

**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 прикладной математики

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

«Практикум по методам машинного обучения»

Шифр: 01.03.02

Направление подготовки: «Прикладная математика и информатика»

Профиль: «Искусственный интеллект и анализ данных»

Квалификация (степень) выпускника: бакалавр

Лист согласования

Составители:

1. Верещагин Сергей Дмитриевич, к. ф.-м.н., доцент
2. Верещагин Михаил Дмитриевич, к. ф.-м.н., доцент
3. Мищук Богдан Ростиславович, к. ф.-м.н., доцент

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

Протокол № 4 от «24» января 2023 г.

Председатель Ученого совета ОНК
«Институт высоких технологий»

Профессор, д.ф.-м.н.

А.В. Юров

Руководитель ОПОП ВО

Е.П. Ставицкая

Содержание

1. Наименование дисциплины «Практикум по методам 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 в 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: «Глубокое машинное обучение».

Целью курса «Практикум по методам машинного обучения» - сформировать у обучающихся практические навыки работы с использованием методов машинного обучения.

В процессе обучения используется язык программирования Python, интерактивная среда разработки Jupiter, программные библиотеки для машинного обучения scikit-learn и другие. Машинное обучение (Machine Learning) — обширный подраздел искусственного интеллекта, изучающий методы построения алгоритмов, способных обучаться. Машинное обучение является основным современным подходом к анализу данных и построению интеллектуальных информационных систем. Методы машинного обучения лежат в основе всех методов компьютерного зрения, активно используются в обработке изображений. В курсе множество практически применимых алгоритмов.

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

| Код компетенции | Результаты освоения образовательной программы (ИДК) | Результаты обучения по дисциплине |
|--------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ПК-5. Способен использовать инструментальные средства для решения задач машинного обучения | ПК-5.1. Осуществляет оценку и выбор инструментальных средств для решения поставленной задачи ПК-5.2. Разрабатывает модели машинного обучения для решения задач ПК-5.3. Создает, поддерживает и использует системы искусственного интеллекта, включающие разработанные модели и методы, с применением выбранных инструментов машинного обучения | Знать основные современные математические модели и алгоритмы машинного обучения Уметь применять современные методы машинного обучения для решения прикладных задач анализа и обработки данных Владеть навыками реализации программных решений прикладных задач анализа и обработки данных на языке Python с использованием библиотек машинного обучения |

3. Место дисциплины в структуре образовательной программы

«Практикум по методам машинного обучения» представляет собой дисциплину Части, формируемой участниками образовательных отношений (Б1.В.07.04) направления подготовки бакалавриата 01.03.02 «Прикладная математика и информатика», профиль «Искусственный интеллект и анализ данных».

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 разделов (тем) дисциплины | Содержание разделов (тем) дисциплин |
|-------|---------------------------------------------|-------------------------------------------------------------------------------------------------|
| 1. | Тема 1. Основные методы машинного обучения. | Метрические методы. Линейные модели. Метод опорных векторов. Решающие деревья. Бустинг. |
| 2. | Тема 2. Градиентный бустинг. | Нейронные сети. Отбор признаков. Поиск закономерностей и аномалий в данных. Методы без учителя. |

6. Рекомендуемая тематика учебных занятий в форме контактной работы

Рекомендуемая тематика учебных занятий лекционного типа (предусматривающих преимущественную передачу учебной информации преподавателями):

Не предусмотрены учебным планом

Рекомендуемая тематика практических занятий:

| № п/п | Наименование разделов (тем) дисциплины | Содержание разделов (тем) дисциплин |
|-------|----------------------------------------|-------------------------------------|
|-------|----------------------------------------|-------------------------------------|

| | | |
|----|---------------------------------------------|----------------------------------------------------------------------------------------------------------|
| 3. | Тема 1. Основные методы машинного обучения. | Метрические методы. Линейные модели. Метод опорных векторов. Решающие деревья. Бустинг. |
| 4. | Тема 2. Градиентный бустинг. | Нейронные сети. Отбор признаков. Поиск закономерностей и аномалий в данных. Методы без учителя. |

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. № 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 |
| Тема 1. Основные методы машинного обучения. | ПК-5. | Решение задач |
| Тема 2. Градиентный бустинг. | ПК-5. | Решение задач |

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

Текущий контроль успеваемости осуществляется путем оценки результатов выполнения заданий практических (семинарских) занятий, самостоятельной работы, предусмотренных учебным планом и посещения занятий/активность на занятиях.

В качестве оценочных средств текущего контроля успеваемости предусмотрены:

решение индивидуальных заданий

Примеры заданий

- 1) Применить алгоритмы линейной регрессии, логистической регрессии и метода опорных векторов для модельных задач.
- 2) Реализовать алгоритмы анализа данных на основе композиции классификаторов и отбора признаков
- 3) Реализовать алгоритм обратного распространения ошибки и обучение нейросети на задаче распознавания рукописных цифр MNIST.
- 4) Решить задачу предсказания опасных событий для страховой компании
- 5) Решить задачу на поиск аномалий в данных
- 6) Показатель X в классах $K1$ и $K2$ распределён нормально с параметрами: в $K1$ математическое ожидание 2, стандартное отклонение 4; в $K2$ математическое ожидание 3, стандартное отклонение 1. Выделить на числовой оси значений показателя X области отнесения байесовским классификатором к классам $K1$ и $K2$. Априорные вероятности классов $K1$ и $K2$ равны 0.6 и 0.4 соответственно.
- 7) Каждый год варан подрастает на $A\%$ от своего веса в начале года. A – случайная величина с известными матожиданием 5 и дисперсией 1 (одна и та же для всех варанов во все годы). В начале жизни каждый варан имеет вес 1. Построить байесовский классификатор для определения возраста варана (полных лет) по его весу, минимизирующий частоту ошибки. Предположить, что распознаваться будут «достаточно» взрослые вараны.
- 8) Выборка объектов из класса 1 и класса 2 определяется таблицами ниже. Указать тупиковые тесты.

| | X1 | X2 | X3 | X4 | | | | | | X1 | X2 | X3 | X4 | |
|------|----|----|----|----|--|--|--|--|--|------|----|----|----|---|
| Об.1 | 0 | 1 | 1 | 0 | | | | | | Об.1 | 0 | 1 | 0 | 0 |
| Об.2 | 0 | 0 | 1 | 1 | | | | | | Об.2 | 1 | 0 | 1 | 0 |
| Об.3 | 1 | 0 | 0 | 1 | | | | | | Об.3 | 0 | 1 | 0 | 0 |
| Об.4 | 1 | 0 | 1 | 1 | | | | | | Об.4 | 1 | 1 | 0 | 0 |

- 9) Тестирование в банке системы распознавания для определения недобросовестных заёмщиков выявило связь между чувствительностью и ложной тревогой, показанную в таблице. Определить, приведёт ли эксплуатация системы к увеличению доходов банка. Определить возможный прирост дохода в расчёте на одну поданную заявку. Известно, что доход банка на одного заёмщика составляет 40000 денежных единиц, потери в результате отказа заёмщика от платежей составляют 120000 единиц. Доля недобросовестных заёмщиков составляет 7%.

| Чувст. | Лож. Тр. |
|--------|----------|
| 0.02 | 0.0001 |
| 0.12 | 0.003 |

| | |
|------|------|
| 0.23 | 0.05 |
| 0.38 | 0.12 |
| 0.47 | 0.16 |
| 0.58 | 0.19 |
| 0.67 | 0.23 |
| 0.78 | 0.34 |
| 0.89 | 0.52 |
| 0.97 | 0.72 |
| 1 | 0.87 |

- 10) В таблице даны значения переменных X и Y для четырёх экспериментов. Найти коэффициент корреляции и значения коэффициентов a и b для оптимальной по методу наименьших квадратов линейной модели $Y=a +b*X$.

| X | Y |
|------|----|
| 0.12 | 52 |
| 0.23 | 37 |
| 0.35 | 17 |
| 0.46 | 2 |

- 11) Рассматривается задача классификации на два класса: положительный и отрицательный. В ходе тестирования классификатора получены следующие результаты: полнота составляет 75%, общая точность составляет 80%. Какие значения может принимать точность?
- 12) Магазин собрал сведения о покупках (транзакции в файле). Были построены ассоциативные правила. Какое правило, содержащее в условии 2 элемента, имеет наибольшую поддержку?
- 13) Государственная избирательная комиссия зафиксировала результаты выборов по партиям и по регионам (таблица в файле). Требуется кластеризовать регионы по правилу k -средних для числа кластеров K от 1 до 12. Для каждого числа кластеров K найти максимальный радиус кластера. Построить график этой величины от K . На основании графика предположить, сколько групп регионов разумно выделить по итогам выборов.
- 14) В алгоритме вычисления оценок написать формулу для числа голосов, если система опорных множеств состоит из всех непустых подмножеств, а функция близости определяется только порогами e_1, \dots, e_n .
- 15) Обоснуйте способ построения всех тупиковых тестов через приведение системы тестовых уравнений к неупрощаемой ДНФ.

8.3. Перечень вопросов и заданий для промежуточной аттестации по дисциплине

- 1) Вывод алгоритма xgBoost. Чем он превосходит градиентный бустинг?
- 2) Задача отбора признаков. Отбор с помощью важности признаков.
- 3) Байесовское решающее правило минимальной цены и его упрощения.
- 4) Генеративные и дискриминативные модели.
- 5) Коллаборативная фильтрация.
- 6) Основные методы кластеризации.
- 7) Методы активного обучения.
- 8) Вывод EM-алгоритма.
- 9) Обнаружение аномалий методами одноклассового SVM.
- 10) Статистическая теория обучения.

8.4. Планируемые уровни сформированности компетенций обучающихся и критерии оценивания

| Уровни | Содержательное описание уровня | Основные признаки выделения уровня (этапы формирования компетенции, критерии оценки сформированности) | Пятибалльная шкала (академическая) оценка | Двухбалльная шкала, зачет | БРС, % освоения (рейтинговая оценка) |
|------------|------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------|---------------------------|--------------------------------------|
| Повышенный | Творческая деятельность | <i>Включает нижестоящий уровень.</i> Умение самостоятельно принимать решение, решать проблему/задачу теоретического и прикладного характера на основе изученных методов, приемов, технологий | отлично | зачтено | 86-100 |
| Базовый | Применение знаний и умений в более широких контекстах учебной и профессиональной деятельности, нежели по образцу с большей степени | <i>Включает нижестоящий уровень.</i> Способность собирать, систематизировать, анализировать и грамотно использовать информацию из самостоятельно найденных теоретических источников и иллюстрировать ими теоретические положения или | хорошо | | 71-85 |

| | | | | | |
|----------------------------------|-------------------------------------------------|---------------------------------------------------------------------------------------|---------------------|------------|----------|
| | самостоятельности и инициативы | обосновывать практику применения | | | |
| Удовлетворительный (достаточный) | Репродуктивная деятельность | Изложение в пределах задач курса теоретически и практически контролируемого материала | удовлетворительно | | 55-70 |
| Недостаточный | Отсутствие признаков удовлетворительного уровня | | неудовлетворительно | не зачтено | Менее 55 |

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

Основная литература

1. Авдеенко, Т. В. Введение в искусственный интеллект и логическое программирование. Программирование в среде Visual Prolog : учебное пособие / Т. В. Авдеенко, М. Ю. Целебровская. - Новосибирск : Изд-во НГТУ, 2020. - 64 с. - ISBN 978-5-7782-4182-4. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1869259> (дата обращения: 04.04.2023). – Режим доступа: по подписке.

Дополнительная литература

1. Пенькова, Т. Г. Модели и методы искусственного интеллекта : учебное пособие / Т. Г. Пенькова, Ю. В. Вайнштейн. - Красноярск : Сиб. федер. ун-т, 2019. - 116 с. - ISBN 978-5-7638-4043-8. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1816605> (дата обращения: 04.04.2023). – Режим доступа: по подписке.
2. Сергеев, Н. Е. Системы искусственного интеллекта. Часть 1: Учебное пособие / Сергеев Н.Е. - Таганрог:Южный федеральный университет, 2016. - 118 с.: ISBN 978-5-9275-2113-5. - Текст : электронный. - URL: <https://znanium.com/catalog/product/991954> (дата обращения: 04.04.2023). – Режим доступа: по подписке.
3. Сопов, Е. А. Многокритериальные нейроэволюционные системы в задачах машинного обучения и человеко-машинного взаимодействия : монография / Е. А. Сопов, И. А. Иванов. - Красноярск : Сиб. федер. ун-т, 2019. - 160 с. - ISBN 978-5-7638-3969-2. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1818898> (дата обращения: 04.04.2023). – Режим доступа: по подписке.

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

- НЭБ Национальная электронная библиотека, диссертации и прочие издания
- ЭБС Консультант студента
- ПРОСПЕКТ ЭБС
- ЭБС ZNANIUM.COM
- ЭБС IBOOKS.RU
- Электронно-библиотечная система (ЭБС) Кантитана (<https://elib.kantiana.ru/>)

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

Программное обеспечение обучения включает в себя:

- система электронного образовательного контента БФУ им. И. Канта – www.lms3.kantiana.ru, обеспечивающую разработку и комплексное использование электронных образовательных ресурсов;
- серверное программное обеспечение, необходимое для функционирования сервера и связи с системой электронного обучения через Интернет;
- корпоративная платформа webinar.ru;
- установленное на рабочих местах обучающихся ПО: Microsoft Windows 7, Microsoft Office Standart 2010, антивирусное программное обеспечение Kaspersky Endpoint Security.
- СУБД PostgreSQL (Свободное ПО, лицензия - Freeware).
- MongoDB (Свободное ПО, лицензия - Freeware).
- Python 2.7.15 (Anaconda2 5.2.0 64-bit)
- Python 3.6.5 (Anaconda3 5.2.0 64-bit)

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.