

**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:	Байесовский выбор моделей
по направлению:	Информатика и вычислительная техника
профиль подготовки:	Математическое моделирование и компьютерные технологии Физтех-школа Прикладной Математики и Информатики кафедра интеллектуальных систем
курс:	4
квалификация:	бакалавр

Семестры, формы промежуточной аттестации:

7 (осенний) - Дифференцированный зачет

8 (весенний) - Экзамен

Аудиторных часов: 60 всего, в том числе:

лекции: 30 час.

семинары: 30 час.

лабораторные занятия: 0 час.

Самостоятельная работа: 45 час.

Подготовка к экзамену: 30 час.

Всего часов: 135, всего зач. ед.: 3

Программу составил: А.А. Адуенко, канд. физ.-мат. наук

Программа обсуждена на заседании кафедры интеллектуальных систем 18.01.2023

Аннотация

В курсе рассматривается байесовский подход к машинному обучению, Рассматриваются понятия: байесовская линейная регрессия, гауссовские процессы и эволюция моделей во времени, методы Монте-Карло по схеме марковских цепей, байесовская оптимизация, понятие графической модели.

1. Цели и задачи

Цель дисциплины

Обучить студентов байесовскому подходу к машинному обучению, сформировать глубокое понимание основных методов и понятий подхода, а также их связи с результатами из классического машинного обучения, теории вероятностей и статистики.

Задачи дисциплины

- Исследование свойств байесовского подхода к машинному обучению;
- сформировать глубокое понимание основных методов и понятий байесовского подхода к машинному обучению.

2. Перечень формируемых компетенций

Освоение дисциплины направлено на формирование следующих компетенций:

Код и наименование компетенции	Индикаторы достижения компетенции
УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.4 Грамотно, логично, аргументированно формирует собственные суждения и оценки
	УК-1.5 Определяет и оценивает практические последствия возможных вариантов решения задачи
ОПК-3 Способен составлять и оформлять научные и (или) технические (технологические, инновационные) отчеты (публикации, проекты)	ОПК-3.2 Владеет на практике методологией составления научно-технических отчетов (проектов)
	ОПК-3.3 Владеет методами визуального и графического представления результатов научной (научно-технической, инновационной технологической) деятельности в виде отчетов, научных публикаций

3. Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- свойства и границы применимости разных методов вывода и уметь выбирать их под задачу.

уметь:

- строить вероятностные модели данных, учитывающие представления о зависимостях между переменными, их распределениях, априорные и экспертные знания.

владеть:

- аппаратом точного и приближенного вывода в байесовских моделях, а также методами построения оптимального и приближенно оптимального байесовского прогноза.

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	Введение.	3	3		3
2	Байесовский классификатор	4	4		4

3	Алгоритмы	4	4		4
4	Методы Монте-Карло	4	4		4
5	Байесовская оптимизация	5	5		10
6	Понятие графической модели.	5	5		10
7	Алгоритмы вывода в графических моделях	5	5		10
Итого часов		30	30		45
Подготовка к экзамену		30 час.			
Общая трудоёмкость		135 час., 3 зач.ед.			

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

Семестр: 7 (Осенний)

1. Введение.

Введение: Напоминание понятий из теории вероятностей и статистики.
Множественное тестирование гипотез и выбор априорного распределения.

2. Байесовский классификатор

Наивный байесовский классификатор, его обобщения и оптимальный прогноз.
Экспоненциальное семейство распределений.
Байесовская линейная регрессия. Обоснованность (evidence).
Байесовская логистическая регрессия и отбор признаков.

3. Алгоритмы

ЕМ-алгоритм и вариационный ЕМ-алгоритм.
Гауссовские процессы и эволюция моделей во времени.
Построение адекватных мультимodelей.

4. Методы Монте-Карло

Методы Монте-Карло по схеме марковских цепей.
Гамильтоновы методы Монте-Карло по схеме марковских цепей.

Семестр: 8 (Весенний)

5. Байесовская оптимизация

Краткое введение. Введение гиперпараметра. Алгоритм автоматической настройки параметров

6. Понятие графической модели.

Понятие графической модели. Представление правдоподобия в виде графа.
Ориентированные графические модели и понятие условной отделимости.
Неориентированные графические модели. Связь и сравнение неориентированных графических моделей с ориентированными.

7. Алгоритмы вывода в графических моделях

Фактор-графы. Алгоритм Sum-Product точного вывода в ациклических моделях.

Скрытые марковские цепи и линейные динамические системы.

Фильтр Калмана.

Алгоритмы вывода в графических моделях с циклами.

Алгоритмы вывода на основании разрезов графов и их применение к сегментации изображений.

Структурное обучение. Сравнение классического и нейросетевого подхода к графическим моделям.

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

Необходимое оборудование для практических занятий: компьютер и мультимедийное оборудование (проектор, маркерная доска, связь с Интернетом).

Необходимое программное обеспечение: система компьютерной алгебры Matlab, систему верстки TeX.

6. Перечень рекомендуемой литературы

Основная литература

1. Байесовские методы статистического оценивания, надежность технических объектов/В. П. Савчук, -М., Наука, 1989

Дополнительная литература

1. David MacKay, 2005, Information Theory, Inference, and Learning Algorithms
2. Christopher Bishop, 2006, Pattern Recognition and Machine Learning
3. David Barber, 2014, Bayesian Reasoning and Machine Learning
4. Daphne Koller and Nir Friedman, 2009, Probabilistic Graphical Models
5. Kevin P. Murphy, 2012, Machine Learning: a Probabilistic Perspective
6. Wainwright, M. J., & Jordan, M. I., 2008, Graphical models, exponential families, and variational inference.
7. Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. Deep learning. MIT press, 2016.

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)

На занятиях используются мультимедийные технологии, включая демонстрацию презентаций. Используются электронные учебники.

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 учебной и научной литературе), подготовку ответов на вопросы, предназначенных для самостоятельного изучения, доказательство отдельных утверждений, свойств;

– подготовку к дифференцированному зачету и экзамену.

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.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

по направлению:	Информатика и вычислительная техника
профиль подготовки:	Математическое моделирование и компьютерные технологии Физтех-школа Прикладной Математики и Информатики кафедра интеллектуальных систем
курс:	4
квалификация:	бакалавр

Семестры, формы промежуточной аттестации:

7 (осенний) - Дифференцированный зачет
8 (весенний) - Экзамен

Разработчик: А.А. Адуенко, канд. физ.-мат. наук

1. Компетенции, формируемые в процессе изучения дисциплины

Код и наименование компетенции	Индикаторы достижения компетенции
УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.4 Грамотно, логично, аргументированно формирует собственные суждения и оценки
	УК-1.5 Определяет и оценивает практические последствия возможных вариантов решения задачи
ОПК-3 Способен составлять и оформлять научные и (или) технические (технологические, инновационные) отчеты (публикации, проекты)	ОПК-3.2 Владеет на практике методологией составления научно-технических отчетов (проектов)
	ОПК-3.3 Владеет методами визуального и графического представления результатов научной (научно-технической, инновационной технологической) деятельности в виде отчетов, научных публикаций

2. Показатели оценивания компетенций

В результате изучения дисциплины «Байесовский выбор моделей» обучающийся должен:

знать:

- свойства и границы применимости разных методов вывода и уметь выбирать их под задачу.

уметь:

- строить вероятностные модели данных, учитывающие представления о зависимостях между переменными, их распределениях, априорные и экспертные знания.

владеть:

- аппаратом точного и приближенного вывода в байесовских моделях, а также методами построения оптимального и приближенно оптимального байесовского прогноза.

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 или в конце занятия по пройденной теме.

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

Список вопросов к дифференцированному зачету:

1. Множественное тестирование гипотез и выбор априорного распределения.
2. Фильтр Калмана.
3. Наивный байесовский классификатор, его обобщения и оптимальный прогноз.
4. Фактор-графы. Алгоритм Sum-Product точного вывода в ациклических моделях.
5. Экспоненциальное семейство распределений.
6. Скрытые марковские цепи и линейные динамические системы.
7. Байесовская линейная регрессия. Обоснованность (evidence).
8. Понятие графической модели. Представление правдоподобия в виде графа.
9. Байесовская логистическая регрессия и отбор признаков.
10. Байесовская оптимизация.

Список вопросов к экзамену:

1. EM-алгоритм и вариационный EM-алгоритм.
2. Понятие графической модели. Представление правдоподобия в виде графа.
3. Гауссовские процессы и эволюция моделей во времени.
4. Алгоритмы вывода в графических моделях с циклами.
5. Построение адекватных мультимоделей.
6. Неориентированные графические модели.

7. Связь и сравнение неориентированных графических моделей с ориентированными.
8. Методы Монте-Карло по схеме марковских цепей.
9. Алгоритмы вывода на основании разрезов графов и их применение к сегментации изображений.
10. Гамильтоновы методы Монте-Карло по схеме марковских цепей.
11. Ориентированные графические модели и понятие условной отделимости.
12. Байесовская оптимизация.
13. Структурное обучение.
14. Сравнение классического и нейросетевого подхода к графическим моделям.

Примеры билетов:

Билет №1

1. Множественное тестирование гипотез и выбор априорного распределения.
2. Байесовская оптимизация.

Билет №2

1. Байесовская логистическая регрессия и отбор признаков.
2. Методы Монте-Карло по схеме марковских цепей.

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

При проведении дифференцированного зачета и экзамена, обучающемуся предоставляется 30 минут на подготовку. Опрос обучающегося не должен превышать двух астрономических часов.

Во время проведения дифференцированного зачета и экзамена, обучающиеся могут пользоваться программой дисциплины, а также справочной литературой и другими материалами.