

Интеллектуальный анализ данных





Понять, что представляет собой интеллектуальный анализ данных, зачем он нужен, каковы его основные направления и области применения, а также изучить основные его методы.





Направления исследований



1

Интеллектуальный анализ данных
(Data Mining)

2

Машинное обучение
(Machine Learning)

3

Обнаружение знаний в базах данных
(Knowledge Discovery in Databases)

4

Распознавание образов
(Pattern Recognition)



Индуктивное формирование баз знаний (ИФБЗ) –

получение общего знания о некоторой совокупности объектов на основании анализа единообразного описания конечного множества отдельных представителей этой совокупности – обучающей выборки (данных).

Некоторые области применения



1

Медицинская и
техническая
диагностика,
научные
исследования,
робототехника,
автоматизация

2

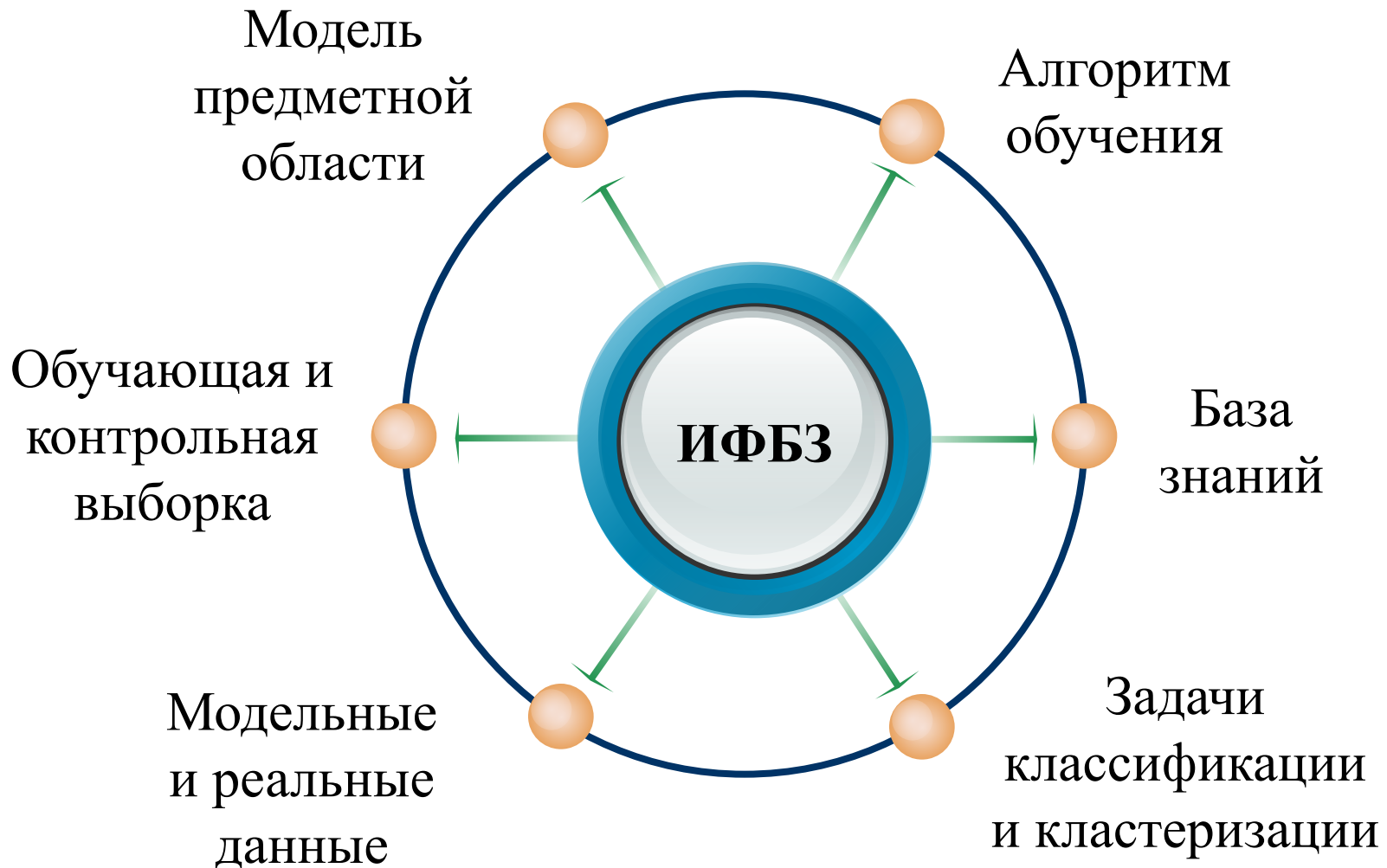
Анализ
информации
в интернете:
поиск, реклама,
популярность,
плагиат,
модерация

3

Военные
приложения:
мониторинг,
распознавание
образов,
прогноз, тактика
и стратегия

4

Маркетинговые
исследования,
финансовые
рынки,
кредиты,
телефония,
банкинг





Модели предметных областей:

проблемно-независимые *vs.* проблемно-ориентированные.

Исследование алгоритмов обучения:

на модельных данных *vs.* на реальных данных.

Результат работы:

детерминированный *vs.* вероятностный.

Объяснение результатов:

хорошо интерпретируемое *vs.* плохо интерпретируемое.



Индуктивная база знаний

формирование

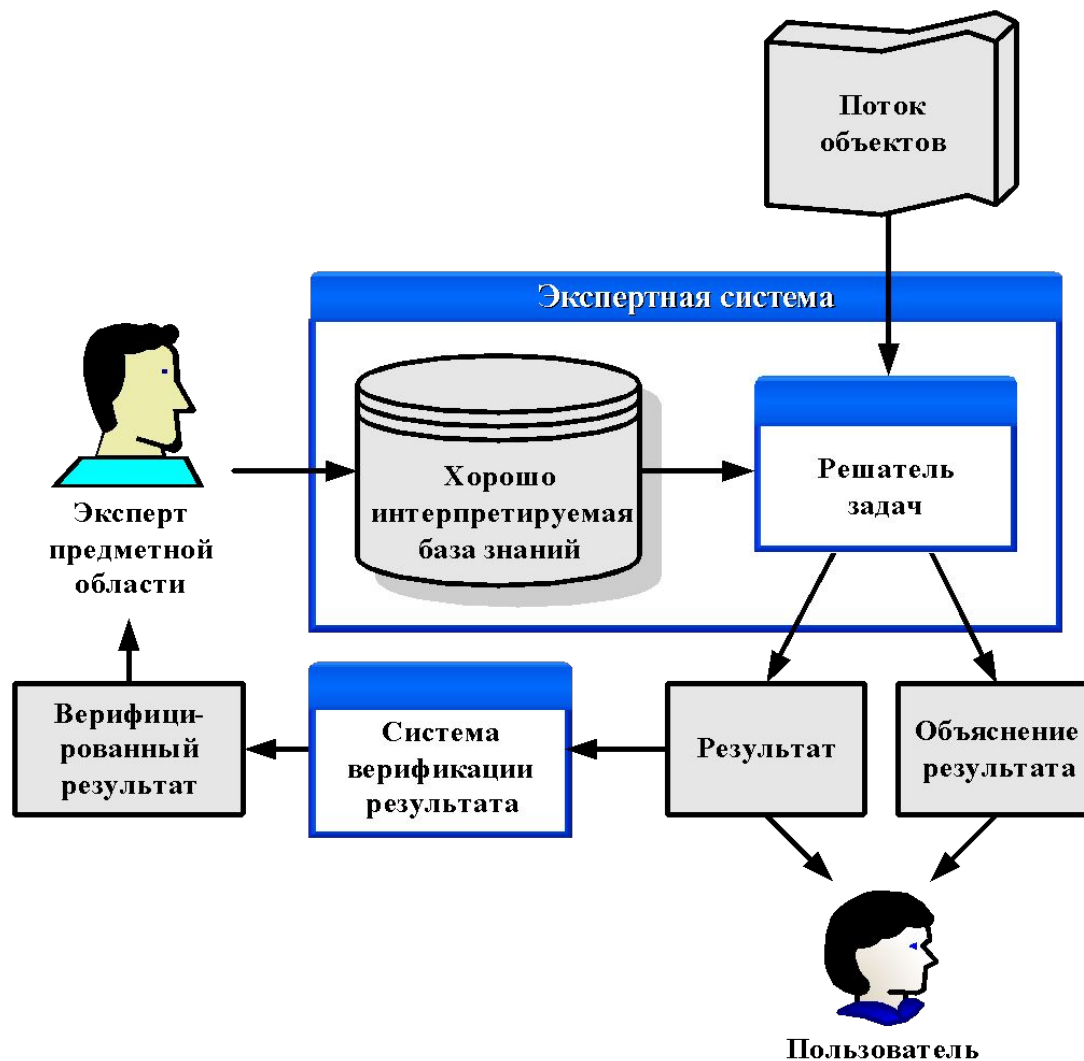
Модель

**Постановка
задачи**

**Алгоритм
обучения**

**Обучающая
выборка**

Экспертное формирование хорошо интерпретируемых баз знаний





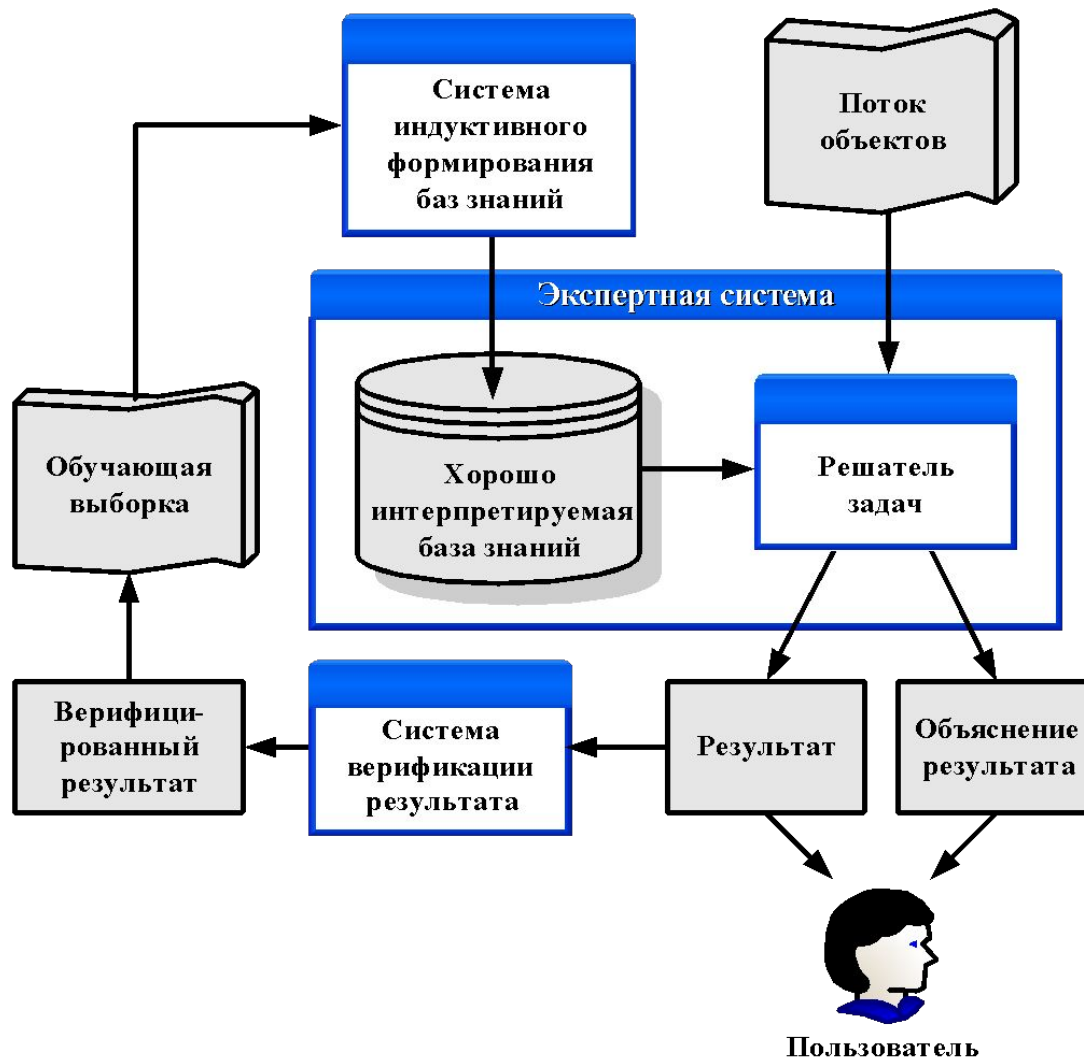
Высокая стоимость и трудозатраты
на создание и сопровождение
экспертных баз знаний

Субъективность, неполнота и
«ручное» формирование
экспертных баз знаний

Плохая интерпретируемость
современных индуктивно
формируемых баз знаний

**Необходимость
нового подхода**

Индуктивное формирование хорошо интерпретируемых баз знаний



Темы для курсовых проектов



1. История развития экспертных систем
2. Обзор современных экспертных систем
3. Проблемно-ориентированные и проблемно-независимые системы интеллектуального анализа данных
4. Современные направления исследований в области индуктивного формирования баз знаний
5. Базы знаний. История развития, современное состояние
6. Обзор редакторов баз знаний
7. Классификация алгоритмов обучения