МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение
высшего образования
МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ГЕОДЕЗИИ И КАРТОГРАФИИ
(МИИГАиК)

Факультет геоинформатики и информационной безопасности
Кафедра информационно-измерительных систем

**ДОКУМЕНТ ПОДПИСАН
ЭЛЕКТРОННОЙ ПОДПИСЬЮ**

Сертификат: 5BFEE1D286E9C95A0C0D2C611E434BD5
Владелец: Степанченко Алексей Леонидович
Действителен: с 23.09.2026 по 17.12.2027

«УТВЕРЖДАЮ»
Проректор
А. Л. Степанченко
18 апреля 2025 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ) «Б1.В.01 МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ В ЗАДАЧАХ ОБРАБОТКИ ПРОСТРАНСТВЕННЫХ ДАННЫХ»

Уровень высшего образования: магистратура

Направление подготовки: 09.04.03 Прикладная информатика

Направленность (профиль) подготовки: Анализ больших пространственных данных

Квалификация (степень) выпускника: магистр

Форма обучения: очная

Год набора (приема на обучение): 2025

Срок получения образования: 2 года

Объем: в зачетных единицах: 4 з.е.
в академических часах: 144 ак.ч.

Москва, 2025

Разработчики:

Доцент кафедры Информационно-измерительных систем, к.т.н.
Гвоздев О. Г.

Преподаватель кафедры Информационно-измерительных систем
Жаров С. И.

Преподаватель кафедры Информационно-измерительных систем
Ширинкин С. Г.

Рабочая программа дисциплины (модуля) составлена в соответствии с требованиями ФГОС ВО по направлению подготовки 09.04.03 Прикладная информатика, утвержденного приказом Минобрнауки от 19.09.2017 № 916

Согласование и утверждение

№	Подразделение	Ответственное лицо	ФИО	Виза	Дата, протокол (при наличии)
1	факультет Геоинформатики и информационной безопасности	Председатель ученого совета факультета	Матерухин А. В.	Рассмотрено	14.10.2024, № 10-24

Содержание

1. Цель и задачи освоения дисциплины (модуля)

2. Планируемые результаты обучения по дисциплине (модулю), соотнесенные с планируемыми результатами освоения образовательной программы

3. Место дисциплины в структуре ОП

4. Объем дисциплины и виды учебной работы

5. Содержание дисциплины
   5.1. Разделы, темы дисциплины и виды занятий
   5.2. Содержание разделов, тем дисциплины и формы текущего контроля

6. Оценочные материалы текущего контроля

7. Оценочные материалы промежуточной аттестации

8. Порядок проведения промежуточной аттестации

9. Материально-техническое и учебно-методическое обеспечение дисциплины
   9.1. Перечень основной и дополнительной учебной литературы
   9.2. Профессиональные базы данных и ресурсы «Интернет», к которым обеспечивается доступ обучающихся
   9.3. Программное обеспечение и информационно-справочные системы, используемые при осуществлении образовательного процесса по дисциплине
   9.4. Оборудование и технические средства обучения

10. Методические указания по освоению дисциплины (модуля)

1. Цель и задачи освоения дисциплины (модуля)

Цель освоения дисциплины - формирование у обучающихся профессиональных компетенций в области применения алгоритмов машинного обучения для анализа и обработки геопространственных данных, включая разработку моделей для задач классификации, сегментации и прогнозирования в GIS-системах, дистанционного зондирования и урбанистики.

Задачи изучения дисциплины:

- изучить базовые методы машинного обучения (классификация, регрессия, кластеризация) с адаптацией к пространственным данным.;
- освоить инструменты предобработки и визуализации геоданных (векторы, растры, координатные системы).;
- научить разрабатывать и оценивать ML-модели для реальных задач, таких как детекция объектов на спутниковых снимках или анализ трафика..

2. Планируемые результаты обучения по дисциплине (модулю), соотнесенные с планируемыми результатами освоения образовательной программы

Компетенции, индикаторы и результаты обучения

ПК-1 Способен выполнять работы по сбору, обработке и анализу больших пространственных данных

ПК-1.2 Обрабатывает большие пространственные данные

Знать:

ПК-1.2/Зн1 Основы и физические принципы регистрации аэрокосмической информации о местности

Уметь:

ПК-1.2/Ум1 Осуществлять выбор аэрокосмических снимков, получаемых различными системами и комплексами дистанционного зондирования, в зависимости от решаемых практических или исследовательских задач

Владеть:

ПК-1.2/Нв1 Методами сбора и оценки качества пространственной информации, имеющей значения для подготовки и реализации проектов в соответствующих сферах профессиональной деятельности

3. Место дисциплины в структуре ОП

Дисциплина (модуль) Б1.В.01 «Методы машинного обучения в задачах обработки пространственных данных» относится к формируемой участниками образовательных отношений части образовательной программы и изучается в семестре(ах): 2.

В процессе изучения дисциплины студент готовится к решению типов задач профессиональной деятельности, предусмотренных ФГОС ВО и образовательной программой.

4. Объем дисциплины (модуля) и виды учебной работы

Период	доемкость сы)	доемкость ЗЕТ)	ые занятия сы)	ие занятия сы)	ьная работа сы)	ая аттестация сы)

обучения	Общая тру (ча	Общая тру (ЗЕ	Лекционни (ча	Практическ (ча	Самостоятел (ча	Промежуточн (ча
2 семестр	144	4	17	34	57	Экзамен (36)
Всего	144	4	17	34	57	36

5. Содержание дисциплины (модуля)

5.1. Разделы, темы дисциплины и виды занятий

(часы промежуточной аттестации не указываются)

Наименование раздела, темы	Всего	Лекционные занятия	Практические занятия	Самостоятельная работа	Планируемые результаты обучения, соотнесенные с результатами освоения программы
Раздел 1. Основы пространственных данных и машинного обучения	**35**	**5**	**11**	**19**	ПК-1.2
Тема 1.1. Структура геопространственных данных	17	3	5	9	
Тема 1.2. Предобработка данных	18	2	6	10	
Раздел 2. Классические методы ML для пространственных задач	**37**	**6**	**11**	**20**	ПК-1.2
Тема 2.1. Классификация и регрессия	19	4	5	10	
Тема 2.2. Кластеризация и снижение размерности	18	2	6	10	
Раздел 3. Продвинутые модели и приложения	**36**	**6**	**12**	**18**	ПК-1.2
Тема 3.1. CNN (U-Net). GNN. PyTorch Geometric	19	4	7	8	
Тема 3.2. Spatial CV. Change detection. Wildfire prediction	17	2	5	10	
Итого	**108**	**17**	**34**	**57**	

5.2. Содержание разделов, тем дисциплин и формы текущего контроля

Раздел 1. Основы пространственных данных и машинного обучения

Тема 1.1. Структура геопространственных данных

Векторные и растровые форматы. CRS. Библиотеки GDAL. GeoPandas.

Текущий контроль

Вид (форма) контроля
Тестирование

Тема 1.2. Предобработка данных

Нормализация. Обработка пропусков. Пространственная интерполяция.

Текущий контроль

Вид (форма) контроля
Тестирование

Раздел 2. Классические методы ML для пространственных задач

Тема 2.1. Классификация и регрессия

Гиперпараметры. Метрики: accuracy, F1. Пример на Landsat.

Текущий контроль

Вид (форма) контроля
Тестирование

Тема 2.2. Кластеризация и снижение размерности

Плотностная кластеризация. Визуализация. Автокорреляция.

Текущий контроль

Вид (форма) контроля
Тестирование

Раздел 3. Продвинутые модели и приложения

Тема 3.1. CNN (U-Net). GNN. PyTorch Geometric

Сегментация растров. Графы дорог. Обучение на GPU.

Текущий контроль

Вид (форма) контроля
Тестирование

Тема 3.2. Spatial CV. Change detection. Wildfire prediction

Блочная валидация. Sentinel-2. MODIS данные.

Текущий контроль

Вид (форма) контроля
Тестирование

6. Порядок проведения промежуточной аттестации

Порядок проведения промежуточной аттестации регламентируется локальными нормативными актами Университета:
- Положением о формах, периодичности, порядке текущего контроля успеваемости и промежуточной аттестации обучающихся по основным образовательным программам высшего образования – программам бакалавриата, программам специалитета, программам магистратуры в федеральном государственном бюджетном образовательном учреждении высшего образования «Московский государственный университет геодезии и картографии», п.6 «Организация и проведение промежуточной аттестации».
- Положением о порядке осуществления образовательного процесса, проведения текущего контроля успеваемости и промежуточной аттестации обучающихся с применением электронного обучения и дистанционных образовательных технологий в федеральном государственном бюджетном образовательном учреждении высшего образования «Московский государственный университет геодезии и картографии», п.8 «Проведение промежуточной аттестации с применением ЭО и ДОТ».

Промежуточная аттестация: Экзамен, 2 семестр.
Экзамен оценивается по пятибалльной шкале:
- зачтено с оценкой «отлично» выставляется, если обучающийся имеет глубокие, исчерпывающие знания в объеме освоенной программы; знание основной (обязательной) литературы; дает правильные и уверенные ответы, свидетельствующие о наличии твердых знаний и навыков в использовании технических средств; полное, четкое, грамотное и логически стройное изложение материала; демонстрирует свободное применение теоретических знаний при анализе практических вопросов;
- зачтено с оценкой «хорошо» выставляется обучающемуся в случае соблюдения тех же требований, но наличии в ответе обучающегося по некоторым перечисленным показателям недостатков принципиального характера, что вызывает замечания или поправки преподавателя;
- зачтено с оценкой «удовлетворительно» выставляется обучающемуся в случае соблюдения тех же требований, но в случае наличия в ответе ошибок, необходимости существенной помощи в виде поправок и наводящих вопросов преподавателя;
- не зачтено («не удовлетворительно») – обучающийся не смог ответить на 2/3 вопросов; не справился с практическим заданием или выполнено менее 50 % практических заданий.

7. Оценочные материалы текущего контроля

Раздел 1. Основы пространственных данных и машинного обучения
Контролируемые ИДК: ПК-1.2

Тема 1.1. Структура геопространственных данных
Форма контроля/оценочное средство: Тестирование
Вопросы/Задания:
1. Машинное обучение — это:
1) система хранения данных
2) метод решения ОДУ
3) область науки, изучающая алгоритмы обучения по данным
4) способ кодирования сигналов
2. Наличие размеченной целевой переменной характерно для:
1) обучения без учителя
2) обучения с подкреплением
3) обучения с учителем
4) кластеризации
3. Задачей классификации пространственных данных является:
1) интерполяция высот DEM

2) определение класса землепокрытия по снимку
3) слияние тайлов
4) spatial join

4. Кригинг высот относится к задаче:

1) кластеризации
2) классификации
3) регрессии
4) отбора признаков

5. Какая метрика подходит для бинарной классификации?

1) Silhouette Score
2) F1-score
3) RMSE
4) ARI

Тема 1.2. Предобработка данных

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Переобучение проявляется как:

1) высокая ошибка на обучении
2) низкая ошибка на обучении и высокая на тесте
3) отсутствие выбросов
4) равенство bias и variance

2. Что помогает снизить переобучение?

1) увеличение случайных признаков
2) L2-регуляризация
3) уменьшение тестовой выборки
4) добавление шума в метки

3. Обучающая выборка служит для:

1) итоговой проверки
2) визуализации
3) настройки параметров модели
4) хранения метаданных

4. Тестовая выборка используется для:

1) оптимизации гиперпараметров
2) оценки обобщающей способности
3) нормализации признаков
4) балансировки классов

5. Для сегментации объектов на спутниковых изображениях часто применяют:

1) k-NN
2) DBSCAN
3) Random Forest
4) U-Net

Раздел 2. Классические методы ML для пространственных задач

Контролируемые ИДК: ПК-1.2

Тема 2.1. Классификация и регрессия

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. В k-NN класс объекта определяется:

1) градиентным спуском
2) по большинству классов ближайших объектов

3) центроидами кластеров
4) глубиной дерева

2. Дерево решений строится путём:

1) свёртки признаков
2) разбиений по информативности
3) усреднения моделей
4) снижения размерности

3. При росте сложности модели:

1) смещение растет, разброс падает
2) смещение падает, разброс растет
3) обе величины растут
4) обе падают

4. Глубина дерева — это:

1) метрика качества
2) параметр ошибки
3) гиперпараметр
4) функция потерь

5. Какая метрика измеряет перекрытие масок при сегментации?

1) Accuracy
2) Precision
3) Intersection over Union (IoU)
4) Recall

Тема 2.2. Кластеризация и снижение размерности

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. IoU предпочтительнее Accuracy при сегментации, потому что:

1) игнорирует ложноположительные
2) оценивает истинное перекрытие областей
3) зависит от классов
4) всегда выше

2. Линейная регрессия минимизирует:

1) энтропию
2) сумму квадратов ошибок
3) Silhouette
4) log-loss

3. Пространственное разрешение снимка — это:

1) число спектральных каналов
2) динамический диапазон
3) площадь поверхности, приходящаяся на пиксель
4) точность GPS

4. Нормализация признаков нужна для:

1) увеличения размерности
2) приведения признаков к одному масштабу
3) удаления классов
4) уменьшения bias

5. Нормализация приводит признак к:

1) диапазону [0;1]
2) лог-шкале
3) нулевому среднему и единичной дисперсии
4) бинарному виду

Раздел 3. Продвинутые модели и приложения

Контролируемые ИДК: ПК-1.2

Тема 3.1. CNN (U-Net). GNN. PyTorch Geometric

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Обучение модели — это:

1) визуализация данных
2) настройка параметров по обучающей выборке
3) сжатие датасета
4) генерация карт

2. Бэггинг отличается от бустинга тем, что:

1) использует одну модель
2) обучает модели параллельно и усредняет их
3) минимизирует энтропию
4) фокусируется на сложных примерах

3. Ансамблевым методом на деревьях является:

1) K-Means
2) Random Forest
3) LASSO
4) t-SNE

4. U-Net — это:

1) дерево решений
2) свёрточная нейронная сеть для сегментации
3) набор центроидов
4) граф кластеров

5. Одна эпоха обучения нейросети — это:

1) шаг оптимизатора
2) полный проход по обучающим данным
3) обновление λ
4) вычисление IoU

Тема 3.2. Spatial CV. Change detection. Wildfire prediction

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Ошибка обобщения измеряется на:

1) обучающей выборке
2) данных, не участвовавших в обучении
3) наборе гиперпараметров
4) ансамбле

2. L2-регуляризация при мультиколлинеарности:

1) увеличивает variance
2) снижает дисперсию коэффициентов
3) удаляет признаки
4) меняет классы

3. Пространственная автокорреляция означает, что:

1) данные несбалансированны
2) близко расположенные объекты имеют схожие значения
3) классы несвязаны
4) признаки независимы

4. Случайный лес — это:

1) одно глубокое дерево
2) ансамбль случайно построенных деревьев
3) k-средних
4) U-Net

5. В градиентном бустинге каждая новая модель аппроксимирует:

1) данные заново
2) метки классов
3) веса деревьев
4) градиент функции потерь текущего ансамбля

8. Оценочные материалы промежуточной аттестации

2 семестр, Экзамен
Контролируемые ИДК: ПК-1.2
Вопросы/Задания:

1. Дайте определение машинного обучения.

2. В чём отличие обучения с учителем от обучения без учителя?

3. Приведите пример задачи классификации на пространственных данных.

4. Приведите пример задачи регрессии на пространственных данных.

5. Назовите три метрики качества для бинарной классификации.

6. Объясните, что такое переобучение.

7. Укажите два способа профилактики переобучения.

8. Что такое обучающая выборка?

9. Для чего используется тестовая выборка?

10. Назовите четыре алгоритма, часто применяемые для классификации или сегментации спутниковых изображений.

11. Объясните принцип метода k-ближайших соседей.

12. Как строится дерево решений?

13. Как увеличение сложности модели влияет на bias и variance?

14. Что называют гиперпараметрами модели?

15. Объясните различие метрик Accuracy и Intersection over Union (IoU).

16. Почему IoU предпочтительнее Accuracy при оценке сегментации снимков?

17. Сформулируйте идею линейной регрессии.

18. Что такое пространственное разрешение снимка?

19. Зачем нормализовать спектральные каналы спутникового снимка?

20. Что называют нормализацией признаков?

21. Чем отличается процесс обучения от процесса предсказания?

22. В чём разница между бэггингом и бустингом?

23. Приведите пример ансамблевого метода на деревьях решений.

24. Что такое U-Net и зачем он применяется при обработке снимков?

25. Что называют эпохой обучения нейронной сети?

26. Чем отличается ошибка обучения от ошибки обобщения?

27. Как L2-регуляризация помогает при мультиколлинеарности признаков?

28. Что такое пространственная автокорреляция и почему её учитывают в ML?

29. Что обозначает термин «размерность данных»?

30. Какую величину аппроксимирует каждая новая модель в градиентном бустинге?

9. Материально-техническое и учебно-методическое обеспечение дисциплины

9.1. Перечень основной и дополнительной учебной литературы

Основная литература

1. Митяков Е. С. Искусственный интеллект и машинное обучение: учебное пособие для вузов / Е. С. Митяков, А. Г. Шмелева, А. И. Ладынин. - Санкт-Петербург: Лань, 2025. - 252 с. - 978-5-507-51465-6. - Текст: электронный // Издательство Лань: [сайт]. - URL: https://e.lanbook.com/book/450827 - Режим доступа: по подписке

2. Ануфриев С. О. Технологии сбора пространственных данных аэрокосмическими методами: учебное пособие для вузов / С. О. Ануфриев, О. С. Ануфриев. - Санкт-Петербург: Лань, 2024. - 144 с. - 978-5-507-49736-2. - Текст: электронный // Издательство Лань: [сайт]. - URL: https://e.lanbook.com/book/427799 - Режим доступа: по подписке

3. Баланов А. Н. Машинное обучение и искусственный интеллект: учебное пособие для вузов / А. Н. Баланов. - Санкт-Петербург: Лань, 2024. - 172 с. - 978-5-507-49194-0. - Текст: электронный // Издательство Лань: [сайт]. - URL: https://e.lanbook.com/book/414920 - Режим доступа: по подписке

Дополнительная литература

1. Зольников И. Д. Введение в геоинформационные системы и дистанционное зондирование: учебно-методическое пособие / И. Д. Зольников, Н. В. Глушкова. - Введение в геоинформационные системы и дистанционное зондирование - Новосибирск: Новосибирский государственный университет, 2023. - 88 с. - 978-5-4437-1498-1. - Текст: электронный // IPR SMART: [сайт]. - URL: https://www.iprbookshop.ru/134567.html - Режим доступа: по подписке

2. Красиков И. И. Геоинформационные системы в лесном хозяйстве: учебное пособие / И. И. Красиков. - Красноярск: СибГУ им. академика М. Ф. Решетнёва, 2018. - 86 с. - Текст: электронный // Издательство Лань: [сайт]. - URL: https://e.lanbook.com/book/147522 - Режим доступа: по подписке

3. Обработка данных магниторазведки: обратные задачи и машинное обучение: учебное пособие / Д. С. Киселев, А. М. Гриф, М. Г. Персова, Ю. Г. Соловейчик. - Новосибирск: НГТУ, 2021. - 76 с. - 978-5-7782-4556-3. - Текст: электронный // Издательство Лань: [сайт]. - URL: https://e.lanbook.com/book/306473 - Режим доступа: по подписке

9.2. Профессиональные базы данных и ресурсы «Интернет», к которым обеспечивается доступ обучающихся

Профессиональные базы данных

Не используются.

Ресурсы «Интернет»

1. https://www.elibrary.ru/ - НЭБ eLIBRARY
2. http://citforum.ru - IT-портал CITForum
3. https://znanium.ru/ - Электронно-библиотечная система "Znanium"
4. https://e.lanbook.com/ - Электронная библиотечная система "Лань"

9.3. Программное обеспечение и информационно-справочные системы, используемые при осуществлении образовательного процесса по дисциплине

Перечень программного обеспечения
(обновление производится по мере появления новых версий программы)

Не используется.

Перечень информационно-справочных систем
(обновление выполняется еженедельно)

Не используется.

9.4. Оборудование и технические средства обучения

Компьютерный класс

Маркерная доска - 0 шт.
Моноблок - 64 шт.
Проектор - 4 шт.

Смешанный тип

Автоматизированное рабочее место (АРМ) - 1 шт.
Доска настенная одноэлементная - 2 шт.
Интерактивная панель - 1 шт.
Проектор - 1 шт.

10. Методические указания по освоению дисциплины (модуля)

С целью успешного освоения данного предмета студенту необходимо выполнять домашние работы по указанным преподавателям темам и представленным на занятиях материалам. Для более успешного выполнения работы студенту предлагается использование учебно-методических материалов, соответствующих указанной теме.

Самостоятельная работа студентов описывается и регулируется методическими

рекомендациями по дисциплине.

Описание возможностей изучения дисциплины лицами с ОВЗ и инвалидами

При необходимости обучения студентов-инвалидов и лиц с ограниченными возможностями здоровья аудиторные занятия могут быть заменены или дополнены изучением полнотекстовых лекций, презентаций, видео- и аудиоматериалов в электронной информационно-образовательной среде (ЭИОС) университета. Индивидуальные задания подбираются в адаптированных к ограничениям здоровья формах (письменно или устно, в форме презентаций). Выбор методов обучения зависит от их доступности для инвалидов и лиц с ограниченными возможностями здоровья.

Рабочая программа дисциплины для инвалидов и лиц с ограниченными возможностями здоровья разрабатывается индивидуально с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья таких обучающихся.