

Банк заданий по специальной части вступительного испытания в магистратуру

Задание №6 экзаменационного билета (5 баллов)

Тема 1. Базы и банки данных. Системы управления базами данных.

Задание 6.1

Основные понятия баз данных: база данных, база знаний, банк данных. Основные требования, предъявляемые к банкам данных. Компоненты банка данных. Классификация баз данных.

Задание 6.2

Система управления базами данных (СУБД). Основные компоненты СУБД. Основные функции СУБД. Классификация СУБД. Примеры СУБД.

Задание 6.3

Модели организации баз данных. Реляционная модель. Иерархическая модель. Сетевая модель. Объектно-ориентированная модель.

Задание 6.4

Реляционная модель данных. Понятия: тип данных, домен, атрибут, кортеж, отношение, первичный ключ, внешний ключ. Фундаментальные свойства отношений.

Задание 6.5

Различные представления о данных в базах данных. Концептуальная и логическая модели. Уровни представления данных: внешний, концептуальный, внутренний. Основные этапы проектирования баз данных.

Задание 6.6

Концептуальное моделирование данных. Модель «сущность-связь». Основные понятия: сущность, атрибуты, класс и экземпляр сущности. Связь и типы связей. ER-диаграмма. Пример ER-диаграммы.

Задание 6.7

Логическое проектирование реляционной базы данных. Методика перехода от ER-модели к реляционным отношениям. Представление в реляционной схеме взаимно исключающих связей.

Задание 6.8

Нормализация. Цель нормализации. Функциональная зависимость. Первая и вторая нормальные формы. Свойства нормальных форм.

Задание 6.9

Нормализация. Цель нормализации. Транзитивная функциональная зависимость. Третья нормальная форма и нормальная форма Бойса-Кодда.

Задание 6.10

Распределенные базы данных (РБД). Общие принципы и критерии распределенности РБД по Дейту. Основные методы поддержки распределенных данных.

Пример правильного ответа на задание 6.10

Под распределенной базой данных (РБД) понимается набор логически связанных между собой разделяемых данных, которые физически распределены по разным узлам компьютерной сети. Распределенная база данных предполагает хранение и выполнение функций управления данными в нескольких узлах и передачу данных между этими узлами в процессе выполнения запросов.

Разбиение данных в распределенной базе данных может достигаться путём хранения различных таблиц на разных компьютерах или даже хранения разных частей и фрагментов одной таблицы на разных компьютерах.

При распределенной обработке работа с базой (представление данных, их обработка и др.) ведётся на компьютере клиента, а поддержание базы в актуальном состоянии – на сервере. При этом такие БД обычно располагаются на нескольких серверах – различных узлах компьютерной сети, а некоторые данные могут дублироваться.

Доступ пользователей к распределенным базам данных и их администрирование осуществляются с помощью системы управления распределенной базой данных (СУРБД).

Система управления распределенными базами данных (СУРБД)- это система управления базами данных, расположенными в нескольких узлах информационной сети.

В СУРБД используется комбинация централизованного и локального способов хранения данных. Прозрачность РБД заключается в том, что с точки зрения конечного пользователя она должна вести себя точно также, как централизованная.

Логически единая БД разделяется на фрагменты, каждый из которых хранится на одном компьютере, а все компьютеры соединены линиями связи.

Каждый из этих фрагментов работает под управлением своей СУБД.

Критерии распределенности и свойства, которым должна удовлетворять РБД по Дейту:

1. Локальная автономность. Управление данными на каждом из узлов распределенной системы выполняется локально. База данных, расположенная на одном из узлов, является неотъемлемым компонентом распределенной системы;

2. Отсутствие опоры на центральный узел. В идеальной системе все узлы равноправны и независимы, а расположенные на них базы являются равноправными поставщиками данных в общее пространство данных. База данных на каждом из узлов самодостаточна - она включает полный собственный словарь данных и полностью защищена от несанкционированного доступа;

3. Непрерывное функционирование. Возможность непрерывного доступа к данным вне зависимости от их расположения и вне зависимости от операций, выполняемых на локальных узлах;

4. Прозрачность расположения. Независимость от расположения;

5. Прозрачная фрагментация. Независимость от фрагментации;

6. Прозрачное тиражирование. Независимость от репликации (репликация — одна из техник масштабирования баз данных);

7. Обработка распределенных запросов. Возможность выполнения операций выборки над распределенной базой данных, сформулированных в рамках обычного запроса на языке SQL;

8. Обработка распределенных транзакций;

9. Независимость от оборудования;

10. Независимость от операционных систем;

11. Прозрачность сети. В распределенной системе возможны любые сетевые протоколы;

12. Независимость от баз данных. СУРБД должна быть способной функционировать поверх различных локальных СУБД (требование гетерогенности).

Основные методы поддержки распределенных данных.

1. Фрагментация – разбиение БД или таблицы на несколько частей и хранение этих частей на разных узлах РБД.
2. Репликация – создание и хранение копий одних и тех же данных на разных узлах РБД.
3. Распределенные ограничения целостности – ограничения, для проверки выполнения которых требуется обращение к другому узлу РБД.
4. Распределенные запросы – это запросы на чтение, обращающиеся более чем к одному узлу РБД.
5. Распределенные транзакции – команды на изменение данных, обращающиеся более чем к одному узлу РБД.

Задание №7 экзаменационного билета (5 баллов)

Тема 2. Экономические аспекты анализа бизнес-процессов.

Задание 7.1

Контур финансовой операции. Частичные платежи. Потребительский кредит. Нарращение процентов в потребительском кредите.

Задание 7.2

Формула наращения. Начисление процентов в смежных календарных периодах. Переменные ставки. Начисление процентов при дробном числе лет.

Задание 7.3

Нарращение процентов m раз в году. Номинальная ставка. Эффективная ставка. Общий и смешанный метод наращения.

Задание 7.4

Операции со сложной учетной ставкой. Учет по сложной учетной ставке. Номинальная и эффективная учетные ставки. Нарращение по сложной учетной ставке.

Задание 7.5

Определение срока ссуды. Определение размера процентной ставки. При наращении по сложной годовой ставке процентов. Определение размера процентной ставки по номинальной ставке. Определение размера процентной ставки при дисконтировании по сложным учетным ставкам.

Задание 7.6

Средние процентные ставки. Средние простые процентные ставки. Средние сложные процентные ставки. Усреднение ставок в однородных операциях.

Задание 7.7

Эквивалентность простых процентных ставок. Эквивалентность простых и сложных ставок. Эквивалентность сложных ставок.

Задание 7.8

Финансовая эквивалентность обязательств. Консолидирование (объединение) задолженности. Определение размера консолидированного платежа

Задание 7.9

Аннуитеты. Обыкновенные и полагающие аннуитеты. Определение платежей аннуитета и процентной ставки

Задание 7.10

Инвестиции. Чистый приведенный доход. Срок окупаемости. Функция риска.

Пример правильного ответа на задание 7.10

Инвестиции – это расходование ресурсов в расчете на получение доходов в будущем по истечении достаточно длительного периода времени. Любая инвестиция подвержена риску в том смысле, что надежда на получение дохода может и не оправдаться. Различают 2 вида инвестиций: финансовые и реальные. Первые представляют собой вложения капитала в долгосрочные финансовые активы (акции, облигации); вторые – в развитие материально-технической базы предприятий производственной и непроизводственной сферы.

Чистый приведенный доход (ЧПД) – текущая стоимость денежных потоков за вычетом текущей стоимости денежных оттоков. Расчет данной величины предусматривает

дисконтирование денежных потоков, т. е. все доходы и затраты приводятся к одному моменту времени. ЧПД – это обобщенный конечный результат инвестиционной деятельности в абсолютном измерении.

При разовой инвестиции ЧПД рассчитывается таким образом:

$$D = \sum_{k=1}^n \frac{R_k}{(1+i)^k} - I$$

где R_k – годовой доход в k -м году; n – число лет, в течение которых поступают доходы; i – процент дисконтирования; I – величина начальных инвестиций.

Если проект предполагает не разовую инвестицию, а последовательные инвестирования финансовых ресурсов течение m лет, то формула будет такой:

$$D = \sum_{k=1}^n \frac{R_k}{(1+i)^k} - \sum_{j=1}^m \frac{I_j}{(1+i)^j}$$

где I_j – инвестиции в j -м году.

ЧПД оценивает, на сколько приведенный доход перекрывает приведенные затраты. Если $D \leq 0$, то проект не имеет прибыли.

Срок окупаемости – минимальное целое число лет, необходимых для возмещения стартовых инвестиций. Найти его можно из формулы:

$$\sum_{k=1}^n \frac{R_k}{(1+i)^k} - \sum_{j=1}^m \frac{I_j}{(1+i)^j} \geq 0$$

Если ожидаемый ЧПД удовлетворяет условию

$$D \geq D_{min},$$

где D_{min} – минимальная допустимая величина дохода, то проект может быть принят.

Поскольку при расчете ЧПД всегда сталкиваются с неопределенностью, например, при нахождении ставки дисконтирования i и величины ожидаемого дохода R , то для оценки возможности того, что действительный доход окажется ниже критического D_{min} , вводят функцию риска F_r .

Если известны вероятностные характеристики D , то функцию риска F_r можно определить как вероятность того, что доход D окажется меньше D_{min} :

$$F_r(D) = P(D < D_{min}) = F(D)$$

где $F(D)$ – функция распределения вероятностей случайной величины D . Эта функция является интегралом от функции плотности $f(x)$ распределения вероятностей случайной величины:

$$F(D) = \int_D f(x) dx$$

Для равномерно распределенной случайной величины на интервале $[0; D_{max}]$

$$d(x) = \frac{1}{D_{max}},$$

где D_{max} – максимально возможный доход. Эта функция указывает на равную вероятность получения любого дохода из этого интервала. В данном случае выражение для функции риска будет выглядеть следующим образом:

$$F_r(D_{min}) = \int_0^{D_{min}} \frac{1}{D_{max}} dx = \frac{1}{D_{max}} \cdot \int_0^{D_{min}} dx = \frac{D_{min}}{D_{max}}$$

При этом риск возрастает линейно от 0 (при ожидании минимально возможного дохода) до 1 (при ожидании максимально возможного дохода).

Задание №8 экзаменационного билета (10 баллов)

Тема 3. Информационная безопасность.

Задание 8.1

Виды, источники и классификация угроз информационной безопасности. Актуальный банк данных угроз информационной безопасности (согласно сведениям Федеральной службы по техническому и экспортному контролю России).

Задание 8.2

Правовые нормы информационной безопасности. Нормативная и законодательная база Российской Федерации, обеспечивающая функционирование систем защиты информации.

Задание 8.3

Организационно-административные методы обеспечения информационной безопасности. Достоинства и недостатки организационно-административных методов защиты информации.

Задание 8.4

Аппаратно-программные методы защиты информации. Примеры и краткая характеристика аппаратно-программных средств обеспечения информационной безопасности.

Задание 8.5

Криптографические методы защиты информации. Симметричные и асимметричные криптосистемы.

Задание 8.6

Электронная цифровая подпись (ЭЦП). Администрирование ЭЦП.

Задание 8.7

Идентификация, аутентификация и авторизация пользователей автоматизированных информационных систем. Возможные и наиболее эффективные на текущий момент способы реализации идентификации и аутентификации пользователей.

Задание 8.8

Организация и ведение работ по обеспечению безопасности персональных данных при их обработке в информационных системах персональных данных.

Задание 8.9

Модели политик информационной безопасности. Характерные особенности дискреционной (избирательной) и мандатной (полномочной) политик безопасности.

Задание 8.10

Политика информационной безопасности. Основные требования стандарта ГОСТ Р ИСО/МЭК 17799-2005, используемого при разработке политики информационной безопасности.

Пример правильного ответа на Задание 8.10.

Политика информационной безопасности (ПИБ) – это совокупность правил, процедур, практических методов и руководящих принципов в области информационной безопасности (ИБ), используемых организацией в своей деятельности.

Целью разработки ПИБ организации является создание единой системы взглядов и понимания целей, задач и принципов обеспечения ИБ.

ПИБ неразрывно связана с развитием предприятия, его стратегическим планированием, она определяет общие принципы и порядок обеспечения ИБ в организации. ПИБ тесно интегрируется в работу предприятия на всем этапе его существования. Все решения, принимаемые на предприятии, должны учитывать её требования.

Основные направления разработки ПИБ:

- определение того, какие данные и насколько серьезно необходимо защищать;
- определение того, кто и какой ущерб может нанести организации;
- вычисление рисков и определение схемы уменьшения их до приемлемой величины.

Основные этапы разработки ПИБ:

- исследование текущего состояния информационной среды и ИБ организации;
- анализ полученных сведений по результатам исследования;
- формирование плана работ по разработке ПИБ;
- разработка ПИБ организации.

Согласно отечественному стандарту ГОСТ Р ИСО/МЭК 17799-2005, ПИБ должна устанавливать ответственность руководства, а также излагать подход организации к управлению информационной безопасностью. В соответствии с указанным стандартом, необходимо, чтобы ПИБ предприятия как минимум включала:

- определение информационной безопасности, её общих целей и сферы действия, а также раскрытие значимости безопасности как инструмента, обеспечивающего возможность совместного использования информации;

- изложение целей и принципов информационной безопасности, сформулированных руководством;

- краткое изложение наиболее существенных для организации политик безопасности, принципов, правил и требований, например, таких как: соответствие законодательным требованиям и договорным обязательствам; требования в отношении обучения вопросам безопасности; предотвращение появления и обнаружение вредоносного ПО; управление непрерывностью бизнеса; ответственность за нарушения политики безопасности;

- определение общих и конкретных обязанностей сотрудников в рамках управления информационной безопасностью, включая информирование об инцидентах нарушения информационной безопасности;

- ссылки на документы, дополняющие политику ИБ, например, более детальные политики и процедуры для конкретных информационных систем, а также правила безопасности, которым должны следовать пользователи.

ПИБ компании должна быть утверждена руководством, издана и доведена до сведения всех сотрудников в доступной и понятной форме.

Для того чтобы ПИБ эффективно работала необходимо, чтобы она была:

- непротиворечивой – разные документы не должны по разному описывать подходы к одному и тому же процессу обработки информации;

- не запрещала необходимые действия (в таком случае неизбежные массовые нарушения приведут к дискредитации ПИБ среди пользователей);

- не налагала невыполнимых обязанностей и требований.

В организации должно быть назначено лицо, ответственное за ПИБ, отвечающее за её эффективную реализацию и регулярный пересмотр.

Задание №9 экзаменационного билета (20 баллов)

Тема 4. Системы, основанные на знаниях. Элементы нечеткой логики.

Задание 9.1

Данные. Знания. Классификация знаний. Методы извлечения знаний. Модели представления знаний.

Задание 9.2

Машина вывода. Цикл работы интерпретатора. Стратегии управления выводом.

Задание 9.3

Определение экспертной системы (ЭС), обобщенная структура, области применения. Классификация систем, основанных на знаниях.

Задание 9.4

Этапы проектирования ЭС. Факты, свидетельствующие о целесообразности разработки ЭС. Характеристики подходящих для ЭС задач.

Задание 9.5

Нечеткие множества. Функция совместимости. Операции над нечеткими множествами и их геометрическая интерпретация. Понятие о треугольных нормах. Методы построения функций принадлежности.

Задание 9.6

Нечеткие отношения. Свертка двух нечетких отношений. Нечеткая импликация.

Задание 9.7

Нечеткие выводы. Этапы логического вывода. Алгоритмы нечетких выводов Mamdani, Tsukamoto и их геометрическая интерпретация.

Задание 9.8

Нечеткие выводы. Этапы логического вывода. Алгоритмы нечетких выводов Sugeno, Larsen и их геометрическая интерпретация.

Задание 9.9

Методы приведения к четкости. Пример.

Задание 9.10

Нечеткая и лингвистическая переменные. Нечеткие числа. Целесообразность применения систем нечеткой логики. Недостатки систем нечеткой логики.

Пример правильного ответа на Задание 9.10

Нечеткая переменная характеризуется тройкой:

$\langle \alpha, X, A \rangle$, где α - наименование переменной, X - универсальное множество (область определения α), A - нечеткое множество на X , описывающие ограничения ($\mu(x)$ - значения нечеткой переменной α).

Лингвистической переменной называется набор (кортеж)

$\langle \beta, T, X, G, M \rangle$, где β - наименование лингвистической переменной; T - терм-множество значений лингвистической переменной, представляющее собой наименование нечетких переменных, областью определения которых является множество X ; G - синтаксическая процедура, позволяющая оперировать элементами T , в частности генерировать новые термы (значения), получая расширенные термы- множества лингвистической переменной; M -

семантическая процедура, позволяющая превращать каждое новое значение лингвистической переменной, образованное процедурой G в нечеткую переменную, то есть сформировать соответствующее нечеткое множество.

Пример. Пусть эксперт определяет толщину выпускаемого изделия с помощью понятий “малая толщина”, “средняя толщина”, “большая толщина”, при чем минимальная толщина=10 мм, а максимальная толщина = 80 мм. Формализация такого описания может быть проведена, с помощью лингвистической переменной $\langle \beta, T, X, G, M \rangle$,

где β - толщина изделия;

T - { малая толщина, средняя, большая }

X =[10 мм, 80 мм]

G - процедура образования новых термов с помощью связок и или и модификаторов типов (“очень”, “не”, “слегка” и так далее)

M – процедура задания на X =[10,80] нечетких подмножеств A_1 - “малая толщина”, A_2 - “средняя толщина”, A_3 - “большая толщина”, а также нечетких множеств для термов из $G(T)$ в соответствии с правилами трансляции связок «И», «ИЛИ», «ОЧЕНЬ» и др.

Системы с нечеткой логикой целесообразно применять в следующих случаях:

- для сложных процессов, когда нет простой математической модели;
- если экспертные знания об объекте или о процессе можно сформулировать только в лингвистической форме;
- требуемый результат может быть получен каким-либо другим стандартным путем;
- когда для объекта или процесса уже найдена адекватная и легко исследуемая математическая модель.

Основными недостатками систем с нечеткой логикой является то, что:

- исходный набор постулированных нечетких правил формулируется экспертом - человеком и может быть неполным и/или противоречивым;
- вид и параметры функции принадлежности, описывающие входные и выходные переменные, выбираются субъективно и могут не вполне отражать действительность.

Задание №10 экзаменационного билета (20 баллов)

Тема 5. Технологии искусственных нейронных сетей в интеллектуальном анализе данных.

Задание 10.1

Понятие искусственной нейронной сети (ИНС). Задачи, решаемые с помощью ИНС.
Структура биологического и искусственного нейрона.

Задание 10.2

Классификации нейронных сетей по топологии, по принципу структуры нейронов, по виду сигналов, по синхронности состояния и рекомендуемые области их применения.

Задание 10.3

Архитектура и рекомендуемые области применения монотонных сетей, сетей без обратных связей, сетей с обратными связями, сетей Хопфилда

Задание 10.4

Виды функций активации. Примеры постановки задач для ИНС. Этапы построения ИНС.
Принципы выбора ИНС.

Задание 10.5

Персептроны. Назначение, обобщенная схема, виды персептронов, принципы работы.
Основные теоремы о персептронах. Достоинства и недостатки персептронных систем.

Задание 10.6

Эволюционные алгоритмы. Генетические операторы. Мягкие вычисления. Понятие агентно-ориентированного подхода в искусственный интеллект.

Задание 10.7

Определение и примеры треугольной нормы и треугольной конормы.

Задание 10.8

Гибридные ИНС: определение, структура (рассмотреть на примере системы. Имеющей базу знаний из двух правил).

Задание 10.9

Методы реализации компьютерного зрения. Сверточные искусственные нейронные сети: определение, структура, области применения.

Задание 10.10

Требования к обучающей выборке. Математическая постановка процесса обучения НС.
Виды алгоритмов обучения НС.

Пример правильного ответа на Задание 10.10

Требования к обучающей выборке:

1. Количество обучающих примеров должно составлять десятки, а лучше сотни штук.
2. Выборка не должна быть составлена предвзято, в ней должны присутствовать все возможные варианты развития (так, сеть, которой в процессе обучения не предъявляли ситуации кризиса, никогда эти ситуации не распознает).

Математическая постановка процесса обучения нейронной сети.

Пусть, в процессе функционирования нейронная сеть формирует выходной сигнал O (ответ), в соответствии с входным сигналом X , реализуя некоторую функцию $G(X)$: $O=G(X)$.

Если архитектура сети задана, то вид функции G определяется значениями синаптических весов и смещений в сети.

Предположим, решением некоторой задачи является функция $Y=F(X)$ заданная парами входных и выходных векторов данных: $(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)$, для которых $Y_k=F(X_k)$ при $k=1 \dots N$.

Обучения состоит в поиске (синтезе) функции G близкой к F в смысле некоторой функции ошибки E . Если выбрано множество обучающих примеров (X_k, Y_k) , $k=1 \dots N$, и способ вычисления функции ошибки E , то обучение сети превращается в задачу многомерной оптимизации, имеющую очень большую размерность. Поскольку функция E может иметь произвольный вид, то обучение нейронной сети с математической точки зрения представляет собой в общем случае многоэкстремальную невыпуклую задачу оптимизации.

Виды алгоритмов обучения НС.

Данные алгоритмы чаще всего делят на две большие группы – обучение с учителем и обучение без учителя. В рамках этих групп реализуются:

- алгоритмы локальной оптимизацией с вычислением частных производных 1-го порядка (градиентный алгоритм «метод скорейшего спуска», методы с одномерной и двумерной оптимизацией целевой функции в направлении антиградиента, метод сопряженных градиентов);

- алгоритмы локальной оптимизации с вычислением производных 1-го и 2-го порядка (метод Ньютона, квазиньютоновские методы, метод Левенберга-Марквардта и др.);

- стохастические алгоритмы оптимизации (поиск в случайном направлении, имитация отжига, метод Монте_Карло);

- алгоритмы глобальной оптимизации (метод перебора значения переменных, от которых зависит целевая функция E);

- генетические алгоритмы – основаны на моделировании развития биологических популяций.