МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ «ТУЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

Институт прикладной математики и компьютерных наук

Кафедра информационной безопасности

Утверждено на за заседании кафедры
«Информационная безопасность»

«20» января 2023г., протокол №5

Заведующий кафедрой
_____ А.А.Сычугов

РАБОЧАЯ ПРОГРАММА

учебной дисциплины (модуля)

ИЗВЛЕЧЕНИЕ ЗНАНИЙ В ИНФОРМАЦИОННЫХ СИСТЕМАХ

Уровень профессионального образования: *высшее образование – магистратура*

Направление подготовки: *09.03.02 Информационные системы и технологии*
Профиль: *Информационные системы*

Квалификация (степень) выпускника: бакалавр

Форма обучения: очная

Идентификационный номер образовательной программы: 090302-01-19

Тула 2023

ЛИСТ
согласования рабочей программы дисциплины *(модуля)*

Рабочая программа учебной дисциплины *«Представление знаний в информационных системах»* разработана *профессором М.Ю. Богатыревым* и обсуждена на заседании кафедры информационной безопасности института прикладной математики и компьютерных наук (протокол заседания кафедры №_5_ от «_20_» __января_ 2023 г.)

Разработчик рабочей программы дисциплины _____.
личная подпись(и)

СОГЛАСОВАНО:
(согласуется в случае реализации дисциплины (модуля) для образовательных программ (профилей, специализаций) других кафедр)

Заведующий кафедрой _____ _____ _____ _____
наименование кафедры *личная подпись* *расшифровка подписи* *дата*

Заведующий кафедрой _____ _____ _____ _____
наименование кафедры *личная подпись* *расшифровка подписи* *дата*

Заведующий кафедрой _____ _____ _____ _____
наименование кафедры *личная подпись* *расшифровка подписи* *дата*

СОДЕРЖАНИЕ

1 Цель и задачи освоения учебной дисциплины (модуля) 4
2 Место учебной дисциплины (модуля) в структуре образовательной программы 4
3 Перечень планируемых результатов обучения по дисциплине (модулю) 4
4 Содержание и структура учебной дисциплины (модуля) 5
4.1 Содержание разделов учебной дисциплины (модуля) 5
4.2 Распределение часов по семестрам и видам занятий 5
4.3 Темы, выносимые на лекционные занятия 7
4.4 Лабораторные работы 8
4.5 Практические занятия (семинары) 8
4.6 Самостоятельная работа студента 8
5 Образовательные технологии 9
6 Материально-техническое обеспечение дисциплины (модуля) 9
6.1 Требования к аудиториям (помещениям, местам) для проведения занятий 9
6.2 Требования к оборудованию рабочих мест преподавателя и обучающихся 9
6.3 Требования к специализированному оборудованию 9
6.4 Требования к программному обеспечению учебного процесса 9
7 Порядок проведения текущего контроля и промежуточных аттестаций. Шкалы оценок 9
7.1 Шкала академических оценок освоения дисциплины (модуля) 9
7.2 Система оценки достижений обучающегося по дисциплине (модулю) 10
7.3 Система оценки компетенций или их элементов, сформированных у обучающихся в ходе освоения дисциплины (модуля) 10
8 Оценочные средства для текущего контроля успеваемости и промежуточных аттестаций обучающихся 11
9 Учебно-методическое обеспечение дисциплины (модуля) 12
9.1 Основная литература 12
9.2 Дополнительная литература 12
9.3 Периодические издания 13
9.4 Интернет-ресурсы 13
9.5 Методические указания к лабораторным занятиям 13
9.6 Методические указания к практическим занятиям 13
9.7 Методические указания к курсовому проектированию и другим видам самостоятельной работы 13

1 Цель и задачи освоения учебной дисциплины (модуля)

Целями освоения дисциплины «Представление знаний в информационных системах» являются:

- приобретение студентами знаний в области современных информационных технологий извлечения знаний;
- получение практического опыта эксплуатации систем извлечения знаний.

Задачами освоения дисциплины являются:

- изучение современных моделей знаний, применяемых в информационных системах.
- изучение алгоритмов работы систем извлечения знаний.
- приобретение практических навыков работы с инструментальными средствами инженерии знаний.

Успешное изучение дисциплины возможно при наличии у студентов компьютерной грамотности и владения основными понятиями и методами теории систем, основами кибернетики, языками программирования, принципами построения баз данных.

2 Место учебной дисциплины (модуля) в структуре образовательной программы

Данная дисциплина относится к вариативной части В.06.

Для успешного освоения учебной дисциплины необходимы знания, умения и навыки, сформированные дисциплинами образовательной программы 09.03.02 «Информационные системы и технологии», профиля «Информационные системы»: Б.04 «Математика», Б.10 «Базы данных», Б.16 «Программирование» и Б.17 «Теория информационных систем», которые совместно формируют целостную картину знаний в данной области информационных технологий.

Знания, умения и навыки, сформированные при изучении данной учебной дисциплины, необходимы для успешного освоения следующих дисциплин: В.01 «Интеллектуальные технологии», В.03 «Технология системного моделирования».

3 Перечень планируемых результатов обучения по дисциплине (модулю)

Процесс изучения дисциплины направлен на формирование элементов следующих компетенций в соответствии с ФГОС ВО по данному направлению подготовки:

- общекультурных компетенций (ОК):

- способностью совершенствовать и развивать свой интеллектуальный и общекультурный уровень (ОК-1);
- способностью к самостоятельному обучению новым методам исследования, к изменению научного и научно-производственного профиля своей профессиональной деятельности (ОК-3);
- способностью заниматься научными исследованиями (ОК-4);

- общепрофессиональных компетенций (ОПК):

- способностью воспринимать математические, естественнонаучные, социально-экономические и профессиональные знания, умением самостоятельно приобретать, развивать и применять их для решения нестандартных задач, в том числе, в новой или незнакомой среде и в междисциплинарном контексте (ОПК-1);

- профессиональных компетенций (ПК):

- знанием методов научных исследований и владение навыками их проведения (ПК-2);

- знанием методов оптимизации и умение применять их при решении задач профессиональной деятельности (ПК-3);
- владением существующими методами и алгоритмами решения задач распознавания и обработки данных (ПК-4);

В результате изучения дисциплины обучающийся должен:

Знать:

1. структуру, параметры, математические формы современных моделей знаний (ОПК-1);
2. принцип действия систем извлечения знаний (ПК-2);
3. теоретические основы методов извлечения знаний (ПК-4);

Уметь:

1. применять методы извлечения знаний для решения практических задач (ПК-2);
2. пользоваться методами и программами для решения задач извлечения знаний (ПК-2);

Иметь навыки:

1. разработки современного ПО (ПК-4);
2. применения современного ПО в прикладных областях (ПК-2);
3. навыками разработки прикладных алгоритмов анализа данных и знаний (ПК-4);
4. навыками анализа результатов обработки данных (ОК-1, ПК-4).

4 Содержание и структура учебной дисциплины (модуля)

4.1 Содержание разделов учебной дисциплины (модуля)

1. **Введение**. Краткая характеристика дисциплины, ее цели, задачи, объем, содержание, порядок изучения материала, связь с другими дисциплинами учебного плана и место в подготовке по специальности. Теоретическая и практическая составляющие. Формы самостоятельной работы. Характеристика учебной литературы. Контрольные мероприятия.

1.1. Основные понятия баз знаний, систем извлечения знаний, систему правления базами знаний.

2. **Знания и данные**.

2.1. Свойства знаний и отличие знаний от данных.
2.2. Модели знаний: модели в виде объектов, модели в виде правил.
2.3. Числовые и нечисловые данные как источники знаний.
2.4. Базы данных и базы знаний

3. Модели знаний

3.1. Ассоциативные правила.
3.2. Деревья решений.
3.3. Кластерные модели знаний.
3.4. Семантические сети.
3.5. Онтологии.

4. Системы искусственного интеллекта, основанные на знаниях.

4.1. Прикладные ИИ – системы, основанные на знаниях.
4.2. Экспертные системы. Области применения и решаемые ими задач.
4.3. Системы общения на естественном языке и речевой ввод-вывод.

4.4. Системы поддержки принятия решений.
4.5. Применение систем извлечения знаний в сети Internet.

5. Ассоциативные правила

5.1. Построение ассоциативных правил. Алгоритм «Априори».
5.2. Другие алгоритмы построения ассоциативных правил
5.3. Применение ассоциативных правил.

6. Деревья решений.

6.1. Построение деревьев решений. Алгоритм «ID3».
6.2. Другие алгоритмы построения деревьев решений (C4.5)
6.3. Применение деревьев решений.

7. Семантические сети.

7.1. Основные понятия семантических сетей.
7.2. Типы отношений в семантических сетях.
7.3. Концептуальные графы и их применение.

8. Системы понимания текста Text Mining

8.1. Особенности и виды текстовых данных. Понятие контекста.
8.2. Задачи Text Mining
8.3. Извлечение ключевых слов и словосочетаний из текстов.
8.4. Классификация текстовых документов.
8.5. Кластеризация текстовых документов.

9. **Анализ формальных понятий в задачах извлечения знаний**.

9.1. Формальный контекст
9.2. Решетки понятий.
9.3. Методы построения решеток понятий
9.4. Применение решеток понятий в системах Text Mining.

10. **Эволюционные алгоритмы в задачах извлечения знаний**.

10.1. Понятие эволюционного алгоритма.
10.2. Генетические алгоритмы. Генетические операторы.
10.3. Преимущества и недостатки генетических алгоритмов.

4.2 Распределение часов по семестрам и видам занятий

Общая трудоемкость дисциплины (модуля) составляет 4 зачетных единиц (144 академических часа), в том числе:[1]

Номер семестра	Формы промежуточной аттестации	Общий объем в зачетных единицах	Общий объем в академических часах	Объем контактной работы в академических часах						Объем самостоятельной работы в академических часах
				Лекционные занятия	Практические (семинарские) занятия	Лабораторные работы	Клинические практические занятия	Консультации	Промежуточная аттестация	
Очная форма обучения										
6	ДЗ	6	180	32	-	32	-	-	0,25	115,75
Итого	–	6	180	32	-	32	-	-	0.25	115.75

4.3 Темы, выносимые на лекционные занятия

№№ лекций	№№ разделов дисциплины (модуля), выносимых на лекции	Кол-во академических часов
Очная форма обучения		
1-й семестр		
1.	1-2	2
2.	3-4	2
3.	5.1,5.2	2
4.	6.1- 6.3, 7.1- 7.3	2
5.	8.1	2
6.	8.2	2
7.	8.3	2
8.	8.4	2
9.	8.5	2
10.	9.1	2
11.	9.2	2
12.	9.3	2
13.	9.4	2
14.	10.1	2
15.	10.2	2
16.	10.3	2
	Итого	32

4.4 Лабораторные работы

№ ЛР	№№ разделов дисциплины (модуля)	Наименование лабораторных работ	Кол-во академических часов
Очная форма обучения			
1-й семестр			
1.	2	Вводное занятие. Техника безопасности.	2
2.	3	Моделирование множеств списками	4
3.	6	Исследование алгоритмов построения ассоциативных правил	4
4.	7	Исследование алгоритмов построения деревьев решений	4
5.	9.2	Построение и исследование решеток понятий	4
6.	3.5	Построение и исследование онтологий	4
7.	11	Исследование свойств генетических алгоритмов	4
8.	8.3, 9.5, 10	Применение генетических алгоритмов в задаче кластеризации	4
9.		Итоговое занятие в семестре	2
		Итого	32

4.5 Практические занятия (семинары)

Учебным планом не предусмотрено.

4.6 Самостоятельная работа студента

№ п/п	Наименование видов самостоятельной работы	Трудоемкость (в академических часах)	Методические материалы
Очная форма обучения			
1-й семестр			
	Подготовка к выполнению и к защите лабораторных работ	18	[1] (см. п. 9.5)
	Самостоятельная дополнительная проработка отдельных тем дисциплины.	24	[1] (см. п. 9.7)
	Курсовое проектирование	20	
	Подготовка к текущим и промежуточной аттестациям	34	[1-4] (см. п. 9.1, 9.2)
	Итого	96	

4.7. Примерная тематика курсовых проектов (работ)

Семестр №1

1. Исследование алгоритма Apriori на категорийных данных.
2. Исследование алгоритма Apriori на числовых данных.
3. Применение алгоритма Apriori к текстовым данным.
4. Исследование алгоритма C4.5 на категорийных данных.
5. Исследование алгоритма C4.5 на числовых данных.
6. Исследование явления элитизма генетических алгоритмов
7. Исследование строящих шаблонов генетических алгоритмов.
8. Генерация множества импликаций решеток понятий

5 Образовательные технологии

Интерактивные образовательные технологии, используемые в аудиторных занятиях

Семестр	Вид занятия (Л, ПР, ЛР)	Используемые интерактивные образовательные технологии	Количество аудиторных часов
Очная форма обучения			
	Л	Презентации.	12
	ПР	Структурированная групповая дискуссия - это совместное обсуждение и анализ проблемной ситуации, вопроса или задачи.	6
	ЛР	Компьютерные симуляции	12
		Итого	30

6 Материально-техническое обеспечение дисциплины (модуля)

6.1 Требования к аудиториям (помещениям, местам) для проведения занятий

Для проведения лекционных занятий по дисциплине требуется аудитория оснащенная видеопроектором, настенным экраном.

Для проведения практических и лабораторных занятий требуется компьютерный класс.

6.2 Требования к оборудованию рабочих мест преподавателя и обучающихся

Рабочее место преподавателя должно быть оснащено видеопроектором, ноутбуком.

6.3 Требования к специализированному оборудованию

Специализированное оборудование не требуется

6.4 Требования к программному обеспечению учебного процесса:

- математический программный пакет *Mathematica*;
- свободно распространяемая программа для проведения тестирования *MyTest*.

7 Порядок проведения текущего контроля и промежуточных аттестаций. Шкалы оценок

7.1 Шкала академических оценок освоения дисциплины (модуля)

Виды оценок	Оценки			
Академическая оценка по 100-балльной шкале (экзамен, дифференцированный зачет, зачет)	0…39	40…60	61…80	81…100
Академическая оценка по 4-балльной шкале (экзамен, дифференцированный зачет)	Неудовлетворительно	Удовлетворительно	Хорошо	Отлично
Академическая оценка по 2-балльной шкале (зачет)	Не зачтено	Зачтено		

7.2 Система оценки достижений обучающегося по дисциплине (модулю)

№ п/п	Виды учебных мероприятий *(указываются мероприятия согласно учебному плану)*	Наименование учебных мероприятий	Максимальное количество баллов за мероприятие*
		1-й семестр	
1	Посещение лекционных занятий *(баллы начисляться за весь курс лекций)*	Лекция №1	0,5
		Лекция №2	0,5
		Лекция №3	0,5
		Лекция №4	0,5
		Лекция №5	0,5
		Лекция №6	0,5
2	Выполнение и защита лабораторных работ	Лабораторная работа №1	4
		Лабораторная работа №2	4
		Лабораторная работа №3	4
		Лабораторная работа №4	4
		Лабораторная работа №5	4
		Лабораторная работа №6	4
3	Работа на практических занятиях, семинарах, коллоквиумах и т.п.	Практическое занятие №1	2
		Практическое занятие №2	2
		Практическое занятие №3	2
		Практическое занятие №4	2
		Практическое занятие №5	2
		Практическое занятие №6	2
	Контрольные мероприятия	Тесты	
4	Самостоятельная работа студента	Подготовка реферата	15
5	Рубежный контроль	*Первый рубежный контроль*	30
		Второй рубежный контроль	30
6	Промежуточная аттестация	*Экзамен (зачет, дифференцированный зачет)*	40 (100**)
7	Выполнение КР(КП)	*Защита КР(КП)*	100

** Максимальная сумма баллов по пунктам 1-4 таблицы должна равняться 60*
*** В случае отказа обучающегося от результатов текущих аттестаций*

7.3 Система оценки компетенций или их элементов, сформированных у обучающихся в ходе освоения дисциплины (модуля)

Код формируемой компетенции (элементов компетенции)	Коды знаний, умений, навыков, необходимых для формирования компетенции (в соответствии с п.3)	Вид контрольного мероприятия для каждого результата обучения	Форма контрольного задания	№№ контрольных заданий из фонда оценочных средств

ОК-4	Навыки разработки современного ПО	Выполнение КР	Защита КР	-
	навыки анализа результатов обработки данных	Выполнение КР	Защита КР	-
ОПК-1	знание математических форм современных моделей знаний	защита ЛР №1, 2,3	тестирование;	вопросы №№ 1-3, 7,10 из теста №1;
ПК-2	знания принципа действия систем извлечения знаний	защита ЛР № 5-7	тестирование;	вопросы №№ 1-5, из теста №2;
	умение применять методы извлечения знаний для решения практических задач	защита ЛР № 6-7	тестирование;	вопросы №№ 1-15, из теста №2;
	навыки применения современного ПО в прикладных областях	Практическое занятие №4, 5	тестирование;	вопросы №№ 1-15, из теста №2;
	умение пользоваться методами и программами для решения задач извлечения знаний	защита ЛР № 6-7	тестирование;	вопросы №№ 1-15, из теста №3;
ПК-4	навыки разработки прикладных алгоритмов анализа данных и знаний	Выполнение ККР	Защита ККР	-

8 Оценочные средства для текущего контроля успеваемости и промежуточных аттестаций обучающихся

Фонд оценочных средств для текущего контроля успеваемости и промежуточной аттестации обучающихся по дисциплине (модулю) описан в приложении к рабочей программе учебной дисциплины (модуля).

Примеры оценочных средств для текущего контроля успеваемости

Задание #1

Вопрос:

Data Mining - это процесс обнаружения в сырых данных

Выберите несколько из 5 вариантов ответа:

1) ранее сформулированных гипотез
2) неочевидных закономерностей
3) практических закономерностей
4) объективных закономерностей
5) большого количества закономерностей

Задание #2

Вопрос:

Набор называют часто встречающимся (frequent), если:

Выберит е один из 3 вариант ов от вет а:
1) его поддержка выше определенного пользователем минимального значения

2) его поддержка ниже определенного пользователем максимального значения
3) его поддержка равна определенному пользователем значению

Задание #3
Вопрос:
В ходе решения каких из перечисленных задач устанавливаются закономерности между событиями, связанными во времени?

Выберит е один из 3 вариант ов от вет а:
1) задачи поиска ассоциативных правил

2) задачи поиска последовательных ассоциативных правил

3) задачи анализа отклонений

Примеры оценочных средств для промежуточных аттестаций

Экзаменационный билет № 6

1. Алгоритм «Априори» и его развитие.

2. Сходства и различия моделей знаний

Экзаменационный билет № 8

1. Модели знаний в виде правил.

2. Задача рубрикации текстовых данных и ее решение.

Экзаменационный билет № 9

1. Морфологический, синтаксический и семантический анализ текстовых данных. Современные системы Text Mining.
2. Деревья решений и их построение.

9 Учебно-методическое обеспечение дисциплины (модуля)

9.1 Основная литература

1. Барсегян, А.А Технологии анализа данных: Data Mining, Visual Mining, Text Mining,OLAP: учеб. пособие / А.А. Барсегян [и др.] .— 2-е изд., перераб. и доп. — СПб.: БХВ-Петербург, 2007 — 384с.

2. Люгер Дж. Ф. Искусственный интеллект: стратегии и методы решения сложных проблем / Д.Ф. Люгер; пер.с англ. Н.И. Галагана, К.Д. Протасовой, Н.Н. Куссуль; под ред. Н.Н. Куссуль .— 4-е изд. — М. : Вильямс, 2005 .— 864с.

9.2 Дополнительная литература

3. Гаврилова, Т.А. Базы знаний интеллектуальных систем : Учеб. пособие для вузов / Т.А. Гаврилова, В.Ф. Хорошевский .— СПб. и др. : Питер, 2000 .— 384с.
4. Джексон, П. Введение в экспертные системы: Учеб. пособие: Пер. с англ. / П. Джексон .— 3-е изд. — М. : Вильямс, 2001 .— 624с.

9.3 Периодические издания

1. Журнал "Открытые системы": Изд-во" Открытые системы ", — ISSN 0868-6157.
2. HARD,n SOFT : Научно-популярный компьютерный журнал : журнал / ООО"Золотая коллекция" ..— М. : Издательский дом "HARD·n SOFT— ЧЗПИ .— ежемесячно.
3. PC Magazine. Персональный компьютер сегодня.— М. : ЗАО "СК Пресс", .— ежемесячно .— ISSN 0869-4257.

9.4 Интернет-ресурсы

1. Электронный читальный зал «БИБЛИОТЕХ»: учебники авторов ТулГУ по всем дисциплинам. – Режим доступа: https://tsutula.bibliotech.ru, по паролю. Загл. с экрана.
2. ЭБС издательства «Юрайт». Режим доступа: http://biblio-online.ru, по паролю. Загл. с экрана.
3. Научная электронная библиотека eLibrary – библиотека электронной периодики. - Режим доступа: http:// eLibrary.ru, по паролю. Загл. с экрана.
4. Единое окно доступа к образовательным ресурсам: портал [Электронный ресурс]. -Режим доступа: http:// window.edu.ru. Загл. с экрана.
5. КонсультантПлюс: справочная правовая система. Компания «Консультант Плюс» - Версия Проф, сетевая. -Режим доступа: Компьютерная сеть НБ ТулГУ, свободный. Загл. с экрана.
6. ЭБС book.ru. – Режим доступа: http://www.book.ru/, по паролю. Загл. с экрана.

9.5 Методические указания к лабораторным занятиям

Богатырев М.Ю. Методические указания к лабораторным работам по курсу «Представление знаний в информационных системах». Электронная версия. Тула, Тульский государственный университет, 2011 г. (ресурс кафедры)

9.6 Методические указания к практическим занятиям

Богатырев М.Ю. Методические указания к практическим занятиям по курсу «Представление знаний в информационных системах». Электронная версия. Тула, Тульский государственный университет, 2013 г. (ресурс кафедры)

9.7 Методические указания к курсовому проектированию и другим видам самостоятельной работы

Богатырев М.Ю. Методические указания к ссамостоятельной работе по курсу «Представление знаний в информационных системах». Электронная версия. Тула, Тульский государственный университет, 2011 г. (ресурс кафедры)