

**Федеральное государственное автономное образовательное
учреждение высшего образования
«Московский физико-технический институт
(национальный исследовательский университет)»**

УТВЕРЖДЕНО

**Директор физтех-школы
прикладной математики и
информатики**

А.М. Райгородский

Рабочая программа дисциплины (модуля)

по дисциплине:	Методы математической статистики в задачах моделирования и распознавания образов
по направлению:	Информатика и вычислительная техника
профиль подготовки:	Прикладная математика и информатика Физтех-школа Прикладной Математики и Информатики кафедра вычислительных технологий и моделирования в геофизике и биоматематике
курс:	1
квалификация:	магистр

Семестр, формы промежуточной аттестации: 1 (осенний) - Дифференцированный зачет

Аудиторных часов: 30 всего, в том числе:

лекции: 30 час.

семинары: 0 час.

лабораторные занятия: 0 час.

Самостоятельная работа: 60 час.

Всего часов: 90, всего зач. ед.: 2

Программу составил: Е.В. Дмитриев, канд. физ.-мат. наук

Программа обсуждена на заседании кафедры вычислительных технологий и моделирования в геофизике и биоматематике 15.03.2024

Аннотация

Дисциплина "Методы математической статистики в задачах моделирования и распознавания образов" изучает применение методов математической статистики для анализа и обработки данных в задачах моделирования и распознавания образов. Она объединяет понятия и инструменты из областей математической статистики, машинного обучения и компьютерного зрения. В курсе рассмотрены основные подходы, методы и алгоритмы описания классов, нахождения решающих функций, выбора информативной системы признаков в случае малой неопределенности исходных данных (детерминистский подход) и в случае большой неопределенности исходных данных вероятностного характера (статистический подход).

1. Цели и задачи

Цель дисциплины

- освоение студентами основных идей и методов матричного спектрального анализа и их применения в задачах численного исследования устойчивости нестационарных физических, технических и физиологических систем.

Задачи дисциплины

- формирование базовых знаний в области прикладных методов математической статистики, используемых для интеллектуализации обработки информации;
- обучение студентов принципам построения статистических моделей;
- обучение студентов принципам создания новых методов распознавания образов и решения обратных задач, основанных на методах математической статистики.
- ознакомление с основными задачами, возникающими при обработке данных дистанционного аэрокосмического зондирования Земли;
- проведение общетеоретической и практической подготовки специалистов в области прикладной математики и физики и обеспечивающей технологические основы современных методов решения задач
- формирование подходов к выполнению исследований студентами в области решения практических задач в рамках выпускных работ на степень магистра.

2. Перечень формируемых компетенций

Освоение дисциплины направлено на формирование следующих компетенций:

Код и наименование компетенции	Индикаторы достижения компетенции
ОПК-2 Имеет представление об актуальных проблемах науки и техники в области информатики и вычислительной техники, способен на научном языке формулировать профессиональные задачи	ОПК-2.1 Имеет представление о современном состоянии исследований в рамках тематической области своей профессиональной деятельности
	ОПК-2.2 Способен оценивать актуальность исследований в области информатики и вычислительной техники и их практическую значимость
	ОПК-2.3 Владеет профессиональной терминологией, используемой в современной научно-технической литературе, обладает навыками устного и письменного изложения результатов научной деятельности в рамках профессиональной коммуникации
ПК-2 Понимает и способен применить в научно-исследовательской и прикладной деятельности основные законы естествознания, современный математический аппарат и алгоритмы	ПК-2.1 Знает основы научно-исследовательской деятельности в области информационных технологий, владеет знанием основ философии и методологии науки; знанием методов научных исследований и навыками их проведения
	ПК-2.2 Умеет применять полученные знания в области фундаментальных научных основ теории информации и решать стандартные задачи в собственной научно-исследовательской деятельности

математический аппарат и алгоритмы, современные информационно-коммуникационные технологии	ПК-2.3 Имеет практический опыт научно-исследовательской деятельности в области информационно-коммуникационных технологий
	ПК-2.4 Владеет методами и алгоритмами решения задач цифровой обработки сигналов, использования сети Интернет, аннотирования, реферирования, библиографического поиска, опыт работы с научными источниками
ПК-1 Готов к включению в профессиональное сообщество; способен проводить под научным руководством локальные исследования на основе существующих методов в конкретной области профессиональной деятельности	ПК-1.1 Знает принципы построения научной работы, методы сбора и анализа полученного материала, способы аргументации; владеет навыками подготовки научных обзоров, публикаций, рефератов и библиографий по тематике проводимых исследований на русском и английском языке
	ПК-1.2 Умеет решать научные задачи с пониманием существующих подходов к верификации моделей программного обеспечения в связи с поставленной целью и в соответствии с выбранной методикой
	ПК-1.3 Имеет практический опыт выступлений и научной аргументации при анализе объекта научной и профессиональной деятельности

3. Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- место и роль общих вопросов науки в научных исследованиях;
- современные проблемы физики, математики;
- теоретические модели фундаментальных процессов и явлений в физике и ее приложениях;
- новейшие открытия естествознания;
- постановку проблем физического моделирования;
- о взаимосвязях и фундаментальном единстве естественных наук.

уметь:

- эффективно использовать на практике теоретические компоненты науки: понятия, суждения, умозаключения, законы;
- представить панораму универсальных методов и законов современного естествознания;
- абстрагироваться от несущественных влияний при моделировании реальных физических ситуаций.

владеть:

- математическими методами интеллектуальной обработки информации.

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

4.1. Разделы дисциплины (модуля) и трудоемкости по видам учебных занятий

№	Тема (раздел) дисциплины	Трудоемкость по видам учебных занятий, включая самостоятельную работу, час.			
		Лекции	Семинары	Лаборат. работы	Самост. работа
1	Базовые понятия и основные задачи математической статистики. Методы получения точечных и интервальных оценок.	6			15
2	Базовые одномерные и многомерные статистические модели.	8			15

3	Классификация без обучения. Модель гауссовских смесей. Графовые методы. Метод k-средних. Иерархический кластерный анализ.	8			15
4	Введение в методы получения и первичной обработки аэрокосмической информации. Приложения тематической обработки аэрокосмических изображений.	8			15
Итого часов		30			60
Подготовка к экзамену		0 час.			
Общая трудоёмкость		90 час., 2 зач.ед.			

4.2. Содержание дисциплины (модуля), структурированное по темам (разделам)

Семестр: 1 (Осенний)

1. Базовые понятия и основные задачи математической статистики. Методы получения точечных и интервальных оценок.

Формулировка основных задач и понятий математической статистики. Стандартные и специальные оценки параметров распределений. Обоснование и свойства точечных и интервальных оценок.

Формулировка и обоснование различных подходов по оценке эффективности статистических оценок.

Постановка базовых задач проверки статистических гипотез. Критерий отношения правдоподобия.

Формулировка и обоснование параметрических методов проверки статистических гипотез. Критерии согласия.

Формулировка и обоснование непараметрических методов проверки статистических гипотез. Формулировка и обоснование методов повторных выборок.

2. Базовые одномерные и многомерные статистические модели.

Методы одномерных линейной и нелинейной регрессии. Многомерные линейные модели. Связь одномерных и многомерных оценок.

Обоснование и сравнение методов редукции измерений и статистической регуляризации.

Свойства и приложения псевдообратных матриц в задачах построения оптимальных линейных оценок.

Применение теории гауссовских случайных процессов для построения статистических моделей.

Проблемы переподгонки и проклятия размерности. Обоснование и приложения пошаговых оптимальных оценок. Разложение случайных векторов по базисам EOF и SVD, канонический корреляционный анализ.

3. Классификация без обучения. Модель гауссовских смесей. Графовые методы. Метод k-средних. Иерархический кластерный анализ.

Основополагающие методы классификации без обучения: метод ФорЭл, метод k-средних и иерархический кластерный анализ. Графовые методы кластеризации. Формулировка и приложения модели гауссовских смесей, EM алгоритм.

Обоснование и приложения метрических классификаторов. Основные принципы Байесовской классификации.

Обоснование и приложения метода опорных векторов.

Построение и обучение решающих деревьев. Рандомизация случайных деревьев.

Обоснование и приложения ансамблевых методов обучаемой классификации. Решение задач многоклассовой обучаемой классификации на основе бинарных классификаторов.

Использование сверточных нейронных сетей в задачах распознавания образов.

4. Введение в методы получения и первичной обработки аэрокосмической информации. Приложения тематической обработки аэрокосмических изображений.

Описание и сравнение существующих систем дистанционного зондирования. Описание и сравнение существующих систем отображения снимков и тематической обработки данных. Методы трансформации, сшивки и географической привязки аэрокосмических изображений. Картографические проекции. Принципы калибровки. Методы коррекции шумов. Стандартные спектральные преобразования. Повышение контрастности изображений.

Принципы интерпретации снимков, характеристики изображений в различных спектральных диапазонах съемки. Принципы построения пространства спектральных и текстурных образов объектов. Практические примеры применения методов тематической обработки наземных измерений и аэрокосмических изображений.

5. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)

Аудитория с проектором.

6.Перечень рекомендуемой литературы

Основная литература

1. Козодеров В. В., Кондранин Т. В., Дмитриев Е. В. Методы обработки многоспектральных и гиперспектральных аэрокосмических изображений // Москва: МФТИ, 2013. - 224 с.
2. Хайкин С. Нейронные сети: полный курс.— 2-е изд., испр. — Санкт-Петербург : Диалектика, 2019. — 1104 с. - ISBN 978-5-907144-22-4.
3. Чавро А.И., Дымников В.П. Методы математической статистики в задачах физики атмосферы // Москва: ИВМ РАН, 2000. - 210 с.

Дополнительная литература

1. Козодеров В. В., Дмитриев Е. В., Каменцев В. П. Когнитивные технологии дистанционного зондирования в природопользовании: монография. - Тверь: ТвГУ. 2016. - 280 с.
2. Румшинский Л. З. Математическая обработка результатов эксперимента. – М.: Наука, 1971. - 192 с.
3. Пытьев Ю. П. Математические методы анализа эксперимента. – М.: Высшая школа, 1989.
4. Дьяконов В. П. MATLAB: Полный самоучитель / В. П. Дьяконов.— М. : ДМК Пресс, 2014.— 768 с. - ISBN 978-5-97060-082-5.
5. Шовенгерт Р.А. Дистанционное зондирование. Методы и модели обработки изображений // Москва: Техносфера, 2010. - 560 с.

7. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)

Не используются

8. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень необходимого программного обеспечения и информационных справочных систем (при необходимости)

Не требуется.

9. Методические указания для обучающихся по освоению дисциплины (модуля)

Студент, изучающий дисциплину, должен с одной стороны, овладеть общим понятийным аппаратом, а с другой стороны, должен научиться применять теоретические знания на практике. В результате изучения дисциплины студент должен знать основные определения, понятия, аксиомы, алгоритмы.

Успешное освоение курса требует напряжённой самостоятельной работы студента. В программе курса приведено минимально необходимое время для работы студента над темой. Самостоятельная работа включает в себя:

- чтение и конспектирование рекомендованной литературы;
- проработку учебного материала (по конспектам лекций, учебной и научной литературе), подготовку ответов на вопросы, предназначенных для самостоятельного изучения, доказательство отдельных утверждений, свойств;
- подготовку к дифференцированному зачету.

Руководство и контроль за самостоятельной работой студента осуществляется в форме индивидуальных консультаций.

Важно добиться понимания изучаемого материала, а не механического его запоминания. При затруднении изучения отдельных тем, вопросов, следует обращаться за консультациями к лектору.

Обязательные задания: одно задание по задачам из 1-4 (основная литература).

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

по направлению:	Информатика и вычислительная техника
профиль подготовки:	Прикладная математика и информатика Физтех-школа Прикладной Математики и Информатики кафедра вычислительных технологий и моделирования в геофизике и биоматематике
курс:	<u>1</u>
квалификация:	магистр
Семестр, формы промежуточной аттестации: 1 (осенний) - Дифференцированный зачет	
Разработчик:	Е.В. Дмитриев, канд. физ.-мат. наук

1. Компетенции, формируемые в процессе изучения дисциплины

Код и наименование компетенции	Индикаторы достижения компетенции
ОПК-2 Имеет представление об актуальных проблемах науки и техники в области информатики и вычислительной техники, способен на научном языке формулировать профессиональные задачи	ОПК-2.1 Имеет представление о современном состоянии исследований в рамках тематической области своей профессиональной деятельности
	ОПК-2.2 Способен оценивать актуальность исследований в области информатики и вычислительной техники и их практическую значимость
	ОПК-2.3 Владеет профессиональной терминологией, используемой в современной научно-технической литературе, обладает навыками устного и письменного изложения результатов научной деятельности в рамках профессиональной коммуникации
ПК-2 Понимает и способен применить в научно-исследовательской и прикладной деятельности основные законы естествознания, современный математический аппарат и алгоритмы, современные информационно-коммуникационные технологии	ПК-2.1 Знает основы научно-исследовательской деятельности в области информационных технологий, владеет знанием основ философии и методологии науки; знанием методов научных исследований и навыками их проведения
	ПК-2.2 Умеет применять полученные знания в области фундаментальных научных основ теории информации и решать стандартные задачи в собственной научно-исследовательской деятельности
	ПК-2.3 Имеет практический опыт научно-исследовательской деятельности в области информационно-коммуникационных технологий
	ПК-2.4 Владеет методами и алгоритмами решения задач цифровой обработки сигналов, использования сети Интернет, аннотирования, реферирования, библиографического поиска, опыт работы с научными источниками
ПК-1 Готов к включению в профессиональное сообщество; способен проводить под научным руководством локальные исследования на основе существующих методов в конкретной области профессиональной деятельности	ПК-1.1 Знает принципы построения научной работы, методы сбора и анализа полученного материала, способы аргументации; владеет навыками подготовки научных обзоров, публикаций, рефератов и библиографий по тематике проводимых исследований на русском и английском языке
	ПК-1.2 Умеет решать научные задачи с пониманием существующих подходов к верификации моделей программного обеспечения в связи с поставленной целью и в соответствии с выбранной методикой
	ПК-1.3 Имеет практический опыт выступлений и научной аргументации при анализе объекта научной и профессиональной деятельности

2. Показатели оценивания компетенций

В результате изучения дисциплины «Методы математической статистики в задачах моделирования и распознавания образов» обучающийся должен:

знать:

- место и роль общих вопросов науки в научных исследованиях;
- современные проблемы физики, математики;
- теоретические модели фундаментальных процессов и явлений в физике и ее приложениях;
- новейшие открытия естествознания;
- постановку проблем физического моделирования;
- о взаимосвязях и фундаментальном единстве естественных наук.

уметь:

- эффективно использовать на практике теоретические компоненты науки: понятия, суждения, умозаключения, законы;
- представить панораму универсальных методов и законов современного естествознания;
- абстрагироваться от несущественных влияний при моделировании реальных физических ситуаций.

владеть:

- математическими методами интеллектуальной обработки информации.

3. Перечень типовых (примерных) вопросов, заданий, тем для подготовки к текущему контролю

С целью контроля освоения обучающимися учебного материала проводится устный опрос в начале занятия по теме прошлой лекции или в конце занятия по пройденной теме.

4. Перечень типовых (примерных) вопросов и тем для проведения промежуточной аттестации обучающихся

1. Системы дистанционного зондирования и обработки снимков.
2. Методы пространственной и спектральной коррекции снимков.
3. Интерпретация снимков, спектральные и текстурные образы объектов.
4. Методы сшивки и повышения контрастности изображений.
5. Спектральная и радиометрическая калибровка. Радиометрическая и радиационная коррекция.
6. Многозональные индексы.
7. Стандартные и специальные точечные оценки. Эффективность точечных оценок.
8. Параметрические и непараметрические методы получения интервальных оценок.
9. Критерий отношения правдоподобия.
10. Критерии согласия.
11. Непараметрические методы проверки статистических гипотез.
12. Методы повторных выборок.
13. Одномерная линейная и нелинейная регрессия.
14. Метод редукции измерений
15. Метод статистической регуляризации.

Критерии оценивания

Оценка отлично 10 баллов - выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины, проявляющему интерес к данной предметной области, продемонстрировавшему умение уверенно и творчески применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений.

Оценка отлично 9 баллов - выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины и умение уверенно применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений.

Оценка отлично 8 баллов - выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины и умение уверенно применять их на практике при решении конкретных задач, правильное обоснование принятых решений, с некоторыми недочетами.

Оценка хорошо 7 баллов - выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но недостаточно грамотно обосновывает полученные результаты.

Оценка хорошо 6 баллов - выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности.

Оценка хорошо 5 баллов - выставляется студенту, если он в основном знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач достаточно большое количество неточностей.

Оценка удовлетворительно 4 балла - выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, недостаточно правильные формулировки базовых понятий, нарушения логической последовательности в изложении программного материала, но при этом он освоил основные разделы учебной программы, необходимые для дальнейшего обучения, и может применять полученные знания по образцу в стандартной ситуации.

Оценка удовлетворительно 3 балла - выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, допускающему ошибки в формулировках базовых понятий, нарушения логической последовательности в изложении программного материала, слабо владеет основными разделами учебной программы, необходимыми для дальнейшего обучения и с трудом применяет полученные знания даже в стандартной ситуации.

Оценка неудовлетворительно 2 балла - выставляется студенту, который не знает большей части основного содержания учебной программы дисциплины, допускает грубые ошибки в формулировках основных принципов и не умеет использовать полученные знания при решении типовых задач.

Оценка неудовлетворительно 1 балл - выставляется студенту, который не знает основного содержания учебной программы дисциплины, допускает грубейшие ошибки в формулировках базовых понятий дисциплины и вообще не имеет навыков решения типовых практических задач.

5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности

Во время проведения дифференцированного зачёта обучающиеся могут пользоваться программой дисциплины, а также справочной литературой, вычислительной техникой, конспектами лекций.

Дифференцированный зачёт проводится путем организации специального опроса, проводимого в устной форме.