

Министерство науки и высшего образования Российской Федерации

федеральное государственное бюджетное образовательное учреждение
высшего образования
РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ ГИДРОМЕТЕОРОЛОГИЧЕСКИЙ
УНИВЕРСИТЕТ

Кафедра Прикладной информатики

Рабочая программа дисциплины

Обработка и анализ геоданных

Основная профессиональная образовательная программа
высшего образования по направлению подготовки

09.04.03 Прикладная информатика

Направленность (профиль):

Прикладные геоинформационные системы управления

Уровень:

Магистратура

Форма обучения

Очная

Согласовано
Руководитель ОПОП

 **Истомин Е.П.**

Утверждаю

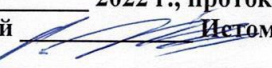
Проректор по УР  **Н.О. Верещагина**

Рекомендована решением
Ученого совета института Информационных
систем и геотехнологий

28 09 2022 г., протокол № 10

Рассмотрена и утверждена на заседании кафедры

28 06 2022 г., протокол № 06

Зав. кафедрой  **Истомин Е.П.**

Авторы-разработчики:

 **Колбина О.Н.**

Санкт-Петербург 2022

Рассмотрено и рекомендовано к использованию в учебном процессе на ____/____
учебный год без изменений*

Протокол заседания кафедры _____ от __. __.20 №__

Рассмотрено и рекомендовано к использованию в учебном процессе на
_____/____ учебный год с изменениями (см. лист изменений)**

Протокол заседания кафедры _____ от __. __.20 №__

*Заполняется при ежегодном пересмотре программы, если в неё не внесены изменения

** Заполняется при ежегодном пересмотре программы, если в неё внесены изменения

1. Цель и задачи освоения дисциплины

Цель освоения дисциплины - формирование у студентов необходимой теоретической базы и практических навыков, которые позволят всесторонне и системно понимать современные проблемы прикладной математики и информатики, проблемы обработки и анализа информации, а также разрабатывать и анализировать концептуальные и теоретические модели при решении научных и прикладных задач в области информационных технологий.

Главная задача курса - сформировать целостное представление о современных проблемах анализа и обработки больших данных, помочь овладеть опытом разработки и анализа концептуальных и теоретических моделей прикладных задач анализа больших данных с применением моделей Data Mining..

Задачи:

- сформировать целостное представление о современных проблемах анализа и обработки геоданных;
- помочь овладеть опытом разработки и анализа концептуальных и теоретических моделей решения прикладных задач анализа геоданных.

2. Место дисциплины в структуре основной профессиональной образовательной программы

Дисциплина относится к обязательной части блока дисциплин 1. Изучается во 2 и 3 семестре, параллельно с такими дисциплинами как:

Геоинформационные технологии;

Геоинформационное управление ресурсами/ Геоинформационное сопровождение оценки территорий;

Научно-исследовательская практика;

Технология беспроводных сетей/Облачные вычисления.

В последствие изучаются такие дисциплины как:

Научно-исследовательская практика;

Интеллектуализация геоинформационных систем;

Разработка и сопровождение требований к геоинформационным системам;

Преддипломная практика;

Языки современных бизнес-приложений.

3. Перечень планируемых результатов обучения

Процесс изучения дисциплины направлен на формирование компетенции ПК-5.

Таблица 1.

Профессиональные компетенции

Код и наименование общепрофессиональной компетенции	Код и наименование индикатора достижения общепрофессиональной компетенции	Результаты обучения
ПК-5. Способен обеспечивать разработку баз данных	ПК-5.1 Обеспечивает соответствия баз данных ИС и процесса их разработки принятым в организации или проекте стандартам и технологиям ПК-5.2 Использует инструменты и методы проектирования структур	<i>Знать:</i> основы методологии проектирования баз данных; NoSQL базы данных, принципы работы с большими данными; методы обработки геоданных.

	баз данных ПК-5.3 Применяет современный отечественный и зарубежный опыт в профессиональной деятельности	<i>Уметь:</i> проектировать базы геоданных; применять в работе NoSQL; работать с большими данными. <i>Владеть:</i> приемами работы с большими данными; навыками работы и организации баз геоданных.
--	--	--

4. Структура и содержание дисциплины

4.1. Объем дисциплины

Объем дисциплины составляет 6 зачетные единицы, 216 академических часа.

Таблица 2.

Объем дисциплины по видам учебных занятий в академических часах

Объём дисциплины	Всего часов
	Очная форма обучения
Объем дисциплины	216
Контактная работа обучающихся с преподавателем (по видам аудиторных учебных занятий) – всего:	
в том числе:	-
лекции	28
занятия семинарского типа:	
практические занятия	
лабораторные занятия	42
Самостоятельная работа (далее – СРС) – всего:	146
в том числе:	-
курсовая работа	
контрольная работа	
Вид промежуточной аттестации	зачет/экзамен

4.2. Структура дисциплины

Таблица 3.

Структура дисциплины для очной формы обучения

№	Раздел дисциплины	С е м е с	Виды учебной работы, в т.ч. самостоятельная работа студентов, час.	Формы текущего контроля успеваемости	Формируемые компетенции	Индикаторы достижения компетенций
---	-------------------	-----------------------	--	--------------------------------------	-------------------------	-----------------------------------

		т р	Ле кц ии	Лаб ора тор ные	СР С			
1	Хранение и первичная обработка данных	1	14	14	80	Опрос Отчеты по лабораторным работам	ПК-5	ПК-5.1., ПК-5.2., ПК-5.3.
2	Базы геоданных как составная часть ГИС	2	14	28	66	Опрос Отчеты по лабораторным работам	ПК-5	ПК-5.1., ПК-5.2., ПК-5.3.
	ИТОГО		28	42	146	-	-	-

4.3. Содержание разделов дисциплины

Таблица 4.

Содержание разделов дисциплины

№ раздела	Наименование раздела дисциплины	Содержание
1	Хранение и первичная обработка данных	<ol style="list-style-type: none"> 1. Виды и источники данных. Загрузка и разделение данных. Объединение данных из разных источников. Очистка данных и заполнение пропусков. Контроль диапазонов. Пространственные данные. 2. Первичная обработка данных. Сглаживание и нормировка данных. Преобразование данных. Визуализация данных. Формы представления количественных и качественных данных. Когнитивная визуализация данных. 3. Хранение и доступ к данным. Виды баз данных. Реляционные СУБД. 4. NoSQL базы данных. Большие данные.
2	Базы геоданных как составная часть ГИС	<ol style="list-style-type: none"> 1. Реализации модели данных: файловые системы, СУБД. 2. Ключевые концепции базы данных. 3. Персональные и многопользовательские базы геоданных. 4. Пространственные индексы, типы пространственных SQL-запросов. 5. Проектирование геобД. 6. Структурные элементы геобД. 7. Таблицы, домены, отношения, топологии, геометрические сети, др. 8. Топология: отношения между пространственными объектами, кластерный допуск, ранги слоев, топологические правила, ошибки и исключения.

4.4. Содержание занятий семинарского типа

Таблица 5.

Содержание лабораторных занятий для очной формы обучения

№ темы дисциплины	Тематика лабораторных занятий	Всего часов
1	Работа с источниками данных, их преобразование и визуализация.	10
1,2	NoSQL базы данных	4
2	Проектирование геобД	28

5. Перечень учебно-методического обеспечения самостоятельной работы обучающихся по дисциплине

Методические рекомендации по организации самостоятельной работы обучающихся.

6. Оценочные средства для текущего контроля успеваемости и промежуточной аттестации по итогам освоения дисциплины

Учет успеваемости обучающегося по дисциплине осуществляется по 100-балльной шкале. Максимальное количество баллов по дисциплине за один семестр – 100:

- максимальное количество баллов за выполнение всех видов текущего контроля -60;
- максимальное количество баллов за посещение лекционных занятий - 10;
- максимальное количество баллов за прохождение промежуточной аттестации – 30.

6.1. Текущий контроль

Типовые задания, методика выполнения и критерии оценивания текущего контроля по разделам дисциплины представлены в Фонде оценочных средств по данной дисциплине.

6.2. Промежуточная аттестация

Форма промежуточной аттестации по дисциплине – **зачет/экзамен.**

Форма проведения зачета/экзамена: *устно по билетам.*

Перечень вопросов для подготовки к зачету:

ПК-5

1. Пирамида знаний.
2. Задача анализа данных (Постановка, история возникновения, онтология методов)
3. Методика анализа данных.
4. Подготовка исходного набора данных.
5. Моделирование неопределённости. Вероятностные рассуждения.
6. Условные Байесовские сети.
7. Интеллектуальный анализ данных (определение и классификация задач).
8. Knowledge Discovery in Databases (KDD).
9. Задача классификации (постановка и методы решения).
10. Задача поиска ассоциативных правил (постановка и методы решения).
11. Задача секвенциального анализа (постановка и методы решения).
12. Задача кластеризации (постановка и методы решения).
13. Оперативная аналитическая обработка (OLAP) и многомерная модель данных.
14. Экспертные системы (определение, назначение, модели представления знаний, достоинства).

15. Визуальный анализ данных.
16. Задача анализа временных рядов (определение временного ряда, классификация рядов динамики, разложение по компонентам).
17. Аномальные уровни временных рядов причины их возникновения и методы их устранения.
18. Задача дисперсионного анализа (постановка и методы решения).
19. Задача корреляционного анализа (постановка и методы решения).
20. Алгоритм Naïve Bayes.
21. Алгоритм C4.5.
22. Алгоритм Apriori.
23. Алгоритм C-means.
24. Основные статистические показатели временных рядов.
25. Метод Ирвина.
26. Метод проверки разности средних уровней.
27. Метод Фостера – Стюарта.
28. Методы сглаживания временных рядов.
29. Метод Четверикова.
30. Основные статистические показатели временных рядов.

Перечень вопросов для подготовки к экзамену:

ПК-5

1. Перечислить топологические правила, которым должны следовать линейные объекты слоя?
2. Перечислить топологические правила, которым должны следовать границ водоемов?
3. Перечислить топологические правила, которым должны следовать линейные объекты слоя ?
4. Перечислить топологические правила, которым должны следовать полигональные объекты слоя?
5. Перечислите топологические правила, которым должны следовать линейные объекты слоя?
6. Определить пространственное разрешение (размер пиксела в надир) для съемочной системы Landsat с высотой 705 км и мгновенным углом зрения прибора 0.00244 градусов?

6.3. Балльно-рейтинговая система оценивания

Таблица 6.

Распределение баллов по видам учебной работы

Вид учебной работы, за которую ставятся баллы	Баллы
Сдача лабораторных работ	0-40
Опрос	0-30
Промежуточная аттестация	0-30
ИТОГО	0-100

Таблица 7.

Распределение дополнительных баллов

Дополнительные баллы (баллы, которые могут быть добавлены до 100)	Баллы
Участие в НИРС	0-8
Участие в Олимпиаде	0-5
Активность на учебных занятиях	0-2

ИТОГО	0-15
--------------	-------------

Таблица 8.

Балльная шкала итоговой оценки на зачете

Оценка	Баллы
Зачтено	60-100
Незачтено	0-59

Таблица 9.

Балльная шкала итоговой оценки на экзамене

Оценка	Баллы
Отлично	85-100
Хорошо	65-84
Удовлетворительно	40-64
Неудовлетворительно	0-39

7. Методические указания для обучающихся по освоению дисциплины

Методические рекомендации ко всем видам аудиторных занятий, а также методические рекомендации по организации самостоятельной работы, в том числе по подготовке к текущему контролю и промежуточной аттестации представлены в Методических рекомендациях для обучающихся по освоению дисциплины «Обработка и анализ геоданных».

8. Учебно-методическое и информационное обеспечение дисциплины

8.1. Перечень основной и дополнительной учебной литературы

Основная литература

- 1) Флах, П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных [Электронный ресурс] — Электрон. дан. — Москва : ДМК Пресс, 2015. — 400 с. — Режим доступа: <https://e.lanbook.com/book/69955> .
- 2) Миркин, Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — М. : Издательство Юрайт, 2018. — 174 с. — (Серия : Авторский учебник). — ISBN 978-5-9916-5009-0. — Режим доступа : www.biblio-online.ru/book/46A41F93-BC46-401C-A30E-27C0FB60B9DE.

Дополнительная литература

- 1) Большаков, В. Д. Теория ошибок наблюдений [Текст] : учебник / Большаков В.Д. - 2-е изд., перераб. доп. - Москва : Недра, 1983. - 222 с.
- 2) Горбунов, М. Е. Обработка и контроль качества данных эксперимента COSMIS по радиозатменному зондированию атмосферы Земли [Текст] / М. Е. Горбунов // Исследование Земли из космоса. - 2009. - № 4. - С. 36-46.

8.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"

1. Сайт специалистов в области ГИС и ДЗЗ <http://gis-lab.info/>
2. База данных систем координат European Petroleum Survey Group (EPSG) - <http://www.epsg.org>.
3. База с описанием различных систем координат и проекций - <http://spatialreference.org/>

4. Сайт открытой ГИС QGIS - <http://www.qgis.org/>
5. Сайт ГИС ArcGIS - <http://www.dataplus.ru/>
6. Сайт ESTI MAP - представитель производителя ГИС MapInfo - <http://www.mapinfo.ru/>
7. Сайт PostGIS - <http://postgis.refractory.net/>
8. Сайт ИТЦ СканЭкс - <http://www.scanex.com>
9. Сайт геологической службы США - <http://www.usgs.gov/>

8.3. Перечень программного обеспечения

1. Открытая версия Q-ГИС
2. MongoDB – демонстрационная версия
3. OpenOffice

8.4. Перечень информационных справочных систем

Не используется

8.5. Перечень профессиональных баз данных

Электронно-библиотечная система eLibrary

9. Материально-техническое обеспечение дисциплины

Материально-техническое обеспечение программы соответствует действующим санитарно-техническим и противопожарным правилам и нормам и обеспечивает проведение всех видов аудиторных занятий и самостоятельной работы студентов.

Учебная аудитория для проведения занятий лекционного типа – укомплектована специализированной (учебной) мебелью, набором демонстрационного оборудования.

Учебная лаборатория прикладных информационных технологий – укомплектована специализированной (учебной) мебелью, компьютерами, служащими для работы с информацией.

Учебная аудитория для групповых и индивидуальных консультаций - укомплектована специализированной (учебной) мебелью, техническими средствами обучения, служащими для представления учебной информации.

Учебная аудитория для текущего контроля и промежуточной аттестации - укомплектована специализированной (учебной) мебелью, техническими средствами обучения, служащими для представления учебной информации.

Помещение для самостоятельной работы – укомплектовано специализированной (учебной) мебелью, оснащено компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечено доступом в электронную информационно-образовательную среду организации.

Помещение для хранения и профилактического обслуживания учебного оборудования.

10. Особенности освоения дисциплины для инвалидов и лиц с ограниченными возможностями здоровья

Обучение обучающихся с ограниченными возможностями здоровья при необходимости осуществляется на основе адаптированной рабочей программы с использованием специальных методов обучения и дидактических материалов, составленных с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья таких обучающихся (обучающегося).

При определении формы проведения занятий с обучающимся-инвалидом учитываются рекомендации, содержащиеся в индивидуальной программе реабилитации инвалида, относительно рекомендованных условий и видов труда.

При необходимости для обучающихся из числа инвалидов и лиц с ограниченными

возможностями здоровья создаются специальные рабочие места с учетом нарушенных функций и ограничений жизнедеятельности.

11. Возможность применения электронного обучения и дистанционных образовательных технологий

Дисциплина может реализовываться с применением электронного обучения и дистанционных образовательных технологий.