



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ ФЕ-
ДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО
ОБРАЗОВАНИЯ
«Дальневосточный федеральный университет»

ИНСТИТУТ МАТЕМАТИКИ И КОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ
Департамент программной инженерии и искусственного интеллекта

**РАЗРАБОТКА СИСТЕМЫ, ОСНОВАННОЙ НА ЗНАНИЯХ,
ДЛЯ ПРЕДМЕТНОЙ ОБЛАСТИ
«КЛАССИФИКАЦИЯ ЗВЁЗД»**

ОТЧЕТ ПО ЛАБОРАТОРНОЙ РАБОТЕ
по дисциплине «Методы и технологии интеллектуализации программных систем»
по образовательной программе подготовки бакалавров по направлению
09.03.04 «Программная инженерия»

Выполнил:

студент гр. Б9121-09.03.04

_____ Ермак В.М.

Руководитель:

профессор департамента ПИиИИ

_____ Артемьева И.Л.,

д.т.н., профессор

г. Владивосток
2025

Оглавление

Введение.....	4
1 Анализ предметной области	5
1.1 Анализ множества задач профессиональной деятельности.....	5
1.2 Анализ знаний предметной области	6
1.3 Анализ смысла ситуаций	7
1.4 Связь между знаниями и действительностью	8
2 Построение модели предметной области	9
2.1 Построение модели непримитивной онтологии.....	9
2.2 Построение модели знаний.....	12
2.3 Построение модели ситуации.....	14
3 Проект системы, основанной на знаниях	15
3.1 Спецификация задачи	15
3.2 Архитектура системы.....	16
3.3 Use-case диаграмма.....	18
3.4 Проект интерфейса редактора знаний	19
3.5 Проект интерфейса системы ввода исходных данных	25
3.6 Проект системы вывода результата и объяснений	28
4 Реализация системы.....	29
4.1 Общая структура реализации	29
4.2 Реализация логического классификатора	30
4.3 Реализация интерфейсов.....	31
4.4 Нейросетевая классификация.....	33
4.5 Особенности реализации	35
Заключение	37

Список литературы	38
-------------------------	----

Введение

Система, основанная на знаниях, представляет собой программный комплекс, который накапливает знания специалистов в определённой области и помогает пользователям с различным уровнем квалификации эффективно решать задачи. Основными компонентами таких систем являются редактор знаний, решатель задач, подсистема вывода и объяснения результатов, а также система ввода исходных данных.

Одной из популярных задач в таких системах является классификация, которая заключается в определении класса, к которому принадлежит объект, на основе его известных характеристик. В данном случае задача заключается в классификации звёзд.

С развитием информационных технологий и автоматизацией множества процессов, рынок программного обеспечения продолжает стремительно расти. Если раньше приложения ограничивались несколькими специализированными областями, то сегодня их роль охватывает практически все сферы жизни и бизнеса. Особенно актуальной становится классификация объектов в астрономии, в частности классификация звёзд.

Цель курсового проекта – разработать проект системы, основанной на знаниях, для классификации звёзд. В систему будут включены следующие компоненты: редактор знаний, решатель задач, подсистема вывода и объяснения результатов, а также система ввода исходных данных. Для достижения этой цели необходимо решить следующие задачи:

1. Провести анализ предметной области «классификация звёзд»;
2. Разработать модель предметной области, включая классификацию различных классов звёзд;
3. Разработать проект системы для классификации звёзд, которая будет включать все необходимые компоненты.

1 Анализ предметной области

1.1 Анализ множества задач профессиональной деятельности

1.1.1 Описание задачи предметной области

В данной предметной области решается задача классификации звёзд. Объектом задачи являются звезды, которые необходимо отнести к определенному классу (типу). Результатом решения задачи является определение класса звезды на основе её характеристик. Для определения класса звезды анализируются следующие параметры:

1. Температура (в Кельвинах);
2. Цвет (например, голубой, белый, жёлтый и т.д.);
3. Показатель цвета (B-V);
4. Абсолютная звёздная величина (M_V).

Задача будет решаться для ограниченного множества типов звёзд, и значения свойств «Температура», «Цвет», «Показатель цвета», «Абсолютная звёздная величина» будут представлены числовыми или категориальными значениями.

Значения всех свойств известны на момент решения задачи и используются для классификации звёзд.

1.1.2 Кодирование значений свойств

Для кодирования значения свойства «Температура» используется диапазон температур, например, для класса О звезды имеют температуру более 30 000 К, для класса В – от 10 000 до 30 000 К и так далее.

Для кодирования значения свойства «Цвет» используются категории, такие как «Голубой», «Белый», «Жёлтый», «Оранжевый», «Красный».

Для кодирования значения свойства «Показатель цвета» применяются числовые значения, например, для класса О показатель цвета может быть $-0,3$, для класса В $-0,2$ и так далее.

Для кодирования значения свойства «Абсолютная звёздная величина» используются значения в диапазоне от $-5,7$ до $+16$, в зависимости от класса звезды.

1.1.3 Классы «Типы звёзд»

1. O
2. B
3. A
4. F
5. G
6. K
7. M

1.2 Анализ знаний предметной области

Каждому типу звезды могут быть сопоставлены различные характеристики, которые имеют значение для её классификации. Множество таких характеристик образует описание конкретной звезды. Если звезда не соответствует одному из заранее определённых типов, она не будет классифицирована.

1.2.1 Термины для описания знаний

Утверждения, связанные с классификацией звёзд, могут быть интерпретированы следующим образом: «Каждое значение свойства, входящего в описание типа звезды, является допустимым значением для этого типа звезды».

Основные термины для описания знаний:

Типы звёзд – понятие, соответствующее множеству классов звёзд (например, «O», «B», «A» и т.д.);

Свойства – понятие, соответствующее множеству характеристик звёзд (например, «Температура», «Цвет», «Показатель цвета», «Абсолютная звёздная величина»);

Возможные значения – понятие, соответствующее допустимым значениям свойств звёзд (например, для свойства «Температура» это могут быть значения от 30 000 К до 2 500 К и т.д.);

Описание свойств типа – понятие, соответствующее множеству характеристик, которые могут быть присущи каждому типу звезды.

Значение для типа – понятие, которое отражает допустимые значения для свойств звезды, например, «Температура > 30 000 К» для типа О.

1.2.2 Ограничения целостности знаний

Для каждого типа звезды множество свойств, принадлежащих его описанию, является непустым.

Для каждого типа звезды, для каждого свойства, характерного для данного типа, множество значений этого свойства является непустым подмножеством множества возможных значений этого свойства.

1.3 Анализ смысла ситуаций

Ситуация – это процесс определения типа звезды на основе её характеристик. В каждой ситуации рассматривается только одна звезда, для которой известны её характеристики. Каждая звезда может принадлежать только одному типу, и на основе этих характеристик определяется класс звезды.

Пример ситуации: у нас есть звезда с характеристиками: «Температура» – 25 000 К, «Цвет» – голубой, «Показатель цвета» – $-0,3$, «Абсолютная звёздная величина» – $-4,1$. В результате, звезда будет отнесена к типу «В», с соответствующими характеристиками.

1.3.1 Термины для описания ситуаций

Терминами для описания ситуаций являются «Экземпляр звезды», «Описание свойств экземпляра», «Значение для экземпляра».

Экземпляр звезды – объект, значение которого берется из множества классов «Типы звёзд» (например, «О», «В», «А»).

Описание свойств экземпляра – понятие, которое обозначает множество характеристик, описывающих данную звезду (например, «Температура», «Цвет», «Показатель цвета»).

Значение для экземпляра – понятие, которое описывает значения характеристик для конкретной звезды (например, «25 000 К», «Голубой», «-0,3»).

1.3.2 Ограничения целостности ситуаций

«Описание свойств экземпляра» является непустым множеством.

Для каждого элемента из «описания свойств экземпляра» должно быть задано «значение для экземпляра».

«Значения для экземпляра» должны принадлежать «возможным значениям» соответствующих «свойств».

Если экземпляр звезды соотнесён с типом, то его «описание свойств экземпляра» должно совпадать с «описанием свойств типа» для данного типа звезды.

Для каждого свойства из описания значение должно входить в «значения для типа» соответствующего типа звезды.

1.4 Связь между знаниями и действительностью

«Описание свойств экземпляра» должно соответствовать «описанию свойств типа», если экземпляр отнесён к определённом типу.

Каждое «значение для экземпляра» должно находиться в пределах «возможных значений» соответствующего «свойства».

Если задан «тип звезды», то все значения экземпляра должны соответствовать «значениям для типа», описанным в базе знаний.

2 Построение модели предметной области

2.1 Построение модели непримитивной онтологии

В основе проектирования информационной системы лежит моделирование предметной области. Для того чтобы создать адекватную модель, которая будет отражать все аспекты функционирования системы классификации типов звёзд, необходимо иметь системное представление модели, которая имитирует структуру и функционирование исследуемой области. Под моделью предметной области понимается система, имитирующая структуру или функционирование исследуемой области и отвечающая основному требованию – быть адекватной ей.

2.1.1 Определение вспомогательных терминов

(1.1) Множество значений: обозначается как $(\{\} N) \cup ([I) \cup ([R)$. Этот термин описывает множество областей возможных значений всех свойств звёзд. Например, такие области могут включать множества имён (для качественных свойств), целочисленные или вещественные интервалы (для количественных свойств).

2.1.2 Определение системы понятий знаний

Описание сортов терминов для описания знаний

(2.1) сорт свойства: $\{\} N \setminus \emptyset$

термин "свойства" обозначает конечное множество названий свойств типов звёзд, таких как температура, цвет, абсолютная звёздная величина, спектральный класс.

(2.2) сорт типы звёзд: $\{\} N \setminus \emptyset$

термин "типы звёзд" обозначает конечное множество названий типов звёзд, например, звезды главной последовательности, красные гиганты, белые карлики.

(2.3) сорт возможные значения: (свойства \rightarrow множество значений)

термин "возможные значения" обозначает функцию, которая сопоставляет каждому свойству звезды возможные значения этого свойства.

(2.4) сорт описание свойств типа: (типы звёзд \rightarrow ($\{\}$ свойства))

термин "описание свойств типа" обозначает функцию, которая сопоставляет каждому типу звезды подмножество множества свойств, описывающих этот тип звезды.

(2.5) сорт значение для типа: ($\{(v: (\times \text{ типы звёзд, свойства})) \mid \pi(2, v) \in \text{описание свойств типа } (\pi(1, v))\} \rightarrow \text{множество значений}$)

термин "значение для типа" обозначает функцию, которая сопоставляет каждому типу звезды свойство и его значение.

Ограничения целостности знаний

(3.1) (v : типы звёзд) (описание свойств типа (v) $\neq \emptyset$)

для каждого типа звезды множество свойств, принадлежащих описанию свойств этого типа, является непустым множеством. То есть для любого типа звезды должны быть определены хотя бы одно свойство.

(3.2) ($v1$: типы звёзд) ($v2$: описание свойств типа ($v1$)) (значение для типа ($v1$, $v2$) $\neq \emptyset$) & (значение для типа ($v1$, $v2$) \subseteq возможные значения ($v2$))

для любого типа звезды, для любого свойства, характерного для данного типа, множество значений этого свойства должно быть непустым подмножеством множества возможных значений этого свойства. То есть для каждого типа звезды и его свойства необходимо указать хотя бы одно допустимое значение, и это значение должно быть из предустановленного множества возможных значений для данного свойства.

2.1.3 Определение системы понятий действительности

Описание сортов терминов для описания ситуаций

(4.1) сорт экземпляр звезды: $\{\}$ типы звёзд

Термин "экземпляр звезды" обозначает конкретный экземпляр типа звезды.

(4.2) (v : свойства) сорт v : возможные значения

Ограничения целостности ситуаций

(5.1) описание свойств экземпляра $\neq \emptyset$

Множество свойств, описывающих звезду, не должно быть пустым.

(5.2) ($\forall v \in \text{описание свойств экземпляра}$) значение для экземпляра(v) $\neq \emptyset$

Для каждого свойства из описания экземпляра должно быть задано значение.

(5.3) ($\forall v \in \text{описание свойств экземпляра}$) значение для экземпляра(v) \in возможные значения(v)

Значения всех характеристик экземпляра должны находиться в пределах допустимых значений свойств.

(5.4) если экземпляр звезды \in типы звёзд, то описание свойств экземпляра = описание свойств типа (экземпляр звезды)

Если экземпляр отнесён к определённому типу, то он должен быть описан теми же свойствами, что и этот тип.

(5.5) если экземпляр звезды \in типы звёзд, то ($\forall v \in \text{описание свойств экземпляра}$) значение для экземпляра(v) \in значение для типа (экземпляр звезды, v)

Все значения характеристик экземпляра должны входить в допустимые значения соответствующего типа звезды.

2.1.4 Связь между знаниями и действительностью

(6.1) ($\forall v \in \text{описание свойств экземпляра}$) значение для экземпляра(v) \in возможные значения(v)

Каждое значение свойства должно быть допустимым в соответствии с определением свойства в базе знаний.

(6.2) если экземпляр звезды \in типы звёзд, то ($\forall v \in \text{описание свойств типа(экземпляр звезды)}$) значение для экземпляра(v) \in значение для типа(экземпляр звезды, v)

Если экземпляру звезды задан тип, то его значения должны соответствовать допустимым значениям свойств данного типа.

(6.3) если экземпляр звезды \in типы звёзд, то описание свойств экземпляра = описание свойств типа(экземпляр звезды)

Набор характеристик экземпляра должен совпадать с описанием типа, если экземпляр отнесён к типу.

2.2 Построение модели знаний

(7.1) свойства \equiv {температура, цвет, показатель цвета, абсолютная звёздная величина}

рассматриваются четыре свойства для звёзд: температура, цвет, показатель цвета, абсолютная звёздная величина.

(7.2) типы звёзд \equiv {O, B, A, F, G, K, M}

рассматриваются типы звёзд: O, B, A, F, G, K, M.

(7.3) возможные значения $\equiv (\lambda(v: \{\text{температура, цвет, показатель цвета, абсолютная звёздная величина}\}) / (v \in \{\text{температура}\} \Rightarrow R[2500, 30000]) (v \in \{\text{цвет}\} \Rightarrow \{\text{голубой, бело-голубой, белый, жёлто-белый, жёлтый, оранжевый, красный}\}) (v \in \{\text{показатель цвета}\} \Rightarrow R[-0.3, 1.5]) (v \in \{\text{абсолютная звёздная величина}\} \Rightarrow R[-5.7, 16]) /)$

возможными значениями свойства «Температура» являются значения в пределах от 2500 К до 100000 К; свойства «Цвет» – голубой, бело-голубой, белый, жёлто-белый, жёлтый, оранжевый, красный; свойства «Показатель цвета» – значения от –0.3 до 1.5, свойства «Абсолютная звёздная величина» – значения от –5.7 до 16.

(7.4) описание свойств типа $\equiv (\lambda(v: \{O, B, A, F, G, K, M\}) / (v \in \{O, B, A, F, G, K, M\} \Rightarrow \{\text{температура, цвет, показатель цвета, абсолютная звёздная величина}\}) /)$

у всех типов звёзд одинаковое описание свойств типа, состоящее из температуры, цвета, показателя цвета и абсолютной звёздной величины.

(7.5) значение для типа $\equiv (\lambda(v: \{<O, \text{температура}>, <O, \text{цвет}>, <O, \text{показатель цвета}>, <O, \text{абсолютная звёздная величина}>, <B, \text{температура}>, <B, \text{цвет}>, <B, \text{показатель цвета}>, <B, \text{абсолютная звёздная величина}>, <A, \text{температура}>, <A, \text{цвет}>, <A, \text{показатель цвета}>, <A, \text{абсолютная звёздная величина}>, <F, \text{температура}>, <F, \text{цвет}>, <F, \text{показатель цвета}>, <F, \text{абсолютная звёздная величина}>, <G, \text{температура}>, <G, \text{цвет}>, <G, \text{показатель цвета}>, <G, \text{абсолютная звёздная величина}>, <K, \text{температура}>, <K, \text{цвет}>, <K, \text{показатель цвета}>, <K, \text{абсолютная звёздная величина}>, <M, \text{температура}>, <M, \text{цвет}>, <M, \text{показатель цвета}>, <M, \text{абсолютная звёздная величина}>\}) /$

$(v \in \{<O, \text{температура}>\} \Rightarrow > 30000), (v \in \{<O, \text{цвет}>\} \Rightarrow \text{голубой}), (v \in \{<O, \text{показатель цвета}>\} \Rightarrow -0.3), (v \in \{<O, \text{абсолютная звёздная величина}>\} \Rightarrow -5.7...-3.3),$

$(v \in \{<B, \text{температура}>\} \Rightarrow 10000-30000), (v \in \{<B, \text{цвет}>\} \Rightarrow \text{бело-голубой}), (v \in \{<B, \text{показатель цвета}>\} \Rightarrow -0.2), (v \in \{<B, \text{абсолютная звёздная величина}>\} \Rightarrow -4.1...+1.5),$

$(v \in \{<A, \text{температура}>\} \Rightarrow 7400-10000), (v \in \{<A, \text{цвет}>\} \Rightarrow \text{белый}), (v \in \{<A, \text{показатель цвета}>\} \Rightarrow 0), (v \in \{<A, \text{абсолютная звёздная величина}>\} \Rightarrow +0.7...+3.1),$

$(v \in \{<F, \text{температура}>\} \Rightarrow 6000-7400), (v \in \{<F, \text{цвет}>\} \Rightarrow \text{жёлто-белый}), (v \in \{<F, \text{показатель цвета}>\} \Rightarrow +0.4), (v \in \{<F, \text{абсолютная звёздная величина}>\} \Rightarrow +2.6...+4.6),$

$(v \in \{<G, \text{температура}>\} \Rightarrow 5000-6000), (v \in \{<G, \text{цвет}>\} \Rightarrow \text{жёлтый}), (v \in \{<G, \text{показатель цвета}>\} \Rightarrow +0.6), (v \in \{<G, \text{абсолютная звёздная величина}>\} \Rightarrow +4.4...+6.0),$

$(v \in \{<K, \text{температура}>\} \Rightarrow 3800-5000), (v \in \{<K, \text{цвет}>\} \Rightarrow \text{оранжевый}), (v \in \{<K, \text{показатель цвета}>\} \Rightarrow +1.0), (v \in \{<K, \text{абсолютная звёздная величина}>\} \Rightarrow +5.9...+9.0),$

$(v \in \{<M, \text{температура}>\} \Rightarrow 2500\text{--}3800), (v \in \{<M, \text{цвет}>\} \Rightarrow \text{красный}), (v \in \{<M, \text{показатель цвета}>\} \Rightarrow +1.5), (v \in \{<M, \text{абсолютная звёздная величина}>\} \Rightarrow +9.0\text{...}+16) /)$

2.3 Построение модели ситуации

2.3.1 Формальная постановка задачи

входные данные \equiv {описание свойств экземпляра, значение для экземпляра}

выходные данные \equiv {экземпляр звезды}

Связь задаёт онтология и модель знаний.

2.3.2 Пример модели ситуации

(8.1) описание свойств экземпляра \equiv {температура, цвет, показатель цвета, абсолютная звёздная величина}

(8.2) значение для экземпляра $\equiv (\lambda(v: \{\text{температура, цвет, показатель цвета, абсолютная звёздная величина}\} / (v = \text{температура} \Rightarrow 5000) (v = \text{цвет} \Rightarrow \text{жёлтый}) (v = \text{показатель цвета} \Rightarrow 0.5) (v = \text{абсолютная звёздная величина} \Rightarrow +5) /)$

(8.3) экземпляр звезды $\equiv G$

3 Проект системы, основанной на знаниях

3.1 Спецификация задачи

Спецификация задачи для системы, основанной на знаниях, имеет вид: $TS = \langle S, \Delta, \Pi, \Psi \rangle$, где:

S – модель предметной области. В данном контексте это описание характеристик звезд, таких как температура, цвет, показатель цвета и абсолютная звёздная величина, и их связь с типами звезд.

Δ – спецификация входных данных задачи, которая состоит из двух частей:

$\Delta 1$ – описание имен входных параметров задачи, например, температура (К), цвет, показатель цвета, абсолютная звёздная величина.

$\Delta 2$ – описание связи между входными параметрами и терминами предметной области. В данном случае это связи, такие как: температура (К) влияет на тип звезды (О, В, А и т. д.), и абсолютная звёздная величина связана с типами звёзд.

Π – спецификация выходных данных задачи, которая включает:

$\Pi 1$ – описание выходных параметров задачи, например, определение типа звезды (О, В, А и т. д.).

$\Pi 2$ – описание связи между выходными параметрами и терминами предметной области. Тип звезды (О, В, А и т. д.) связан с диапазоном температур и показателей цвета.

Ψ – ограничения на значения терминов для описания ситуаций. Для данной задачи это ограничения на диапазоны температур, показателей цвета или абсолютной звёздной величины для каждого типа звезды. Например, для типа А температура может быть от 7400 до 10 000 К, и показатель цвета равен 0.

Пример спецификации задачи для классификации звезд

$\Delta 1 = \{\text{температура, цвет, показатель цвета, абсолютная звёздная величина}\},$

$\Delta 2$ образуют следующие предложения:

температура \rightarrow тип звезды,

показатель цвета \rightarrow тип звезды,

абсолютная звёздная величина \rightarrow тип звезды.

$\Pi_1 = \{\text{тип звезды}\},$

Π_2 образуют следующие предложения:

тип звезды $\subseteq \{O, B, A, F, G, K, M\},$

$(v: \text{тип звезды}) \rightarrow (\text{температура}(v) \in \text{диапазон для типа звезды}),$

$(v: \text{тип звезды}) \rightarrow (\text{показатель цвета}(v) \in \text{диапазон для типа звезды}).$

$\Psi = \{\text{температура} > 30\,000 \text{ для типа } O, \text{ температура } 2500\text{--}3800 \text{ для типа } M, \text{ и т.д.}\}.$

Алгоритм решения задачи

Получаем входные данные: температуру, цвет, показатель цвета, абсолютную звёздную величину.

Для каждого типа звезды проверяем, в какой диапазон температуры, показателя цвета и абсолютной звёздной величины попадают данные.

На основе полученных данных, используя заранее заданные диапазоны, определяем тип звезды.

Выводим тип звезды как результат классификации.

Пример алгоритма:

Если температура $> 30\,000$ К, то тип звезды – O.

Если температура $10\,000\text{--}30\,000$ К, то тип звезды – B.

Если температура $7400\text{--}10\,000$ К, то тип звезды – A.

И так далее для всех типов.

Этот процесс продолжается до тех пор, пока не будет найден тип звезды, соответствующий всем входным данным.

3.2 Архитектура системы

Была построена контекстная диаграмма верхнего уровня системы [Рисунок 1].



Рисунок 1 – Контекстная диаграмма системы

На вход принимаются знания ПО и исходные данные. Результатом работы системы является результат решения задачи классификации звезды в понятном пользователю виде с обоснованием решения.

Ниже представлена архитектурно-контекстная диаграмма ИС для определения типа звезды [Рисунок 2].

Компонентами системы являются:

1. Редактор знаний и база знаний;
2. Система ввода исходных данных;
3. Решатель задач;
4. Система вывода результата и объяснений.

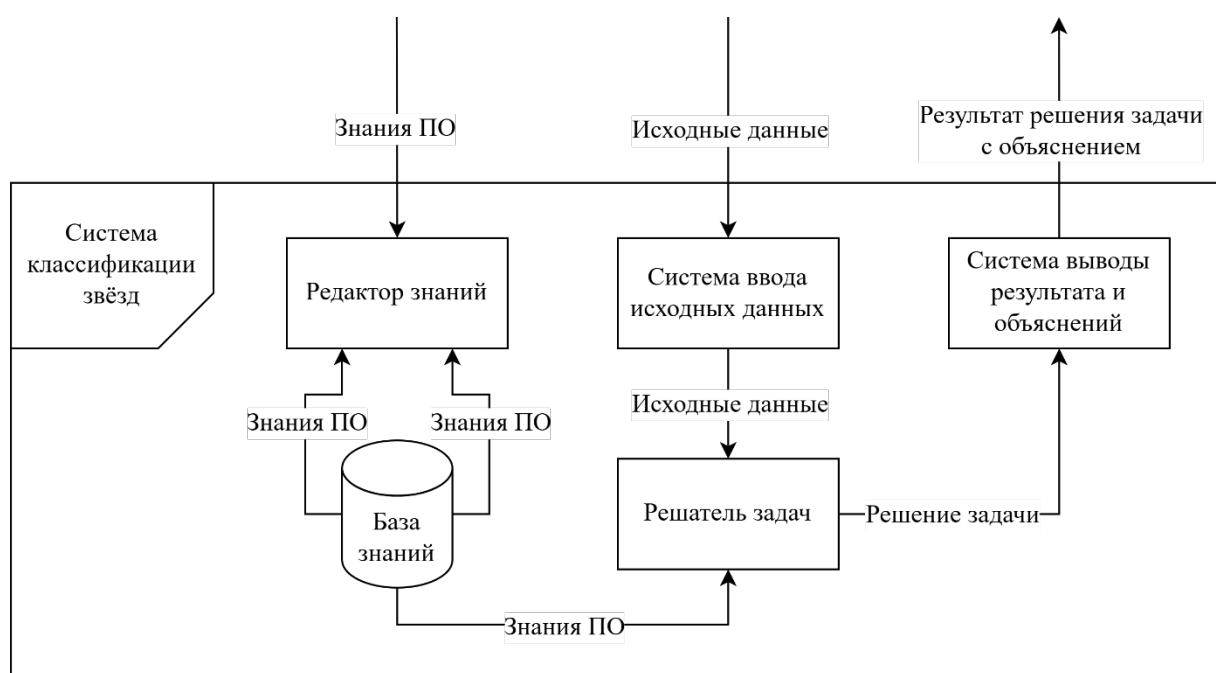


Рисунок 2 – Архитектурно-контекстная диаграмма системы

Редактор знаний предназначен для создания и изменения базы знаний, в которую специалисты ПО вносят свои знания. Эта база служит для хранения информации о программном обеспечении, добавленной через редактор.

Через пользовательский интерфейс в систему вводятся как данные от пользователей, так и знания программного обеспечения для решения задач. На этом этапе также происходит проверка входных данных на правильность и полноту.

Система вывода результатов и объяснений используется для представления решения задачи в доступной форме, с детальным объяснением полученного результата.

Решатель задач предоставляет результаты решения, основываясь на введенных пользователем данных.

3.3 Use-case диаграмма

Была составлена use-case диаграмма ИС для классификации звезд [Рисунок 3].

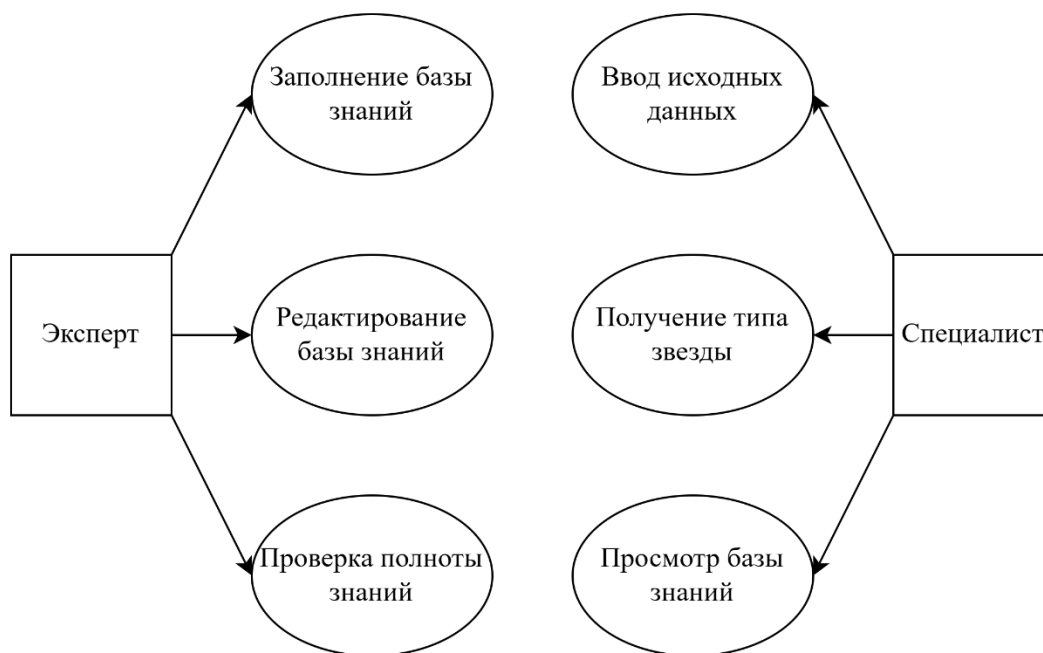


Рисунок 3 – Use-case диаграмма

На данной диаграмме видно, что у системы есть два типа пользователей:

Эксперт. Эксперт может использовать систему для создания модели, то есть добавлять новые типы звёзд, а также редактировать модель, включая изменение свойств и типов звёзд, а также добавление новых свойств.

Специалист. Специалист использует систему для ввода исходных данных, редактирования их, а также получения результатов решения задачи с объяснением.

3.4 Проект интерфейса редактора знаний

3.4.1 Определение отношения частичного порядка на множестве терминов для описания знаний

- Термины «типы звёзд» и «свойства» не зависят от других терминов.
- Термину «возможные значения» предшествует термин «свойства».
- Термину «описание свойств типа» предшествуют термины «типы звёзд» и «свойства».
- Термину «значение для типа» предшествуют термины «типы звёзд», «описание свойств типа», «возможные значения».

Представим эти отношения частичного порядка на множестве терминов для описания знаний в виде графа [Рисунок 4]. В вершинах графа находятся термины, между двумя вершинами существует ребро от первой ко второй, если вторая вершина определяется через первую.

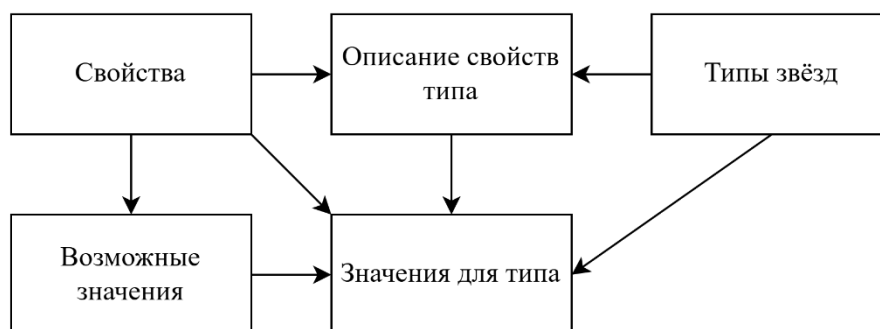


Рисунок 4 – Граф зависимостей терминов для описания знаний

3.4.2 Сценарий диалога с экспертом при формировании базы знаний

Ниже представлен возможный сценарий диалога с экспертом в редакторе базы знаний [Рисунок 5].

Редактор базы знаний	
Типы звёзд	
Свойства	
Возможные значения	
Описание свойств типа	
Значения для типа	
Проверка полноты знаний	

Рисунок 5 – Интерфейс редактора базы знаний

Возможные действия эксперта:

1. Ввод значений термина «типы звёзд»;
2. Ввод значений термина «свойства»;
3. Выбор свойств для термина «описание свойств типа»
4. Ввод множества возможных значений для каждого свойства;
5. Выбор значений из множества возможных значений для каждого типа звёзд, каждого свойства;
6. Проверка полноты знаний.

Окно «Типы звёзд» редактора базы знаний

Для добавления нового типа звёзд необходимо ввести его название в текстовое поле, затем нажать кнопку «Добавить». Тип звёзд не добавится, если поле будет пустым. Для удаления типа звёзд необходимо нажать на кнопку «–» рядом с его названием [Рисунок 6].

Редактор базы знаний	
Типы звёзд	<div> <input type="checkbox"/> O </div> <div> <input type="checkbox"/> B </div> <div> <input type="checkbox"/> A </div> <div> <input type="checkbox"/> F </div>
Свойства	
Возможные значения	
Описание свойств типа	
Значения для типа	
<div> Введите название типа звезды <div>Добавить</div> </div>	
Проверка полноты знаний	

Рисунок 6 – Окно «Типы звёзд»

Окно «Свойства» редактора базы знаний

Для добавления нового свойства звезды необходимо ввести его название в текстовое поле, затем нажать кнопку «Добавить». Свойство не добавится, если поле будет пустым. Для удаления свойства необходимо нажать на кнопку «–» рядом с его названием [Рисунок 7].

Редактор базы знаний	
Типы звёзд	<div> <input type="checkbox"/> Температура </div> <div> <input type="checkbox"/> Цвет </div> <div> <input type="checkbox"/> Показатель цвета </div> <div> <input type="checkbox"/> Абсолютная звёздная величина </div>
Свойства	
Возможные значения	
Описание свойств типа	
Значения для типа	
<div> Введите название свойства <div>Добавить</div> </div>	
Проверка полноты знаний	

Рисунок 7 – Окно «Свойства»

Окно «Возможные значения» редактора базы знаний

Для добавления нового возможного значения для свойства необходимо выбрать свойства, затем выбрать тип значений («числовой» или «перечислимый»), после чего для числовых значений ввести диапазон, указав, являются ли числа в диапазоне целыми [Рисунок 8], а для перечислимых значений необходимо ввести название значения в текстовое поле [Рисунок 9]. По нажатии кнопки «Добавить» значение будет добавлено. Новое возможное значение не будет добавлено, если поле не будет заполнено. Для удаления возможного значения свойства необходимо нажать на кнопку «-» рядом с его названием.

Редактор базы знаний

Типы звёзд

Свойства

Возможные значения

Описание свойств типа

Значения для типа

Проверка полноты знаний

< A3B

Температура

Цвет >

Значения

☐ Перечислимые ☒ Числовые

[2500; 100000]

☒ Только целые числа

От До

Рисунок 8 – Окно «Возможные значения» для ввода числовых значений

Редактор базы знаний	
Типы звёзд	<div> <div>< Температура</div> <div>Цвет</div> <div>Показатель цвета ></div> </div>
Свойства	
Возможные значения	Значения <div> <input checked="" type="radio"/> Перечислимые <input type="radio"/> Числовые </div> <div> <input type="checkbox"/> Голубой <input type="checkbox"/> Белый <input type="checkbox"/> Жёлтый <input type="checkbox"/> Красный </div> <div> <input type="text" value="Введите значение"/> <input type="button" value="Добавить"/> </div>
Описание свойств типа	
Значения для типа	
Проверка полноты знаний	

Рисунок 9 – Окно «Возможные значения» для ввода перечислимых значений

Окно «Описание свойств вида» редактора базы знаний

Для добавления нового свойства в описание свойств типа, выбранного в списке слева, необходимо выбрать в списке справа необходимые свойства с помощью флажков [Рисунок 10]. Чтобы выбрать добавить сразу все свойства в описание свойств типа, нужно нажать «Выбрать все». Чтобы убрать свойства из описания свойств типа, нужно повторно нажать на соответствующее свойство.

Редактор базы знаний	
Типы звёзд	<div> <div>Типы звёзд</div> <div>Свойства</div> </div>
Свойства	
Возможные значения	<div> <div> <input type="checkbox"/> Выбрать все <input checked="" type="checkbox"/> Температура <input checked="" type="checkbox"/> Цвет <input checked="" type="checkbox"/> Показатель цвета <input checked="" type="checkbox"/> A3B </div> </div>
Описание свойств типа	
Значения для типа	
Проверка полноты знаний	

Рисунок 10 – Окно «Описание свойств типа»

Окно «Значение для вида» редактора базы знаний

Для добавления значения свойства для типа звезды необходимо выбрать тип звезды, затем выбрать свойство, после чего ввести требуемый диапазон для числового свойства [Рисунок 11] или выбрать значения в случае перечислимого типа [Рисунок 12]. Изменение значения будет сохранено автоматически.

The screenshot shows a window titled 'Редактор базы знаний' (Knowledge Base Editor). On the left is a sidebar with a list of items: 'Типы звёзд' (Star types), 'Свойства' (Properties), 'Возможные значения' (Possible values), 'Описание свойств типа' (Description of type properties), 'Значения для типа' (Values for type), and 'Проверка полноты знаний' (Check for knowledge completeness). The 'Значения для типа' item is selected and highlighted in blue. The main area of the window is divided into two vertical panes. The left pane, titled 'Свойства' (Properties), contains a list of properties: 'Температура' (Temperature), 'Цвет' (Color), 'Показатель цвета' (Color index), and 'АЗВ' (AZV). The 'Температура' property is selected and highlighted in blue. The right pane, titled 'Значения' (Values), shows a range input with 'От' (From) set to '2500,0' and 'До' (To) set to '100000,0'. Navigation arrows are visible on the left and right sides of the main area, and a '< O' button is at the top left and 'A >' at the top right.

Рисунок 11 – Окно «Значение для типа» для числовых значений

The screenshot shows the same 'Редактор базы знаний' window. In the sidebar, 'Значения для типа' is still selected. In the main area, the left pane titled 'Типы звёзд' (Star types) contains a list of types: 'Температура' (Temperature), 'Цвет' (Color), 'Показатель цвета' (Color index), and 'АЗВ' (AZV). The 'Цвет' type is selected and highlighted in blue. The right pane, titled 'Свойства' (Properties), shows a list of color options with checkboxes: 'Выбрать всё' (Select all), 'Голубой' (Blue), 'Белый' (White), 'Жёлтый' (Yellow), and 'Красный' (Red). The 'Белый' (White) option is selected, indicated by a blue square next to its checkbox. The same navigation arrows and top buttons are present.

Рисунок 12 – Окно «Значение для вида» для перечислимых значений

Окно «Проверка полноты знаний»

После завершения редактирования базы знаний необходимо провести проверку полноты заданной информации. В случае незаполненных полей будет выведена ошибка и перечислены ссылки на соответствующие поля [Рисунок 13].



Рисунок 13 – Окно «Проверка полноты знаний»

3.5 Проект интерфейса системы ввода исходных данных

3.5.1 Определение отношения частичного порядка на множестве терминов для описания действительности

На терминах для описания действительности вводится структура отношения частичного порядка. Эту структуру можно описать графом [Рисунок 14]. В вершинах графа находятся термины, между двумя вершинами существует ребро от первой ко второй, если вторая вершина определяется через первую.



Рисунок 14 – Граф зависимостей терминов для описания действительности

3.5.2 Сценарий диалога при вводе исходных данных

Для каждого свойства с известным значением необходимо ввести значение, затем нажать кнопку «Определить тип звезды». Для числового и перечислимого типа данных ввод отличается [Рисунок 15, Рисунок 16]. После нажатия кнопки специалисту будет выведен на экран тип звезды, удовлетворяющий введённым данным.

Рисунок 15 – Окно «Ввод исходных данных» для числовых значений

Ввод исходных данных

< O

Свойства

Температура

Цвет

Показатель цвета

АЗВ

В

Значение

☐ Голубой

☒ Белый

☐ Жёлтый

☐ Красный

A >

Итого

Температура

11337,6

Цвет

Белый

Показатель цвета

-0,2

АЗВ

-3,1

Посмотреть базу знаний

Определить тип звезды

Рисунок 16 – Окно «Ввод исходных данных» для перечислимых значений

Если пользователь хочет посмотреть знания системы, то он может нажать кнопку «Посмотреть базу знаний», которая откроет новое окно [Рисунок 17]. Пользователь может переключать типы звёзд и просматривать значения различных свойств типа звёзд. В окне отображаются все свойства, заданные в знаниях. Если свойство не определено для выбранного типа звезды, то будет выведено соответствующее сообщение. Нажатие на кнопку «Вернуться к вводу исходных данных» откроет окно «Ввод исходных данных» со всеми введенными ранее значениями.

Просмотр базы знаний

< O

Свойства

Температура

Цвет

Показатель цвета

АЗВ

В

Значения

От 10000,0

До 30000,0

A >

Вернуться к вводу исходных данных

Рисунок 17 – Окно «Просмотр базы знаний»

3.6 Проект системы вывода результата и объяснений

После нажатия на кнопку «Определить тип звезды» в окне «Ввод исходных данных» [Рисунок 15, Рисунок 16], пользователь получает одно из следующих сообщений:

Подходящие типы звёзд: «название типа звезды». Другие типы звёзд опровергнуты по следующим причинам: тип звезды «название типа звезды» опровергнут, так как значение «значение» характеристики «название характеристики» не соответствует описанию типа звезды.

Тип звезды не определён. Звезда не относится к главной последовательности, или знания о данном типе звёзд не занесены в систему. Обратитесь к эксперту для разрешения проблемы. Все типы звёзд опровергнуты по следующим причинам: тип звезды «название типа звезды» опровергнут, так как значение «значение» характеристики «название характеристики» не соответствует описанию типа звезды.

4 Реализация системы

4.1 Общая структура реализации

Система, основанная на знаниях для классификации звёзд, реализована на языке программирования Python 3.10+ с использованием фреймворка PyQt6 для создания графического пользовательского интерфейса. Архитектура построена по модульному принципу: каждый компонент системы выделен в отдельный модуль, что повышает удобство разработки, тестирования и поддержки.

Основные компоненты проекта:

1. `main.py` – точка входа в приложение. Содержит запуск GUI и основного окна (MainWindow).

2. `main_window.py` – основное окно программы с вкладками. Реализует переключение между разделами: редактирование знаний, ввод данных и просмотр базы знаний.

3. `knowledge_base.py` – класс KnowledgeBase отвечает за загрузку и сохранение базы знаний из JSON-файла, а также за операции добавления/удаления типов звёзд, свойств, описаний и значений.

4. `knowledge_base.json` – файл базы знаний, содержащий все данные, необходимые для классификации: типы звёзд, свойства, допустимые значения, описание свойств для каждого типа и т.д.

Интерфейсные модули:

1. `knowledge_editor.py` – редактор базы знаний (для эксперта);
2. `data_input.py` – форма ввода характеристик звезды (для специалиста);
3. `knowledge_viewer.py` – просмотр существующих знаний.

Классификаторы:

4. `classifier.py` – логический классификатор, использующий точные диапазоны и значения;

5. `neural_classifier.py` – нейросетевой классификатор, использующий модель TensorFlow/Keras;

6. `ai_learn.py` – отдельный скрипт для генерации обучающей выборки и обучения модели.

Дополнительные вспомогательные модули:

1. `validators.py` – функции для проверки и приведения введённых значений;
2. `theme.py` – стили оформления интерфейса (цвета, оформление кнопок);
3. `result_output.py` – компонент для вывода результатов и объяснений пользователю.

4.2 Реализация логического классификатора

Логический классификатор реализован в модуле `classifier.py` и используется в системе для определения типа звезды на основе введённых пользователем характеристик и знаний, представленных в базе `knowledge_base.json`.

Принцип работы: метод исключения.

Классификатор проверяет все типы звёзд из базы знаний и исключает те, для которых хотя бы одно свойство не соответствует заданным условиям. Как только найдено первое несоответствие, тип звезды исключается из рассмотрения. При этом все ранее совпавшие свойства сохраняются и отображаются пользователю в качестве пояснений.

Таким образом, система определяет только те типы, для которых все свойства полностью соответствуют заданным значениям. Такой подход делает классификацию строгой и однозначной.

4.2.1 Обработка свойств

Числовые свойства (например, температура) проверяются на принадлежность к допустимому диапазону.

Перечислимые свойства (например, цвет) проверяются на вхождение в список допустимых значений.

При совпадении значение сохраняется как положительный результат.

При несоответствии фиксируется причина отказа, и дальнейшая проверка для текущего типа прекращается.

4.2.2 Формат результата

Метод классификации возвращает:

1. список подходящих типов звёзд;
2. объяснения причин исключения неподходящих типов;
3. совпавшие свойства для каждого типа, даже если он был исключён.

Пример

Если звезда имеет температуру 5800 К и цвет "Белый", то тип G будет исключён, так как:

1. температура совпадает с допустимым диапазоном [5000, 6000];
2. но цвет "Белый" не входит в список допустимых значений (ожидался "Жёлтый").

Пользователь увидит:

Температура = 5800, в пределах [5000, 6000]

Цвет = Белый, ожидалось одно из ['Жёлтый']

4.3 Реализация интерфейсов

Пользовательский интерфейс системы реализован с использованием библиотеки PyQt6 и представляет собой оконное приложение с вкладками. Основной компонент интерфейса — это главное окно (MainWindow), в котором размещён элемент QTabWidget с тремя вкладками:

1. Редактор знаний (для эксперта);
2. Ввод данных (для специалиста);
3. Просмотр базы знаний (для всех пользователей).

Каждая вкладка реализована в виде отдельного виджета с полной изоляцией логики и представления.

4.3.1 Интерфейс 1: Редактор знаний

Доступен для эксперта. Реализован в модуле knowledge_editor.py. Содержит следующие разделы (реализованные через меню):

1. Типы звёзд — позволяет добавлять и удалять спектральные классы (например, O, B, A и т.д.).

2. Свойства — позволяет создавать новые свойства звёзд (например, Температура, Цвет).

3. Возможные значения — задаёт допустимые диапазоны (числовые) или списки значений (перечислимые).

4. Описание свойств типа — определяет, какие свойства описывают каждый тип звезды.

5. Значения для типа — определяет допустимые диапазоны или значения для каждого свойства конкретного типа.

6. Проверка полноты знаний — осуществляет автоматическую валидацию базы знаний, выявляя незаполненные поля, конфликты и нарушения ограничений целостности.

Интерфейс снабжён кнопками добавления, удаления и сохранения. Валидация ввода встроена непосредственно в поля ввода.

4.3.2 Интерфейс 2: Ввод данных

Реализован в модуле `data_input.py`. Предназначен для ввода характеристик звезды специалистом и запуска классификации. Содержит:

1. Форму для ввода свойств: числовые значения (например, температура) и выпадающие списки (например, цвет);

2. Флажок «Использовать нейросеть», позволяющий выбрать между логическим и нейросетевым классификатором;

3. Кнопку «Определить тип звезды», запускающую классификацию;

Результат классификации отображается в отдельном компоненте `ResultOutputWidget` в виде форматированного текста: подходящие типы, объяснения, совпадения и исключения.

4.3.3 Интерфейс 3: Просмотр базы знаний

Реализован в `knowledge_viewer.py`. Позволяет выбрать тип звезды из списка и просмотреть значения всех его свойств в режиме только для чтения. Отображаются:

1. Названия свойств;
2. Значения в виде диапазонов (для числовых) или списков (для перечислимых);
3. Если значение отсутствует, отображается предупреждение.

Данный интерфейс удобен для специалистов, которым необходимо понимать текущие настройки системы без возможности их изменять.

4.4 Нейросетевая классификация

Для повышения гибкости и устойчивости к неточным или неполным данным в системе реализован альтернативный механизм классификации звёзд с использованием искусственной нейронной сети. Он дополняет логический классификатор и позволяет выполнять предсказания даже при отклонениях от жёстких правил.

4.4.1 Архитектура нейросети

Нейросеть построена в модуле `ai_learn.py` с использованием библиотеки TensorFlow/Keras. Её структура:

1. Входной слой: количество нейронов соответствует числу всех нормализованных признаков (4 числовых + кодированные категориальные).
2. Два скрытых слоя по 64 нейрона с:
 - 2.1. Dense (полносвязные слои),
 - 2.2. BatchNormalization — стабилизация обучения,
 - 2.3. ReLU — функция активации,
 - 2.4. Dropout(0.3) — регуляризация.
3. Выходной слой: Dense с softmax — выдаёт вероятностное распределение по классам звёзд.

4.4.2 Подготовка обучающих данных

Обучение модели осуществляется на основе базы знаний, без привлечения сторонних астрономических датасетов. Это возможно благодаря тому, что в базе чётко указаны диапазоны и допустимые значения для каждого свойства каждого типа звезды.

Алгоритм генерации обучающей выборки:

1. Загружаются:

- 1.1. все свойства (числовые и перечислимые),
- 1.2. диапазоны допустимых значений из `knowledge_base.json`.

2. Для каждого типа звезды:

- 2.1. Создаётся 2000 случайных примеров (параметр `samples_per_class`).
- 2.2. Числовые свойства случайным образом выбираются в пределах заданных диапазонов.

2.3. Категориальные свойства — случайным образом из допустимых значений.

2.4. Каждое значение нормализуется в диапазон `[0, 1]` (для чисел) или кодируется индексом (для перечислимых значений).

3. Класс звезды сохраняется как метка.

Пример нормализации:

```
def normalize(val, min_val, max_val):  
    return (val - min_val) / (max_val - min_val)
```

4.4.3 Обучение и сохранение модели

Обучение происходит с использованием:

1. One-hot encoding меток (`CategoricalCrossentropy`);
2. Adam — адаптивный оптимизатор;
3. Эпохи: 100;
4. Размер батча: 32.

После завершения обучения:

1. Модель сохраняется в файл `ai/star_model.keras`;

2. Кодировщик меток LabelEncoder сериализуется в ai/label_encoder.pkl.

Эти файлы затем используются в рабочей системе при запуске нейросетевого классификатора (neural_classifier.py).

4.4.4 Использование в системе

При выборе опции "Использовать нейросеть" в интерфейсе DataInputWidget:

1. Значения нормализуются и кодируются.
2. Подготовленный вектор передаётся в модель.
3. Модель возвращает:
 - 3.1. Метку (тип звезды),
 - 3.2. Уверенность в предсказании (вероятность).

Результат выводится в виде:

Тип звезды (нейросеть): G (уверенность: 94.56%)

4.5 Особенности реализации

Реализация системы учитывает ряд ключевых инженерных решений, обеспечивающих её надёжность, расширяемость и удобство использования.

Все компоненты системы разделены на логические модули:

1. Интерфейсы (ввод данных, редактор знаний, просмотр),
2. Логика классификации (логическая и нейросетевая),
3. Работа с базой знаний,
4. Вспомогательные утилиты (валидация, стили, преобразование данных).

Такой подход упрощает сопровождение, тестирование и последующее расширение функциональности.

Программа ориентирована на две основные роли:

1. Эксперт — создаёт и редактирует базу знаний (типы звёзд, свойства, диапазоны);
2. Специалист — вводит характеристики звезды и получает классификацию.

Интерфейс каждого пользователя адаптирован под его задачи.

Система реализует:

1. Логический классификатор, исключающий неподходящие типы на основе жёстких правил;

2. Нейросетевой классификатор, способный учитывать погрешности и делать вероятностные предсказания.

Это делает систему более гибкой в использовании и надёжной при неидеальных данных.

Редактор знаний включает встроенную проверку корректности:

1. Все свойства и значения должны быть заданы;

2. Диапазоны не должны выходить за допустимые границы;

3. Исключается конфликтная или неполная информация.

Нейросеть обучается на искусственно сгенерированных данных, соответствующих описаниям из базы знаний. Это исключает необходимость поиска и подготовки внешних датасетов, делает обучение воспроизводимым и управляемым.

Интерфейс снабжён:

1. Валидацией ввода,

2. Подсказками и пояснениями,

3. Функцией генерации случайных данных для тестирования,

4. Возможностью быстро просматривать и проверять знания системы.

Заключение

В рамках курсового проекта был разработан проект системы, основанной на знаниях, для классификации звёзд. Для достижения поставленной цели были выполнены следующие задачи:

Проведён анализ предметной области «классификация звёзд»;

Разработана модель предметной области, включая классификацию различных классов звёзд;

Разработан проект системы для классификации звёзд, включающий редактор знаний, решатель задач, подсистему вывода и объяснения результатов, а также систему ввода исходных данных.

Таким образом, все цели данного курсового проекта были достигнуты, и созданный проект может быть использован в качестве основы для реализации системы, предназначенной для классификации звёзд.

Список литературы

1. Кононович Э. В., Мороз В. И. Общий курс астрономии. — 2-е, исправленное. — М.: УРСС, 2004. — 544 с.;
2. Артемьева И.Л. Методы системного анализа и моделирования. Методические указания по выполнению самостоятельной работы и индивидуальных заданий. — 2019.