

**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:	Байесовское муьтимоделирование
по направлению:	Информатика и вычислительная техника
профиль подготовки:	Прикладная математика и информатика Физтех-школа Прикладной Математики и Информатики кафедра интеллектуальных систем
курс:	1
квалификация:	магистр

Семестры, формы промежуточной аттестации:

1 (осенний) - Дифференцированный зачет

2 (весенний) - Дифференцированный зачет

Аудиторных часов: 105 всего, в том числе:

лекции: 45 час.

семинары: 60 час.

лабораторные занятия: 0 час.

Самостоятельная работа: 75 час.

Всего часов: 180, всего зач. ед.: 4

Программу составил: О.Ю. Бахтеев, канд. физ.-мат. наук

Программа обсуждена на заседании кафедры интеллектуальных систем 03.04.2024

Аннотация

Рассматриваются особенности различных подходов к решению задач выбора, агрегации и композиции моделей как специального, так и общего вида. Задачей курса является, в том числе, получение навыков подготовки докладов по неизвестной студенту теме из области машинного обучения.

1. Цели и задачи

Цель дисциплины

- получение теоретических знаний о методах вероятностного выбора моделей и мультимodelей, с фокусом на методах выбора моделей глубокого обучения.

Задачи дисциплины

- формирование базовых знаний в области теоретических и прикладных аспектов выбора, агрегации и композиции моделей;
- изучение слушателями статистических и алгоритмических основ;
- знакомство с практическими приложениями.

2. Перечень формируемых компетенций

Освоение дисциплины направлено на формирование следующих компетенций:

Код и наименование компетенции	Индикаторы достижения компетенции
УК-1 Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	УК-1.1 Анализирует проблемную ситуацию как систему, выявляя ее составляющие и связи между ними
	УК-1.2 Осуществляет поиск вариантов решения поставленной проблемной ситуации на основе доступных источников информации
	УК-1.3 Разрабатывает стратегию достижения поставленной цели как последовательность шагов, предвидя результат каждого из них и оценивая их влияние на внешнее окружение планируемой деятельности и на взаимоотношения участников этой деятельности
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:

- основные теоретические аспекты задачи выбора моделей и смежных задач, включая задачи оптимизации параметров модели, выбора структуры модели, задачи автоматического построения ансамблей моделей и переноса знаний между моделями.

уметь:

- читать и понимать научные работы, посвященные задачам из области выбора моделей машинного обучения.

владеть:

- основным математическим аппаратом, используемым в области выбора моделей машинного обучения.

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	Байесовский выбор моделей	15	10		20
2	Модели	15	10		20
3	Оптимизация моделей	15	10		20
4	Иерархические модели		10		5
5	Случайные процессы в задаче выбора моделей		10		5
6	Многозадачное обучение		10		5
Итого часов		45	60		75
Подготовка к экзамену		0 час.			
Общая трудоёмкость		180 час., 4 зач.ед.			

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

Семестр: 1 (Осенний)

1. Байесовский выбор моделей

Основные свойства распределений, используемых в Байесовском выводе. Байесовский выбор моделей. Принцип минимальной длины описания.

2. Модели

Вероятностные метрические пространства. Генеративные и дискриминативные модели. Вероятностные графические модели. Вариационный вывод.

3. Оптимизация моделей

Оптимизация гиперпараметров. Метаоптимизация. Гауссовские процессы в задаче выбора модели.

Семестр: 2 (Весенний)

4. Иерархические модели

Иерархические модели. Выбор структуры модели.

5. Случайные процессы в задаче выбора моделей

Случайные процессы в задаче выбора моделей. Метод Монте-Карло в задаче выбора моделей. Выбор ансамбля моделей. Задача дистилляции. Задача переноса информации.

6. Многозадачное обучение

Многозадачное обучение. Построение латентного пространства для моделей. Мультиагентные модели.

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

Учебная аудитория, оснащенная компьютером и мультимедийным оборудованием (проектор, звуковая система).

6. Перечень рекомендуемой литературы

Основная литература

1. Введение в методы байесовского статистического вывода [Текст] = An introduction to bayesian statistical inference for economists/Дж. Хей, -М., Финансы и статистика, 1987

Дополнительная литература

- MacKay D. J. C. et al. Information theory, inference and learning algorithms. – Cambridge university press, 2003.
- Bishop C. M., Nasrabadi N. M. Pattern recognition and machine learning. – New York : springer, 2006. – Т. 4. – №. 4. – С. 738.
- Goodfellow I., Bengio Y., Courville A. Deep learning. – MIT press, 2016.

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

<http://www.machinelearning.ru/wiki/index.php>
<https://intelligent-systems-phystech.github.io/ru/>

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, методы доказательств.

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.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

по направлению:	Информатика и вычислительная техника
профиль подготовки:	Прикладная математика и информатика Физтех-школа Прикладной Математики и Информатики кафедра интеллектуальных систем
курс:	<u>1</u>
квалификация:	магистр

Семестры, формы промежуточной аттестации:

- 1 (осенний) - Дифференцированный зачет
- 2 (весенний) - Дифференцированный зачет

Разработчик: О.Ю. Бахтеев, канд. физ.-мат. наук

1. Компетенции, формируемые в процессе изучения дисциплины

Код и наименование компетенции	Индикаторы достижения компетенции
УК-1 Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	УК-1.1 Анализирует проблемную ситуацию как систему, выявляя ее составляющие и связи между ними
	УК-1.2 Осуществляет поиск вариантов решения поставленной проблемной ситуации на основе доступных источников информации
	УК-1.3 Разрабатывает стратегию достижения поставленной цели как последовательность шагов, предвидя результат каждого из них и оценивая их влияние на внешнее окружение планируемой деятельности и на взаимоотношения участников этой деятельности
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 «Байесовское мультимоделирование» обучающийся должен:

знать:

- основные теоретические аспекты задачи выбора моделей и смежных задач, включая задачи оптимизации параметров модели, выбора структуры модели, задачи автоматического построения ансамблей моделей и переноса знаний между моделями.

уметь:

- читать и понимать научные работы, посвященные задачам из области выбора моделей машинного обучения.

владеть:

- основным математическим аппаратом, используемым в области выбора моделей машинного обучения.

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

Вопросы к дифференцированному зачёту:

1. Оценки центральной тенденции распределений, свойства оценок, полученных методом максимального правдоподобия.
2. Основные свойства распределений, используемых в Байесовском выводе.
3. Байесовский выбор моделей

4. Принцип минимальной длины описания, связь с Колмогоровской сложностью.
5. Вероятностные метрические пространства
6. Генеративные и дискриминативные модели
7. Свойства и взаимосвязь дивергенции Кульбака-Лейблера, расстояния Хеллингера, дисергенции Йенсена-Шеннона.
8. Представление вероятностных графических моделей, plate notation.
9. Регуляризованный автокодировщик как вероятностная модель.
10. Вариационный вывод с применением оператора оптимизации.
11. Оптимизация моделей.
12. Оптимизация гиперпараметров.
13. Метаоптимизация.
14. Гауссовские процессы в задаче выбора модели.
15. Отличие гиперпараметров от метапараметров, методы их оптимизации.
16. Градиентная оптимизация гиперпараметров.
17. Иерархические модели.
18. Выбор структуры модели.
19. Графовое представление структуры модели, распределения на структуре.
20. Сэмплирование по Гиббсу в задаче оптимизации ограниченной машины Больцмана.
21. Дистилляция разнородных моделей.
22. Перенос информации и смещение в наборе данных.
23. Методы представления моделей в векторном пространстве.
24. Случайные процессы в задаче выбора моделей.
25. Метод Монте-Карло в задаче выбора моделей.
26. Выбор ансамбля моделей.
27. Задача дистилляции.
28. Задача переноса информации.
29. Многозадачное обучение.
30. Построение латентного пространства для моделей.
31. Мультиагентные модели.

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 зачета обучающиеся могут пользоваться программой дисциплины, а также справочной литературой, вычислительной техникой, конспектами лекций.

Дифференцированный зачет проводится путем организации специального опроса, проводимого в письменной и устной формах.