

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Горбунев Алексей Александрович

Должность: Заместитель начальника университета по учебной работе

Дата подписания: 06.06.2025 14:28:29

Уникальный программный ключ:

286e49ee1471d400cc1f45539d51ed7bbf0e9cc7

ФГБОУ ВО «Санкт-Петербургский университет ГПС МЧС России»

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
АНАЛИЗ ДАННЫХ, МОНИТОРИНГ В ЭКОЛОГИИ И
ПРИРОДОПОЛЬЗОВАНИИ**

**Бакалавриат по направлению подготовки
05.03.06 «Экология и природопользование»**

**Направленность (профиль)
«Экологическая безопасность и экологический мониторинг»**

Санкт-Петербург

1. Цели и задачи дисциплины

Цели освоения дисциплины:

получение компетенций для решения задач в профессиональной деятельности в области экологии, природопользования и охраны природы с использованием информационных технологий.

Перечень компетенций, формируемых в процессе изучения дисциплины

Компетенции	Содержание
ОПК-5	Способен понимать принципы работы информационных технологий и решать стандартные задачи профессиональной деятельности в области экологии, природопользования и охраны природы с использованием информационно - коммуникационных, в том числе геоинформационных технологий

Задачи дисциплины:

изучение методов работы с данными, их области применения, достоинств и недостатков;

овладение основными методами статистического анализа данных и методов математического моделирования применительно к сфере экологии и природопользования;

формирование культуры представления, описания, интерпретации и оценки выводов над данными.

2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

Индикаторы достижения компетенции	Планируемые результаты обучения по дисциплине
ОПК – 5.1 Знает принципы работы с информацией, в том числе с пространственной для решения задач в области профессиональной деятельности	Знает: Особенности решения профессиональных задач на основе применения информационных технологий, методов вычислительной техники
ОПК – 5.2 Умеет осуществлять поиск, хранение, обработку и анализ информации из различных источников и представлять её в требуемом формате	Владеет: Основными современными средствами реализации информационных технологий для анализа данных

3. Место дисциплины в структуре ОПОП

Дисциплина относится к обязательной части основной профессиональной образовательной программы бакалавриата по направлению подготовки 05.03.06 Экология и природопользование направленность (профиль) «Экологическая безопасность и экологический мониторинг»

4. Структура и содержание дисциплины

Общая трудоемкость дисциплины составляет 5 зачетных единиц, 180 часов.

4.1 Распределение трудоемкости дисциплины по видам работ для очной формы обучения

Вид учебной работы	Трудоемкость		
	з.е.	час.	по семестрам
			4
Общая трудоемкость дисциплины по учебному плану	5	180	180
Контактная работа		82	82
Лекции		36	36
Практические занятия		44	44
Лабораторные работы			
Консультации перед экзаменом		2	2
Самостоятельная работа		62	62
Курсовая работа			
Зачёт			
Зачёт с оценкой			
Экзамен		36	36

4.2 Тематический план, структурированный по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий для очной формы обучения

№ п/п	Номер и наименование тем	Всего часов	Количество часов по видам занятий, в том числе практическая подготовка*			Консультация	Контроль	Самостоятельная работа
			Лекции	Практические занятия	Лабораторные работы			
№ семестра 4								
1	Тема 1. Введение в анализ данных	26	4	10				12
2	Тема 2. Корреляционный и регрессионный анализ	32	8	10				14
3	Тема 3. Теория временных рядов	42	12	12				18
4	Тема 4. Интеллектуальный анализ данных	42	12	12				18
Консультация		2				2		
Экзамен		36					36	
Итого		180	36	44		2	36	62

4.3. Содержание дисциплины для очной формы обучения

Тема 1. Введение в анализ данных

Лекции. Основные понятия и определения. Специфика работы с информацией. Свойства информации. Информация и данные. Виды и методы анализа данных. Этапы решения задачи анализа данных и их взаимосвязи.

Практические занятия. Загрузка и разделение данных. Объединение данных из разных источников. Очистка данных и заполнение пропусков. Контроль диапазонов. Инструменты для обработки данных. Визуализация данных. Формы представления количественных и качественных данных. Когнитивная визуализация данных. Первичная обработка данных.

Самостоятельная работа. Преобразование данных. Обработка временных рядов. NoSQL базы данных. Большие данные. Хранение и доступ к данным. Виды баз данных. Реляционные СУБД.

Рекомендуемая литература:

основная [3];

дополнительная [1].

Тема 2. Корреляционный и регрессионный анализ

Лекции. Корреляция, коэффициент корреляции, индекс корреляции, оценка значимости коэффициента корреляции, индексов корреляции, доверительные интервалы, проверка гипотез, типы ошибок, области принятия и непринятия.

Модель парной регрессии, модель множественной регрессии. Нахождение уравнения линейной регрессии. Система нормальных уравнений. Свойства оценок по методу наименьших квадратов. Уравнения нелинейной регрессии.

Практические занятия. Выборочный коэффициент корреляции. Ранговая корреляция. Применение корреляционно-регрессионного анализа в области экологии, природопользования и охраны природы. Разработка многофакторных регрессионных моделей.

Самостоятельная работа. Статистические критерии: виды, особенности, области применения. Приложение теории статистических решений в практико-ориентированных задачах. Приложение корреляционно-регрессионного анализа в практико-ориентированных задачах. Нелинейные функции регрессии. Приложение методов математической статистики в практико-ориентированных задачах.

Рекомендуемая литература:

основная [3];

дополнительная [1].

Тема 3. Теория временных рядов

Лекции. Виды временных рядов. Их построение. Составляющие элементы временного ряда. Показатели временного ряда и методы их исчисления. Предварительная обработка временного ряда. Методы распознавания типа тренда и оценки его параметров.

Аддитивная и мультипликативная модели временного ряда. Тренд, сезонность. Адаптивные модели: экспоненциальное сглаживание, модели Хольта- Уинтерса и Тейла-Вейджа.

Полиномиальные модели и многократное сглаживание. Доверительная вероятность и точность прогноза. Сезонные модели.

Практические занятия. Сглаживание временного ряда. Разработка и исследование моделей с постоянными параметрами адаптации. Модель авторегрессии первого и второго порядка. Модель скользящего среднего первого и второго порядка. Модель авторегрессии-скользящего среднего.

Самостоятельная работа. Методы распознавания типа колебаний и оценки параметров колеблемости. Методы изучения и измерения устойчивости уровней ряда и тренда. Изучение динамики комплекса взаимосвязанных признаков. схемой с дрейфующими коэффициентами. Модели с адаптивными параметрами адаптации. Адаптивные комбинированные модели. Моделирование взаимосвязанных временных рядов.

Рекомендуемая литература.

основная [1, 2]

дополнительная [1]

Тема 4. Интеллектуальный анализ данных

Лекции. Анализ многомерных данных: OLAP технологии. Многомерная модель данных. Понятия OLAP-куба, операции, выполняемые над гиперкубом. Архитектура OLAP систем.

Понятие интеллектуального анализа данных (DM). Требования, предъявляемые к новым знаниям. Задачи Data Mining. Стадии Data Mining. Модели Data Mining: классификация, кластеризация, поиск ассоциативных правил, регрессионные модели. Сферы применения технологии интеллектуального анализа данных.

Постановка задачи классификации с обучением. Логистическая регрессия. Понятие о деревьях решений. Метрики качества классификации (точность/специфичность, ROC-кривая, площадь под кривой).

Практические занятия. Решения задач классификации. Бинарная и многоклассовая классификация. Вероятностная постановка задачи классификации. Метрики качества классификации (точность/специфичность, ROC-кривая, площадь под кривой).

Кластерный анализ. Общие сведения. Метод К-средних. Кластерная обработка данных. Базовые метрики качества, используемые в задачах кластеризации.

Самостоятельная работа. Основные понятия теории машинного обучения: проблемы, решаемые методами машинного обучения, модели машинного обучения (геометрические, вероятностные, логические), признаки. Постановка задач обучения по прецедентам. Объекты и признаки. Типы шкал: бинарные, номинальные, порядковые, количественные. Концептуальное обучение: пространство гипотез, поиск в пространстве гипотез, обучаемость, оценка качества решения задачи.

Рекомендуемая литература:

основная [2, 4];

дополнительная [2].

5. Методические рекомендации по организации изучения дисциплины

При реализации программы дисциплины «Анализ данных в экологии и природопользовании» используются лекционные и практические занятия.

Общими целями занятий являются:

- обобщение, систематизация, углубление, закрепление теоретических знаний по конкретным темам дисциплины;
- формирование умений применять полученные знания на практике, реализация единства интеллектуальной и практической деятельности;
- выработка при решении поставленных задач профессионально значимых качеств: самостоятельности, ответственности, точности, творческой инициативы.

Целями лекции являются:

- систематизированные основы научных знаний по дисциплине, раскрывать состояние и перспективы развития соответствующей области науки и техники;
- концентрировать внимание обучающихся на наиболее сложных и узловых вопросах;
- стимулировать их активную познавательную деятельность и способствовать формированию творческого мышления.

В ходе практического занятия обеспечивается процесс активного взаимодействия обучающихся с преподавателем; приобретаются практические навыки и умения. Цели практического занятия: выработка практических умений и приобретение навыков, закрепление пройденного материала по соответствующей теме дисциплины.

Самостоятельная работа обучающихся направлена на углубление и закрепление знаний, полученных на лекциях и других занятиях, выработку навыков самостоятельного активного приобретения новых, дополнительных знаний, подготовку к предстоящим занятиям.

6. Оценочные материалы по дисциплине

Текущий контроль успеваемости обеспечивает оценивание хода освоения дисциплины, проводится в соответствии с содержанием дисциплины по видам занятий в форме опроса, тестирования, решения практических задач.

Промежуточная аттестация обеспечивает оценивание промежуточных и окончательных результатов обучения по дисциплине, проводится в форме экзамена.

6.1 Примерные оценочные материалы:

6.1.1. Вопросы для текущего контроля

Типовые вопросы для опроса:

1. Дайте определение понятию «данные»
2. Чем дискретные данные отличаются от непрерывных?
3. Что такое анализ данных?
4. Назовите этапы анализа данных.
5. Что такое доверительный интервал?
6. На основе какого метода вычисляются коэффициенты уравнения регрессии?
7. В чем специфика построения моделей по временным рядам данных
8. Дайте определение OLAP системы
9. Какие задачи решаются Data Mining?

Типовые задания для тестирования:

1. Какие существуют виды временных рядов

- 1: ряд Тейлора и ряд Эйлера
- 2: возрастающий временной ряд и убывающий временной ряд
- 3: интервальный временной ряд и моментный временной ряд
- 4: арифметический ряд и геометрический ряд
- 5: арифметическая прогрессия и геометрическая прогрессия

2. В каких формах может задаваться временной ряд

- 1: в виде возрастающей последовательности
- 2: в виде числовой прогрессии
- 3: в виде геометрической прогрессии
- 4: в табличной и графической форме
- 5: в виде убывающей последовательности

3. Что такое корреляция?

- 1: закономерность развития какого-либо явления
- 2: статистическая взаимосвязь двух или более случайных величин
- 3: независимость двух или более случайных величин
- 4: рост количества исследуемых явлений
- 5: степень статистической зависимости между двумя множествами

данных

4. Временной ряд состоит из следующих составляющих

- 1: трендовая и циклическая составляющая
- 2: циклическая и сезонная составляющая
- 3: вероятная и прогнозная составляющая
- 4: детерминированная и случайная составляющая

5. Кластеризация - это ...

1: задача машинного обучения, в которой метки объектов принимают ограниченное число значений, например, ранг пожара

2: задача машинного обучения, в которой метки объектов принимают любое численное значение, например, размер ущерба

3: задача машинного обучения, заключающаяся в объединении похожих объектов в однородные группы

6. Определение ранга пожара на основе анализа первичных данных является задачей ...

- 1: регрессии
- 2: классификации
- 3: кластеризации

Типовые задания для практических задач:

1. На территории лесопарка на учетной площади проводилось измерение диаметров стволов подроста вяза гладкого.

Диаметр стволов вяза, см: 1,4; 1,1; 0,8; 1,7; 0,9; 2,0; 1,8; 0,9; 0,5; 1,4; 1,1; 1,8; 0,8; 1,1; 2,7; 2,5; 1,2; 1,0; 1,2; 2,0; 1,6; 2,6; 2,4; 0,9; 1,6; 1,9; 2,0; 2,1; 2,6; 2,8; 1,9; 2,0; 1,6; 1,6; 1,2; 4,2; 1,9; 2,1; 1,8; 3,2; 2,1; 2,1; 2,2; 0,8; 1,1; 2,7; 2,5; 1,2; 1,0; 1,2; 1,9; 2,1; 1,8; 3,2; 1,5; 4,2; 2,3; 1,5; 2,4; 3,2; 1,1; 2,7; 2,5; 1,2; 1,0; 1,2; 1,1; 2,7; 2,5; 1,2; 1,0; 1,2; 4,2; 1,9; 2,1; 1,8; 3,2; 2,1; 2,1; 2,2; 0,8; 1,1; 2,7; 2,5; 1,2; 4,2; 1,9; 2,1; 1,8; 3,2; 2,1; 2,1; 2,2; 0,8; 1,1; 2,7; 2,5; 1,2; 1,1; 2,7; 2,5; 1,2; 1,0; 1,2; 2,0; 1,6;

2,6; 2,4; 0,9; 1,6; 1,9; 2,0; 2,1; 2,6; 1,1; 2,7; 2,5; 1,2; 1,0; 1,2; 2,0; 1,6; 2,6; 2,4; 0,9; 1,6; 1,9; 2,0; 2,1; 2,6; 1,1; 2,7; 2,5; 1,2; 1,0; 1,2; 2,0; 1,6; 2,6; 2,4; 0,9; 1,6; 1,9; 2,0; 2,1; 2,6; 1,1; 2,7; 2,5; 1,2; 1,0; 1,2; 2,0; 1,6; 2,6; 2,4; 0,9; 1,6; 1,9; 2,0; 2,1; 2,6; 1,1; 2,7; 2,5; 1,2; 1,0; 1,2; 2,0; 1,6; 2,6; 2,4; 0,9; 1,6; 1,9; 2,0; 2,1; 2,1; 2,6; 1,1; 2,7; 2,5; 1,2; 1,0; 1,2; 3,2; 1,5; 4,2; 2,3; 1,5; 2,4; 3,2; 1,1; 2,7; 2,5; 1,2; 1,0; 1,2; 2,6.

С помощью процедуры "Описательная статистика" вычислите следующие показатели: объем выборки, среднюю арифметическую, стандартное отклонение, верхний и нижний квартили, медиану. Чему равен коэффициент вариации изучаемого признака?

2. Имеются данные по проективному покрытию травостоя на территории лесопарка.

Проективное покрытие, %: 50, 40, 50, 60, 60, 50, 40, 50, 40, 40, 30, 30, 60, 70, 70, 60, 70, 50, 40, 40, 60, 60, 60, 50, 30, 30, 40, 50, 70, 70, 60, 60, 60, 70, 50, 60, 40, 50, 10, 30, 70, 40, 60, 40, 30, 20, 40, 70, 50, 40, 40, 60, 60, 60, 50, 30, 30, 40, 50, 70, 70, 60, 60, 60, 70, 50, 60, 70, 40, 60, 40, 30, 20, 40, 70, 50, 40, 40, 60, 60, 60, 50, 30, 30, 40, 50, 30, 30, 40, 50, 70, 40, 60, 40, 30, 20, 40, 70, 50, 40, 40, 60, 60, 60, 50, 30, 30, 40, 50, 70, 40, 60, 40, 30, 20, 40, 70, 50, 40, 40, 60, 60, 60, 50, 30, 30, 40, 50, 70, 40, 60, 40, 30, 20, 40, 70, 50, 40, 40, 60, 60, 60, 50, 30, 30, 40, 50, 70, 40, 60, 40, 30, 20, 40, 70, 50, 40, 40, 60, 60, 70, 70, 60, 70, 50, 40, 40, 70, 40, 60, 40, 30, 20, 40, 70, 50, 40, 40, 60, 60, 60, 50, 60.

При помощи "Описательной статистики" вычислите следующие показатели: объем выборки, среднюю арифметическую, ошибку средней, стандартное отклонение, 95-% доверительный интервал, размах. Чему равен коэффициент вариации признака?

3. На 10 участках леса были проведены измерения диаметра стволов деревьев, высоты кустарников, проективного покрытия травостоя и плотности населения птиц-дуплогнездников.

Необходимо выявить, имеются ли корреляции между:

а) высотой кустарников и проективным покрытием травостоя; б) диаметром стволов деревьев и плотностью гнездования птиц.

По итогам проведения корреляционного анализа сделать вывод о характере корреляции (линейная или нет), ее знаке и значимости.

6.1.2. Вопросы для промежуточной аттестации

Примерный перечень вопросов, выносимых на экзамен

1. Понятие анализа данных. Методология анализа данных.
2. Подходы к моделированию.
3. Принципы анализа данных.
4. Типы данных. Методы анализа данных.
5. Этапы статистического исследования.

6. Задачи и методы статистического анализа данных.
7. Корреляционный анализ данных.
8. Поле корреляции, его построение и применение.
9. Выборочный коэффициент корреляции и его свойства.
10. Понятие ранговой корреляции.
11. Основные понятия и методы регрессионного анализа.
12. Применение метода наименьших квадратов в регрессионном анализе. Уравнение парной линейной регрессии.
13. Нелинейные модели регрессии.
14. Виды временных рядов. Их построение
15. Составляющие элементы временного ряда.
16. Показатели временного ряда и методы их исчисления
17. Предварительная обработка временного ряда
18. Методы распознавания типа тренда и оценки его параметров
19. Статистические методы анализа факторов
20. Экспоненциальное сглаживание.
21. Начальное значение экспоненциального сглаживания.
22. Постоянная экспоненциального сглаживания.
23. Полиномиальные модели и многократное сглаживание.
24. Доверительная вероятность и точность прогноза.
25. Сезонные модели.
26. Основные понятия машинного обучения.
27. Обучение с учителем и обучение без учителя.
28. Классы задач машинного обучения: регрессия, классификация, кластерный анализ.
29. Постановка задачи классификации с обучением.
30. Линейные алгоритмы классификации.
31. Логистическая регрессия.
32. Понятие о деревьях решений.
33. Метрики качества классификации
34. Постановка задачи кластерного анализа.
35. Метод К-средних
36. Понятие о методах машинного обучения в задачах поиска аномалий.
37. Метрики качества, используемые в задачах кластеризации.

6.2 Шкала оценивания результатов промежуточной аттестации и критерии выставления оценок

Форма контроля	Показатели оценивания	Критерии выставления оценок	Шкала оценивания
экзамен	правильность и полнота ответа	дан правильный, полный ответ на поставленный вопрос, показана совокупность осознанных знаний по дисциплине, доказательно раскрыты основные положения вопросов; могут быть допущены недочеты, исправленные самостоятельно в процессе ответа.	отлично
		дан правильный, недостаточно полный ответ на поставленный вопрос, показано умение выделить существенные и несущественные признаки, причинно-следственные связи; могут быть допущены недочеты, исправленные с помощью преподавателя.	хорошо
		дан недостаточно правильный и полный ответ; логика и последовательность изложения имеют нарушения; в ответе отсутствуют выводы.	удовлетворительно
		ответ представляет собой разрозненные знания с существенными ошибками по вопросу; присутствуют фрагментарность, нелогичность изложения; дополнительные и уточняющие вопросы не приводят к коррекции ответа на вопрос	неудовлетворительно

7. Ресурсное обеспечение дисциплины

7.1. Лицензионное и свободно распространяемое программное обеспечения:

Перечень лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства:

1. Astra Linux Common Edition релиз Орел - операционная система общего назначения. Лицензия №217800111-ore-2.12-client-6196

2. Astra Linux Special Edition - операционная система общего назначения. Лицензия №217800111-alse-1.7-client-medium-x86_64-0-14545

3. Astra Linux Special Edition - операционная система общего назначения. Лицензия №217800111-alse-1.7-client-medium-x86_64-0-14544

7.2. Профессиональные базы данных и информационные справочные системы

1. Портал открытых данных Российской Федерации <https://data.gov.ru/> (свободный доступ);
2. Федеральный портал «Российское образование» <http://www.edu.ru> (свободный доступ);
3. Система официального опубликования правовых актов в электронном виде <http://publication.pravo.gov.ru> (свободный доступ);
4. Электронная библиотека университета <http://elib.igps.ru> (авторизованный доступ);
5. Электронно-библиотечная система «ЭБС IPR BOOKS» <http://www.iprbookshop.ru> (авторизованный доступ).
6. Электронно-библиотечная система "Лань" <https://e.lanbook.com> (авторизованный доступ).

7.1. Литература

Основная литература:

1. Матвеев А.В. Методы моделирования и прогнозирования / А.В. Матвеев. – СПб.: Санкт-Петербургский университет ГПС МЧС России, 2022. – 231 с. Режим доступа: <http://elib.igps.ru/?5&type=card&cid=ALSFR-6782836a-aa47-45c0-a375-56944375dd26&remote=false>
2. Бутырский Е.Ю. Математическое моделирование систем и процессов / Е.Ю. Бутырский, А.В. Матвеев. – СПб.: Санкт-Петербургский университет ГПС МЧС России, 2022. – 733 с. Режим доступа: <http://elib.igps.ru/?3&type=card&cid=ALSFR-af3b3441-07fe-40ad-8218-cb3af5b35666&remote=false>
3. Мельниченко А.С. Математическая статистика и анализ данных: учебное пособие / Мельниченко А.С.. — Москва : Издательский Дом МИСиС, 2018. — 45 с. — ISBN 978-5-906953-62-9. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/78563.html>
4. Ракитский А.А. Методы машинного обучения: учебно-методическое пособие / Ракитский А.А.. — Новосибирск : Сибирский государственный университет телекоммуникаций и информатики, 2018. — 32 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/90591.html>

Дополнительная литература:

1. Заборский Б.В., Крюкова М.С., Медведева О.М. Теория вероятностей и математическая статистика: учебное пособие. – СПб УГПС МЧС России, 2017. Режим доступа: <http://elib.igps.ru/?5&type=card&cid=ALSFR-193f0c9f-5f9c-43f0-8a80-2eaff3aaf1f3&remote=false>

2. Сараев П.В. Методы машинного обучения: методические указания и задания к лабораторным работам по курсу / Сараев П.В.. — Липецк : Липецкий государственный технический университет, ЭБС АСВ, 2017. — 48 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/83183.html>

7.4 Материально-техническое обеспечение дисциплины

Для проведения и обеспечения занятий используются помещения, которые представляют собой учебные аудитории для проведения учебных занятий, предусмотренных программой бакалавриата, оснащенные оборудованием и техническими средствами обучения: автоматизированное рабочее место преподавателя, маркерная доска, мультимедийный проектор, документ-камера, посадочные места обучающихся.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа к электронной информационно-образовательной среде университета.

Авторы: кандидат технических наук, доцент Матвеев Александр Владимирович.