Методические указания для обучающихся по прохождению самостоятельной работы

В ходе выполнения самостоятельной работы, обучающийся выполняет работу по заданию и при методическом руководстве преподавателя, но без его непосредственного участия.

В процессе выполнения самостоятельной работы, у обучающегося формируется целесообразное планирование рабочего времени, которое позволяет им развивать умения и навыки в усвоении и систематизации приобретаемых знаний, обеспечивает высокий уровень успеваемости в период обучения, помогает получить навыки повышения профессионального уровня.

Методическими материалами, направляющими самостоятельную работу обучающихся являются учебно-методический материал по дисциплине.

Методические указания для обучающихся по прохождению промежуточной аттестации

Промежуточная аттестация обучающихся предусматривает оценивание промежуточных и окончательных результатов обучения по дисциплине. Она включает в себя:

– экзамен – форма оценки знаний, полученных обучающимся в процессе изучения всей дисциплины или ее части, навыков самостоятельной работы, способности применять их для решения практических задач. Экзамен, как правило, проводится в период экзаменационной сессии и завершается аттестационной оценкой «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».

Система оценок при проведении промежуточной аттестации осуществляется в соответствии с требованиями Положений «О текущем контроле успеваемости и промежуточной аттестации студентов ГУАП, обучающихся по программам высшего образования» и «О модульно-рейтинговой системе оценки качества учебной работы студентов в ГУАП».

Лист внесения изменений в рабочую программу дисциплины

Дата внесения изменений и дополнений. Подпись внесшего изменения	Содержание изменений и дополнений	Дата и № протокола заседания кафедры	Подпись зав. кафедрой