



Основная профессиональная образовательная программа  
02.03.02 Фундаментальная информатика и информационные технологии  
Фундаментальная информатика и информационные технологии

Министерство науки и высшего образования Российской Федерации

**ИВАНОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ**

Кафедра прикладной математики и компьютерных наук

ОДОБРЕНО:

Руководитель ОП

 Е.В. Соколов  
(подпись)

« 19 » июня 20 19 г.

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

Уровень высшего образования:	бакалавриат
Квалификация выпускника:	бакалавр
Направление подготовки:	02.03.02 Фундаментальная информатика и информационные технологии
Направленность (профиль) образовательной программы:	Фундаментальная информатика и информационные технологии

Иваново



## **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 правоотношений (Б1.В).

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

Для освоения данной дисциплины обучающийся должен:

Знать основные положения и концепции в области программирования, архитектуру языков программирования, теории коммуникации, основную терминологию, быть знакомым с содержанием Единого Реестра Российских программ (ОПК-2.1). Знать методы теории алгоритмов, методы системного и прикладного программирования (ОПК-3.1). Знать типовые решения, библиотеки программных модулей, шаблоны, классы объектов, используемые при разработке программного обеспечения; методы и средства проектирования программного обеспечения, методы и средства проектирования программных интерфейсов (ПК-3.1).

Уметь анализировать типовые языки программирования, составлять программы (ОПК 2.2). Уметь соотносить знания в области программирования, интерпретацию прочитанного, определять и создавать информационные ресурсы средств тестирования систем (ОПК 3.2). Уметь использовать существующие типовые решения и шаблоны проектирования программного обеспечения; применять методы и средства проектирования программного обеспечения, структур данных, программных интерфейсов (ПК-3.2).

Иметь практический опыт решения задач анализа, интеграции различных типов программного обеспечения (ОПК-2.3). Иметь практический опыт применения разработки программного обеспечения (ОПК-3.3). Иметь практический опыт проектирования структур данных, проектирования программных интерфейсов (ПК-3.3).

## **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 прикладной математики и (или) информационных технологий

### **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) информационных технологий (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 прикладной математики и (или) информационных технологий (ПК-1.3).



Основная профессиональная образовательная программа  
02.03.02 Фундаментальная информатика и информационные технологии  
Фундаментальная информатика и информационные технологии

#### 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.

№ п/п	Разделы (темы) дисциплины	Семестр	Виды занятий, их объем (в ак. часах, по очной форме обучения)		Формы текущего контроля успеваемости (по очной форме обучения)
			Занятия лекционного типа	Занятия семинарского типа	Формы промежуточной аттестации
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



Основная профессиональная образовательная программа  
02.03.02 Фундаментальная информатика и информационные технологии  
Фундаментальная информатика и информационные технологии

---

- Методы, применяемые для решения задач классификации
- Точность классификации: оценка уровня ошибок
- Оценивание классификационных методов
- Задача кластеризации
- Оценка качества кластеризации
- Процесс кластеризации
- 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. 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> (дата обращения: 30.06.2019). – Библиогр. в кн. – ISBN 978-5-8353-1770-7. – Текст : электронный.

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> (дата обращения: 30.06.2019). – Библиогр. в кн. – Текст : электронный.

Дополнительная литература:

1. Агалаков, С.А. Статистические методы анализа данных : [16+] / С.А. Агалаков ; Министерство образования и науки РФ, Омский государственный университет им. Ф. М. Достоевского. – Омск : ОмГУ им. Ф.М. Достоевского, 2017. – 92 с. : табл., граф., схем., ил. – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=562918> (дата обращения: 30.06.2019). – Библиогр. в кн. – ISBN 978-5-7779-2187-1. – Текст : электронный.

2. Жуковский, О.И. Информационные технологии и анализ данных / О.И. Жуковский ; Министерство образования и науки Российской Федерации, Томский Государственный Университет



Основная профессиональная образовательная программа  
02.03.02 Фундаментальная информатика и информационные технологии  
Фундаментальная информатика и информационные технологии

---

Систем Управления и Радиоэлектроники (ТУСУР). – Томск : Эль Контент, 2014. – 130 с. : схем., ил. – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=480500> (дата обращения: 30.06.2019). – Библиогр.: с. 126. – ISBN 978-5-4332-0158-3. – Текст : электронный.

3. Чубукова, И.А. Data Mining / И.А. Чубукова. – 2-е изд., испр. – Москва : Интернет-Университет Информационных Технологий, 2008. – 383 с. – (Основы информационных технологий). – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=233055> (дата обращения: 30.06.2019). – ISBN 978-5-94774-819-2. – Текст : электронный.

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://www.biblioclub.ru)

Электронная библиотека ИвГУ <http://lib.ivanovo.ac.ru>

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

Программное обеспечение: операционная система Windows, офисный пакет Microsoft Office и(или) LibreOffice, Интернет-браузер Internet Explorer и(или) Microsoft Edge и(или) Yandex Browser, кроссплатформенная среда разработки Code::Blocks, интегрированная среда разработки приложений Visual Studio 2013.

## **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  
02.03.02 Фундаментальная информатика и информационные технологии  
Фундаментальная информатика и информационные технологии

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

Программа рассмотрена и утверждена на заседании кафедры прикладной математики и компьютерных наук

« 13 » июня 20 19 г., протокол № 11

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

Согласовано:

Руководитель ОП \_\_\_\_\_  
(подпись)

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

Согласовано:

Руководитель ОП \_\_\_\_\_  
(подпись)

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

Согласовано:

Руководитель ОП \_\_\_\_\_  
(подпись)