

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ**
федеральное государственное автономное образовательное учреждение высшего
образования
«Балтийский федеральный университет имени Иммануила Канта»
Высшая школа компьютерных наук и искусственного интеллекта

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Байесовские методы машинного обучения

Шифр: 02.03.02

**Направление подготовки: Фундаментальная информатика и информационные
технологии**

Профиль: Программная инженерия в искусственном интеллекте

Квалификация (степень) выпускника: бакалавр

Лист согласования

Составитель: Верещагин Михаил Дмитриевич, PhD, директор Высшей школы компьютерных наук и искусственного интеллекта

Рабочая программа утверждена на заседании
Ученого совета ОНК «Институт высоких технологий»

Протокол № 33 от «27» октября 2023 г.

Председатель Ученого совета ОНК «Институт высоких технологий»

Профессор, д.ф.-м.н.

А.В. Юров

Директор высшей школы компьютерных наук
и искусственного интеллекта

М.Д. Верещагин

Руководитель ОПОП ВО

С.С. Головин

Содержание

1. Наименование дисциплины «Байесовские методы машинного обучения».
2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы.
3. Место дисциплины в структуре образовательной программы.
4. Виды учебной работы по дисциплине.
5. Содержание дисциплины, в том числе практической подготовки в рамках дисциплины, структурированное по темам.
6. Рекомендуемая тематика учебных занятий в форме контактной работы.
7. Методические рекомендации по видам занятий
8. Фонд оценочных средств
 - 8.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы в рамках учебной дисциплины
 - 8.2. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений и навыков и (или) опыта деятельности процессе текущего контроля
 - 8.3. Перечень вопросов и заданий для промежуточной аттестации по дисциплине
 - 8.4. Планируемые уровни сформированности компетенций обучающихся и критерии оценивания
9. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины
11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине.
12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

1. Наименование дисциплины: «Байесовские методы машинного обучения».

Целью курса «Байесовские методы машинного обучения» является обучение основам байесовского подхода и его приложениям в иерархических моделях, линейных и обобщенных линейных моделях, смешанных моделях и различных типах принятия решений. Студентам предложено изучить основы байесовского метода, а также современные приложения этого подхода с помощью языка R и библиотек семплирования JAGS и STAN.

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

Код и содержание компетенции.	Результаты освоение образовательной программы (ИДК)	Результаты обучения по дисциплине
ПК-1. Способен классифицировать и идентифицировать задачи искусственного интеллекта, выбирать адекватные методы и инструментальные средства решения задач искусственного интеллекта	ПК-1.1. Классифицирует и идентифицирует задачи систем искусственного интеллекта в зависимости от особенностей проблемной и предметной областей ПК-1.2. Выбирает методы и инструментальные средства искусственного интеллекта для решения задач в зависимости от особенностей проблемной и предметной областей ПК-1.3. Собирает исходную информацию и формирует требования к решению задач с использованием методов искусственного интеллекта	Знать: <ul style="list-style-type: none">• основные понятия, концепции и проблемы байесовского анализа и его приложений в различных моделях, основные методы семплирования из апостериорного распределения; Уметь: <ul style="list-style-type: none">• выбирать подходящие под конкретную задачу априорные распределения для латентных переменных, применять на практике методы семплирования латентных переменных, делать статистические выводы на основе выборки из апостериорного распределения и интерпретировать полученные результаты; Владеть: <ul style="list-style-type: none">• методами сбора и подготовки данных с помощью пакетов программ на языке R, навыками проведения байесовского анализа в среде R с помощью

		библиотек JAGS и STAN, методами проверки адекватности выборки из апостериорного распределения.
--	--	--

3. Место дисциплины в структуре образовательной программы

«Байесовские методы машинного обучения» представляет собой дисциплину Части, формируемой участниками образовательных отношений (Б1.В.ДВ.04) направления подготовки бакалавриата 02.03.02 «Фундаментальная информатика и информационные технологии», профиль «Программная инженерия в искусственном интеллекте».

4. Виды учебной работы по дисциплине.

Виды учебной работы по дисциплине зафиксированы учебным планом основной профессиональной образовательной программы по указанному направлению и профилю, выражаются в академических часах. Часы контактной работы и самостоятельной работы обучающегося и часы, отводимые на процедуры контроля, могут различаться в учебных планах ОПОП по формам обучения. Объем контактной работы включает часы контактной аудиторной работы (лекции/практические занятия/ лабораторные работы), контактной внеаудиторной работы (контроль самостоятельной работы), часы контактной работы в период аттестации. Контактная работа, в том числе может проводиться посредством электронной информационно-образовательной среды университета с использованием ресурсов сети Интернет и дистанционных технологий.

5. Содержание дисциплины, структурированное по темам (разделам)

Исходя из рамок, установленных учебным планом по трудоемкости и видам учебной работы по дисциплине, преподаватель самостоятельно выбирает тематику занятий по формам и количеству часов проведения контактной работы: лекции и иные учебные занятия, предусматривающие преимущественную передачу учебной информации преподавателем и (или) занятия семинарского типа (семинары, практические занятия, практикумы, лабораторные работы, коллоквиумы и иные аналогичные занятия), и (или) групповые консультации, и (или) индивидуальную работу обучающихся с преподавателем, в том числе индивидуальные консультации (по курсовым работам/проектам – при наличии курсовой работы/проекта по данной дисциплине в учебном плане). Рекомендуемая тематика занятий максимально полно реализуется в контактной работе с обучающимися очной формы обучения. В случае реализации образовательной программы в заочной / очно-заочной форме трудоемкость дисциплины сохраняется, однако объем учебного материала в значительной части осваивается обучающимися в форме самостоятельной работы. При этом требования к ожидаемым образовательным результатам обучающихся по данной дисциплине не зависят от формы реализации образовательной программы

№ п/п	Наименование раздела	Содержание раздела
1	Введение в курс. Байесовские рассуждения.	Байесовский подход к теории вероятностей. Оценка параметров в байесовском и частотном подходе.

		Примеры байесовских рассуждений.
2	Сопряжённые распределения, аналитический байесовский вывод	Сопряжённые распределения. Примеры. Экспоненциальный класс распределений, его свойства.
3	Байесовский выбор модели	Решение задачи выбора модели по Байесу. Обоснованность модели. Полный байесовский вывод.
4	Байесовская проверка гипотез	Понятие и задача оптимизации параметров байесовской процедуры проверки многоальтернативных гипотез с их предварительной иерархической селекцией.
5	Метод релевантных векторов для задачи регрессии	Вероятностная модель линейной регрессии. Метод релевантных векторов для задачи регрессии.
6	Метод релевантных векторов для задачи классификации	Логистическая регрессия. Метод релевантных векторов для задачи классификации.
7	EM-алгоритм. Байесовский метод главных компонент	EM-алгоритм в общем виде. Примеры применения.
8	Вариационный вывод	Вариационный подход для приближенного байесовского вывода.
9	Байесовская модель разделения гауссиан	Вариационная линейная регрессия. Задача уменьшения размерности в данных. Вероятностная модель главных компонент, ее обучение с помощью метода максимального правдоподобия и EM-алгоритма.
10	Латентное размещение Дирихле (LDA)	Распределение Дирихле. Свойства накопления и нейтральности. Генерация выборки из Дирихле через гамма-распределения и через stick-breaking.
11	Методы Монте Карло по схеме марковский цепей (MCMC)	Байесовская модель разделения смеси гауссиан. Вариационный вывод для неё. Тематическая модель LDA. Обучение и вывод в модели. Методы MCMC для оценки статистик вероятностных распределений. Теоретические свойства марковских цепей. Схема Метрополиса-Хастингса и схема Гиббса. Примеры использования.
12	Гауссовские процессы для регрессии и классификации	Гауссовские процессы для задачи регрессии. Подбор параметров ковариационной функции. Гауссовские процессы для задачи классификации.
13	Непараметрические байесовские методы. Процессы Дирихле	Процессы Дирихле. Представление процесса Дирихле с помощью процесса китайского ресторана. Схема Гиббса для разделения смеси распределений с процессом Дирихле. Процессы Дирихле. Представление процесса Дирихле с помощью stick-breaking. Вариационный вывод для разделения смеси распределений с процессом Дирихле.

6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

Рекомендуемая тематика учебных занятий *лекционного* типа (предусматривающих

преимущественную передачу учебной информации преподавателями):

№ п/п	Наименование раздела	Тема лекции
1	Введение в курс. Байесовские рассуждения.	Лекция 1 Введение в курс. Байесовские рассуждения.
2	Сопряжённые распределения, аналитический байесовский вывод	Лекция 2 Сопряжённые распределения, аналитический байесовский вывод
3	Байесовский выбор модели	Лекция 3 Байесовский выбор модели
4	Байесовская проверка гипотез	Лекция 4 Байесовская проверка гипотез
5	Метод релевантных векторов для задачи регрессии	Лекция 5 Метод релевантных векторов для задачи регрессии
6	Метод релевантных векторов для задачи классификации	Лекция 6 Метод релевантных векторов для задачи классификации
7	EM-алгоритм. Байесовский метод главных компонент	Лекция 7 EM-алгоритм. Байесовский метод главных компонент
8	Вариационный вывод	Лекция 8 Вариационный вывод
9	Байесовская модель разделения гауссиан	Лекция 9 Байесовская модель разделения гауссиан
10	Латентное размещение Дирихле (LDA)	Лекция 10 Латентное размещение Дирихле (LDA)
11	Методы Монте Карло по схеме марковский цепей (MCMC)	Лекция 11 Методы Монте Карло по схеме марковский цепей (MCMC)
12	Гауссовские процессы для регрессии и классификации	Лекция 12 Гауссовские процессы для регрессии и классификации
13	Непараметрические байесовские методы. Процессы Дирихле	Лекция 13 Непараметрические байесовские методы. Процессы Дирихле

Рекомендуемая тематика практических занятий:

1. Точечные операции обработки изображений
2. Алгебраические и геометрические операции над изображениями
3. Морфологические преобразования изображений
4. Генерация признаков на основе линейных преобразований
5. Дискретное преобразование Фурье
6. Вейвлет-преобразование изображений
7. Генерация признаков формы на основе анализа границ изображения
8. Генерация признаков формы на основе построения и анализа скелетов изображения
9. Нейронные сети и распознавание изображений

Требования к самостоятельной работе обучающихся

1. Работа с лекционным материалом, предусматривающая проработку конспекта лекций и учебной литературы, по всем темам из п. 6 настоящей рабочей программы.
2. Выполнение домашнего задания, предусматривающего решение задач, выполнение упражнений, выдаваемых на практических занятиях, по всем темам из п. 6 настоящей рабочей программы.

Руководствуясь положениями статьи 47 и статьи 48 Федерального закона от 29

декабря 2012 г. N 273-ФЗ «Об образовании в Российской Федерации» научно-педагогические работники и иные лица, привлекаемые университетом к реализации данной образовательной программы, пользуются предоставленными академическими правами и свободами в части свободы преподавания, свободы от вмешательства в профессиональную деятельность; свободы выбора и использования педагогически обоснованных форм, средств, методов обучения и воспитания; права на творческую инициативу, разработку и применение авторских программ и методов обучения и воспитания в пределах реализуемой образовательной программы и отдельной дисциплины.

Исходя из рамок, установленных учебным планом по трудоемкости и видам учебной работы по дисциплине, преподаватель самостоятельно выбирает тематику занятий по формам и количеству часов проведения контактной работы: лекции и иные учебные занятия, предусматривающие преимущественную передачу учебной информации преподавателем и (или) занятия семинарского типа (семинары, практические занятия, практикумы, лабораторные работы, коллоквиумы и иные аналогичные занятия), и (или) групповые консультации, и (или) индивидуальную работу обучающихся с преподавателем, в том числе индивидуальные консультации (по курсовым работам/проектам – при наличии курсовой работы/проекта по данной дисциплине в учебном плане).

Рекомендуемая тематика занятий максимально полно реализуется в контактной работе с обучающимися очной формы обучения. В случае реализации образовательной программы в заочной / очно-заочной форме трудоемкость дисциплины сохраняется, однако объем учебного материала в значительной части осваивается обучающимися в форме самостоятельной работы. При этом требования к ожидаемым образовательным результатам обучающихся по данной дисциплине не зависят от формы реализации образовательной программы.

7. Методические рекомендации по видам занятий

Лекционные занятия.

В ходе лекционных занятий обучающимся рекомендуется выполнять следующие действия. Вести конспектирование учебного материала. Обращать внимание на категории, формулировки, раскрывающие содержание тех или иных явлений и процессов, научные выводы и практические рекомендации по их применению. Задавать преподавателю уточняющие вопросы с целью уяснения теоретических положений, разрешения спорных ситуаций.

Желательно оставить в рабочих конспектах поля, на которых во внеаудиторное время можно сделать пометки из рекомендованной литературы, дополняющие материал прослушанной лекции, а также подчеркивающие особую важность тех или иных теоретических положений.

Практические и семинарские занятия.

На практических и семинарских занятиях в зависимости от темы занятия выполняется поиск информации по решению проблем, практические упражнения, контрольные работы, выработка индивидуальных или групповых решений, итоговое обсуждение с обменом знаниями, участие в круглых столах, разбор конкретных ситуаций, командная работа, представление портфолио и т.п.

Самостоятельная работа.

Самостоятельная работа осуществляется в виде изучения литературы, эмпирических данных по публикациям и конкретных ситуаций из практики, подготовке индивидуальных работ, работа с лекционным материалом, самостоятельное изучение отдельных тем дисциплины; поиск и обзор литературы и электронных источников; чтение и изучение учебника и учебных пособий.

8. Фонд оценочных средств

8.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы в рамках учебной дисциплины

Основными этапами формирования указанных компетенций при изучении обучающимися дисциплины являются последовательное изучение содержательно связанных между собой тем учебных занятий. Изучение каждой темы предполагает овладение обучающимися необходимыми компетенциями. Результат аттестации обучающихся на различных этапах формирования компетенций показывает уровень освоения компетенций.

Контролируемые разделы (темы) дисциплины	Индекс контролируемой компетенции (или её части)	Оценочные средства по этапам формирования компетенций
		текущий контроль по дисциплине
Введение в курс. Байесовские рассуждения.	ПК-1	Лабораторная работа
Сопряжённые распределения, аналитический байесовский вывод	ПК-1	Лабораторная работа
Байесовский выбор модели	ПК-1	Лабораторная работа
Байесовская проверка гипотез	ПК-1	Тесты
Метод релевантных векторов для задачи регрессии	ПК-1	Лабораторная работа
Метод релевантных векторов для задачи классификации	ПК-1	Лабораторная работа
EM-алгоритм. Байесовский метод главных компонент	ПК-1	Лабораторная работа
Вариационный вывод	ПК-1	Тесты
Байесовская модель разделения гауссиан	ПК-1	Лабораторная работа
Латентное размещение Дирихле (LDA)	ПК-1	Лабораторная работа
Методы Монте Карло по схеме марковский цепей (MCMC)	ПК-1	Лабораторная работа
Гауссовские процессы для регрессии и классификации	ПК-1	Лабораторная работа
Непараметрические байесовские методы. Процессы Дирихле	ПК-1	Тесты

8.2. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений и навыков и (или) опыта деятельности в процессе текущего контроля

Выполнение заданий на практических (семинарски) занятиях

Примеры задач:

Рассмотрим модель посещаемости студентами одного курса лекции. Пусть аудитория данного курса состоит из студентов профильной кафедры, а также студентов других кафедр. Обозначим через a количество студентов, распределившихся на профильную кафедру, а через b — количество студентов других кафедр на курсе. Пусть студенты профильной кафедры посещают курс с некоторой вероятностью p_1 , а студенты остальных кафедр — с вероятностью p_2 . Обозначим через c количество студентов на данной лекции. Тогда случайная величина $c|a, b$ есть сумма двух случайных величин, распределенных по биномиальному закону $B(a, p_1)$ и $B(b, p_2)$ соответственно. Пусть далее на лекции по курсу ведется запись студентов. При этом каждый студент записывается сам, а также, быть может, записывает своего товарища, которого на лекции на самом деле нет. Пусть студент записывает своего товарища с некоторой вероятностью p_3 . Обозначим через d общее количество записавшихся на данной лекции. Тогда случайная величина $d|c$ представляет собой сумму c и случайной величины, распределенной по биномиальному закону $B(c, p_3)$. Для завершения задания вероятностной модели осталось определить априорные вероятности для a и для b . Пусть обе эти величины распределены равномерно в своих интервалах $[a_{\min}, a_{\max}]$ и $[b_{\min}, b_{\max}]$. Таким образом, мы определили следующую вероятностную модель:

Модель 1

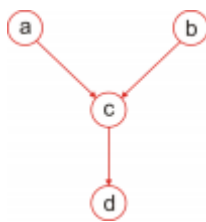
$$p(a, b, c, d) = p(d|c)p(c|a, b)p(a)p(b),$$

$$d|c \sim c + B(c, p_3),$$

$$c|a, b \sim B(a, p_1) + B(b, p_2),$$

$$a \sim R[a_{\min}, a_{\max}],$$

$$b \sim R[b_{\min}, b_{\max}].$$



Графическая модель для вероятностной модели 1

Рассмотрим несколько упрощенную версию модели 1. Известно, что биномиальное распределение $B(n, p)$ при большом количестве испытаний и маленькой вероятности успеха может быть с высокой точностью приближено пуассоновским распределением $Poiss(\lambda)$ с $\lambda = np$. Известно также, что сумма двух пуассоновских распределений с параметрами λ_1 и λ_2 есть пуассоновское распределение с параметром $\lambda_1 + \lambda_2$. Таким образом, мы можем сформулировать вероятностную модель, которая является приближенной версией модели 1:

Модель

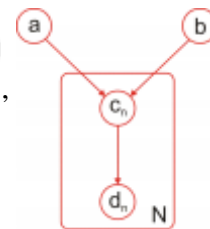
2

$$\begin{aligned}
p(a,b,c,d) &= p(d|c)p(c|a,b)p(a)p(b), \\
d|c &\sim c + B(c, p_3), \\
c|a,b &\sim \text{Pois}(ap_1 + bp_2), \\
a &\sim R[a_{\min}, a_{\max}], \\
b &\sim R[b_{\min}, b_{\max}].
\end{aligned}$$

Рассмотрим теперь модель посещаемости нескольких лекций курса. Будем считать, что посещаемости отдельных лекций являются независимыми. Тогда:
Модель 3

$$p(a,b,c_1, \dots, c_N, d_1, \dots, d_N) = \prod_{n=1}^N p(d_n|c_n)p(c_n|a,b)p(a)p(b)$$

$$\begin{aligned}
d_n|c_n &\sim c_n + B(c_n, p_3), \\
c_n|a,b &\sim B(a, p_1) + B(b, p_2), \\
a &\sim R[a_{\min}, a_{\max}], \\
b &\sim R[b_{\min}, b_{\max}].
\end{aligned}$$



Графическая модель для вероятностной модели 3

По аналогии с моделью 2 можно сформулировать упрощенную модель для модели 3:
Модель 4

$$p(a,b,c_1, \dots, c_N, d_1, \dots, d_N) = \prod_{n=1}^N p(d_n|c_n)p(c_n|a,b)p(a)p(b)$$

$$\begin{aligned}
d_n|c_n &\sim c_n + B(c_n, p_3), \\
c_n|a,b &\sim \text{Pois}(ap_1 + bp_2), \\
a &\sim R[a_{\min}, a_{\max}], \\
b &\sim R[b_{\min}, b_{\max}].
\end{aligned}$$

Вариант 1

Рассматривается модель с параметрами $a_{\min} = 15, a_{\max} = 30, b_{\min} = 250, b_{\max} = 350, p_1 = 0.5, p_2 = 0.05, p_3 = 0.5$. Провести на компьютере следующие исследования:

1. Найти математические ожидания и дисперсии априорных распределений для всех параметров a, b, c, d .
2. Пронаблюдать, как происходит уточнение прогноза для величины c по мере прихода новой косвенной информации. Для этого построить графики и найти мат.ожидание и дисперсию

для распределений $p(c), p(c|b), p(c|a,b), p(c|a,b,d)$ при параметрах a, b, d , равных мат. ожиданиям своих априорных распределений, округленных до ближайшего целого.

3. Определить, какая из величин a, b, d вносит больший вклад в уточнение прогноза для величины c (в смысле дисперсии распределения). Для этого убедиться в том, что $D[c|d] < D[c|b]$ и $D[c|d] < D[c|a]$ для любых допустимых значений a, b, d . Найти множество точек (a, b) таких, что $D[c|b] < D[c|a]$. Являются ли множества $\{(a, b) | D[c|b] < D[c|a]\}$ и $\{(a, b) | D[c|b] \geq D[c|a]\}$ линейно разделимыми?
4. Провести временные замеры по оценке всех необходимых распределений $p(c), p(c|a), p(c|b), p(c|d), p(c|a,b), p(c|a,b,d), p(d)$.
5. Провести исследования из пп. 1-4 для точной модели 1 и сравнить результаты с аналогичными для модели 2. Привести пример оценки параметра, в котором разница между моделью 1 и 2 проявляется в большой степени.

Взять в качестве диапазона допустимых значений для величины c интервал $[0, a_{max} + b_{max}]$, а для величины d — интервал $[0, 2^*(a_{max} + b_{max})]$

При оценке выполнения задания будет учитываться эффективность программного кода. В частности, временные затраты на расчет отдельного распределения не должны превышать одной секунды.

Вариант 2

Рассматривается модель 2 параметрами $a_{min} = 15, a_{max} = 30, b_{min} = 250, b_{max} = 350, p_1 = 0.5, p_2 = 0.05, p_3 = 0.5$. Провести на компьютере следующие исследования:

1. Найти математические ожидания и дисперсии априорных распределений для всех параметров a, b, c, d .
2. Пронаблюдать, как происходит уточнение прогноза для величины b по мере прихода новой косвенной информации. Для этого построить графики и найти мат. ожидание и дисперсию для распределений $p(b), p(b|a), p(b|a,d)$ при параметрах a, d , равных мат. ожиданиям своих априорных распределений, округленных до ближайшего целого.
3. Определить, при каких соотношениях параметров p_1, p_2 изменяется относительная важность параметров a, b для оценки величины c . Для этого найти множество точек $\{(p_1, p_2) | D[c|b] < D[c|a]\}$ при a, b , равных мат. ожиданиям своих априорных распределений, округленных до ближайшего целого. Являются ли множества $\{(p_1, p_2) | D[c|b] < D[c|a]\}$ и $\{(p_1, p_2) | D[c|b] \geq D[c|a]\}$ линейно разделимыми?
4. Провести временные замеры по оценке всех необходимых распределений $p(c), p(c|a), p(c|b), p(b|a), p(b|a,d), p(d)$.
5. Провести исследования из пп. 1-4 для точной модели 1 и сравнить результаты с аналогичными для модели 2. Привести пример оценки параметра, в котором разница между моделью 1 и 2 проявляется в большой степени.

Взять в качестве диапазона допустимых значений для величины c интервал $[0, a_{max} + b_{max}]$, а

для величины d — интервал $[0, 2^*(a_{max} + b_{max})]$

При оценке выполнения задания будет учитываться эффективность программного кода. В частности, временные затраты на расчет отдельного распределения не должны превышать одной секунды.

Вариант 3

Рассматривается модель с параметрами $a_{min} = 15, a_{max} = 30, b_{min} = 250, b_{max} = 350, p_1 = 0.5, p_2 = 0.05, p_3 = 0.5, N = 50$.

Провести на компьютере следующие исследования:

1. Найти математические ожидания и дисперсии априорных распределений для всех параметров a, b, c_n, d_n .
2. Реализовать генератор выборки d_1, \dots, d_N из модели при заданных значениях параметров a, b .
3. Пронаблюдать, как происходит уточнение прогноза для величины b по мере прихода новой косвенной информации. Для этого построить графики и найти мат.ожидание и дисперсию для распределений $p(b), p(b|d_1), \dots, p(b|d_1, \dots, d_N)$, где выборка d_1, \dots, d_N сгенерирована из модели при параметрах a, b , равных мат.ожиданиям своих априорных распределений, округленных до ближайшего целого и 2) $d_1 = \dots = d_N$, где d_n равно мат.ожиданию своего априорного распределения, округленного до ближайшего целого. Провести аналогичный эксперимент, если дополнительно известно значение a . Сравнить результаты двух экспериментов.
4. Провести временные замеры по оценке всех необходимых распределений $p(c_n), p(d_n), p(b|d_1, \dots, d_n), p(b|a, d_1, \dots, d_n)$.
5. Провести исследования из пп. 1-4 для точной модели 3 и сравнить результаты с аналогичными для модели 4.

Взять в качестве диапазона допустимых значений для величины c интервал $[0, a_{max} + b_{max}]$, а для величины d — интервал $[0, 2^*(a_{max} + b_{max})]$

При оценке выполнения задания будет учитываться эффективность программного кода. В частности, временные затраты на расчет отдельного распределения не должны превышать одной секунды.

8.3. Перечень вопросов и заданий для промежуточной аттестации по дисциплине

Вопросы для итогового контроля

Примерный перечень вопросов к экзамену:

1. Байесовский подход к теории вероятностей. Оценка параметров в байесовском и частотном подходе. Примеры байесовских рассуждений.
2. Сопряжённые распределения. Примеры. Экспоненциальный класс распределений, его свойства.

3. Решение задачи выбора модели по Байесу. Обоснованность модели. Полный байесовский вывод.
4. Вероятностная модель линейной регрессии. Метод релевантных векторов для задачи регрессии.
5. Логистическая регрессия. Метод релевантных векторов для задачи классификации.
6. EM-алгоритм в общем виде. Примеры применения.
7. Вариационный подход для приближенного байесовского вывода.
8. Вариационная линейная регрессия.
9. Задача уменьшения размерности в данных. Вероятностная модель главных компонент, ее обучение с помощью метода максимального правдоподобия и EM-алгоритма.
10. Распределение Дирихле. Свойства накопления и нейтральности. Генерация выборки из Дирихле через гамма-распределения и через stick-breaking.
11. Байесовская модель разделения смеси гауссиан. Вариационный вывод для неё.
12. Тематическая модель LDA. Обучение и вывод в модели.
13. Методы MCMC для оценки статистик вероятностных распределений. Теоретические свойства марковских цепей.
14. Схема Метрополиса-Хастингса и схема Гиббса. Примеры использования.
15. Гауссовские процессы для задачи регрессии. Подбор параметров ковариационной функции.
16. Гауссовские процессы для задачи классификации.
17. Процессы Дирихле. Представление процесса Дирихле с помощью процесса китайского ресторана. Схема Гиббса для разделения смеси распределений с процессом Дирихле.
18. Процессы Дирихле. Представление процесса Дирихле с помощью stick-breaking. Вариационный вывод для разделения смеси распределений с процессом Дирихле.

Пример экзаменационного билета:

1. Вероятностная модель линейной регрессии. Метод релевантных векторов для задачи регрессии.
2. Схема Метрополиса-Хастингса и схема Гиббса. Примеры использования.

8.4. Планируемые уровни сформированности компетенций обучающихся и критерии оценивания

Уровни	Содержательно е описание уровня	Основные признаки выделения уровня (этапы формирования компетенции, критерии оценки сформированнос ти)	Пятибалльная шкала (академическая) оценка	Двухб а льна я шкал а, зачет	БРС, % освоения (рейтингов ая оценка)
Повышенны й	Творческая деятельность	Включает нижестоящий	отлично	зачтен о	86-100

		уровень. Умение самостоятельно принимать решение, решать проблему/задачу теоретического и прикладного характера на основе изученных методов, приемов, технологий			
Базовый	Применение знаний и умений в более широких контекстах учебной и профессиональной деятельности, нежели по образцу с большей степени самостоятельности и инициативы	Включает нижестоящий уровень. Способность собирать, систематизировать, анализировать и грамотно использовать информацию из самостоятельно найденных теоретических источников и иллюстрировать ими теоретические положения или обосновывать пр	хорошо		71-85
Удовлетворительный (достаточный)	Репродуктивная деятельность	Изложение в пределах задач курса теоретически и практически контролируемого материала	удовлетворительно		55-70
Недостаточный	Отсутствие признаков удовлетворительного уровня		Не удовлетворительно	Не зачтено	Менее 55

9. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины.

Основная литература

1. Гупал, В. М. Методы распознавания сложных систем. Байесовская процедура - оптимальная процедура распознавания : монография / В. М. Гупал. - Москва : Компания Спутник+, 2005. - 78 с. - Текст : электронный. - URL: <https://znanium.com/catalog/product/358812> (дата обращения: 18.11.2023). – Режим доступа: по подписке.
2. Дауни А. Б. Байесовские модели / пер. с англ. В. А. Яроцкого. / А.Б. Дауни. - Москва : ДМК Пресс, 2018. - 182 с. - ISBN 978-5-97060-664-3. - URL: <https://ibooks.ru/bookshelf/385149/reading> (дата обращения: 18.11.2023). - Текст: электронный.

Дополнительная литература

1. Кельберт, М. Я. Вероятность и статистика в примерах и задачах. Том I. Основные понятия теории вероятностей и математической статистики: Учебное пособие / М. Я. Кельберт, Ю. М. Сухов, - 2-е изд. - :, 2017. - 486 с.: ISBN 978-5-4439-2326-0. - Текст : электронный. - URL: <https://znanium.com/catalog/product/958605> (дата обращения: 18.11.2023). – Режим доступа: по подписке.

10. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины (модуля).

19. НЭБ Национальная электронная библиотека, диссертации и прочие издания
20. eLIBRARY.RU Научная электронная библиотека, книги, статьи, тезисы докладов конференций
21. Гребенников Электронная библиотека ИД журналы
22. ЭБС Консультант студента
23. ПРОСПЕКТ ЭБС
24. ЭБС ZNANIUM.COM
25. ЭБС IBOOKS.RU
26. РГБ Информационное обслуживание по МБА
27. БЕН РАН
28. Электронно-библиотечная система (ЭБС) Кантиана (<https://elib.kantiana.ru/>)

11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине.

Программное обеспечение обучения включает в себя:

6. система электронного образовательного контента БФУ им. И. Канта – www.lms.kantiana.ru, обеспечивающую разработку и комплексное использование электронных образовательных ресурсов;
7. серверное программное обеспечение, необходимое для функционирования сервера и связи с системой электронного обучения через Интернет;
8. корпоративная платформа;
9. установленное на рабочих местах обучающихся ПО: Microsoft Windows 7, Microsoft

Office Standart 2010, антивирусное программное обеспечение Kaspersky Endpoint Security.

10. Visual Studio Community Интегрированная среда разработки ПО. Свободно-распространяемое ПО
11. PyCharm Community Интегрированная среда разработки ПО. Свободно-распространяемое ПО
12. Anaconda Интегрированная среда разработки ПО. Свободно-распространяемое ПО

12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине.

Для проведения занятий лекционного типа, практических и семинарских занятий используются специальные помещения (учебные аудитории), оборудованные техническими средствами обучения – мультимедийной проекционной техникой. Для проведения занятий лекционного типа используются наборы демонстрационного оборудования.

Для проведения лабораторных работ, (практических занятий – при необходимости) используются специальные помещения (учебные аудитории), оснащенные специализированным лабораторным оборудованием: персональными компьютерами с возможностью выхода в интернет и с установленным программным обеспечением, заявленным в п.11.

Для проведения групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации используются специальные помещения (учебные аудитории), оборудованные специализированной мебелью (для обучающихся), меловой / маркерной доской.

Для организации самостоятельной работы обучающимся предоставляются помещения, оснащенные компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду университета.

Для обучения инвалидов и лиц с ограниченными возможностями здоровья университетом могут быть представлены специализированные средства обучения, в том числе технические средства коллективного и индивидуального пользования.