Министерство науки и высшего образования Российской Федерации

Федеральное государственное автономное образовательное учреждение   
высшего образования «Новосибирский национальный исследовательский  
государственный университет» (Новосибирский государственный университет, НГУ)

**Факультет информационных технологий**

СОГЛАСОВАНО

Декан ФИТ НГУ

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ М.М. Лаврентьев

«23» июля 2020 г.

РАБОЧАЯ ПРОГРАММА Дисциплины

|  |
| --- |
| **Методы машинного обучения** |

Направление подготовки: 09.03.01 ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА

Направленность (профиль): Компьютерные науки и системотехника

Форма обучения: очная

Год обучения: 3 семестр: 5

|  |  |  |
| --- | --- | --- |
| **№** | **Вид деятельности** | **Семестр** |
| **5** |
| **1** | Лекции, час. | 32 |
| **2** | Практические занятия, час. | 32 |
| **3** | Лабораторные занятия, час. |  |
| **4** | Занятий в контактной форме без учета промежуточной аттестации, час, из них | 66 |
| **5** | в электронной форме, час. |  |
| **6** | из них аудиторных занятий, час. | 64 |
| **7** | из них в активной и интерактивной форме, час. | 48 |
| **8** | консультаций, час. | 2 |
| **9** | Самостоятельная работа, час. | 112 |
| **10** | в том числе на выполнение письменных работ, час | 50 |
| **11** | Форма аттестации (экзамен, зачет, дифференцированный зачет), час | Э 2 |
| **12** | Всего зачетных единиц[[1]](#footnote-1) | 5 |

Новосибирск 2020

Рабочая программа дисциплины составлена на основании федерального государственного образовательного стандарта (ФГОС) высшего образования - бакалавриат по направлению подготовки 09.03.01 ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА.

Федеральный государственный образовательный стандарт (ФГОС) высшего образования - бакалавриат по направлению подготовки 09.03.01 ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА введен в действие приказом Минобрнауки от 19.09.2017 № 929.

Место дисциплины в структуре учебного плана: Блок 1 Дисциплины (модули); часть, формируемая участниками образовательных отношений; обязательная дисциплина.

Рабочая программа дисциплины утверждена решением Ученого совета факультета информационных технологий от 22.07.2020, протокол № 77.

Программу разработал:

профессор кафедры теоретической кибернетики ММФ,

доктор технических наук С.Н. Постовалов

ассистент кафедры Информационной биологии ФЕН,

кандидат биологических наук Д.В. Антонец

доцент кафедры систем информатики ФИТ,

кандидат физико-математических наук Д.С. Мигинский

Заведующий кафедрой систем информатики ФИТ,

доктор физико-математических наук М.М. Лаврентьев

Ответственный за образовательную программу:

Доцент кафедры систем информатики ФИТ,

кандидат физико-математических наук Д.С. Мигинский

**Аннотация к рабочей программе дисциплины**

**«Методы машинного обучения»**

Дисциплина «Методы машинного обучения» реализуется в рамках образовательной программы высшего образования – программы бакалавриата 09.03.01 ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА, направленность (профиль): Компьютерные науки и системотехника по очной форме обучения на русском языке.

**Место в образовательной программе:**

Дисциплина «Методы машинного обучения» реализуется в 5 семестре рамках части, формируемой участниками образовательных отношений, Блока 1 дисциплин (модулей) и является обязательной дисциплиной. Изучение данной дисциплины базируется на следующих дисциплинах предыдущих семестров: «Императивное программирование», «Декларативное программирование», «Объектно-ориентированное программирование», «Модели вычислений», «Теория параллелизма», «Разработка программно-аппаратного комплекса для решения научных и прикладных задач (групповой проект)».

Результаты освоения дисциплины «Методы машинного обучения» используются при работе в рамках практик и для написания выпускной квалификационной работы, в дисциплине по выбору «Методы глубокого обучения».

Дисциплина «Методы машинного обучения» направлена на формирование компетенций:

Способен разрабатывать требования и проектировать программное обеспечение (ПКС-1), в части следующих индикаторов достижения компетенции:

ПКС-1.1 уметь применять современные методы проектирования программного обеспечения, позволяющие вести разработку программных систем средней и высокой сложности

ПКС- 1.2 уметь применять методы проектирования предметной области в модели «сущность-связь» и разрабатывать логическую и физическую модель базы данных

ПКС- 1.3 уметь применять программные компоненты среды программирования, используемые для формирования интерфейса "человек - электронно-вычислительная машина"

ПКС-1.4 владеть основными приемами функционального и логического программирования

ПКС-1.5 уметь использовать программные средства для решения прикладных задач

ПКС-1.6 Способен на основе знания первых принципов информатики и широкой эрудиции в моделях и методах с ней связанных проектировать программно-аппаратные средства для решения практических задач на основе как неформального технического задания, так и формальных спецификаций

Способен выполнять работы и управлять работами по созданию (модификации) и сопровождению ИС, автоматизирующих задачи организационного управления и бизнес-процессы (ПКС-2), в части следующих индикаторов достижения компетенции:

ПКС-2.1 уметь применять современные инструментальные средства для разработки компонентов аппаратно-программных комплексов и баз данных

ПКС-2.2 уметь применять современные технологии программирования для разработки компонентов аппаратно-программных комплексов и баз данных

Способен осуществлять концептуальное, функциональное и логическое проектирование систем среднего и крупного масштаба и сложности (ПКС-3) в части следующих индикаторов достижения компетенции:

ПКС-3.1 проводить эксперименты по заданной методике и анализировать результаты

ПКС-3.2 проводить измерения и наблюдения, составлять описания проводимых исследований, готовить данные для составления обзоров, отчетов и научных публикаций

ПКС-3.3 знать инструментальные средства, применяемые для контроля принимаемых проектных решений

ПКС-3.4 Уметь применять различные формализмы для моделирования параллельных систем, а так же для спецификации и верификации их свойств

ПКС-3.5 уметь подтверждать корректность работы программной системы путем организации модульного тестирования и представления результатов тестов

ПКС-3.6 понимает природу и иерархическую сущности абстракций, а также роль и знание математических моделей в разработке программных и аппаратных технологий

ПКС-3.7 умеет использовать логические и алгебраические формализмы при характеризации технологических аспектов, возникающих в процессе разработки программных и программно-аппаратных комплексов

ПКС-3.8 умеет анализировать научно-технические публикации и определять дальнейшее направление исследования в рамках заданной тематики

**Перечень основных разделов дисциплины:**

Дисциплина «Методы машинного обучения» предусматривает проведение лекций и практических занятий в интерактивной форме.

Содержание дисциплины охватывает круг вопросов, связанных с основами интеллектуального анализа данных, включая преобразование и очистку данных, работу с пропущенными значениями, основные способы визуализации данных, корреляционный анализ, различные методы отбора признаков, задачи снижения размерности данных, кластеризации, классификации, регрессии. Студенты освоят работу со специализированными программными библиотеками для визуализации и анализа данных и научатся применять полученные знания для решения практических задач, в том числе, загружать данные, сохраненные в разных форматах, выбирать и группировать нужные записи по заданным критериям, строить предсказательные модели и оценивать их качество.

Общий объем дисциплины – 5 зачетных единиц (180 часов)

**Правила аттестации по дисциплине.**

В соответствии с учебным планом устанавливаются следующие формы контроля: текущий контроль успеваемости в форме портфолио (выполнения заданий) и промежуточный контроль в форме экзамена (5 семестр)

Результаты оцениваются по шкале «неудовлетворительно», «удовлетворительно», «хорошо», «отлично». Оценки «отлично», «хорошо», «удовлетворительно» означают успешное прохождение промежуточной аттестации.

**Учебно-методическое обеспечение дисциплины.**

Учебно-методические материалы по дисциплине «Методы машинного обучения» выложены на странице курса в сети Интернет: <https://classroom.google.com/c/MTQ0Mzk0OTQyMzk2?cjc=xw5mml7> Адрес курса сообщается студентам на первом занятии

**1. Внешние требования к дисциплине**

Таблица 1.1

|  |  |
| --- | --- |
| ***Компетенция* ПКС-1 Способен разрабатывать требования и проектировать программное обеспечение,** ***в части следующих индикаторов достижения компетенции:*** | |
| ПКС-1.1 | уметь применять современные методы проектирования программного обеспечения, позволяющие вести разработку программных систем средней и высокой сложности |
| ПКС- 1.2 | уметь применять методы проектирования предметной области в модели «сущность-связь» и разрабатывать логическую и физическую модель базы данных |
| ПКС- 1.3 | уметь применять программные компоненты среды программирования, используемые для формирования интерфейса "человек - электронно-вычислительная машина" |
| ПКС-1.4 | владеть основными приемами функционального и логического программирования |
| ПКС-1.5 | уметь использовать программные средства для решения прикладных задач |
| ПКС-1.6 | Способен на основе знания первых принципов информатики и широкой эрудиции в моделях и методах с ней связанных проектировать программно-аппаратные средства для решения практических задач на основе как неформального технического задания, так и формальных спецификаций |
| ***Компетенция* ПКС-2 Способен выполнять работы и управлять работами по созданию (модификации) и сопровождению ИС, автоматизирующих задачи организационного управления и бизнес-процессы,** ***в части следующих индикаторов достижения компетенции:*** | |
| ПКС-2.1 | уметь применять современные инструментальные средства для разработки компонентов аппаратно-программных комплексов и баз данных |
| ПКС-2.2 | уметь применять современные технологии программирования для разработки компонентов аппаратно-программных комплексов и баз данных |
| ***Компетенция* ПКС-3 Способен осуществлять концептуальное, функциональное и логическое проектирование систем среднего и крупного масштаба и сложности,** ***в части следующих индикаторов достижения компетенции:*** | |
| ПКС-3.1 | проводить эксперименты по заданной методике и анализировать результаты |
| ПКС-3.2 | проводить измерения и наблюдения, составлять описания проводимых исследований, готовить данные для составления обзоров, отчетов и научных публикаций |
| ПКС-3.3 | знать инструментальные средства, применяемые для контроля принимаемых проектных решений |
| ПКС-3.4 | Уметь применять различные формализмы для моделирования параллельных систем, а так же для спецификации и верификации их свойств |
| ПКС-3.5 | уметь подтверждать корректность работы программной системы путем организации модульного тестирования и представления результатов тестов |
| ПКС-3.6 | понимает природу и иерархическую сущности абстракций, а также роль и знание математических моделей в разработке программных и аппаратных технологий |
| ПКС-3.7 | умеет использовать логические и алгебраические формализмы при характеризации технологических аспектов, возникающих в процессе разработки программных и программно-аппаратных комплексов |
| ПКС-3.8 | умеет анализировать научно-технические публикации и определять дальнейшее направление исследования в рамках заданной тематики |

**2. Требования к результатам освоения дисциплины**

Таблица 2.1

|  |  |  |  |
| --- | --- | --- | --- |
| **Результаты изучения дисциплины по уровням освоения (иметь представление, знать, уметь, владеть)** | **Формы организации занятий** | | |
| **Лекции** | **Практики / семинары** | **Самостоятельная работа** |
| ПКС-1.1 уметь применять современные методы проектирования программного обеспечения, позволяющие вести разработку программных систем средней и высокой сложности | | | |
| 1 Уметь применять современные методы проектирования при разработке программных систем машинного обучения | + | + | + |
| ПКС- 1.2 уметь применять методы проектирования предметной области в модели «сущность-связь» и разрабатывать логическую и физическую модель базы данных | | | |
| 2 Уметь применять модели искусственного нейрона и нейронных сетей для задач машинного обучения | + |  | + |
| ПКС- 1.3 уметь применять программные компоненты среды программирования, используемые для формирования интерфейса "человек - электронно-вычислительная машина" | | | |
| 3Уметь визуализировать данные, в том числе с помощью методов снижения размерности | + | + | + |
| ПКС-1.4 владеть основными приемами функционального и логического программирования | | | |
| 4 Уметь применять и комбинировать различные вычислительные модели машинного обучения | + | + | + |
| ПКС-1.5 уметь использовать программные средства для решения прикладных задач | | | |
| 5 Уверенно проводить разведочный анализ данных, проводить предобработку и очистку данных, работать с пропущенными значениями | + | + | + |
| ПКС-1.6 Способен на основе знания первых принципов информатики и широкой эрудиции в моделях и методах с ней связанных проектировать программно-аппаратные средства для решения практических задач на основе как неформального технического задания, так и формальных спецификаций | | | |
| 6 Уметь обоснованно выбирать наиболее подходящие алгоритмы решения задач машинного обучения и оценивать качество построенных моделей | + | + | + |
| ПКС-2.1 уметь применять современные инструментальные средства для разработки компонентов аппаратно-программных комплексов и баз данных | | | |
| 7 Уметь применять базовые инструменты анализа данных и решения задач машинного обучения, реализованных в библиотеках pandas и sklearn, для разработки систем машинного обучения. |  | + | + |
| ПКС-2.2 уметь применять современные технологии программирования для разработки компонентов аппаратно-программных комплексов и баз данных | | | |
| 8 Уметь составлять конвейеры для предобработки данных, построения и подбора оптимальных гиперпараметров моделей |  | + | + |
| ПКС-3.1 проводить эксперименты по заданной методике и анализировать результаты | | | |
| 9 Уметь анализировать результаты экспериментов с помощью методов корреляционного и регрессионного анализа, решения задач классификации и кластерного анализа | + | + | + |
| ПКС-3.2 проводить измерения и наблюдения, составлять описания проводимых исследований, готовить данные для составления обзоров, отчетов и научных публикаций | | | |
| 10 Уметь составлять описания результатов работы систем машинного обучения, их использования в научных отчетах и публикациях |  | + | + |
| ПКС-3.3 знать инструментальные средства, применяемые для контроля принимаемых проектных решений | | | |
| 11 Знать особенности работы со специализированными программными библиотеками языка программирования Python для анализа данных и решения задач машинного обучения |  | + | + |
| ПКС-3.4 Уметь применять различные формализмы для моделирования параллельных систем, а также для спецификации и верификации их свойств | | | |
| 12 Уметь применять различные формализмы для моделирования параллельных систем машинного обучения |  | + | + |
| ПКС-3.5 уметь подтверждать корректность работы программной системы путем организации модульного тестирования и представления результатов тестов | | | |
| 13 Уметь оценивать качество построенной системы машинного обучения по контрольным выборкам, с помощью скользящего экзамена и кросс-валидации | + | + | + |
| ПКС-3.6 понимает природу и иерархическую сущности абстракций, а также роль и знание математических моделей в разработке программных и аппаратных технологий | | | |
| 14 Понимать роль и знание математических моделей в разработке программных и аппаратных технологий машинного обучения | + |  |  |
| ПКС-3.7 умеет использовать логические и алгебраические формализмы при характеризации технологических аспектов, возникающих в процессе разработки программных и программно-аппаратных комплексов | | | |
| 15 Уметь использовать логические и алгебраические формализмы при разработке систем машинного обучения | + | + | + |
| ПКС-3.8 умеет анализировать научно-технические публикации и определять дальнейшее направление исследования в рамках заданной тематики | | | |
| 16 Уметь работать с современными публикациями в области машинного обучения, следить за основными тенденциями в развитии методов машинного обучения |  |  | + |

**3. Содержание и структура учебной дисциплины**

Таблица 3.1

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Темы лекций** | **Активные формы, час.** | **Часы** | **Ссылки на результаты обучения** | | |
| **Семестр: 5** | | | | | |
| Лекция 1. Введение в машинное обучение и анализ данных Цели и задачи анализа данных. Основные подходы, этапы анализа, понятия и обозначения. Методы генерации признаков. Риск неправильного прогнозирования, минимизация эмпирического риска. Виды ошибок. Проблема переобучения. Оценивание качества по контрольной выборке, процедура скользящего экзамена, кросс-валидация. Анализ ROC кривых. Характеристика AUC. | 2 | 4 | | | 1,5,6,13,14 |
| Лекция 2. Методы регрессионного анализа. Модель множественной линейной регрессии. МНК - оценки параметров модели. Проверка статистической значимости модели, коэффициент детерминации. Обобщения линейной модели. Проблема мультиколлинеарности. Регуляризация модели. Гребневая регрессия, метод LASSO. Дерево регрессии. Логистическая регрессия. Непараметрическая регрессия, оценка Надарая-Уотсона. | 2 | 4 | | | 9,14 |
| Лекция 3. Байесовское распознавание образов. Задача распознавания образов. Дискриминантная (решающая) функция, оптимальная (байесовская) решающая функция. Задача оценивания распределений по выборке. Оптимальная решающая функция при многомерных нормальных распределениях. Построение решающих функций в пространстве бинарных, номинальных переменных. Наивный байесовский классификатор. Метод ближайших соседей kNN | 2 | 4 | | | 9,14 |
| Лекция 4. Линейные классификаторы и их обобщения. Классификация с помощью линейных функций. Линейный дискриминант Фишера. Метод опорных векторов. Случай линейно неразделимых классов, Kernel trick, теорема Мерсера. Основы нейросетевого подхода в распознавании образов. Модель кибернетического нейрона, персептрон, алгоритм Розенблатта. Обучение двухслойного персептрона, метод обратного распространения ошибок. | 2 | 4 | | | 2,14 |
| Лекция 5. Методы распознавания, основанные на логических закономерностях. Подходы к определению закономерностей: эвристический, информационный, статистический. Алгоритмы КОРА, ТЕМП. Поиск ассоциативных правил. Алгоритм APRIORI. Решающие списки. Деревья решений. Методы построения деревьев решений, рекурсивный алгоритм. Процедура усечения (pruning). Критерии качества деревьев решений. Алгоритмы ID3, C4.5, CART . | 2 | 4 | | | 9,15 |
| Лекция 6. Коллективный подход в распознавании образов. Комитет решающих функций. Процедура AdaBoost и ее сходимость. Градиентный бустинг. Методы генерации базовых решений, процедуры bagging и boosting, алгоритм случайных подпространств. Решающий лес. Мощность и корреляция случайного леса. Out -of-bag оценка качества леса | 2 | 4 | | | 4,14 |
| Лекция 7. Кластерный анализ. Постановка задачи кластерного анализа, этапы ее решения. Кластерные структуры. Расстояния между кластерами. Иерархические алгоритмы кластерного анализа. Алгоритмы k-means, kernel k-means, fuzzy C-means, DBSCAN. Непараметрические алгоритмы кластерного анализа. Алгоритм mean-shift, самоорганизующаяся карта Кохонена. Оценивание качества группировки, внешние и внутренние индексы качества: Rand index, NMI, Dunn index, Gamma index. Методы ансамблевого кластерного анализа. | 2 | 4 | | | 9,14 |
| Лекция 8. Отбор информативных признаков. Пошаговый отбор информативных признаков в задачах классификации и регрессии. Метод случайного поиска с адаптацией. Анализ главных компонент. Метод t-SNE. | 2 | 4 | | 3,14 | |
| **Итого** | **16** | **32** | |  | |

Таблица 3.2

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Темы практических занятий** | **Активные формы, час.** | | **Часы** | **Ссылки на результаты обучения** | | **Учебная деятельность** |
| **Семестр: 5** | | | | | | |
| 1. Введение в предметную область. Примеры использования методов машинного обучения для решения прикладных задач. Повторение основ программирования на языке Python. | 4 | 4 | | | 1 | Примеры использования методов машинного обучения для практических задач. Краткий обзор синтаксиса языка Python. Встроенные операции и функции, типы и структуры данных. |
| 1. Знакомство со специализированными библиотеками языка программирования Python для научных расчетов и анализа данных. NumPy, SciPy, pandas. | 4 | 4 | | | 7. 11 | Библиотеки NumPy и SciPy. Матрицы. Разреженные матрицы. Индексирование, срезы. Объединение массивов. Библиотека pandas. Запросы к таблицам: выборка строк/столбцов по заданным критериям. Модификация элементов таблицы. Добавление строк/столбцов. Группировка и агрегирование. Объединение таблиц (различные виды join). Многомерные данные: мультииндексы. Операции stack-unstack. Построение сводных таблиц (pivot tables). |
| 1. Знакомство с различными методами предобработки данных, описательными статистиками и основными способами визуализации данных, методами снижения размерности. Метод главных компонент. Важность нормировки данных. Предобработка данных. Работа с пропущенными значениями. | 4 | 4 | | | 3, 5, 7 | Описательные статистики. Обзор библиотек matplotlib, seaborn, bokeh. Базовые типы визуализации данных. Знакомство с библиотекой scikit-learn (sklearn). Предобработка данных. Метод главных компонент. Работа с пропущенными значениями. |
| 1. Основы машинного обучения и основные типы задач. Классификация задач машинного обучения. | 2 | 2 | | | 7 | Дальнейшее знакомство студентов с пакетом sklearn. Основные функции. Работа с данными из набора MNIST (рукописные цифры). Работа с синтетическими данными. |
| 1. Обучение на неразмеченных данных. Нормировка данных. Кластеризация. Иерархическая кластеризация. Метод K-средних, DBSCAN и др. Обзор методов кластеризации, реализованных в библиотеке sklearn. | 2 | 2 | | | 3, 9 | Использование методов снижения размерности и методов кластеризации в задаче распознавания рукописных цифр (MNIST). Работа с синтетическими данными. |
| 1. Задачи обучения с учителем. Разделение данных на обучающие и тестовые. Определение переобученности модели. Критерии оценки качества полученных моделей. | 4 | 4 | | | 6, 13 | Примеры задач обучения с учителем. Важность определения целевой метрики качества. Сравнение различных метрик качества моделей. Работа с несбалансированными наборами. |
| 1. Постановка задачи регрессии. Линейный регрессионный анализ. Отбор признаков, коллинеарность, влиятельные наблюдения, анализ остатков. Непараметрическая регрессия (ядерное сглаживание). L1 и L2 регуляризация. Метрики качества. | 4 | 4 | | | 7, 8 | Объединение алгоритмов, реализованных в sklearn, в цепочки и конвейеры с помощью класса Pipeline. Реализация регрессионных и классификационных моделей с помощью sklearn. Работа с синтетическими данными. Самостоятельная реализация метода градиентного спуска. |
| 1. Постановка задачи классификации, обзор основных методов ее решения. Бинарная и многоклассовая классификация. Логистическая регрессия. Решающие деревья. Метрики качества классификации (точность и специфичность, ROC-кривая, площадь под кривой). | 4 | 4 | | | 7, 9,10,15 | Реализация классификационных моделей с помощью sklearn. Реализация моделей на основе метода k-ближайших соседей. Метод логистической регрессии. Реализация метода градиентного спуска. Реализация решающего дерева. |
| 1. Ансамбли алгоритмов машинного обучения. Агрегирование моделей. Ансамбли решающих деревьев. Метод случайного леса. Градиентный бустинг. | 4 | 4 | | | 4, 8, 12 | Реализация моделей с помощью метода градиентного бустинга, метода случайного леса. Блендинг и стеккинг. Методы отбора признаков. Оптимизация гиперпараметров. |
| **Итого** | **32** | **32** | | |  |  |

**4. Самостоятельная работа бакалавров**

Таблица 4.1

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **№** | **Виды самостоятельной работы** | | **Ссылки на результаты обучения** | **Часы на выполнение** | **Часы на консультации** |
| **Семестр: 5** | | | | | |
| 1 | Самостоятельная работа с учебным материалом: основной учебной литературой, с дополнительной литературой | | 1,2,3,4,5,6,7,8,9,10,11,12,13, 15,16 | 36 |  |
| Изучение предлагаемых алгоритмов и структур данных, анализ и детальное изучение представленных технологий программирования. Учебно-методические материалы по дисциплине «Методы машинного обучения» выложены на странице курса в сети Интернет | | | | |
| 2 | Подготовка к практическим работам, к текущему контролю знаний и промежуточной аттестации | | 1,2,3,4,5,6,7,8,9,10,11,12,13, 15,16 | 50 |  |
| Разбор решенных задач, самостоятельное решение задач | | | | |
| 3 | Подготовка к экзамену | | 1,2,3,4,5,6,7,8,9,10,11,12,13, 15,16 | 26 | 2 |
| Повторение теоретического материала по вопросам, совпадающим с темами лекций | | | | |
| **Итого** | |  | | **112** | **2** |

**5. Образовательные технологии**

В ходе реализации учебного процесса по дисциплине проводятся лекционные и практические занятия. Темы, рассматриваемые на лекциях и изучаемые самостоятельно, закрепляются на практических занятиях, по вопросам, вызывающим затруднения, проводятся консультации.

В ходе реализации учебного процесса по дисциплине применяются лекционные и практические занятия, а также применяются следующие интерактивные формы обучения (таблица 5.1).

Таблица 5.1

|  |  |  |  |
| --- | --- | --- | --- |
| **1** | | Лекция в форме дискуссии | ПКС-1, ПКС-2, ПКС-3 |
| **Формируемые умения:**  Студент изучают основные приемы, применяемые при решении задач классификации и анализа данных, в том числе, построение оптимального байесовского решающего правила, построение линейного дискриминанта Фишера, построение опорных векторов, обучение искусственного нейрона, построение логических закономерностей и деревьев решений. Учатся обоснованно выбирать наиболее подходящие алгоритмы решения задач машинного обучения и оценивать качество построенных моделей. | | | |
| **Краткое описание применения:** В лекциях дается обзор тем изучаемой дисциплины, определения основных понятий, терминов приводятся доказательства утверждений по теории машинного обучения. Часть времени отдельных лекций отводится для интерактивного обсуждения понятий, моделей и методов машинного обучения, разбором конкретных ситуаций, разбор возможных вариантов решений. | | | |
| **2** | Портфолио | | ПКС-1, ПКС-2, ПКС-3 |
| **Формируемые умения:**  При решении практических и теоретических задач из домашних заданий студенты закрепляют знания и навыки, полученные в ходе лекций и семинарских занятий. Учатся использовать инструменты анализа данных и решения задач машинного обучения, реализованные в библиотеках pandas и sklearn. Учатся проводить разведочный анализ данных, предобработку и очистку данных. Учатся визуализировать данные, в том числе, с использованием методов снижения размерности. Учатся обоснованно выбирать наиболее подходящие алгоритмы решения задач машинного обучения и оценивать качество построенных моделей. | | | |
| **Краткое описание применения:** бакалавры ведут портфолио (оценки за выполненные заданий, групповых проектов), которое является основой для проведения аттестации по дисциплине | | | |

Для организации и контроля самостоятельной работы бакалавров, а также проведения консультаций применяются информационно-коммуникационные технологии (таблица 5.2).

Таблица 5.2

|  |  |
| --- | --- |
| Информирование | Адрес почты – сообщается бакалаврам на первом занятии. |
| Консультирование | Адрес почты – сообщается бакалаврам на первом занятии. |
| Контроль | Адрес почты – сообщается бакалаврам на первом занятии. |
| Размещение учебных материалов | <https://classroom.google.com/u/1/c/MzcyNzIyMjY1NDla> |

**6. Правила аттестации бакалавров по учебной дисциплине**

По дисциплине «Методы машинного обучения» проводится текущая и промежуточная аттестация (итоговая по дисциплине).

**Текущая аттестация** по дисциплине «Методы машинного обучения»:

**Критерии формирования оценки**

Оценивание знаний и умений производится в 5-балльной системе в соответствии с оценочной шкалой (см выше). Оценке "удовлетворительно" - 3 балла, оценке "хорошо" – 4 балла, оценке "отлично" – 5 баллов.

Результаты оцениваются по шкале «неудовлетворительно», «удовлетворительно», «хорошо», «отлично». Оценки «отлично», «хорошо», «удовлетворительно» означают успешное прохождение промежуточной аттестации.

В таблице 6.1 представлено соответствие форм аттестации заявляемым требованиям к результатам освоения дисциплины.

Таблица 6.1

|  |  |  |  |
| --- | --- | --- | --- |
| **Коды компетенций ФГОС** | **Результаты обучения** | **Формы аттестации** | |
| **семестр 5** | |
| портфолио | экзамен |
| **ПКС-1** | ПКС-1.1 уметь применять современные методы проектирования программного обеспечения, позволяющие вести разработку программных систем средней и высокой сложности | **+** | **+** |
| **ПКС-1** | ПКС- 1.2 уметь применять методы проектирования предметной области в модели «сущность-связь» и разрабатывать логическую и физическую модель базы данных | **+** | **+** |
| **ПКС-1** | ПКС- 1.3 уметь применять программные компоненты среды программирования, используемые для формирования интерфейса "человек - электронно-вычислительная машина" | **+** | **+** |
| **ПКС-1** | ПКС-1.4 владеть основными приемами функционального и логического программирования | **+** | **+** |
| **ПКС-1** | ПКС-1.5 уметь использовать программные средства для решения прикладных задач | **+** | **+** |
| **ПКС-1** | ПКС-1.6 Способен на основе знания первых принципов информатики и широкой эрудиции в моделях и методах с ней связанных проектировать программно-аппаратные средства для решения практических задач на основе как неформального технического задания, так и формальных спецификаций | **+** | **+** |
| **ПКС-2** | ПКС-2.1 уметь применять современные инструментальные средства для разработки компонентов аппаратно-программных комплексов и баз данных | **+** | **+** |
| **ПКС-2** | ПКС-2.2 уметь применять современные технологии программирования для разработки компонентов аппаратно-программных комплексов и баз данных | **+** | **+** |
| **ПКС-3** | ПКС-3.1 проводить эксперименты по заданной методике и анализировать результаты | **+** | **+** |
| **ПКС-3** | ПКС-3.2 проводить измерения и наблюдения, составлять описания проводимых исследований, готовить данные для составления обзоров, отчетов и научных публикаций | **+** | **+** |
| **ПКС-3** | ПКС-3.3 знать инструментальные средства, применяемые для контроля принимаемых проектных решений | **+** | **+** |
| **ПКС-3** | ПКС-3.4 Уметь применять различные формализмы для моделирования параллельных систем, а так же для спецификации и верификации их свойств | **+** | **+** |
| **ПКС-3** | ПКС-3.5 уметь подтверждать корректность работы программной системы путем организации модульного тестирования и представления результатов тестов | **+** | **+** |
| **ПКС-3** | ПКС-3.6 понимает природу и иерархическую сущности абстракций, а также роль и знание математических моделей в разработке программных и аппаратных технологий | **+** | **+** |
| **ПКС-3** | ПКС-3.7 умеет использовать логические и алгебраические формализмы при характеризации технологических аспектов, возникающих в процессе разработки программных и программно-аппаратных комплексов | **+** | **+** |
| **ПКС-3** | ПКС-3.8 умеет анализировать научно-технические публикации и определять дальнейшее направление исследования в рамках заданной тематики | **+** | **+** |

Требования к структуре и содержанию портфолио, оценочные средства, а также критерии оценки сформированности компетенций и освоения дисциплины в целом, представлены в Фонде оценочных средств, являющемся приложением 1 к настоящей рабочей программе дисциплины.

**7. Литература**

1. Крутиков, В.Н. Анализ данных : учебное пособие / В.Н. Крутиков, В.В. Мешечкин ; Кемеровский государственный университет. – Кемерово : Кемеровский государственный университет, 2014. – 138 с. : ил.– URL: <http://biblioclub.ru/index.php?page=book&id=278426>– Библиогр. в кн. – ISBN 978-5-8353-1770-7. – Текст : электронный.
2. Каган, Е.С. Прикладной статистический анализ данных : учебное пособие : [16+] / Е.С. Каган ; Кемеровский государственный университет. – Кемерово : Кемеровский государственный университет, 2018. – 235 с. : ил., табл. – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=573550> (дата обращения: 25.09.2020). – Библиогр.: с. 184-186. – ISBN 978-5-8353-2413-2. – Текст : электронный.
3. Шелудько, В.М. Основы программирования на языке высокого уровня Python : учебное пособие / В.М. Шелудько ; Министерство науки и высшего образования РФ, Южный федеральный университет, Инженерно-технологическая академия. – Ростов-на-Дону ; Таганрог : Южный федеральный университет, 2017. – 147 с. : ил. – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=500056>. – Библиогр. в кн. – ISBN 978-5-9275-2649-9. – Текст : электронный
4. Глебов, В.И. Практикум по математической статистике: проверка гипотез с использованием Excel, MatCalc, R и Python : [16+] / В.И. Глебов, С.Я. Криволапов ; Финансовый университет при Правительстве Российской Федерации. – Москва : Прометей, 2019. – 87 с. : ил. URL: <http://biblioclub.ru/index.php?page=book&id=576035>– Библиогр. в кн. – ISBN 978-5-907100-66-4. – Текст : электронный.

*Интернет-ресурсы*

Таблица 7.1

|  |  |  |
| --- | --- | --- |
| № п/п | Наименование Интернет-ресурса | Краткое описание |
| 1 | <http://www.machinelearning.ru/> | Большая коллекция материалов по машинному обучению на русском языке. |
| 2 | [http://anaconda.org](http://anaconda.org/) | Дистрибутив Python с большинством необходимых библиотек. |
| 3 | <http://scipy.org/> | Библиотека для научных вычислений для языка программирования Python. |
| 4 | <http://pandas.pydata.org/> | Библиотека для анализа данных pandas. |
| 5 | <http://scikit-learn.org/stable/user_guide.html> | Документация библиотеки sklearn. |
| 6 | <http://scikit-learn.org/stable/tutorial/index.html> | Примеры решения некоторых задач. |
| 7 | [http://kaggle.com](http://kaggle.com/) | Платформа для проведения конкурсов по решению задач машинного обучения. Содержит обучающие ресурсы с примерами решений задач и их обсуждением. |
| 8 | <http://archive.ics.uci.edu/ml/> | Коллекция данных и задач. |
| 9 | <https://stepik.org/course/67> | Курс «Программирование на Python» по основам программирования на языке Python. |
| 10 | <https://www.coursera.org/learn/machine-learning> | Курс по основам машинного обучения от Эндрю Ына (Andrew Ng). Преподается на английском языке. |
| 11 | <https://ru.coursera.org/learn/vvedenie-mashinnoe-obuchenie> | Курс по основам машинного обучения с использованием Python + pandas + sklearn. Преподается на русском языке. Преподаватель: Константин Вячеславович Воронцов. |

**8. Учебно-методическое и программное обеспечение дисциплины**

**8.1. Учебно-методическое обеспечение**

**Учебно-методическое обеспечение дисциплины.**

Учебно-методические материалы по дисциплине «Методы машинного обучения» выложены на странице курса в сети Интернет:

<https://classroom.google.com/c/MTQ0Mzk0OTQyMzk2?cjc=xw5mml7>

**8.2. Программное обеспечение**

Для обеспечения реализации дисциплины используется стандартный комплект программного обеспечения (ПО), включающий регулярно обновляемое лицензионное ПО Windows и MS Office.

Перечень специализированного программного обеспечения для изучения дисциплины представлен в таблице 8.1.

Специализированное программное обеспечение Таблица 8.1

|  |  |  |
| --- | --- | --- |
| **№** | **Наименование ПО** | **Назначение** |
| 1 | Python 3.7.1 (Anaconda3 2018.12 64-bit) | Среда разработки приложений |
| 2 | Notepad++ | Программа для работы с текстовыми файлами |
| 3 | RStudio 1.1.4 | Среда разработки программного обеспечения |

**9. Профессиональные базы данных и информационные справочные системы**

1. Полнотекстовые журналы Springer Journals за 1997-2015 г., электронные книги (2005-2016 гг.), коллекция научных биомедицинских и биологических протоколов SpringerProtocols, коллекция научных материалов в области физических наук и инжиниринга SpringerMaterials, реферативная БД по чистой и прикладной математике zbMATH.
2. Электронная библиотека диссертаций Российской государственной библиотеки (ЭБД РГБ)
3. Электронные ресурсы Web of Science Core Collection (Thomson Reuters Scientific LLC.), Journal Citation Reports + ESI
4. БД Scopus (Elsevier)

**10. Материально-техническое обеспечение**

Таблица 10.1

|  |  |  |
| --- | --- | --- |
| **№** | **Наименование** | **Назначение** |
| 1 | Презентационное оборудование (мультимедиа-проектор, экран, компьютер для управления) | Для проведения лекционных занятий |
| 2 | Компьютерный класс (с выходом в Internet)  с компьютерами на базе процессоров не хуже P-IV 3Hz, 2Gb RAM | Для проведения практических занятий и организации самостоятельной работы обучающихся |

Материально-техническое обеспечение образовательного процесса по дисциплине для обучающихся из числа лиц с ограниченными возможностями здоровья осуществляется согласно «Порядку организации и осуществления образовательной деятельности по образовательным программам для инвалидов и лиц с ограниченными возможностями здоровья в Новосибирском государственном университете».

**Лист актуализации рабочей программы дисциплины  
«Методы машинного обучения»**

|  |  |  |  |
| --- | --- | --- | --- |
| № | Характеристика внесенных изменений (с указанием пунктов документа) | Дата и №  протокола Ученого совета ФИТ | Подпись  ответственного |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

1. С учетом выделенных часов на промежуточную аттестацию [↑](#footnote-ref-1)