

Министерство науки и высшего образования Российской Федерации  
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ (НИ ТГУ)

Институт экономики и менеджмента

УТВЕРЖДАЮ:  
Директор Института  
экономики и менеджмента

  
\_\_\_\_\_ Е.В. Нехода  
Институт  
экономики и менеджмента

« 20 » \_\_\_\_\_ 04 \_\_\_\_\_ 20 23 г.

Рабочая программа дисциплины

**Бизнес-аналитика и Data Mining**

по направлению подготовки

**38.04.02 Менеджмент**

Направленность (профиль) подготовки:

**«Менеджмент»**

Форма обучения

**Очная**

Квалификация

**Магистр**

Год приема

**2023**

Код дисциплины в учебном плане: Б1.В.ДВ.03.02

СОГЛАСОВАНО:

Руководитель ОП

\_\_\_\_\_ М.В. Герман

Председатель УМК

\_\_\_\_\_ М.В. Герман

Томск – 2023

## **1. Цель и планируемые результаты освоения дисциплины (модуля)**

Целью освоения дисциплины является формирование следующих компетенций:

– ПК-2 – Разработка и усовершенствование процессной архитектуры организации.

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

– ПК-2.1. Систематизирует информацию о процессной архитектуре организации.

## **2. Задачи освоения дисциплины**

– Освоить математический аппарат, лежащий в основе методов интеллектуального анализа данных.

– Научиться применять аналитические платформы и библиотеки для решения практических бизнес-задач, связанных с анализом данных.

– Анализировать бизнес-процессы организации, планировать изменения бизнес-процессов в организации и обосновывать экономическую эффективность предлагаемых изменений.

## **3. Место дисциплины (модуля) в структуре образовательной программы**

Дисциплина относится к части образовательной программы, формируемой участниками образовательных отношений, предлагается обучающимся на выбор.

## **4. Семестр(ы) освоения и форма(ы) промежуточной аттестации по дисциплине**

Семестр 3, зачет.

## **5. Входные требования для освоения дисциплины**

Для успешного освоения дисциплины требуются компетенции, сформированные в ходе освоения образовательных программ предшествующего уровня образования.

Для успешного освоения дисциплины требуются результаты обучения по следующим дисциплинам: «Математический анализ», «Линейная алгебра», «Теория вероятностей», «Математическая статистика», «Методы оптимизации», «Базы данных».

## **6. Язык реализации**

Русский

## **7. Объем дисциплины (модуля)**

Общая трудоемкость дисциплины составляет 3 з.е., 108 часов, из которых:

– лекции: 8 ч.;

– практические занятия: 20 ч.;

Практическая подготовка – 20 ч.

Объем самостоятельной работы студента определен учебным планом.

## **8. Содержание дисциплины (модуля), структурированное по темам**

Тема 1. Введение в бизнес-аналитику и DataMining.

Основные понятия и терминология анализа данных. Отраслевой стандарт CRISP-DM. Классификация задач машинного обучения. Классификация аналитических систем. Примеры бизнес-задач, для решения которых может быть использовано машинное обучение.

Тема 2. Базы и хранилища данных.

Реляционные базы данных. Нормализация данных. Языки запросов QBE и SQL. Хранилища данных. Технология OLAP.

Тема 3. Задача регрессии.

Постановка задачи. Методы решения задачи регрессии на основе линейных регрессионных моделей, нейронных сетей и деревьев решений. Оценка качества регрессионных моделей.

Тема 4. Задача классификации.

Постановка задачи классификации. Методы решения задачи классификации на основе логистической регрессии, нейронных сетей, деревьев решений. Методы оценки качества классификаторов.

## **9. Текущий контроль по дисциплине**

Текущий контроль по дисциплине проводится путем контроля посещаемости и выполнения домашних заданий и фиксируется в форме контрольной точки не менее одного раза в семестр.

## **10. Порядок проведения и критерии оценивания промежуточной аттестации**

**Зачет** проводится в письменной форме по билетам и в форме защиты проекта, демонстрирующего освоение компетенции – ПК-2 (Разработка и усовершенствование процессной архитектуры организации), индикатор достижения ПК-2.1 (Систематизирует информацию о процессной архитектуре организации).

Билет содержит один теоретический вопрос. Продолжительность зачета 1,5 часа.

Примерный перечень теоретических вопросов:

1. Перечислите основные элемента стандарта CRISP-DM. Приведите примеры бизнес-задач, для решения которых требуются базы и хранилища данных. Дайте классификацию задач DataMining.

2. Дайте определение реляционной базы данных. Опишите назначение процедуры нормализации данных. Опишите назначение и возможности языков запросов QBE и SQL.

3. Дайте определение хранилища данных. Опишите назначение моделей данных «звезда» и «снежинка». Опишите назначение и возможности OLAP-систем.

4. Сформулируйте задачу регрессии и кратко перечислите подходы к её решению. Опишите алгоритм решения задачи регрессии на основе модели множественной линейной регрессии.

5. Сформулируйте задачу классификации и кратко перечислите подходы к её решению. Опишите алгоритм решения задачи классификации на основе модели множественной логистической регрессии. Как интерпретируются коэффициенты модели логистической регрессии?

Проект, выносимый на защиту предполагает:

1. Самостоятельный выбор темы проекта.
2. Сбор и подготовку исходных данных по теме проекта.
3. Формулировку проверяемых гипотез.
4. Разработку решения.
5. Анализ результатов и их бизнес-интерпретация.
6. Подготовку презентации и доклад с описанием основных элементов решения и полученных результатов.

Результаты зачета определяются оценками «зачтено», «не зачтено».

Оценка «зачтено» выставляется при выполнении следующих условий:

1. Ответ на теоретический вопрос является полным, содержит постановку проблемы, описание методов её решения с указанием их достоинств, недостатков и условий применимости.

2. Практический проект направлен на решение задачи определения развития организации, является практически полезным. Проект должен содержать оценку текущего состояния проблемы (формулировку в бизнес-терминах), формальную постановку задачи,

описание доступных источников данных (доступность, качество), решение проблемы, описание процедуры внедрения полученного решения, включая оценку бизнес-возможностей организации, необходимых для проведения стратегических изменений, и достигаемого экономического эффекта.

Оценка «не зачтено» выставляется при выполнении любого из следующих условий:

1. Ответ на теоретический вопрос не является полным, не содержит постановку проблемы, описания методов её решения с указанием их достоинств, недостатков и условий применимости.

2. Практический проект, направленный на решение задачи определения развития организации, не содержит оценку текущего состояния проблемы (формулировку в бизнес-терминах), формальную постановку задачи, решения проблемы, описание процедуры внедрения полученного решения, включая оценку бизнес-возможностей организации, необходимых для проведения стратегических изменений, и достигаемого экономического эффекта.

## **11. Учебно-методическое обеспечение**

а) электронный учебный курс по дисциплине в электронном университете «Moodle»;

б) оценочные материалы текущего контроля и промежуточной аттестации по дисциплине.

## **12. Перечень учебной литературы и ресурсов сети Интернет**

а) основная литература:

– Дадян Э. Современные базы данных. Основы. Часть 1 : Учебное пособие / Финансовый университет при Правительстве Российской Федерации. - Москва : ООО "Научно-издательский центр ИНФРА-М", 2017. - 88 с.

– Нестеров С. А. Базы данных : Учебник и практикум для вузов / Нестеров С. А.. - Москва : Юрайт, 2022. - 230 с

– О'Нил К. Data Science : инсайдерская информация для новичков, включая язык R / Кэти О'Нил, Рэйчел Шатт ; [пер. с англ. И. Пальти и др.]. - Санкт-Петербург [и др.] : Питер, 2019. - 362, [2] с

– Флах П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / Петер Флах ; [пер. с англ. А. А. Слинкина]. - Москва : ДМК Пресс, 2015. - 399 с.

– Foundations and novel approaches in data mining / edited by Tsau Young Lin [a. o.]. - Berlin [a. o.] : Springer, 2006. - х, 376 p.: ill. - ( Studies in computational intelligence / ed. by Janusz Kasprzyk ; Vol. 9: )

б) дополнительная литература:

– Келлехер Д. Д. Основы машинного обучения для аналитического прогнозирования : алгоритмы, рабочие примеры и тематические исследования : пер. с англ. / Джон Д. Келлехер, Брайан Мак-Нейми, Аоифе д'Арси. - Санкт-Петербург [и др.] : Диалектика, 2019. - 656 с.

– Алдохина О. И. Информационно-аналитические системы и сети : учебное пособие. Ч. 1 / О. И. Алдохина, О. Г. Басалаева ; Кемеровский гос. ун-т культуры и искусств, Ин-т культуры. - Кемерово : КемГУКИ, 2010. - 147 с.

– Финн В. К. Искусственный интеллект : методология, применение, философия / В. К. Финн ; науч. ред. М. А. Михеенкова ; Российская акад. наук, Всероссийский ин-т науч. и техн. информации. - Москва : Красанд, 2011. - 447 с.

– Абдикеев Н. М. Интеллектуальные информационные системы / Н. М. Абдикеев ; под общ. ред. К. И. Курбакова ; Рос. экон. акад. им. Г. В. Плеханова. - М. : КОС ИНФ, 2003. - 187 с.

– Осипов Г. С. Лекции по искусственному интеллекту / Г. С. Осипов ; Российская акад. наук, Ин-т системного анализа. - Изд. 2-е, испр. и доп.. - Москва : ЛИБРОКОМ, 2013. - 267 с.

в) ресурсы сети Интернет:

- <http://www.machinelearning.ru/wiki/>
- <http://citforum.ru>

### **13. Перечень информационных технологий**

а) лицензионное и свободно распространяемое программное обеспечение:

- Microsoft Office Standart 2013 Russian: пакет программ. Включает приложения: MS Office Word, MS Office Excel, MS Office PowerPoint, MS Office On-eNote, MS Office Publisher, MS Outlook, MS Office Web Apps (Word Excel MS PowerPoint Outlook);
- публично доступные облачные технологии (Google Docs, Яндекс диск и т.п.).

б) информационные справочные системы:

- Электронный каталог Научной библиотеки ТГУ – <http://chamo.lib.tsu.ru/search/query?locale=ru&theme=system>
- Электронная библиотека (репозиторий) ТГУ – <http://vital.lib.tsu.ru/vital/access/manager/Index>
- ЭБС Лань – <http://e.lanbook.com/>
- ЭБС Консультант студента – <http://www.studentlibrary.ru/>
- Образовательная платформа Юрайт – <https://urait.ru/>
- ЭБС ZNANIUM.com – <https://znanium.com/>
- ЭБС IPRbooks – <http://www.iprbookshop.ru/>

### **14. Материально-техническое обеспечение**

Аудитории для проведения занятий лекционного типа.

Аудитории для проведения практических занятий, индивидуальных и групповых консультаций, текущего контроля и промежуточной аттестации.

Помещения для самостоятельной работы, оснащенные компьютерной техникой и доступом к сети Интернет, в электронную информационно-образовательную среду и к информационным справочным системам.

### **15. Информация о разработчиках**

Богданов Александр Леонидович, к.т.н., доцент, ИЭМ ТГУ, доцент кафедры Информационных технологий и бизнес-аналитики.