

МИНОБРНАУКИ РОССИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ НАУКИ  
**ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР**  
**«КОЛЬСКИЙ НАУЧНЫЙ ЦЕНТР РОССИЙСКОЙ АКАДЕМИИ НАУК»**  
(ФИЦ КНЦ РАН)

УТВЕРЖДАЮ  
Начальник управления аспирантуры и магистратуры  
ФИЦ КНЦ РАН  
к.г.ч.-м.н., доцент И.В. Чикирёв

  
\_\_\_\_\_ подпись  
29 июня 2020 г.

**РАБОЧАЯ ПРОГРАММА**

**По дисциплине** Б1.В.02 Технологии хранения данных

указывается цикл (раздел) ОП, к которому относится дисциплина, название дисциплины

**для направления подготовки (специальности)** 09.04.02 Информационные системы и технологии

код и наименование направления подготовки (специальности)

**направленность программы (профиль)** Информационные системы предприятий и учреждений

наименование профиля /специализаций/образовательной программы

**Квалификация выпускника, уровень подготовки**

Магистр

(указывается квалификация (степень) выпускника в соответствии с ФГОС ВО)

Апатиты

2020

## Лист согласования

1 Разработчик:

доцент  
должность

УАиМ

  
подпись

Н.А. Тоичкин  
И.О. Фамилия

2 Рабочая программа рассмотрена и одобрена на заседании учебно-методической комиссии управления аспирантуры и магистратуры 29 июня 2020 г., протокол № 02.

Председатель УМК УАиМ

29.06.2020  
дата

  
подпись

Л.Д. Кириллова  
И.О.Фамилия

## Лист переутверждения

Рабочая программа переутверждена на 2021/2022 учебный год без изменений и дополнений.

Председатель УМК УАиМ  Л.Д. Кириллова

Основание: протокол № 2 от «29» июня 2021 г.

Рабочая программа переутверждена на \_\_\_\_\_/\_\_\_\_\_ учебный год без изменений и дополнений.

Председатель УМК УАиМ \_\_\_\_\_ Л.Д. Кириллова

Основание: протокол № \_\_\_\_\_ от « \_\_\_\_\_ » \_\_\_\_\_ 20 \_\_\_\_\_ г

Рабочая программа переутверждена на \_\_\_\_\_/\_\_\_\_\_ учебный год без изменений и дополнений.

Председатель УМК УАиМ \_\_\_\_\_ Л.Д. Кириллова

Основание: протокол № \_\_\_\_\_ от « \_\_\_\_\_ » \_\_\_\_\_ 20 \_\_\_\_\_ г

Рабочая программа переутверждена на \_\_\_\_\_/\_\_\_\_\_ учебный год без изменений и дополнений.

Председатель УМК УАиМ \_\_\_\_\_ Л.Д. Кириллова

Основание: протокол № \_\_\_\_\_ от « \_\_\_\_\_ » \_\_\_\_\_ 20 \_\_\_\_\_ г

Рабочая программа переутверждена на \_\_\_\_\_/\_\_\_\_\_ учебный год без изменений и дополнений.

Председатель УМК УАиМ \_\_\_\_\_ Л.Д. Кириллова

Основание: протокол № \_\_\_\_\_ от « \_\_\_\_\_ » \_\_\_\_\_ 20 \_\_\_\_\_ г

**Лист изменений, вносимых в РП\* по дисциплине «Проблемно-ориентированные информационные системы»**

В рабочую программу вносятся следующие изменения и дополнения:

1. \_\_\_\_\_  
\_\_\_\_\_
2. \_\_\_\_\_  
\_\_\_\_\_
3. \_\_\_\_\_  
\_\_\_\_\_

Дополнения и изменения внесены и одобрены на заседании учебно-методической комиссии управления аспирантуры и магистратуры ФИЦ КНЦ РАН

от «\_\_\_\_» \_\_\_\_\_ г., протокол № \_\_\_\_\_.

Председатель УМК УАиМ \_\_\_\_\_ Л.Д. Кириллова

---

\* Изменения, вносимые в РП – действия по изменению тематики и перечня лабораторных, практических работ, форм текущего и промежуточного контроля. В случае внесения изменений в РП в части количества часов, РП должна переутверждаться полностью. Лист изменений включается в структуру РП.

## Аннотация рабочей программы дисциплины

| Коды циклов дисциплин, модулей, практик | Название циклов, разделов, дисциплин, модулей, практик   | Краткое содержание (Цель, задачи, содержание разделов дисциплины, реализуемые компетенции, формы промежуточного контроля, формы отчетности)   |
|---|--|---|
| Б1                                      | Дисциплины (модули)                                      |   |
| Б1.В                                    | Часть, формируемая участниками образовательных отношений |   |
| Б1.В.03                                 | Проблемно-ориентированные информационные системы         | <p><b>Цель дисциплины:</b> изучение современных технологий анализа информации и методов машинного обучения и их применение на практике.</p> <p><b>Задачи дисциплины:</b></p> <ul style="list-style-type: none"> <li>– изучить основные методы и алгоритмы машинного обучения;</li> <li>– получить навыки применения алгоритмов машинного обучения в задачах анализа информации;</li> <li>– осуществлять математическую и информационную постановку задач по обработке информации.</li> </ul> <p><b>В результате изучения дисциплины, обучающиеся должны:</b></p> <p><b>знать</b></p> <ul style="list-style-type: none"> <li>– формализацию задачи машинного обучения;</li> <li>– понятие больших данных и их свойства;</li> <li>– постановку задачи классификации и регрессии;</li> <li>– понятие обобщенного метрического классификатора;</li> <li>– алгоритмы метрической классификации;</li> <li>– основные принципы построения логических алгоритмов классификации;</li> <li>– алгоритм построения дерева классификации ID 3;</li> <li>– линейные методы классификации.</li> </ul> <p><b>уметь</b></p> <ul style="list-style-type: none"> <li>– использовать алгоритмы обработки информации для различных приложений;</li> <li>– выполнять постановку задачи машинного обучения;</li> <li>– выбирать методы и средства для решения задач машинного обучения;</li> </ul> |

|  |  |  |
|--|--|--|
|  |  | <p><b><i>владеет</i></b></p> <ul style="list-style-type: none"><li>– инструментальными средствами решения задач машинного обучения;</li><li>– методами интеллектуального анализа информации.</li></ul> <p><b>Содержание разделов дисциплины.</b></p> <ul style="list-style-type: none"><li>• Большие данные и машинное обучение.</li><li>• Метрические методы классификации</li><li>• Логические методы классификации</li><li>• Линейные методы классификации</li></ul> <p><b><i>Реализуемые компетенции:</i></b><br/>ПК-1</p> <p><b><i>Формы отчетности</i></b><br/>Семестр 2 – зачет</p> |
|--|--|--|

## Пояснительная записка

1. **Рабочая программа** составлена в соответствии с требованиями федерального государственного образовательного стандарта по образовательной программе высшего образования – программе магистратуры по направлению подготовки 09.04.02 Информационные системы и технологии, утвержденного приказом Минобрнауки России от 19.09.2017 № 917.

2. Цель дисциплины: изучение современных технологий анализа информации и методов машинного обучения и их применение на практике.

### Задачи дисциплины:

- изучить основные методы и алгоритмы машинного обучения;
- получить навыки применения алгоритмов машинного обучения в задачах анализа информации;
- осуществлять математическую и информационную постановку задач по обработке информации.

3. **Требования к уровню подготовки обучающегося в рамках данной дисциплины.**

Процесс изучения дисциплины (модуля) «Проблемно-ориентированные информационные системы» направлен на формирование элементов следующих компетенций в соответствии с ФГОС ВО 09.04.02 Информационные системы и технологии (уровень магистратуры), представленных в таблице 1.

**Таблица 1 – Компетенции, формируемые в процессе изучения дисциплины «Проблемно-ориентированные информационные системы»**

| № п/п | Код компетенции | Содержание компетенции   |
|-------|-----------------|--|
| 1.    | ПК-1            | Способен проводить экспертизу и оказывать информационно-аналитическую поддержку в решении профессиональных задач в научной деятельности. |

4. **Планируемые результаты обучения по дисциплине (модулю) «Проблемно-ориентированные информационные системы».**

Результаты формирования компетенций и обучения представлены в таблице 2.

**Таблица 2 – Планируемые результаты обучения**

| № п/п | Код компетенции | Компоненты компетенции, степень их реализации | Результаты обучения |
|-------|-----------------|---|---------------------|
|-------|-----------------|---|---------------------|

|    |      |  |  |
|----|------|--|--|
| 1. | ПК-1 | Компоненты компетенции соотносятся с содержанием дисциплины и компетенция реализуется полностью. | <p><b>знать</b></p> <ul style="list-style-type: none"> <li>– формализацию задачи машинного обучения;</li> <li>– понятие больших данных и их свойства;</li> <li>– постановку задачи</li> <li>– классификации и регрессии;</li> <li>– понятие обобщенного метрического классификатора;</li> <li>– алгоритмы метрической классификации;</li> <li>– основные принципы построения логических алгоритмов классификации;</li> <li>– алгоритм построения дерева классификации ID 3;</li> <li>– линейные методы классификации.</li> </ul> <p><b>уметь</b></p> <ul style="list-style-type: none"> <li>– использовать алгоритмы обработки информации для различных приложений;</li> <li>– выполнять постановку задачи машинного обучения;</li> <li>– выбирать методы и средства для решения задач машинного обучения;</li> </ul> <hr/> <p><b>владеть</b></p> <ul style="list-style-type: none"> <li>– инструментальными средствами решения задач машинного обучения;</li> <li>– методами интеллектуального анализа информации.</li> </ul> |
|----|------|--|--|

### 5. Место дисциплины (модуля) «Проблемно-ориентированные информационные системы» в структуре образовательной программы.

Дисциплина относится к части, формируемой участниками образовательных отношений (Б1.В.03) Блока 1– Дисциплины (модули), и преподается во 2-ом семестре.

Перечень дисциплин, усвоение которых необходимо обучающимся для изучения данной дисциплины.

«Информатика» - знание основ информационных технологий.

«Программирование» - знание базовых принципов алгоритмизации и основных алгоритмических конструкций, умение разрабатывать простые программы, знание принципов структурного и объектно-ориентированного подходов.

«Анализ алгоритмов» - знать понятие сложности алгоритма, владеть методикой анализа алгоритмов, знание базовых алгоритмов: алгоритмы сортировки и поиска, алгоритмы работы с графами и деревьями, жадные алгоритмы, рекурсивные алгоритмы и динамическое программирование.

Наименование дисциплин, изучение которых опирается на данную дисциплину.

«Научно-исследовательская работа», «Методы проектирования и исследования информационных процессов и технологий».

## 6. Структура учебной дисциплины (модуля)

**Таблица 3 – Распределение учебного времени дисциплины**  
**Общая трудоемкость дисциплины составляет 2 зачетные единицы, 72 часа**

| Виды учебной нагрузки, часов | Распределение трудоемкости дисциплины по формам обучения |    |   |             |
|------------------------------|--|----|---|-------------|
|                              | Очная  |    |   |             |
|                              | Номер семестра обучения                                  |    |   | Всего Часов |
|                              | 1  | 2  | 3 |             |
| Лекции                       | -  | 8  | - | <b>8</b>    |
| Практические занятия         | -  | 14 | - | <b>14</b>   |
| Лабораторные работы          | -  | -  | - | -           |
| Самостоятельная работа       | -  | 50 | - | <b>50</b>   |
| Подготовка и сдача экзамена  | -  | -  | - | -           |
| Всего часов по дисциплине    | -  | 72 | - | <b>72</b>   |

Формы промежуточного и текущего контроля

|                              |   |   |   |   |
|------------------------------|---|---|---|---|
| Экзамен                      | - | - | - | - |
| Зачет / зачет с оценкой      | - | + | - | + |
| Курсовая работа (проект)     | - | - | - | - |
| Количество РГЗ               | - | - | - | - |
| Количество контрольных работ | - | 1 | - | 1 |
| Количество рефератов         | - | - | - | - |
| Количество эссе              | - | - | - | - |

## 2. Содержание учебной дисциплины (модуля)

**Таблица 4 – Содержание разделов дисциплины (модуля), виды работ**

| № п/п | Содержание разделов (модулей), тем дисциплины | Количество часов, выделяемых на виды учебной подготовки по формам обучения |
|-------|---|--|
|       |   | Очная форма  |
|       |   |  |

|    |  | Объем работы в часах |        |         |
|----|--|----------------------|--------|---------|
|    |  | Лекции               | Практ. | Самост. |
| 1. | <p><b>Тема 1. Большие данные и машинное обучение.</b><br/>           Большие данные. Свойства больших данных. Машинное обучение, формализация задачи машинного обучения. Признаковое описание объекта. Ответы и типы задач машинного обучения. Модель алгоритмов. Метод обучения. Этап обучения и этап применения. Функционалы качества. Сведение задачи обучения к задаче оптимизации. Переобучение и обобщение. Пример переобучения (Рунге). Эмпирические оценки обобщающей способности. Примеры задач машинного обучения: задачи классификации и регрессии; задачи ранжирования. Эксперименты в машинном обучении: эксперименты на реальных и синтетических данных.</p>   | 2                    | 2      | 12      |
| 2. | <p><b>Тема 2. Метрические методы классификации</b><br/>           Формализация задачи. Обобщенный метрический классификатор. Метод ближайшего соседа. Метод k взвешенных ближайших соседей. Метод парзеновского окна. Метод потенциальных функций. Отбор эталонных объектов. Понятие отступа объекта. Типы объектов в зависимости от отступа. Отбор эталонов, алгоритм STOLP. Задача выбора метрики. Жадное добавление признаков.</p>  | 2                    | 4      | 12      |
| 3. | <p><b>Тема 3. Логические методы классификации</b><br/>           Логическая закономерность. Основы вопросы построения логических алгоритмов классификации. Виды закономерностей. Критерии информативности: простые критерии, статистический критерий, энтропийный критерий. Где находятся закономерности в <math>(p, n)</math>-плоскости. Схема локального поиска информативных закономерностей. Определение бинарного решающего дерева. Жадный алгоритм построения дерева ID 3. Варианты критериев ветвления в ID 3. Обработка пропусков, алгоритм обработки пропусков на этапе обучения и этапе классификации. Алгоритм ID3: достоинства и недостатки. Стратегии редукции решающих деревьев. Небрежные решающие деревья. Бинаризация вещественного признака.</p> | 2                    | 4      | 12      |

| №<br>п/п | Содержание разделов (модулей), тем дисциплины   | Количество часов, выделяемых на виды учебной подготовки по формам обучения |           |           |
|----------|---|--|-----------|-----------|
|          |   | Очная форма  |           |           |
|          |   | Объем работы в часах   |           |           |
|          |   | Лекции   | Практ.    | Самост.   |
| 4        | <b>Тема 4. Линейные методы классификации.</b><br>Задача построения разделяющей поверхности.<br>Задача построения разделяющей поверхности.<br>Минимизация эмпирического риска.<br>Непрерывные аппроксимации пороговой функции потерь. Линейный классификатор.<br>Персептрон. Устройство нервной клетки.<br>Линейная модель нейрона МакКаллока-Питтса.<br>Алгоритм Stochastic Gradient. Дельта-правило ADALINE. Правило Хебба. SG: инициализация весов. SG: проблемы переобучения. Принцип максимума правдоподобия. Оптимальная разделяющая гиперплоскость. Метод SVM.<br>Нелинейное обобщение SVM. | 2  | 4         | 14        |
|          | <b>Итого</b>  | <b>8</b>   | <b>14</b> | <b>50</b> |

**Таблица 5 - Соответствие компетенций, формируемых при изучении дисциплины (модуля), и видов занятий с учетом форм контроля**

| Перечень компетенций | Виды занятий |    |    |       |   |     |   |     | Формы контроля                          |
|----------------------|--------------|----|----|-------|---|-----|---|-----|---|
|                      | Л            | ЛР | ПЗ | КР/КП | Р | К/Р | Э | СРС |   |
| ПК-1                 | +            | -  | +  | -/-   | - | +   | - | +   | Практическая работа, контрольная работа |

Примечание: Л – лекции, ЛР – лабораторные работы, ПЗ – практические занятия, КР/КП – курсовая работа (проект), Р – реферат, К/Р – контрольная работа, Э – эссе, СРС – самостоятельная работа студентов.

**Таблица 6 - Перечень лабораторных работ**

| №<br>п/п | Наименование лабораторных работ | Количество часов | Наименование темы по табл. 4 |
|----------|---------------------------------|------------------|------------------------------|
|          | Не предусмотрены                |                  |                              |

**Таблица 7 - Перечень практических работ**

| №<br>п/п | Наименование практических работ                 | Количество часов | Наименование темы по табл. 4 |
|----------|---|------------------|------------------------------|
| 1.       | Предобработка данных в Pandas                   | 2                | 1                            |
| 2.       | Метрические методы классификации в Scikit-learn | 4                | 2                            |

|                    |  |           |   |
|--------------------|--|-----------|---|
| 3.                 | Деревья решений. Важность признаков            | 4         | 3 |
| 4.                 | Линейная классификация. Нормализация признаков | 4         | 4 |
| <b>Итого часов</b> |  | <b>14</b> |   |

### **8. Перечень примерных тем курсовой работы (проекта)**

Не предусмотрены.

### **9. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)**

Перечень учебно-методического обеспечения для самостоятельной работы обучающихся приводится в Методических указаниях к самостоятельной работе по дисциплине «Проблемно-ориентированные информационные системы».

### **10. Фонд оценочных средств (ФОС)**

ФОС входит в состав образовательной программы в качестве самостоятельного документа.

### **11. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля)**

#### **Основная литература:**

1. Архипенков С. Я. , Голубев Д. , Максименко О. Хранилища данных: от концепции до внедрения. М.: Диалог-МИФИ, 2002, 528 с. Режим доступа: [https://biblioclub.ru/index.php?page=book\\_red&id=89285&sr=1](https://biblioclub.ru/index.php?page=book_red&id=89285&sr=1)
2. Чубукова И. А. Data Mining. М.: Интернет-Университет Информационных Технологий, 2008, 383 с. Режим доступа: [https://biblioclub.ru/index.php?page=book\\_red&id=233055&sr=1](https://biblioclub.ru/index.php?page=book_red&id=233055&sr=1)

#### **Дополнительная литература:**

3. Введение в анализ данных с помощью Pandas. Режим доступа: <https://habrahabr.ru/post/196980/>

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

1. ЭБС «Издательство Лань» [Электронный ресурс]: электронная библиотечная система / ООО «Издательство Лань». – Режим доступа: <https://e.lanbook.com/>;
2. Электронно-библиотечная система "Университетская библиотека онлайн". [Электронный ресурс]. – Режим доступа: <http://biblioclub.ru>
3. ЭБС «Университетская библиотека онлайн» [Электронный ресурс]: электронно-периодическое издание; программный комплекс для организации онлайн-доступа к лицензионным материалам / ООО «НексМедиа». – Режим доступа: <https://biblioclub.ru/>.

**13. Перечень информационных технологий и лицензионного программного обеспечения, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем.**

Операционная система 'Windows 10', - лицензия: Win Pro 10 32-bit/64-bit All Lng PK Lic Online DwnLd NR; пакет офисного ПО 'Microsoft Office Pro 2007', - лицензия: Office Professional Plus 2007 License: 43364231; антивирусный пакет 'Kaspersky', - лицензия: номер лицензии: 0E26-201116-120400-323-2233; дистрибутив языков программирования Python и R 'Anaconda', - лицензия: Freemium License..

#### 14. Таблица 8 - Материально-техническое обеспечение дисциплины (модуля)

| № п/п | Наименование оборудованных учебных кабинетов, лабораторий  | Перечень оборудования и технических средств обучения  |
|-------|--|---|
| 1.    | Лекционные аудитории и аудитории для практических занятий. | Мультимедийный проектор BenQ SP890, переносной ноутбук Lenovo 4240-3EG, переносной экран для воспроизведения изображения, ПЭВМ Intel Core i5-3450 CPU 3.2 ГГц ОЗУ 8 Гб DDR3 с ЖК-монитором 19”. |

#### 15. Таблица 9 – Технологическая карта дисциплины (промежуточная аттестация – зачет)

| № п/п                                   | Контрольные точки                   | Зачетное количество баллов |            | График прохождения (недели сдачи)  |
|---|-------------------------------------|----------------------------|------------|--|
|   |                                     | min                        | max        |  |
| <b>Текущий контроль</b>                 |                                     |                            |            |  |
| 1                                       | 2                                   | 3                          | 4          | 5  |
| 1.                                      | Выполнение практической работы № 1  | 11                         | 18         | 2-я неделя   |
| 2.                                      | Выполнение практической работы № 2  | 11                         | 18         | 5-я неделя   |
| 3.                                      | Выполнение практической работы № 3  | 11                         | 18         | 8-я неделя   |
| 4.                                      | Выполнение практической работы № 4  | 11                         | 18         | 11-я неделя  |
| 5.                                      | Выполнение контрольной работы       | 10                         | 18         | 14-я неделя  |
| 6.                                      | Посещение занятий                   | 6                          | 10         | Свыше 75% посещенных занятий – 10, от 75 до 50% - 5, менее 50% - 0                     |
| 6.                                      | <b>Итого за работу в семестре:</b>  | <b>60</b>                  | <b>100</b> |  |
| <b>Промежуточная аттестация – зачет</b> |                                     |                            |            |  |
|   |                                     |                            |            | <i>Оценка «5» - 20 баллов;<br/>Оценка «4» - 15 баллов;<br/>Оценка «3» - 10 баллов.</i> |
|   | <b>Итоговые баллы по дисциплине</b> | <b>60</b>                  | <b>100</b> | Зачетная неделя  |

#### 16. Методические указания для обучающихся по освоению дисциплины

Приступая к изучению дисциплины, обучающемуся необходимо внимательно ознакомиться с тематическим планом занятий, списком рекомендованной литературы.

Следует уяснить последовательность выполнения индивидуальных учебных заданий. Самостоятельная работа обучающегося предполагает работу с учебной и научной литературой. Уровень и глубина усвоения дисциплины зависят от активной и систематической работы на лекциях, изучения рекомендованной литературы, решения задач и выполнения практических работ.

При изучении дисциплины обучающиеся:

- изучают рекомендованную учебную и научно-практическую и литературу;
- выполняют задания, предусмотренные для самостоятельной работы.

Основными видами аудиторной работы студентов являются лекции, практические занятия.

В ходе лекций преподаватель излагает и разъясняет основные, наиболее сложные понятия темы, а также связанные с ней теоретические и практические проблемы, дает рекомендации на семинарское занятие и указания к самостоятельной работе.

В учебном процессе, помимо чтения лекций и аудиторных занятий, используются активные и интерактивные формы (разбор конкретных ситуаций, выполнение практических работ, обсуждение отдельных разделов дисциплины, консультации). В сочетании с внеаудиторной работой это способствует формированию и развитию профессиональных навыков обучающихся.

Качество учебной работы обучающихся оценивается в соответствии с фондом оценочных средств и технологической картой дисциплины.

## **17. Обеспечение образования для инвалидов и лиц с ОВЗ**

Для обеспечения образования инвалидов и лиц с ограниченными возможностями здоровья реализация дисциплины может осуществляться в адаптированном виде, с учетом специфики освоения и дидактических требований, исходя из индивидуальных возможностей и по личному заявлению обучающегося.