

Архитектура ИИАС

Входной ассессмент для программ

«Искусственный интеллект и предиктивная аналитика»

«Искусственный интеллект и бизнес-аналитика в реальном секторе экономики»

Лекция 4

Интеллектуальная информационно-аналитическая система

Интеллектуальная информационно-аналитическая система (ИИАС) призвана рационализировать информационные процессы и решение задач управления во всех сферах экономики и производства.

Функционирование ИИАС обеспечивает более качественное решение задач при улучшении оперативности решения и снижении объема ресурсов.

Создание и эксплуатация ИИАС требуют привлечения значительных ресурсов.

Но не всегда и не везде эффективность ИИАС удовлетворяет установленным требованиям.

Только грамотно выполненный анализ способен обеспечить качество ИИАС требуемого уровня.

Качество ИИАС

На качество ИИАС влияют следующие факторы:

1. Внутренние:

- увеличение масштабов производства;
- усложнение взаимосвязей между субъектами хозяйства;
- увеличение объемов данных;
- усложнение процессов обработки данных.

2. Внешние:

- низкое качество ИИАС;
- экстенсивное развитие ИИАС.

Проблема эффективности ИИАС решается по нескольким направлениям:

1. Подготовка высококвалифицированных специалистов в области создания и эксплуатации ИИАС.
2. Повышение качества проектирование ИИАС.
3. Выпуск специализированной литературы.

Информация

Информация – это сведения, передаваемые людьми устным, письменным или другим способом.

Информация – обмен сведениями между людьми, человеком и автоматом, автоматом и автоматом, обмен сигналами в животном и растительном мире, передача признаков от клетки к клетке, от организма к организму.

Информация – результат логической переработки данных, который используется людьми в общественно-исторической практике путем применения различных форм, методов и средств.

Отличительная черта информации – способность подвергаться обработке в соответствии с потребностями человека, задачами специалиста.

Получение информации

Данные – сведения об объектах реального мира, представленные в регламентной форме.

Форма представления данных может регламентироваться приказами, стандартами, инструкциями, другими нормативными документами соответствующей управляющей системы.

Сведения – характеристики, признаки, свойства объектов.

В методологическом отношении объекты реального мира – предметы и явления (процессы) – воспринимаются через проявляемые ими свойства.

Информация, данные и сведения обладают функциональным свойством отображения объектов реального мира.

Данные → Сведения → Информация

Чем отличаются «данные» от «знаний»?

Исходные данные, требующие обработки обычно представлены в виде таблиц размерностью m строк на n столбцов.

Например, при m и n равным 3 такая таблица будет иметь вид:

$$D = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix},$$

где $a_i = |x_{i1} \ x_{i2} \ x_{i3}|$ – каждый объект реального мира a_i рассматривается в разрезе множества свойств $|x_{i1} \ x_{i2} \ \dots \ x_{in}|$.

Объектами могут быть любые сущности реального мира: живые организмы, сигналы, физические тела, сложные технические системы и т. д. Множество свойств определяется природой объекта и характером решаемой задачи.

Для успешного решения задачи анализа данных свойства должны отражать наиболее важные стороны объекта.

Пример

Например, рассмотрим проблему покупки нового телефона. Таким образом, объектом анализа будет выступать множество телефонов, доступных в некотором магазине. Допустим, что в качестве свойств пользователю важные следующие параметры: стоимость, размер оперативной памяти и емкость батареи.

В настоящий момент в магазине доступны следующие телефоны:

- OakPhone 12, 4Gb, 3000 mAh, 26 000 руб.
- Brave 10i, 3Gb, 3200 mAh, 15 500 руб.
- ReadMe Note 9 Pro, 6 Gb, 4000 mAh, 30 000 руб.

Представление данных

Тогда данные о телефонах можно представить следующей таблицей данных:

$$D^{Phones} = \begin{vmatrix} 4 & 3000 & 26\,000 \\ 3 & 3200 & 15\,500 \\ 6 & 4000 & 30\,000 \end{vmatrix}.$$

Можно сказать, что данные представляют собой совокупность отдельных конкретных фактов об объектах.

В процессе анализа была обнаружена следующая закономерность –

На стоимость телефона в большей степени влияет размер оперативной памяти, а не емкость батареи.

Знания (закономерности)

Обозначим признак «размер оперативной памяти» через x_1 , а признак «стоимость» через x_2 . Тогда обнаруженная закономерность может быть записана в виде следующих логических высказываний:

если $(x_1 < 4)$, тогда $(x_2 < 26\ 000)$;

если $(x_1 < 6)$, тогда $(x_2 < 30\ 000)$.

Полученные высказывания не содержат информации о характеристиках конкретного телефона, но отображают наши знания об обобщенных характеристиках объектов в таблице данных:

влиянии размера оперативной памяти на стоимость телефона.

Знания представляют собой краткое обобщенное описание основного содержания информации, представленной данными.

Ограничения ИИАС

С помощью ИИАС может перерабатываться далеко не вся информация, используемая для управления объектом, поскольку на предприятиях циркулируют огромные информационные потоки, играющие важную роль в принятии решений, обработка которых в их полном объеме с помощью компьютеров невозможна.

Ограничения:

1. Сложность структуризации информации и формализации процессов ее переработки.
2. Недостаточное количество вычислительных устройств (парка ЭВМ).
3. Отсутствие экономической целесообразности и др.

В ИИАС от объекта управления направляется только та часть информации, которую можно систематизировать и обрабатывать с помощью компьютера (10-20%).

Предметная область

Предметная область (ПрО) в контексте АИС – совокупность сведений о структуре, процессах и свойствах объектов управления и АИС, способах взаимосвязи и процессах взаимодействия между ними.

ПрО в значительной мере определяет специфику решения задач построения и функционирования АИС.

ПрО может быть описана в виде некоторой совокупности сведений о ее структуре, основных характеристиках, процессах, протекающих в ней, а также способов решения задач.

Значительная роль в ПрО принадлежит отношениям между объектами.

Цели и задачи ИИАС

Цель ИИАС – обеспечение специалистов информацией для решения задач управления, повышение уровня качества информации, выдаваемой специалистам – пользователям АИС.

Задачи ИИАС – это совокупность методов, средств и процедур, реализация которых обеспечивает достижение цели ИИАС.

Задачи ИИАС

Универсальные:

- выполнение процессов преобразования информации и выдача ее в удобном для восприятия виде;
- экономия ресурсов при выполнении процессов преобразования информации;
- развитие социального статуса работников, занятых в контуре функционирования АИС.

Специальные:

- обеспечение необходимого объема производства продукции;
- обеспечение ритмичности в производстве продукции или услуг предприятия (фирмы);
- проведение мероприятий по обеспечению заданного уровня качества продукции;
- проведение технико-экономического анализа;
- выполнение материально-технического снабжения предприятия;
- обеспечение маркетинговой деятельности предприятия;
- обеспечение организационно-технических мероприятий по развитию предприятия и др.

Функции ИИАС

Функция ИИАС – постоянный набор процедур, выполнение которых обеспечивает реализацию задач ИИАС.

Функции управления:

- планирование и прогнозирование деятельности предприятия;
- нормирование производственной деятельности;
- учет и отчетность;
- контроль производства;
- анализ производственной деятельности.

Информационно-технологические функции:

- сбор сведений об управляемом объекте;
- регистрация данных, передача данных, ввод данных, обработка данных;
- индексирование данных, поиск данных;
- хранение данных;
- актуализация информации, корректировка информации и др.

Структура ИИАС

Структура ИИАС – способ взаимосвязи элементов системы, обеспечивающий ее целостность.

Целостность ИИАС – свойство ИИАС, обеспечивающее устойчивость и функционирование системы в соответствии с ее назначением.

По характеру решаемых задач современные ИИАС можно условно разделить на четыре основных класса:

1. Автоматизированные системы обработки данных – характеризуются большим объемом исходных данных и несложностью алгоритмов их обработки.
2. Автоматизированные информационно-поисковые системы – предназначены для поиска и выдачи информации по запросу пользователя.
3. Автоматизированные системы управления – обеспечивают обработку данных по алгоритму оптимизации решения экономической задачи.
4. Автоматизированные интеллектуальные информационные системы – предназначены для генерации новых знаний, не содержащихся в исходных данных в явном виде.

Обеспечивающая часть структуры ИИАС

Обеспечивающая часть:

- информационное обеспечение (базы данных, базы знаний и т.д.);
- лингвистическое обеспечение (классификаторы, ИПЯ и т.д.);
- техническое обеспечение (ЭВМ, средства передачи данных, средства копирования и т.д.);
- программное обеспечение (ОС, прикладные программы, средства разработки и т.д.);
- математическое обеспечение (модели, методы, алгоритмы и т.д.);
- организационное обеспечение (персонал, проектно-техническая документация и т.д.);
- правовое обеспечение (нормативная документация, стандарты, должностные инструкции и т. д.).

Базы знаний

База знаний – совокупность знаний, организованная по принципам порождения знаний, явно не присутствующих в исходных данных.

Статическая БЗ содержит сведения, отображающие особенности конкретной предметной области и остающиеся неизменными в ходе решения задачи.

Динамическая БЗ применяется для организации сведений, важных для решения конкретной задачи и изменяющихся в процессе ее решения.

Генерация БЗ выполняется на основе механизма интеллектуального анализа данных с помощью набора сведений, правил, аппарата логического вывода и др.

Формирование базы знаний

1. Извлечение знаний – процесс приобретения материализованных знаний из источников информации с помощью некоторой совокупности методов и процедур, позволяющих переходить от знаний в текстовой или структурированной форме к их аналогам, адаптированным для ввода в БЗ.
2. Получение знаний – процесс приобретения вербализуемых и невербализуемых знаний эксперта, основанный на использовании непосредственно им самим или инженером по знаниям приемов, процедур, методов и инструментальных средств.
3. Формирование знаний – процесс автоматического приобретения ИИАС нового и полезного знания из исходной и текущей информации, которое в явном виде эксперты не формируют.
4. Приобретение знаний – процесс, основанный на переносе знаний из разнообразных источников в БЗ путем использования различных методов, моделей, алгоритмов и средств интеграции знаний.

Лингвистическое обеспечение ИИАС

- информационно-поисковый язык (ИПЯ) – упорядоченное множество понятий, терминов определенной предметной области;
- методики индексирования документов – совокупность логических операций по отображению содержания документов и запросов средствами принятого ИПЯ;
- классификаторы – типы, форматы, структуры информационных категорий (данные, показатели, записи, таблицы, файлы, документы, массивы и др.);
- критерий смыслового соответствия (критерий выдачи) документов и (или) поисковых образов документов по различным классам документальной информации, содержащейся в БД.

Информационный поиск

Классификатор – систематизированная совокупность наименований и кодов языковых элементов определенной предметной области. Классификаторы строятся по иерархическому принципу.

Поисковый образ документа (ПОД) – совокупность ключевых слов, кодовых обозначений, отображающих содержание документа, адрес хранения и его системный номер (идентификатор).

Поисковый образ запроса (ПОЗ) – это совокупность ключевых слов, отображающих содержание запроса и условия поиска документов.

Критерий смыслового соответствия – правило, определяющее степень смысловой близости ПОД и ПОЗ и формирующее решение о выдаче данного документа в ответ на запрос пользователя.

Обеспечивающая часть структуры ИИАС

Обеспечивающая часть:

- информационное обеспечение (базы данных, базы знаний и т.д.);
- лингвистическое обеспечение (классификаторы, ИПЯ и т.д.);
- **техническое обеспечение** (ЭВМ, средства передачи данных, средства копирования и т.д.);
- **программное обеспечение** (ОС, прикладные программы, средства разработки и т.д.);
- **математическое обеспечение** (модели, методы, алгоритмы и т.д.);
- **организационное обеспечение** (персонал, проектно-техническая документация и т.д.);
- **правовое обеспечение** (нормативная документация, стандарты, должностные инструкции и т. д.).

Функциональная часть структуры ИИАС

Функциональная часть:

- прогнозирование;
- планирование;
- учет;
- анализ;
- контроль;
- регулирование.

Методы и средства контроля качества обработки информации

Источники ошибок:

- 0,4% ошибок возникают по причине неисправности технических устройств;
- 21,6% ошибок обусловлены недостатками проекта ИС;
- 78% ошибок обусловлено человеческим фактором.

Методы контроля качества:

1. Поддержание целостности БД.
2. Применение системы контроля обработки входных данных.
3. Контроль достоверности выходных данных.

Способы контроля:

1. Ручной, или визуальный, способ.
2. Механизированный способ.
3. Автоматизированный метод контроля.
4. Автоматический метод.

Методы программного контроля информации

Лексический – проверка правильности формата значений реквизитов:

- контрольный разряд;
- код реквизита по словарю;
- класс значений реквизита.

Синтаксический – проверка количества элементов в форматах и порядка их расположения:

- количество символов в значении поля;
- количество значений полей в записи;
- количество записей в файле;
- количество записей в базе данных.

Логический – проверка содержательной взаимосвязи между отдельными значениями:

- соотношения типа $=$, \neq , $>$, $<$, \geq , \leq ;
- соотношения типа $a+b-c=q$.

Арифметический – проверка равенства контрольного значения и значения группового и/или итогового значения.

Требования к методам программного контроля информации

1. Повышение уровня достоверности и полноты информации.
2. Снижение объема временных, трудовых, материальных и финансовых ресурсов, используемых в технологии обработки данных.
3. Адаптация к сравнительно широкому классу обрабатываемых форматов данных.
4. Возможность применения в других системах.
5. Максимальный состав функций лексического, синтаксического, логического и арифметического контроля при минимальном физическом объеме программного модуля.

Интеллектуальный анализ данных

Подходы к интеллектуальному анализу данных в рамках ИИАС:

1. Экспертные знания:

- формирование базы знаний;
- формирование набора правил;
- классический (детерминированный) алгоритм;
- тестирование решения.

2. Машинное обучение:

- сбор данных;
- формирование выборок;
- предобработка данных;
- конструирование признаков;
- обучение модели;
- тестирование модели.

Детали реализации зависят от бизнес-цели проекта, его технологических целей, архитектуры и т. д.