



Основная профессиональная образовательная программа  
02.03.01 Математика и компьютерные науки  
(Математика и компьютерные науки)

Министерство науки и высшего образования Российской Федерации

**ИВАНОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ**

Кафедра информационных технологий и прикладной математики

ОДОБРЕНО:

Руководитель ОП

 Б.Я. Солон

(подпись)

« 1 » сентября 20 21 г.

**Рабочая программа дисциплины**  
**Интеллектуальный анализ данных**

Уровень высшего образования:	бакалавриат
Квалификация выпускника:	бакалавр
Направление подготовки:	02.03.01 Математика и компьютерные науки
Направленность (профиль) образовательной программы:	Математика и компьютерные науки



Основная профессиональная образовательная программа  
02.03.01 Математика и компьютерные науки  
(Математика и компьютерные науки)

---

### **1. Цели освоения дисциплины**

Целью освоения дисциплины является изучение теоретических аспектов технологии Data Mining, получение практических навыков по использованию инструментальных средств Data Mining, формирование профессиональных (ПК-1) компетенций, необходимых для осуществления профессиональной деятельности.

### **2. Место дисциплины в структуре ОП**

Дисциплина относится к части профессионального цикла, формируемой участниками образовательных отношений (Б1.В.07).

Изучению данной дисциплины должно предшествовать освоение дисциплины «Языки программирования» для возможности программной реализации задач.

Для освоения данной дисциплины обучающийся должен:

Знать основные положения и концепции прикладного и системного программирования, архитектуры компьютеров и сетей (в том числе и глобальных), современные языки программирования, технологии создания и эксплуатации программных продуктов и программных комплексов.

Уметь использовать их в профессиональной деятельности.

Иметь практические навыки разработки ПО.

### **3. Планируемые результаты обучения по дисциплине**

#### **3.1. Компетенции, формированию которых способствует дисциплина**

При освоении дисциплины формируются следующие компетенции в соответствии с ФГОС ВО по данному направлению подготовки:

а) профессиональные (ПК):

ПК-1. Способен применять в научно-исследовательской деятельности знания в области математики и (или) компьютерных наук.

#### **3.2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с индикаторами достижения формируемых компетенций**

В результате освоения дисциплины обучающийся должен:

Обладать расширенными знаниями, полученными в области математики и (или) компьютерных наук (ПК-1.1).

Уметь применять полученные знания при решении стандартных задач в собственной научно-исследовательской деятельности (ПК-1.2).

Иметь практический опыт научно-исследовательской деятельности в области математики и (или) компьютерных наук (ПК-1.3).

### **4. Объем и содержание дисциплины**

Объем дисциплины составляет 4 зачетные единицы (144 академических часа).

#### **4.1. Содержание дисциплины по разделам (темам), соотнесенное с видами и трудоемкостью занятий лекционно-семинарского типа**

Объем иной контактной работы и самостоятельной работы обучающегося по дисциплине указан в учебном плане образовательной программы.



Основная профессиональная образовательная программа  
02.03.01 Математика и компьютерные науки  
(Математика и компьютерные науки)

№ п/п	Разделы (темы) дисциплины	Семестр	Виды занятий, их объем (в ак. часах, по оч- ной форме обучения)		Формы текущего контроля успеваемости (по очной форме обучения)  Формы промежуточной ат- тестации
			Занятия лекцион- ного типа	Занятия се- минарского типа	
1.	Основные понятия Data Mining	7	6	4	
2.	Задачи Data Mining	7	30	28	
Итого за семестр:			36	32	Экзамен
Итого по дисциплине:			36	32	

#### 4.2. Развернутое описание содержания дисциплины по разделам (темам)

##### 1. Понятие Data Mining

##### 2. Что такое данные?

Набор данных и их атрибутов

Измерения

Шкалы

Типы наборов данных

Данные, состоящие из записей

Графические данные

Химические данные

Форматы хранения данных

Базы данных. Основные положения

Системы управления базами данных, СУБД

Классификация видов данных

Метаданные

##### 3. Задачи Data Mining. Информация и знания

Задачи Data Mining

Классификация задач Data Mining

Связь понятий

От данных к решениям

От задачи к приложению

Информация

Свойства информации

Требования, предъявляемые к информации

Знания

Сопоставление и сравнение понятий "информация", "данные", "знание"

##### 4. Задачи Data Mining. Классификация и кластеризация

Задача классификации

Процесс классификации

Методы, применяемые для решения задач классификации

Точность классификации: оценка уровня ошибок

Оценивание классификационных методов

Задача кластеризации

Оценка качества кластеризации

Процесс кластеризации



Основная профессиональная образовательная программа  
02.03.01 Математика и компьютерные науки  
(Математика и компьютерные науки)

---

- Применение кластерного анализа
- 5. Задачи Data Mining. Прогнозирование и визуализация
  - Задача прогнозирования
  - Сравнение задач прогнозирования и классификации
  - Прогнозирование и временные ряды
  - Тренд, сезонность и цикл
  - Точность прогноза
  - Виды прогнозов
  - Методы прогнозирования
  - Задача визуализации
- 6. Методы классификации и прогнозирования. Деревья решений
  - Преимущества деревьев решений
  - Процесс конструирования дерева решений
  - Критерий расщепления
  - Большое дерево не означает, что оно "подходящее"
  - Остановка построения дерева
  - Сокращение дерева или отсечение ветвей
  - Алгоритмы
  - Алгоритм CART
  - Алгоритм C4.5
  - Разработка новых масштабируемых алгоритмов
- 7. Методы классификации и прогнозирования. Метод опорных векторов. Метод "ближайшего соседа". Байесовская классификация
  - Метод опорных векторов
  - Линейный SVM
  - Метод "ближайшего соседа" или системы рассуждений на основе аналогичных случаев
  - Преимущества метода
  - Недостатки метода "ближайшего соседа"
  - Решение задачи классификации новых объектов
  - Решение задачи прогнозирования
  - Оценка параметра  $k$  методом кросс-проверки
  - Байесовская классификация
  - Байесовская фильтрация по словам
- 8. Методы классификации и прогнозирования. Нейронные сети
  - Элементы нейронных сетей
  - Архитектура нейронных сетей
  - Обучение нейронных сетей
  - Модели нейронных сетей
  - Персептрон
  - Программное обеспечение для работы с нейронными сетями
  - Классификация нейронных сетей
  - Подготовка данных для обучения
  - Выбор структуры нейронной сети
  - Карты Кохонена
  - Самоорганизующиеся карты (Self-Organizing Maps, SOM)
  - Задачи, решаемые при помощи карт Кохонена
  - Обучение сети Кохонена
- 9. Методы кластерного анализа. Иерархические методы



- Иерархические методы кластерного анализа
- Меры сходства
- Методы объединения или связи
- 10. Методы кластерного анализа. Итеративные методы.
  - Алгоритм k-средних (k-means)
  - Описание алгоритма
  - Проверка качества кластеризации
  - Алгоритм PAM ( partitioning around Medoids)
  - Предварительное сокращение размерности
  - Факторный анализ
  - Итеративная кластеризация в SPSS
  - Процесс кластерного анализа. Рекомендуемые этапы
  - Сложности и проблемы, которые могут возникнуть при применении кластерного анализа
  - Сравнительный анализ иерархических и неиерархических методов кластеризации
  - Новые алгоритмы и некоторые модификации алгоритмов кластерного анализа
  - Алгоритм BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies)
  - Алгоритм WaveCluster
  - Алгоритм CLARA (Clustering LARge Applications)
  - Алгоритмы Clarans, CURE, DBScan
- 11. Методы поиска ассоциативных правил
  - Часто встречающиеся приложения с применением ассоциативных правил:
  - Введение в ассоциативные правила
  - Часто встречающиеся шаблоны или образцы
  - Поддержка
  - Характеристики ассоциативных правил
  - Границы поддержки и достоверности ассоциативного правила
  - Методы поиска ассоциативных правил
  - Разновидности алгоритма Apriori
  - AprioriTid
  - AprioriHybrid
- 12. Способы визуального представления данных. Методы визуализации
  - Визуализация инструментов Data Mining
  - Визуализация Data Mining моделей
  - Методы визуализации
  - Представление данных в одном, двух и трех измерениях
  - Представление данных в 4 + измерениях
  - Параллельные координаты
  - "Лица Чернова"
  - Качество визуализации
  - Представление пространственных характеристик

## 5. Образовательные технологии

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине: технологии смешанного обучения, технология визуализации, интерактивные информационные технологии, технология активного (контекстного) обучения, технология проблемного обучения, технология развития критического мышления.



## **6. Учебно-методическое обеспечение самостоятельной работы обучающихся**

Самостоятельная работа студентов основывается на проработке лекционного материала и подготовке к экзамену. Контроль качества изучения материала обучающимся основывается на проведении устного экзамена по изученным разделам дисциплины.

Предполагается также выполнение лабораторных работ, заключающихся в разработке программных систем. Задания для лабораторных работ размещаются преподавателем в системе «Мой университет» <https://uni.ivanovo.ac.ru>. Контроль качества выполнения лабораторных работ обучающимся проводится путем устной беседы с преподавателем по выполненным заданиям.

## **7. Характеристика оценочных средств для текущего контроля успеваемости и промежуточной аттестации обучающихся по дисциплине**

По итогам обучения проводится устный экзамен. Вопросы к экзамену приведены в Приложении 2 к настоящей рабочей программе.

Критерии оценки к экзамену: «отлично» – положительный ответ на три вопроса из трех, приведенных в билете; «хорошо» – положительный ответ, содержащий небольшие неточности или недочеты, на три теоретических вопроса из трех, приведенных в билете, возможно, с негрубыми ошибками; «удовлетворительно» – положительный ответ на два из заданных трех вопросов, возможно, с негрубыми ошибками; «неудовлетворительно» – нет положительного ответа по крайней мере на два вопроса из трех в билете.

## **8. Учебно-методическое и информационное обеспечение дисциплины**

Основная литература:

1. Замятин, А.В. Интеллектуальный анализ данных / А.В. Замятин ; Министерство образования и науки Российской Федерации, Национальный исследовательский Томский государственный университет. – Томск : Издательский Дом Томского государственного университета, 2016. – 119 с. : табл., граф., схем. – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=464472>

2. Нестеров, С.А. Интеллектуальный анализ данных средствами MS SQL Server 2008 / С.А. Нестеров. – 2-е изд., испр. – Москва : Национальный Открытый Университет «ИНТУИТ», 2016. – 338 с. : ил. – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=429083>

3. Крутиков, В.Н. Анализ данных / В.Н. Крутиков, В.В. Мешечкин ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Кемеровский государственный университет». – Кемерово : Кемеровский государственный университет, 2014. – 138 с. : ил. – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=278426>

Дополнительная литература:

Ресурсы информационно-телекоммуникационной сети «Интернет»:

Техническая документация <https://docs.microsoft.com/ru-ru/>

Система электронной поддержки образовательного процесса «Мой университет» <https://uni.ivanovo.ac.ru>

Профессиональные базы данных и информационно-справочные системы:

ЭБС «Университетская библиотека онлайн» [www.biblioclub.ru](http://lib.ivanovo.ac.ru/index.php/polnotekstovye-resursy/ebs-universitetskaya-biblioteka);  
<http://lib.ivanovo.ac.ru/index.php/polnotekstovye-resursy/ebs-universitetskaya-biblioteka>

Электронная библиотека ИвГУ <http://lib.ivanovo.ac.ru/index.php/polnotekstovye-resursy/elibnew>

Электронный каталог НБ ИвГУ <http://lib.ivanovo.ac.ru/index.php/ek>



Основная профессиональная образовательная программа  
02.03.01 Математика и компьютерные науки  
(Математика и компьютерные науки)

---

Программное обеспечение: операционная система Microsoft Windows, пакет офисных программ Microsoft Office и(или) LibreOffice, интернет-браузер Microsoft Edge и(или) Yandex Browser, редактор исходного кода Visual Studio Code.

## **9. Материально-техническое обеспечение дисциплины**

Учебные аудитории:

- для проведения занятий лекционного типа с комплектом специализированной учебной мебели и техническими средствами обучения, служащими для предоставления учебной информации большой аудитории;
- для проведения занятий семинарского типа, консультаций, текущего контроля и промежуточной аттестации с комплектом специализированной учебной мебели и техническими средствами обучения.

Помещение для самостоятельной работы, оснащенное комплектом специализированной учебной мебели, компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в ЭИОС.

Демонстрационное оборудование и учебно-наглядные пособия для занятий лекционного типа, обеспечивающие тематические иллюстрации: презентации, видеоматериалы, таблицы и схемы, визуализирующие лекционный материал, демонстрационные образцы написания компьютерных программ.



Основная профессиональная образовательная программа  
02.03.01 Математика и компьютерные науки  
(Математика и компьютерные науки)

---

**Автор(ы) рабочей программы дисциплины:** доцент кафедры ИТ и прикладной математики, кандидат экономических наук Степович-Цветкова Г.С.

Программа рассмотрена и утверждена на заседании кафедры информационных технологий и прикладной математики

« 30 » августа 2021 г., протокол № 1

Программа обновлена

протокол заседания кафедры № 1 от « 1 » сентября 2023 г.

Согласовано:

Руководитель ОП  Туртин Д.В.  
(подпись)

Программа обновлена

протокол заседания кафедры № \_\_\_\_\_ от « \_\_\_\_\_ » \_\_\_\_\_ 20 \_\_\_\_ г.

Согласовано:

Руководитель ОП \_\_\_\_\_ Туртин Д.В.  
(подпись)

Программа обновлена

протокол заседания кафедры № \_\_\_\_\_ от « \_\_\_\_\_ » \_\_\_\_\_ 20 \_\_\_\_ г.

Согласовано:

Руководитель ОП \_\_\_\_\_ Туртин Д.В.  
(подпись)