**Интеллектуальные системы**

**1. Понятие интеллектуальной системы и ее основные компоненты.**

**Интеллектуальная система** – программная или аппаратно-программная система, способная решать задачи, традиционно считающиеся творческими и принадлежащие конкретной предметной области, на основе знаний экспертов этой области.

**Основные компоненты интеллектуальной системы:**

**База знаний** – хранилище знаний предметной области, представленных в форме, понятной компьютеру. Содержит факты (данные) и правила (способы манипулирования фактами). Основа любой интеллектуальной системы.

**Механизм логического вывода** – компонент, реализующий процедуры рассуждения на основе имеющихся знаний. Обрабатывает правила и факты для получения новых знаний и принятия решений.

**Подсистема объяснения** – компонент, объясняющий, как система пришла к определенному выводу. Повышает доверие пользователя и помогает отладке системы.

**Интерфейс пользователя** – обеспечивает взаимодействие с системой в удобной для человека форме. Может включать естественно-языковой интерфейс или графические элементы.

**Подсистема приобретения знаний** – позволяет добавлять, изменять и обновлять знания в базе знаний. Может включать инструменты для работы с экспертами или методы машинного обучения.

Интеллектуальные системы способны к адаптации, обучению и объяснению своих решений, что отличает их от обычных информационных систем.

**2. Структура и принцип работы ИНС Хемминга.**

**Нейронная сеть Хемминга** – особый тип рекуррентной нейронной сети, предназначенный для распознавания бинарных или биполярных векторов на основе минимального расстояния Хемминга.

**Структура сети Хемминга:**

**Входной слой** – нейроны, принимающие входной вектор размерности n

**Скрытый слой** – содержит m нейронов, где m – количество эталонных образцов

**Выходной слой** – слой нейронов с обратной связью (слой MAXNET)

**Принцип работы:**

**Инициализация**:

Весовые коэффициенты скрытого слоя настраиваются как половина соответствующих эталонных векторов

Смещения нейронов скрытого слоя устанавливаются как n/2 (половина размерности входного вектора)

**Прямое распространение**:

Входной вектор подается на входной слой

Каждый нейрон скрытого слоя вычисляет скалярное произведение с входным вектором

Выходы скрытого слоя соответствуют мере близости входного вектора к эталонным образцам

**Соревновательный процесс**:

Слой MAXNET выполняет итеративный процесс, в результате которого активным остается только один нейрон с максимальным откликом

Остальные нейроны подавляются

**Применение** ИНС Хемминга: - Распознавание образов - Классификация бинарных векторов - Ассоциативная память - Задачи восстановления зашумленных данных

Преимущества сети Хемминга: быстрая сходимость, гарантированное нахождение ближайшего эталона, устойчивость к шумам.

**3. Продукционная модель представления знаний и принцип работы систем, основанных на правилах.**

**Продукционная модель** – один из основных способов представления знаний в интеллектуальных системах, основанный на правилах вида “ЕСЛИ условие ТОГДА действие”.

**Структура продукционной системы:**

**База знаний**:

**Продукционные правила** – формализованные утверждения типа “ЕСЛИ A ТОГДА B”

**Факты** – набор утверждений, описывающих текущее состояние предметной области

**Рабочая память** – хранит текущее состояние решаемой задачи (факты, промежуточные выводы)

**Механизм логического вывода**:

**Сопоставитель с образцом** – находит правила, условия которых удовлетворяют текущим фактам

**Механизм разрешения конфликтов** – выбирает одно правило из нескольких подходящих

**Интерпретатор правил** – применяет выбранное правило и обновляет рабочую память

**Принцип работы систем, основанных на правилах:**

**Прямой вывод**:

Начинается с имеющихся фактов

Применяет подходящие правила

Добавляет новые факты

Процесс продолжается, пока не будет достигнута цель или не останется применимых правил

**Обратный вывод**:

Начинается с цели (гипотезы)

Ищет правила, которые могут подтвердить гипотезу

Проверяет условия этих правил, которые становятся подцелями

Процесс продолжается рекурсивно, пока подцели не будут сведены к известным фактам

**Преимущества продукционной модели:** - Модульность (правила независимы друг от друга) - Естественность (близость к человеческому мышлению) - Гибкость (легко добавлять и модифицировать правила)

**Недостатки:** - Неэффективность при большом количестве правил - Сложность отслеживания взаимосвязей между правилами - Проблемы при обработке неопределенности

Примеры систем: MYCIN (медицинская диагностика), DENDRAL (химия), CLIPS (инструментальная среда).

**4. Искусственная нейронная сеть прямого распространения и метод ее обучения.**

**Искусственная нейронная сеть прямого распространения (feedforward neural network)** – тип нейронной сети, в которой информация движется только в одном направлении: от входных нейронов через скрытые слои к выходным нейронам, без обратных связей.

**Структура сети прямого распространения:**

**Входной слой** – принимает внешние данные

**Скрытые слои** – выполняют нелинейные преобразования входных данных

**Выходной слой** – формирует результат работы сети

**Функционирование нейрона:** - Получает взвешенную сумму входных сигналов - Применяет функцию активации (сигмоидную, ReLU, tanh и др.) - Передает результат следующему слою

**Методы обучения: алгоритм обратного распространения ошибки (backpropagation)**

**Прямой проход**:

Входные данные подаются на входной слой

Сигналы распространяются через сеть, формируя выходной сигнал

Вычисляется функция ошибки (разница между фактическим и желаемым выходом)

**Обратный проход**:

Ошибка распространяется от выходного слоя к входному

Вычисляются градиенты ошибки по весам

Веса корректируются с использованием градиентного спуска

**Обновление весов**:

Для каждого веса w вычисляется корректировка Δw = -η \* ∂E/∂w

где η – скорость обучения, ∂E/∂w – частная производная ошибки по весу

**Варианты алгоритма обратного распространения:** - Пакетный (batch) – обновление весов после обработки всего обучающего набора - Стохастический – обновление после каждого примера - Mini-batch – обновление после обработки подмножества примеров

**Проблемы обучения:** - Локальные минимумы - Переобучение - Затухание/взрыв градиентов - Выбор гиперпараметров

**Современные улучшения:** - Адаптивные методы обучения (Adam, RMSProp) - Регуляризация (L1, L2, Dropout) - Пакетная нормализация - Остаточные связи (ResNet)

Нейронные сети прямого распространения широко применяются в задачах классификации, регрессии и распознавания образов.

**5. Представление знаний с помощью семантических сетей. Виды семантических сетей и их назначение.**

**Семантическая сеть** – модель представления знаний в виде ориентированного графа, где вершины представляют понятия (объекты, события, свойства), а дуги – отношения между ними.

**Основные элементы семантической сети:** - **Узлы (вершины)** – концепты, понятия, объекты реального мира - **Дуги (ребра)** – отношения между узлами, описывающие семантику связи - **Метки** – имена узлов и дуг, конкретизирующие их смысл

**Виды семантических сетей:**

**Классификационные сети**

Используют отношения типа “является” (IS-A), “имеет свойство” (HAS-PROPERTY)

Организуют понятия в таксономии (иерархии)

Назначение: представление иерархических знаний, категоризация объектов

Пример: биологическая классификация видов

**Функциональные сети**

Описывают процессы, функциональные зависимости, причинно-следственные отношения

Используют отношения “влияет”, “вызывает”, “определяет”

Назначение: моделирование процессов, причинно-следственных связей

Пример: модели бизнес-процессов

**Сценарные сети** (скрипты)

Представляют стандартные последовательности событий

Используют отношения “следует за”, “предшествует”, “часть сценария”

Назначение: представление типовых ситуаций, планирование

Пример: сценарий посещения ресторана

**Ситуационные сети**

Описывают конкретные ситуации и их участников

Используют отношения “агент”, “объект”, “место”, “время”

Назначение: моделирование конкретных ситуаций

Пример: описание события в новостях

**Обобщенные или гибридные сети**

Комбинируют различные типы отношений

Используют механизмы наследования и процедурные вложения

Назначение: комплексное представление знаний

Пример: semantic web (RDF, OWL)

**Преимущества семантических сетей:** - Наглядность и интуитивность представления - Естественная поддержка наследования свойств - Возможность отображения ассоциативных связей - Легкость дополнения и модификации

**Недостатки:** - Сложность формализации правил вывода - Проблемы с представлением процедурных знаний - Потеря эффективности при большом объеме сети

Семантические сети широко используются в системах обработки естественного языка, экспертных системах, базах знаний и проекте Semantic Web.

**6. Формальная постановка задачи кластеризации и методы ее решения.**

**Кластеризация** – задача разделения множества объектов на группы (кластеры) так, чтобы объекты внутри кластера были более похожи друг на друга, чем на объекты из других кластеров.

**Формальная постановка задачи:**

Дано: - Множество объектов X = {x₁, x₂, …, xₙ} - Каждый объект xᵢ описывается вектором признаков xᵢ = (xᵢ₁, xᵢ₂, …, xᵢₘ) - Функция расстояния (меры сходства) d(xᵢ, xⱼ) между объектами

Требуется: - Разбить множество X на непересекающиеся подмножества (кластеры) C = {C₁, C₂, …, Cₖ} - Минимизировать внутрикластерные расстояния - Максимизировать межкластерные расстояния

**Основные методы кластеризации:**

**Центроидные методы**

**Алгоритм k-средних (k-means)**:

Выбор k начальных центров кластеров

Отнесение каждого объекта к ближайшему центру

Пересчет центров как средних значений объектов кластера

Повторение до сходимости

**k-медоидов (k-medoids)**:

Подобен k-means, но центры кластеров – реальные объекты

Преимущества: простота, эффективность, линейная сложность O(n) Недостатки: необходимость задавать k, чувствительность к выбору начальных центров, предположение о сферической форме кластеров

**Иерархические методы**

**Агломеративные** (снизу вверх):

Начало с одноэлементных кластеров

Последовательное объединение ближайших кластеров

Построение дендрограммы

**Дивизимные** (сверху вниз):

Начало с одного кластера, содержащего все объекты

Рекурсивное разделение на подкластеры

Преимущества: не требуют предварительного задания числа кластеров, визуализация результатов через дендрограмму Недостатки: квадратичная или кубическая сложность, нет возможности перераспределить объекты между кластерами

**Плотностные методы**

**DBSCAN** (Density-Based Spatial Clustering of Applications with Noise):

Выделение кластеров как областей с высокой плотностью объектов

Определение шумовых точек

**OPTICS**:

Усовершенствованный DBSCAN с переменной плотностью

Преимущества: нахождение кластеров произвольной формы, устойчивость к шуму, автоматическое определение числа кластеров Недостатки: сложность выбора параметров, проблемы с кластерами разной плотности

**Вероятностные методы**

**EM-алгоритм** (Expectation-Maximization):

Предположение о порождении данных смесью распределений

Итеративное вычисление параметров распределений и вероятностей принадлежности

**Байесовские методы**

Преимущества: мягкая кластеризация (вероятностная принадлежность), теоретическое обоснование Недостатки: чувствительность к начальным значениям, вычислительная сложность

**Спектральная кластеризация**

Построение матрицы сходства

Вычисление собственных векторов матрицы Лапласа графа

Кластеризация в пространстве собственных векторов

Преимущества: хорошо работает с нелинейно разделимыми данными, кластерами сложной формы Недостатки: вычислительная сложность, выбор функции сходства и количества кластеров

**Оценка качества кластеризации:** - Внутренние метрики: индекс силуэта, индекс Дэвиса-Болдина, дисперсионный критерий - Внешние метрики (при наличии эталонного разбиения): чистота кластеров, ARI, NMI

Выбор метода кластеризации зависит от особенностей данных, требуемой формы кластеров и вычислительных ресурсов.