

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ «НАЦИОНАЛЬНЫЙ
ИССЛЕДОВАТЕЛЬСКИЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ «МИСИС»

Институт компьютерных наук Кафедра
Инженерной Кибернетики

Отчет по практической работе

Выполнила:

студентка 2-го курса,

группы БПМ-22-1 Горохова П.А.

Проверил:

профессор, д.э.н. Мишуров С.С.

Москва, 2024

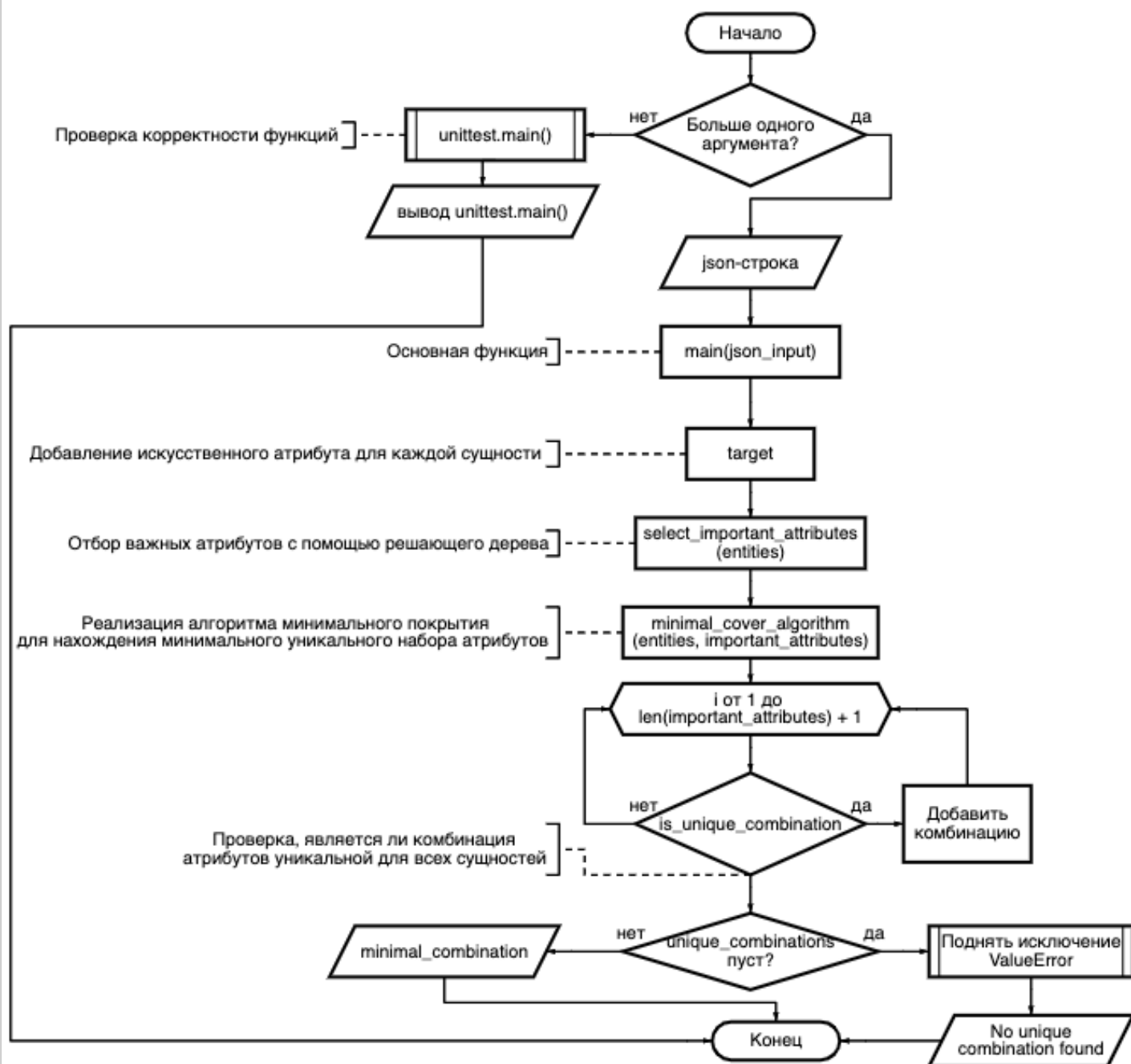


Схема программы по ГОСТу 19.701-90

Описание работы алгоритма

Алгоритм решает задачу выбора минимального уникального набора атрибутов на основе данных, представленных в формате JSON. Программа реализует несколько ключевых функций, которые выполняют отбор важных признаков, проверку уникальности комбинаций атрибутов и нахождение минимального покрытия.

Основные функции

1. `is_unique_combination(entities, attributes)`

Эта функция проверяет, является ли комбинация атрибутов уникальной для всех сущностей.

- Параметры:
 - `entities`: список сущностей в формате словаря.
 - `attributes`: список атрибутов, для которых проверяется уникальность комбинации.
- Возвращаемое значение:
 - `True`, если комбинация уникальна для всех сущностей, иначе `False`.

2. `minimal_cover_algorithm(entities, important_attributes)`

Эта функция реализует алгоритм минимального покрытия для нахождения минимального уникального набора атрибутов.

- Параметры:
 - `entities`: список сущностей в формате словаря.
 - `important_attributes`: список важных атрибутов, отобранных с использованием метода отбора признаков.
- Возвращаемое значение:
 - Минимальная уникальная комбинация атрибутов, которая является оптимальным покрытием.

3. `select_important_attributes(entities)`

Эта функция отбирает важные атрибуты с помощью решающего дерева и метода `pandas.get_dummies`.

- Параметры:
 - `entities`: список сущностей в формате словаря.
- Возвращаемое значение:

- Список важных атрибутов, определенных с использованием метода отбора признаков.

4. main(json_data)

Основная функция программы, которая принимает JSON-строку с данными, выполняет отбор важных атрибутов, находит минимальное покрытие и возвращает CSV-строку с минимальной уникальной комбинацией атрибутов.

- Параметры:
 - json_data: JSON-строка с исходными данными в формате строкового представления.
- Возвращаемое значение:
 - CSV-строка с минимальной уникальной комбинацией атрибутов.

Unit-тестирование

Алгоритм также включает в себя набор юнит-тестов для проверки корректности работы основных функций.

Ссылка на репозиторий: https://github.com/GorokhovaPolina/MISIS_praktika2024