

## Оглавление

Лабораторная работа 1. Решение задач с использованием генетического алгоритма.....	4
Лабораторная работа 2. Использование линейных моделей для решения задач регрессии и классификации с использованием библиотек машинного обучения. 5	
Лабораторная работа 3. Деревья решений и случайный лес. Сравнение методов для решения задачи классификации.....	6
Лабораторная работа 4. Реализация простейших алгоритмов кластеризации в задачах технического зрения.....	7
Лабораторная работа 5. Построение нейронных сетей (с использованием библиотек машинного обучения) для задачи распознавания образов.....	8
Лабораторная работа 6. Построение нейронных сетей для задач классификации .....	9
Лабораторная работа 7 Реализация байесовского подхода для задач фильтрации спама .....	10
Лабораторная работа 8 Классификация текстов .....	11
Список литературы:.....	12

# Лабораторная работа 1.

## Решение задач с использованием генетического алгоритма.

**Цель работы:** Ознакомить студентов с базовыми понятиями и параметрами генетического алгоритма.

**Задание:** Найти решение транспортной задачи. Имеем  $K$  поставщиков, у каждого 1 единица товара,  $p_i$  и  $s_i$  - цена товара и стоимость доставки у  $i$ -го поставщика. (сгенерировать массивы случайно  $p_i$  от 100 до 1000  $s_i$  от 10 до 100). Сформировать поставку не менее  $M$  единиц товара.

### Варианты :

Размер начальной популяции  $MAX\_ENT = 10$  у всех.

### Критерий останова:

1. Выполнение заданного числа итераций
2. Выполнение заданного числа итераций без изменения целевой функции
3. Достижение заданного значения целевой функции.

### Оператор скрещивания:

1. Одноточечное скрещивание
2. Двухточечное скрещивание
3. Однородное скрещивание

### Мутация:

1. Изменение случайно выбранных битов
2. Перестановка случайно выбранных битов местами.

### Виды селекции:

1. Отбор усечением
2. Схема пропорционального отбора
3. Турнирная схема

Вариант	Критерий	Скращивание	Мутация	Селекция
1	1	3	1	1
2	2	1	2	2
3	3	2	1	3
4	1	3	2	1
5	2	1	1	2
6	3	2	2	3
7	1	3	1	1
8	2	1	2	2
9	3	2	1	3
10	1	3	2	1
11	2	1	1	2
12	3	2	2	3
13	1	3	1	1
14	2	1	2	2
15	3	2	1	3
16	1	3	2	1
17	2	1	1	2
18	3	2	2	3
19	1	3	1	1
20	2	1	2	2

**Формирование начальной популяции решений.** Генерация начальной популяции может происходить как случайным образом, так и с помощью некоторого алгоритма.

**Целевая функция** позволяет оценить степень приспособленности данной особи в популяции и характеризует качество получаемого решения. В данной задаче целевая функция – это значение многочлена в точке, соответствующей данной особи. Во время

генетического процесса вычисление целевой функции осуществляется над элементами всей популяции решений. Нужно отметить, что достаточно часто сложность генетических алгоритмов оценивается по количеству вычислений целевой функции.

**Критерий останова.** В качестве критерия останова генетического алгоритма могут выступать следующие условия:

- Выполнение алгоритмом априорно заданного числа итераций.
- Выполнение алгоритмом априорно заданного числа итераций без улучшения целевой функции.
- Достижение некоторого априорно заданного значения целевой функции.

**Оператор селекции.** Выбор решений для следующей популяции (оператор селекции) предназначен для улучшения качества решений в новой популяции, а именно сохранение разнообразия популяции, сохранение лучших решений и удаление из нее недопустимых решений. Обычно выбираются элементы с наибольшей приспособленностью.

Возможны различные варианты операции селекции, основанные на разных схемах отбора:

- Схема пропорционального отбора.

В данной схеме отбора вычисляется значение целевой функции для каждого решения  $F_i$  и определяется среднее значение целевой функции в популяции  $F_{cp}$ . Затем для каждого решения вычисляется отношение  $F_i / F_{cp}$ . Например, если отношение равно 2.36, то данное решение имеет двойной шанс на скрещивание, и будет иметь вероятность равную 0.36 третьего скрещивания. Если же приспособленность равна 0.54, то решение примет участие в единственном скрещивании с вероятностью 0.54.

- Схема отбора на основе рулетки.

Каждому решению выделяется сектор рулетки  $2\pi \cdot F_i / \sum_{j=1}^N F_j$

Решение попадает в новую популяцию, если случайным образом сгенерированное число попадает в этот сектор.

- Турнирный отбор.

Схему турнирного отбора можно описать следующим образом: из популяции, содержащей  $N$  решений, выбирается случайным образом 2 решения и между выбранными решениями проводится турнир. Победившее решение используется для скрещивания.

- Отбор усечением. Данная стратегия использует отсортированную по возрастанию популяцию. Число решений для скрещивания выбирается в соответствии с порогом  $T \in [0; 1]$ . Порог

определяет, какая доля особей, начиная с самой первой (самой приспособленной) будет принимать участие в отборе. Порог можно задавать числом, большим единицы, тогда он будет равен числу решений из текущей популяции, допущенных к отбору.

### **Оператор скрещивания:**

Оператор скрещивания используется для передачи родительских признаков потомкам. Пары для скрещивания выбираются на основе одной из схем селекции, описанных выше. Возможны следующие варианты оператора скрещивания (рис. 2):

- **Одноточечное скрещивание.**  
Выбирается одна точка, и относительно неё решения обмениваются своими частями.
- **Двухточечное скрещивание.**  
Аналогично предыдущему, но точек скрещивания выбирается две.
- **Универсальное скрещивание.**  
С некоторой вероятностью выбирается бит либо одного, либо другого родителя.
- **Однородное скрещивание.**

Каждый ген в потомстве создается посредством копирования соответствующего гена от одного или другого родителя, выбранного согласно случайно сгенерированной маске скрещивания. Если в маске скрещивания стоит 1, то ген копируется от первого родителя, если в маске стоит 0, то ген копируется от второго родителя. Процесс повторяется с новыми родителями для создания второго потомства. Новая маска скрещивания случайно генерируется для каждой пары родителей.

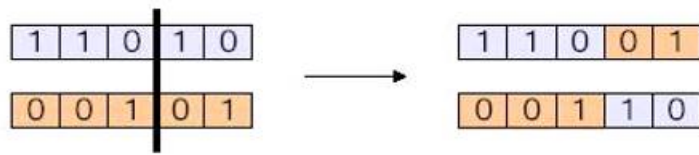
После операции скрещивания новые решения занимают места своих родителей в популяции.

### **Оператор мутации**

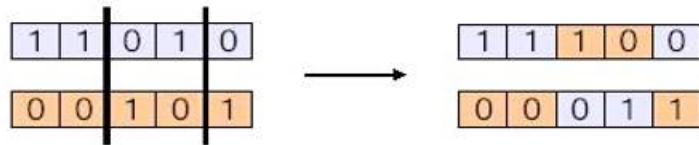
Оператор мутации используется для внесения в решение некоторых новых признаков. Некоторые варианты реализации операции мутации представлены на рисунке 3. Все варианты изменяют биты битовой строки с некоторой вероятностью.

- Изменение случайно выбранного бита.
- Перестановка случайно выбранных битов местами.
- Реверс битовой строки, начиная со случайно выбранного бита.

Одноточечное скрещивание



N-точечное скрещивание

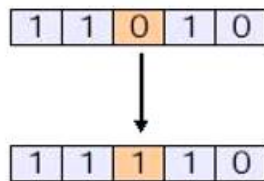


N=2

Универсальное скрещивание



Рис. 2 Варианты оператора скрещивания.



Каждый бит изменяется с вероятностью P



Рисунок 3. Варианты операции мутации