

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ**

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«НОВОСИБИРСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»  
(НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ, НГУ)**

Утверждаю:

«\_\_\_\_\_» \_\_\_\_\_ 201\_\_ г.

**МЕХАНИКО-МАТЕМАТИЧЕСКИЙ ФАКУЛЬТЕТ**

Рабочая программа дисциплины

**БАЗЫ ДАННЫХ И ЭКСПЕРТНЫЕ СИСТЕМЫ**

Направление подготовки

**010200 – «Математика и компьютерные науки»**

Квалификация (степень)

**Бакалавр**

Форма обучения

**Очная**

Новосибирск – 2014 год

## **Аннотация рабочей программы**

Дисциплина «Базы данных и экспертные системы» входит в Базовую часть Профессионального цикла ООП по направлению подготовки 010200 – «Математика и компьютерные науки». Дисциплина реализуется на Механико-математическом факультете Новосибирского государственного университета кафедрой «Дискретной математики и информатики» ММФ НГУ.

Содержание дисциплины представляет собой связующее звено между теоретическими дисциплинами и прикладными. Он дает возможность понять, как работает математика в такой важной области информатики как базы данных, базы знаний и экспертные системы. Курс характеризуется с одной стороны математической строгостью изложения, а с другой стороны практической направленностью.

Дисциплина нацелена на формирование общекультурных компетенций ОК-7, ОК-10, ОК-14, ОК-15, профессиональных компетенций ПК-2 – ПК-10, ПК-13, ПК-14, ПК-16, ПК-26, ПК-27, ПК-29.

Преподавание дисциплины предусматривает следующие формы организации учебного процесса: лекции и самостоятельная работа студента.

Программой дисциплины предусмотрены следующие виды контроля: текущий контроль посещаемости, самостоятельных и индивидуальных работ. Рубежный контроль – в форме зачета.

Общая трудоемкость дисциплины составляет 2 зачетных единиц – 48 часов, аудиторно – 32, остальное – разные виды самостоятельной работы.

## **1. Цели освоения дисциплины**

Целью курса является знакомство студентов с одним из важнейших направлений искусственного интеллекта и прикладной информатики – методами работы со знаниями, включающими: методы машинного обучения (machine learning), обнаружения знаний (knowledge discovery in data bases and data mining), интеллектуального анализа данных и экспертных систем.

Задачи курса – дать базовые понятия, используемых в курсе дисциплин: методов машинного обучения, инженерии знаний, интеллектуального анализа данных и экспертных систем. Показать неразрывную связь теоретических дисциплин, таких как логика, теория алгоритмов, алгебра и функциональный анализ, с прикладными дисциплинами. Показать, как теоретические дисциплины применяются для решения задач в упомянутых прикладных дисциплинах.

Освоение курса поможет студентам ориентироваться в современных разработках в области информационных технологий и успешно решать возникающие перед ними задачи, поможет им лучше понять, как применяются теоретические дисциплины на практике.

Данный курс знакомит студентов с такими понятиями как: величина, числовая система, шкала, закон, экспертная система, реляционная база данных, реляционная таблица, язык SQL, логическая программа, вычисление логической программы, метод извлечения знаний, гипотеза, извлечение знаний из эксперта, база знаний, эксперимент, событие в эксперименте, эксперимент с шумом, проблемой статистической двусмысленности, выводом предсказаний в логическом программировании, реляционным подходом к извлечению знаний.

## **2. Место дисциплины в структуре ООП бакалавриата**

Дисциплина «Базы данных и экспертные системы» входит в Базовую часть Профессионального цикла ООП по направлению подготовки 010200 – «Математика и компьютерные науки».

Дисциплина «Базы данных и экспертные системы» опирается на следующие дисциплины данной ООП:

- Математическая логика (алгебраическая система, высказывания, модель, выполнимость на модели, истинность, язык логики первого порядка, логический вывод, системы аксиом, теория, вероятностная логика);
- Алгебра (алгебраическая система, гомоморфизм, булева алгебра, абелева

группа, дистрибутивная решетка, конечно-порожденная группа, лупа);

- Теория алгоритмов (алгоритм, вычислимость, конструктивизация, конструктивная модель, нумерация, процесс вычисления);
- Функциональный анализ (функциональная зависимость, числовое пространство, непрерывность).

### 3. Компетенции обучающегося, формируемые в результате освоения дисциплины:

- общекультурные компетенции: ОК-6, ОК-7, ОК-8, ОК-9; ОК-10, ОК-11, ОК-12, ОК-13; ОК-14;
- профессиональные компетенции: ПК-1 – ПК-13, ПК-15, ПК-19 – ПК-26.

В результате освоения дисциплины обучающийся должен:

- иметь современное представление о месте информационных технологий среди различных областей математики;
- знать основные понятия, доказывать их свойства, видеть взаимосвязь между понятиями; уметь использовать понятийный аппарат, принципы и методы интеллектуального анализа данных и экспертных систем при решении задач.
- владеть навыками решения задач, предложенными в настоящем курсе.

### 4. Структура и содержание дисциплины

Общая трудоемкость дисциплины составляет 2 зачетных единиц, 48 часов.

№ п/п	Раздел дисциплины	Семестр	Неделя семестра	Виды учебной работы, включая самостоятельную работу студентов и трудоемкость (в часах)						Формы текущего контроля успеваемости (по неделям семестра) Форма промежуточной аттестации (по семестрам)
				Лекция	Практич. работа	Самост. Работа	Контр. Работа	Зачет	Экзамен	
1	Понятие эмпирической системы, величины, числовой системы, шкалы. Пример шкалы.	6	1	2		1				
2	Проблема существования шкалы. Проблема единственности шкалы. Группы допустимых преобразований шкал. Проблема адекватности.	6	2	2		1				
3	Представление законов в Теории Измерений. Класс функций F, удовлетворяющий свойствам аддитивной соединительной структуры. Теорема о числовом представлении аддитивных соединительных структур. Процедура	6	3	2		1				

	перешкалирования величин зависимости $y = f(x,z)$ и её перевода в закон $y = x + z$ .									
4	Теория Физических Структур. Определение физического закона ранга (r,s). Классификация законов.	6	4	2		1				
5	Определение конструктивного числового представления. Проблемы существования конструктивного числового представления. Конструктивное числовое представление процедуры шкалирования экстенсивных величин.	6	5	2		1				
6	Конструктивные измерительные процедуры, тесты и анкеты. Конструктивное числовое представление дистрибутивной решетки.	6	6	2		1				
7	Экспертная система компьютерного познания. Построение логической эмпирической теории. Построение количественной и конструктивной эмпирических теорий.	6	7	2		1				
8	Реляционные базы данных. Реляционные таблицы. Реляционная алгебра. Операции реляционной алгебры.	6	8	2		1				
9	Базисный язык SQL. Подязык запросов. Подязык манипуляции данными. Базисный язык SQL. Подязык определения данных. Операции реляционной алгебры.	6	9	2		1				
10	Логическое программирование. Логическая программа. Подстановки. Алгоритм унификации. Вычисление логической программы. Пространство вычислений. Вычисление запросов. Ответ программы P <sub>r</sub> на запрос. Дерево вывода запроса.	6	10	2		1				
11	Логический анализ методов извлечения знаний. Онтология метода. Класс гипотез метода. Реляционный подход к извлечению знаний. Система Discovery. Свойства реляционного подхода.	6	11	2		1				
12	Проблемы извлечения знаний из эксперта. Создание «полной» и непротиворечивой базы знаний, включающей как экспертные, так и объективные знания. Свойство монотонности и декомпозиция задачи. Цепи Ханселя и построение опросной таблицы.	6	12	2		1				
13	Понятие эксперимента. Определение вероятностного закона. Определение сильнейшего вероятностного закона. Определение эксперимента с шумами. Определение сохраняющей модели шумов. Теорема о совпадении множеств законов и вероятностных законов для данного сохраняющего шума.	6	13	2		1				
14	Проблема статистической двусмысленности. Пример статистической двусмысленности. Дедуктивно-номологическая модель предсказания и индуктивно-статистическая модель предсказания. Вывод предсказаний в логическом программировании. Дерево вывода предсказаний.	6	14	2		1				

15	Семантический вероятностный вывод. Дерево семантического вероятностного вывода. Сравнение деревьев логического вывода предсказаний в логическом программировании и с помощью семантического вероятностного вывода.	6	15	2		1				
16	Определение максимально специфических законов. Требование максимальной специфичности. Теорема о том, что любое максимально специфическое правило удовлетворяет требованию максимальной специфичности. Теорема о непротиворечивости множеств максимально специфических правил.	6	16	2		1				
			17							<b>Зачет</b>
				<b>32</b>		<b>16</b>	<b>24</b>			

## 5. Образовательные технологии

Используется традиционная система обучения, включающая лекции и самостоятельную работу.

Для проверки уровня знаний студентов и их аттестации в конце курса предполагается проведение зачета. При проведении аттестации по предмету студентов важно помнить, что систематичность, объективность, аргументированность – главные принципы, на которых основаны контроль и оценка знаний студентов.

## 6. Учебно-методическое обеспечение самостоятельной работы студентов. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины

### 6.1. Для самостоятельной работы используется учебное пособие:

Витяев Е.Е. Новые информационные технологии работы с данными и знаниями. (Учебное пособие по курсу «Базы данных и экспертные системы»), НГУ, Новосибирск, 2008, 197с

### 6.2. Вопросы для подготовки к зачету.

1. Понятие эмпирической системы, величины, числовой системы, шкалы. Пример шкалы.
2. Проблема существования шкалы. Проблема единственности шкалы. Группы допустимых преобразований шкал. Проблема адекватности.
3. Матричное представление бинарных отношений. Отношения толерантности, эквивалентности, частичного порядка, интервального упорядочения,

полуупорядка, древесного порядка, слабого порядка и их числовые представления.

4. Матрицы упорядочений. Матрица близости. Шкалы положительных разностей, алгебраических разностей, равных конечных промежутков, абсолютных разностей. Матрица объект-признак.
5. Представление законов в Теории Измерений. Класс функций  $F$ , удовлетворяющий свойствам аддитивной соединительной структуры. Теорема о числовом представлении аддитивных соединительных структур. Процедура перешкалирования величин зависимости  $y = \square(x,z)$  и её перевода в закон  $y = x + z$ .
6. Теория Физических Структур. Определение физического закона ранга  $(r,s)$ . Классификация законов.
7. Определение конструктивного числового представления. Проблемы существования конструктивного числового представления. Конструктивное числовое представление процедуры шкалирования экстенсивных величин.
8. Конструктивные измерительные процедуры, тесты и анкеты. Конструктивное числовое представление дистрибутивной решетки.
9. Экспертная система компьютерного познания. Построение логической эмпирической теории. Построение количественной и конструктивной эмпирических теорий.
10. Реляционные базы данных. Реляционные таблицы. Реляционная алгебра. Операции реляционной алгебры.
11. Базисный язык SQL. Подязык запросов. Подязык манипуляции данными. Базисный язык SQL. Подязык определения данных. Операции реляционной алгебры.
12. Логическое программирование. Логическая программа. Подстановки. Алгоритм унификации. Вычисление логической программы. Пространство вычислений. Вычисление запросов. Ответ программы  $Pg$  на запрос. Дерево вывода запроса.
13. Представление реляционных операций в логическом программировании. Программа, запрашивающая информацию у пользователя.
14. Программа, объясняющая, как доказывается цель. Программа вычисления оценок утверждений.
15. Логический анализ методов извлечения знаний. Онтология метода. Класс

гипотез метода.

16. Проблемы извлечения знаний из эксперта. Проблемы и способ сравнения экспертных и объективных знаний, полученных индуктивным выводом или системой discovery. Создание «полной» и непротиворечивой базы знаний, включающей как экспертные, так и объективные знания. Свойство монотонности и декомпозиция задачи. Цепи Ханселя и построение опросной таблицы.
17. Задача обнаружения теории эмпирической системы. Теорема о подправилах. Понятие закона. Теорема о выводимости теории эмпирической системы из множества законов.
18. Понятие эксперимента. Определение двоичного куба результатов эксперимента. Определение закона на множестве всех возможных экспериментов. Теорема о совпадении понятия закона для эмпирической системы и для множества всех экспериментов.
19. Определение вероятности на двоичном кубе. Определение события в эксперименте. Теорема о совпадении определения закона с вероятностным определением. Определение вероятностного закона на  $E_{xp}$  в детерминированном случае. Следствие о совпадении определения закона и вероятностного закона в детерминированном случае.
20. Общее определение закона, основанное на свойстве неупрощаемости. Определение вероятностного закона на  $E_{xp}$ . Определение сильнейшего вероятностного закона. Соотношение между множеством законов и множеством вероятностных законов.
21. Определение эксперимента с шумами. Определение сохраняющей модели шумов. Пример сохраняющего двоичного шума. Теорема о совпадении множеств законов и вероятностных законов для данного сохраняющего шума.
22. Проблема статистической двусмысленности. Пример статистической двусмысленности. Дедуктивно-номологическая модель предсказания и индуктивно-статистическая модель предсказания. Вывод предсказаний в логическом программировании. Дерево вывода предсказаний.
23. Семантический вероятностный вывод. Дерево семантического вероятностного вывода. Сравнение деревьев логического вывода предсказаний в логическом программировании и с помощью семантического вероятностного вывода.

24. Определение максимально специфических законов. Требование максимальной специфичности. Теорема о том, что любое максимально специфическое правило удовлетворяет требованию максимальной специфичности. Теорема о непротиворечивости множеств максимально специфических правил.
25. Реляционный подход к извлечению знаний. Система Discovery. Свойства реляционного подхода.

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

### **7.1. Учебное пособие:**

Витяев Е.Е. Новые информационные технологии работы с данными и знаниями. (Учебное пособие по курсу «Базы данных и экспертные системы»), НГУ, Новосибирск, 2008, 197с

### **7.2. Список основной и дополнительной литературы.**

1. Витяев Е.Е. Извлечение знаний из данных. Компьютерное познание. Модели когнитивных процессов // Новосибирский гос. ун-т. Новосибирск, 2006. 293 с.
2. Стюарт Рассел, Питер Норвиг. Искусственный интеллект: современный подход, Киев, 2006, Издательский дом “Вильямс”, 1409 стр.
3. Пфанцагль И. Теория измерений. М., Мир, 1976.
4. Логический подход к искусственному интеллекту. М., «Мир», 1990
5. Э.В.Попов. Экспертные системы. М., Наука, 1987.

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

- Ноутбук, медиа-проектор, экран.
- Программное обеспечение для демонстрации слайд-презентаций.

Программа составлена в соответствии с требованиями ФГОС ВПО с учетом рекомендаций и ПрООП ВПО по направлению 010200 – «Математика и компьютерные науки», все профили подготовки.

Автор: \_\_\_\_\_ Витяев Евгений Евгеньевич  
д.ф.-м.н., профессор ММФ НГУ  
в.н.с. ИМ СО РАН

Рецензент (ы) \_\_\_\_\_  
\_\_\_\_\_

Программа одобрена на заседании \_\_\_\_\_  
от \_\_\_\_\_ года, протокол № \_\_\_\_\_