

Федеральное государственное бюджетное образовательное учреждение высшего
профессионального образования
Московский государственный университет имени М.В. Ломоносова
Геологический факультет

УТВЕРЖДАЮ

Декан Геологического факультета
академик

_____ /Д.Ю.Пушаровский/
«__» _____ 20 г.

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

Введение в анализ пространственных данных

Авторы-составители: Н.В. Никифоров, Е.С. Хазиева

Уровень высшего образования:

Магистратура (ММ)

Направление подготовки:

05.04.01 «Геология»

Магистерская программа:

Гидрогеология, инженерная геология, геоэкология

Форма обучения:

Очная

Рабочая программа рассмотрена и одобрена
Учебно-методическим Советом Геологического факультета
(протокол № ___ от _____)

Москва

Рабочая программа дисциплины (модуля) разработана в соответствии с самостоятельно установленным МГУ образовательным стандартом (ОС МГУ) для реализуемых основных профессиональных образовательных программ высшего образования по направлению подготовки «Геология» (*программы бакалавриата, магистратуры, реализуемых последовательно по схеме интегрированной подготовки*) в редакции приказа МГУ от 30 декабря 2016 г.

Год (годы) приема на обучение – 2018.

© Геологический факультет МГУ имени М.В. Ломоносова

Программа не может быть использована другими подразделениями университета и другими вузами без разрешения факультета.

Цели и задачи дисциплины

Цель – формирование компетенций в области современных методов анализа данных дистанционного спутникового зондирования Земли и их применение в различных сферах наук о Земле.

Задачи: освоение базового инструментария современного анализа данных, включая интеллектуальные методы обработки пространственной информации.

1. Место дисциплины в структуре ОПОП ВО – вариативная часть, профессиональный цикл, профессиональная дисциплина по выбору, курс – I, семестр – 2.

2. Входные требования для освоения дисциплины, предварительные условия: освоение дисциплин математического цикла в бакалавриате.

3. Результаты обучения по дисциплине, соотнесенные с требуемыми компетенциями выпускников:

ОПК-4. Способность использовать современные вычислительные методы и компьютерные технологии для решения задач профессиональной деятельности (формируется частично).

ПК-1. Способность самостоятельно проводить научные исследования с помощью современного оборудования, информационных технологий, с использованием новейшего отечественного и зарубежного опыта (формируется частично).

Планируемые результаты обучения по дисциплине:

Знать: перечень современных подходов и источников для получения релевантных пространственных данных; основы анализа различных типов пространственных данных (векторных, растровых), включая базовые методы машинного обучения.

Уметь: пользоваться современными инструментами сбора и анализа пространственных данных.

Владеть: навыками разработки высокоуровневого программного кода Java Script для облачных ГИС.

4. Формат обучения – лекционные и семинарские занятия

5. Объем дисциплины составляет 2 з.е., 72 академических часа, в том числе 26 часов, отведенных на контактную работу обучающихся с преподавателем (13 часов – занятия лекционного типа, 13 часов – занятия семинарского типа), 46 часов отведено на самостоятельную работу (в том числе 10 часов на мероприятия промежуточной аттестации). Форма промежуточной аттестации – зачет.

6. Содержание дисциплины, структурированное по темам (разделам) с указанием отведенного на них количества академических часов и виды учебных занятий

Краткое содержание дисциплины (аннотация):

В настоящее время активно развивается направление использования открытых платформ для работы с пространственными данными. Широкое распространение получают облачные решения, позволяющие оперировать Большими Данными (BigData).

В курсе рассматриваются методы извлечения и анализа релевантной информации из каталога пространственных данных облачной ГИС Google Earth Engine (более 8 петабайт). Анализ пространственных данных используется для автоматизации решения прикладных задач мониторинга в области наук о Земле.

Наименование и краткое содержание разделов и тем дисциплины Форма промежуточной аттестации по дисциплине	Всего (часы)	В том числе			Самостоятельная работа обучающегося, часы
		Контактная работа (работа во взаимодействии с преподавателем) Виды контактной работы, часы			
		Занятия лекционного типа	Занятия семинарского типа	Всего	
Раздел 1. Географические данные, геоинформационные системы, дистанционное зондирование Земли.		4		4	
Раздел 2. Google Earth Engine и Java Script для автоматизации решения прикладных задач анализа данных и мониторинга в области наук о Земле		9	13	22	Разработка собственного алгоритма классификации на основе методов машинного обучения и данных дистанционного зондирования, 36 часов
Промежуточная аттестация <u>зачет</u>					10
Итого	72	26			46

Содержание лекционных занятий:

Географические данные и геоинформационные системы
 Основы анализа данных дистанционного зондирования Земли (ДЗЗ)
 Основы интеллектуального анализа данных.
 Обзор облачной платформы Google Earth Engine
 Методы обработки спутниковых данных в Google Earth Engine
 Методы обработки тематических данных в Google Earth Engine
 Методы обработки векторных данных и визуализация статистической информации
 Решение задач автоматической классификации с помощью методов машинного обучения в Google Earth Engine

Содержание семинаров:

Источники данных, типы и форматы пространственных данных;

- Модели представления данных (растровая, векторная);
- Особенности хранения пространственных данных и организация доступа к ним;
- Примеры прикладного использования геопространственных решений в лесном и сельском хозяйстве, строительстве, нефтегазовом секторе, горнорудной промышленности и мониторинге чрезвычайных ситуаций;
- Типы данных дистанционного зондирования;
- Обзор данных мультиспектральной и радарной спутниковой съемки;
- Характеристики данных: пространственное и спектральное разрешение, структура и способы представления;

- Основные методы обработки данных дистанционного зондирования Земли;
- Искусственный интеллект как раздел компьютерных наук;
- Виды машинного обучения (supervised & unsupervised learning);
- Постановка задачи машинного обучения;
- Математическое обоснование алгоритмов машинного обучения на примере Random Forest;
- Прикладные примеры решения задач на основе интеллектуального анализа данных;
- Основные понятия и возможности платформы;
- Базовый синтаксис JavaScript;
- Обзор библиотек и встроенных инструментов для обработки данных
- Примеры алгоритмов и сервисов на основе Google Earth Engine;
- Поиск и отображение данных Landsat – 8 и Sentinel – 2;
- Обзор метаданных спутниковых снимков, фильтрация снимков по метаданным;
- Применение спектральных комбинаций каналов и расчет вегетационных индексов;
- Поиск и отображение тематической пространственной информации (цифровые модели рельефа, данные о населении, климатические показатели);
- Статистическая обработка тематических данных по заданной области интереса;
- Экспорт полученных результатов во внешние хранилища;
- Загрузка векторных (.shp; .kml) и табличных данных;
- Формирование запросов к векторным и табличным данным;
- Посторонние статистических графиков и диаграмм на основе комбинирования типов данных;
- Принцип работы алгоритма Random Forest;
- Принципы создания обучающей выборки;
- Процесс обучения и классификации;
- Назначение индивидуального проекта по анализу данных участникам курса;
- Визуальная оценка качества классификации;
- Разбор факторов влияния параметров обучения и характеристик исходных данных на качество обучения;
- Проверка качества модели на кросс-валидации, критерий R^2 . Сравнение моделей;

Рекомендуемые образовательные технологии:

В ходе курса используется облачная платформа для анализа пространственных данных, позволяющая удаленно подключаться к рабочему проекту. В качестве заключительной проверки знаний используется электронная система тестирования.

7. Фонд оценочных средств (ФОС) для оценивания результатов обучения по дисциплине

7.1. Типовые контрольные задания или иные материалы для проведения текущего контроля успеваемости:

Текущий контроль усвоения дисциплины осуществляется при сдаче каждым студентом заключений по заданиям.

Домашние задания для самостоятельной подготовки студентов

1. Выполнить поиск и отображение космических снимков Landsat-7, Landsat -8, Sentinel-1, Sentinel-2;
2. Выполнить спектральный анализ космических снимков Landsat-7, Landsat -8, Sentinel-1, Sentinel-2;
3. Выполнить пространственный анализ данных цифровой модели рельефа и распределения плотности населения;
4. Выполнить классификацию термокарстовых озер с использованием методов машинного обучения.

Типовые упражнения и расчетные задания

1. Создать переменную в виде постоянного числа и отобразите его в консоли Google Earth Engine;
2. Отобразить последовательно числа от 1 до 100 в панели консоли Google Earth Engine;
3. Выполнить поиск по дате космических снимков Landsat- 8;
4. Выполнить фильтрацию данных по облачности;
5. Отобразить данные в окне карты со следующими параметрами отображения: 5, 4, 3 (ИК, синий, зеленый).

7.2. Типовые контрольные задания или иные материалы для проведения промежуточной аттестации:

Примерный перечень вопросов при промежуточной аттестации:

1. Дайте определение понятию дистанционное зондирование Земли;
2. Дайте определение понятию воздушное лазерное сканирование;
3. Дайте определение понятию информация;
4. Дайте определение понятию пространственные данные;
5. Перечислите типы систем координат;
6. Какой тип проекции чаще всего применяется для отображения на карте России с целью минимизации искажений?
7. Какой тип проекции чаще всего применяется для отображения на карте Антарктиды с целью минимизации искажений?
8. Какой тип проекции чаще всего применяется для отображения на карте всей поверхности Земли с целью минимизации искажений?
9. Охарактеризуйте два основных типа пространственных данных;
10. Сколько спектральных каналов имеет панхроматическое изображение;
11. Сколько спектральных каналов имеет мультиспектральное изображение в видимом диапазоне;
12. Перечислите спектральные каналы, которые используются для расчета вегетационного индекса NDVI;
13. Перечислите спектральные каналы, которые необходимо использовать для отображения снимка в натуральных цветах;
14. Перечислите спектральные каналы, которые чаще всего используются для анализа растительности;
15. Укажите какие типы данных представлены в каталоге Google Earth Engine;
16. Какие виды пользовательских данных можно загрузить дополнительно в каталог Google Earth Engine?
17. Каким образом можно скачать данные из каталога Google Earth Engine?
18. Опишите основной функционал доступный в разделах Inspector и Code Editor;
19. Опишите основной функционал доступный в разделе Console;
20. Опишите основной функционал доступный в разделе Tasks.

Шкала и критерии оценивания результатов обучения по дисциплине

Результаты обучения	Незачет	Зачет
Знания: перечень современных подходов и источников для получения релевантных пространственных данных; основы анализа различных типов пространственных	Фрагментарные знания или отсутствие знаний	Сформированные систематические знания или общие, но не структурированные знания

данных (векторных, растровых), включая базовые методы машинного обучения		
Умения: находить, загружать, обрабатывать релевантные данные дистанционного зондирования Земли для аналитического решения прикладных и научных задач в области наук о Земле, связанных с мониторингом поверхности Земли.	Не систематическое умение или отсутствие умений	Успешное и систематическое умение или в целом успешное, но содержащее отдельные пробелы умение (допускает неточности непринципиального характера)
Владение: навыками разработки высокоуровневого программного кода Java Script для облачных ГИС	Наличие отдельных навыков или отсутствие навыков	Владение навыками, применяемыми при решении задач или в целом сформированные навыки (владения), но используемые не в активной форме

8. Ресурсное обеспечение:

А) Перечень основной и дополнительной литературы

- основная литература:

Google Earth Engine: API Tutorials. URL: <https://developers.google.com/earth-engine/tutorials>

Google Earth Engine: Get Started with Earth Engine. URL: <https://developers.google.com/earth-engine/getstarted>

Электронный репозиторий материалов курса URL: <https://github.com/keyonix>

Б) Перечень лицензионного программного обеспечения:

лицензионное программное обеспечение не требуется

В) Материально-техническое обеспечение:

персональные компьютеры, доступ к сети интернет

9. Язык преподавания – русский.

10. Преподаватели – Н.В. Никифоров, Е.С. Хазиева

11. Авторы программы – Н.В. Никифоров, Е.С. Хазиева