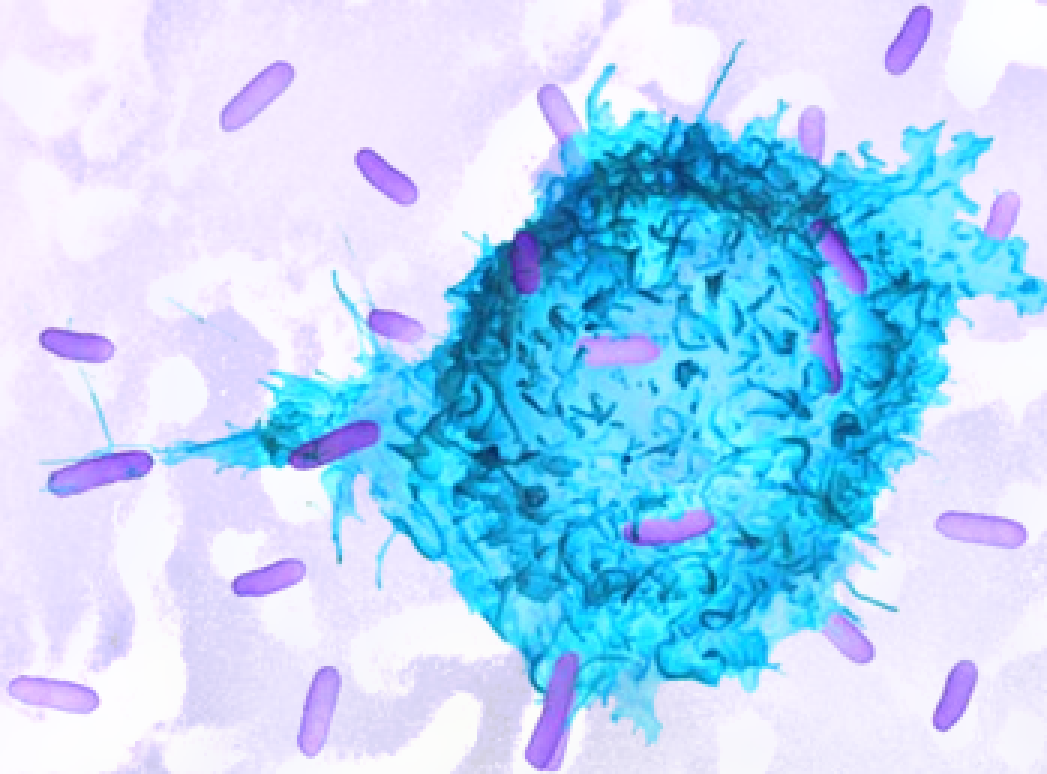


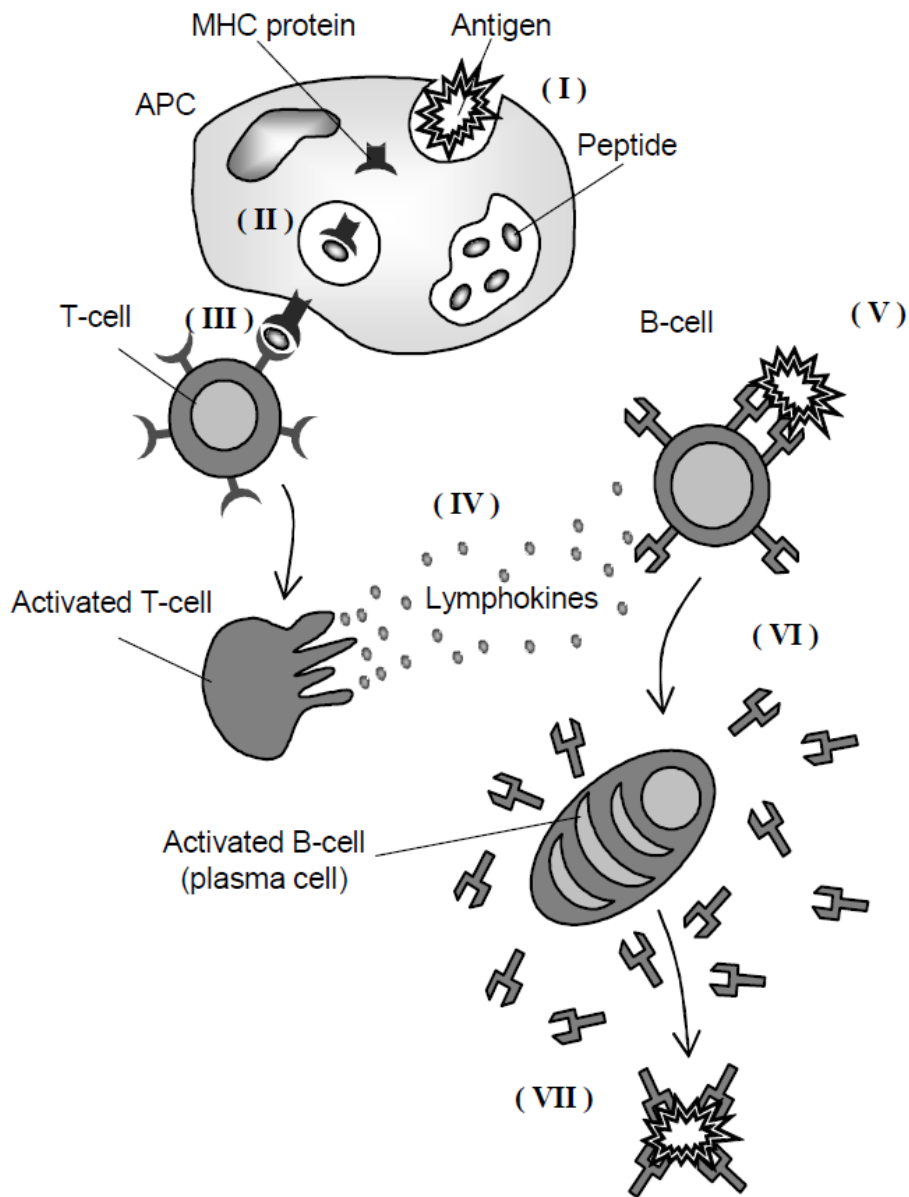
Иммуноподобный алгоритм классификации VALIS



Карпов Пётр Михайлович

PeterKarpov@inversed.ru

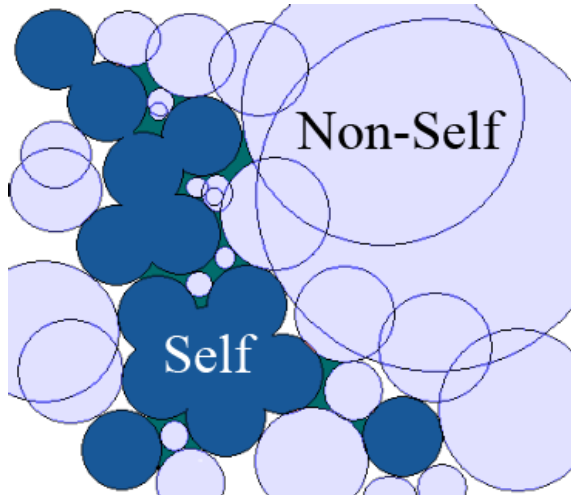
Схема иммунной системы



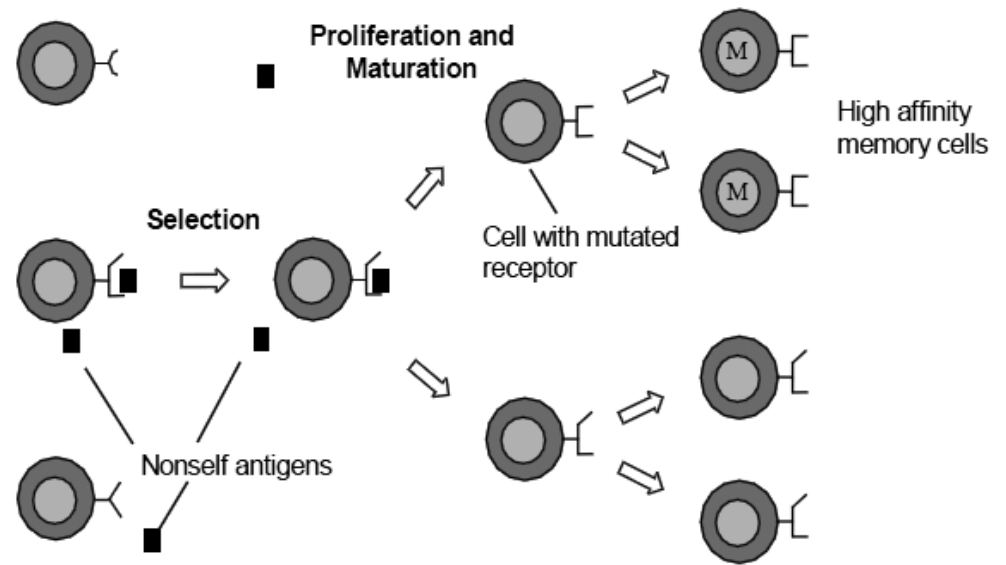
- Выделение антигена
- Активация Т-клеток
- Активация Б-клеток
- Совпадение с антигеном
- Производство антител
- Уничтожение патогена

Иммуноподобные алгоритмы

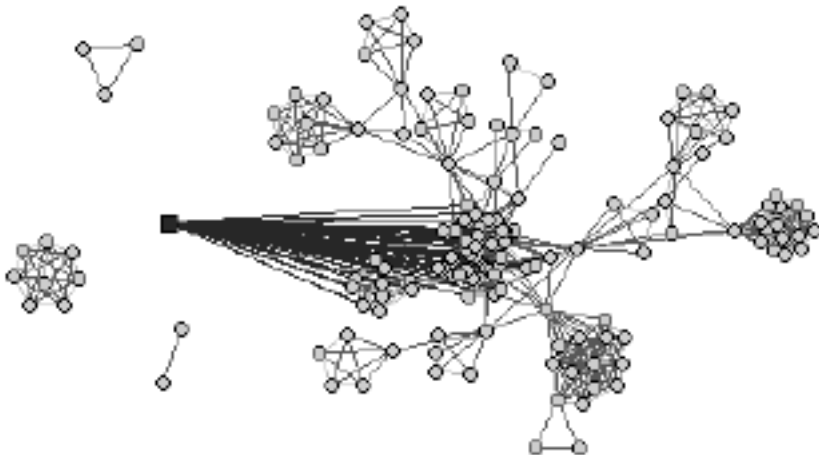
Отрицательный отбор



Клональный отбор



Иммунные сети

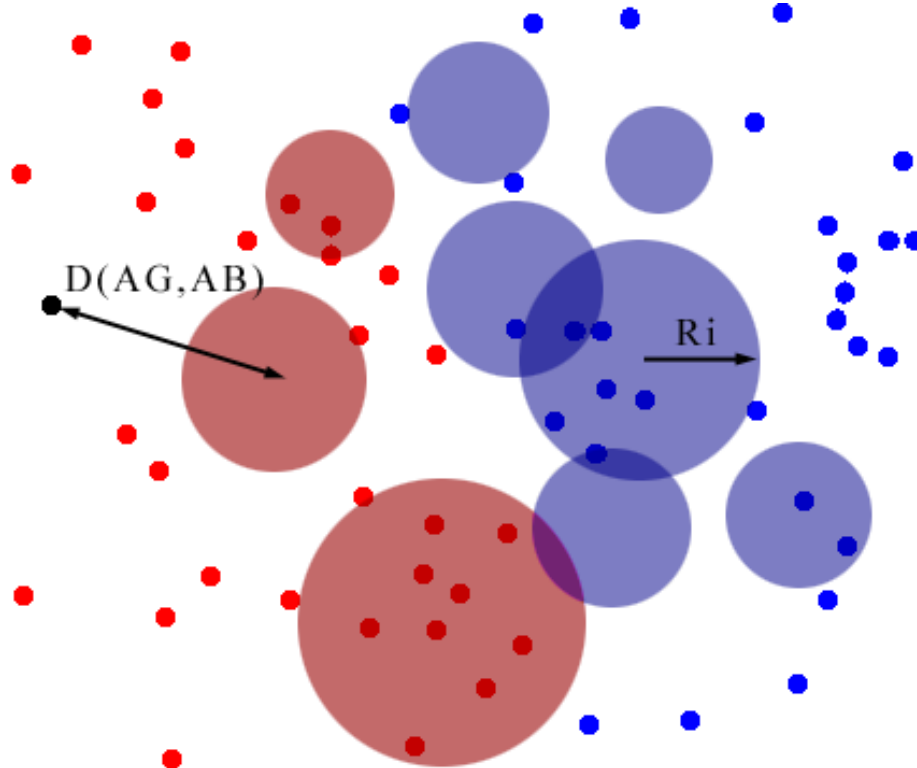


Дендритные клетки,
Теория опасности,

... ..

Аналогия между ИС и задачей классификации

| | | |
|----------------|-------|-------------------------|
| Антиген | ←---→ | Образец данных AG |
| Антитело | ←---→ | Структура данных AB |
| Иммунный ответ | ←---→ | Результат классификации |



Связывание:

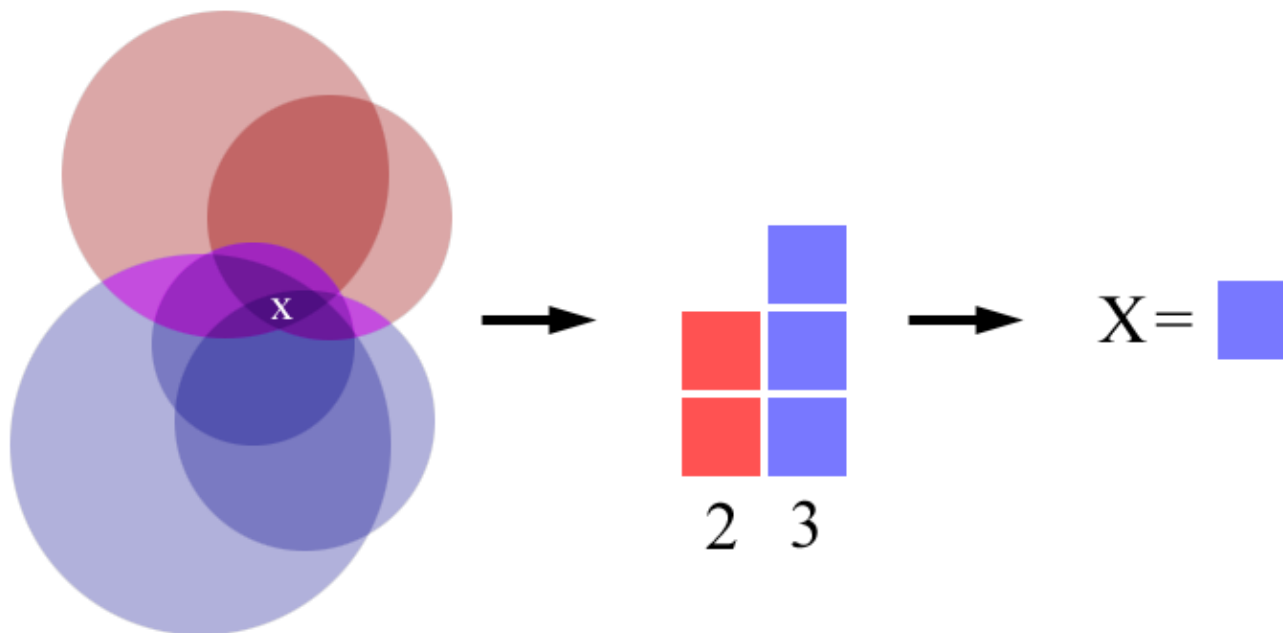
$$D(AB_i, AG) < R_i,$$

D — функция расстояния,

R_i — порог связывания

Классификация путём голосования

VALIS — Vote Allocating Immune System
(имунная система с распределением голосов)



Связанные антитела голосуют за свой класс. Результат классификации определяется максимумом голосов.

Алгоритм обучения

1. Инициализация

Создать популяцию из $PopSize$ случайных антител

2. Предъявление набора данных

Предъявить данные системе, для каждого антитела вычислить:

- › Количество связанных и верно классифицированных антигенов
- › Вероятность правильного ответа $pCorrect$
- › Коэффициент перекрытия $KShare$
- › Приспособленность $f = pCorrect / KShare$

3. Присвоение классов

Определить классы антител по большинству классов связанных с ним антигенов

4. Репродукция

Путём скрещивания и мутации создать $NChildren = LearnRate \times PopSize$ потомков

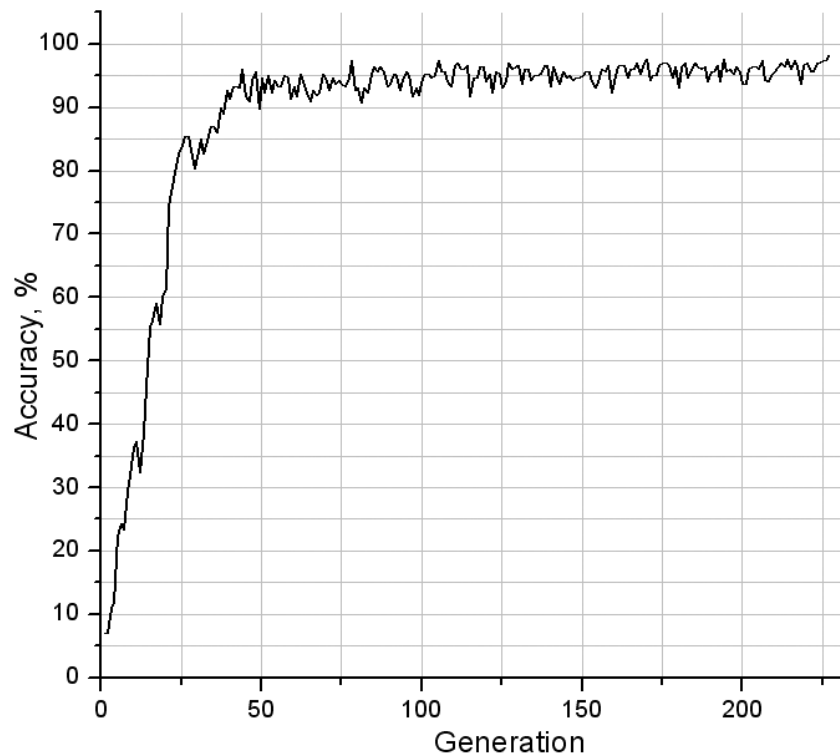
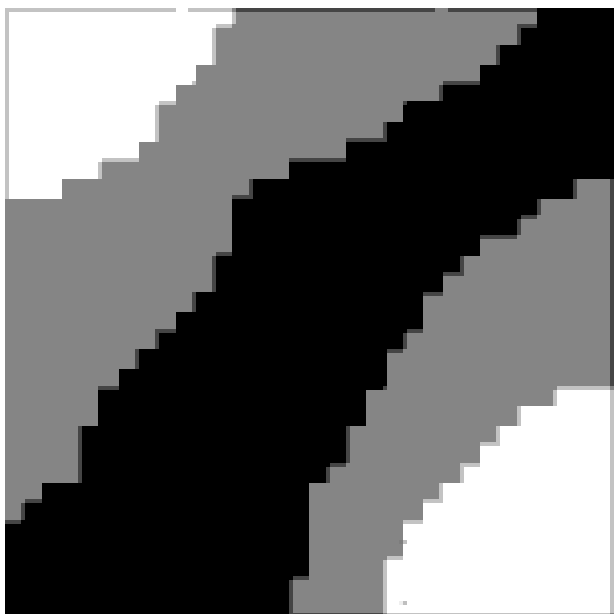
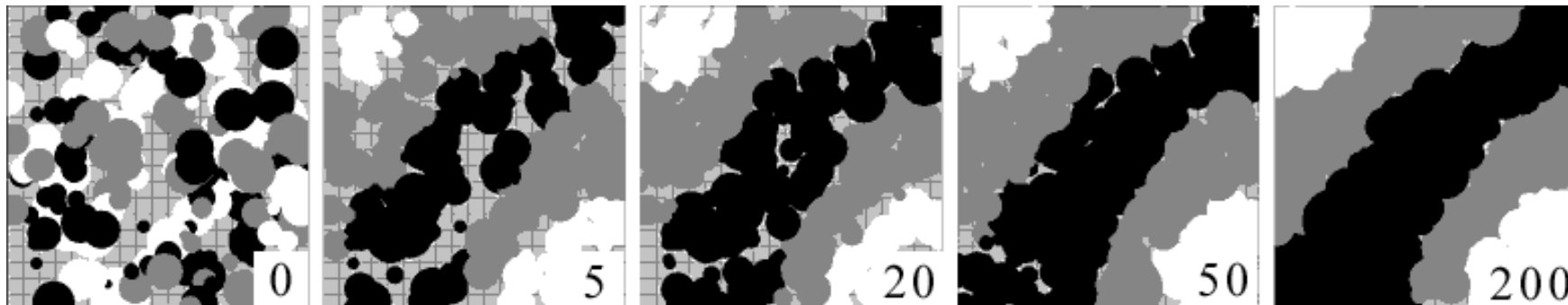
5. Замена

Заменить $NChildren$ худших по приспособленности антител потомками

6. Конец одного поколения

Пока не достигнут критерий остановки, перейти к шагу 2

Тестовая задача



Динамика обучения

Распознавание Символов

Антигены

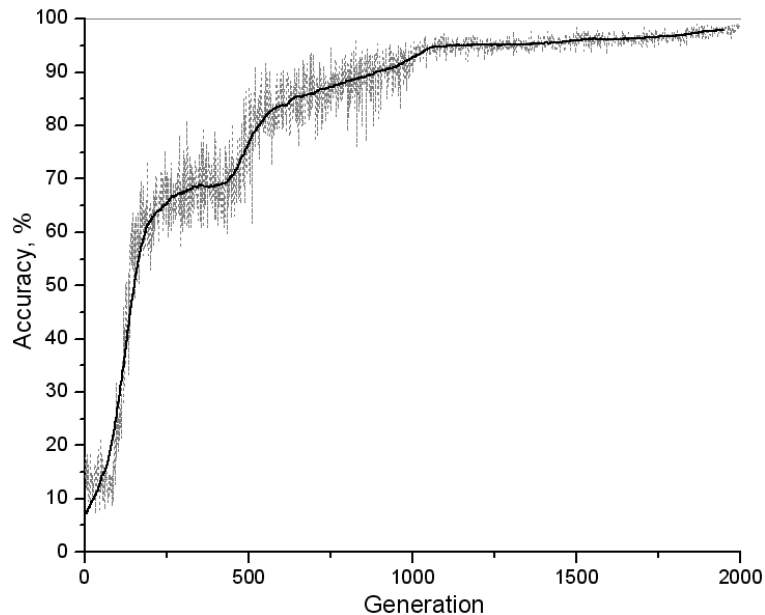
Цифры от 0 до 9 со случайным поворотом и масштабom

Антитела

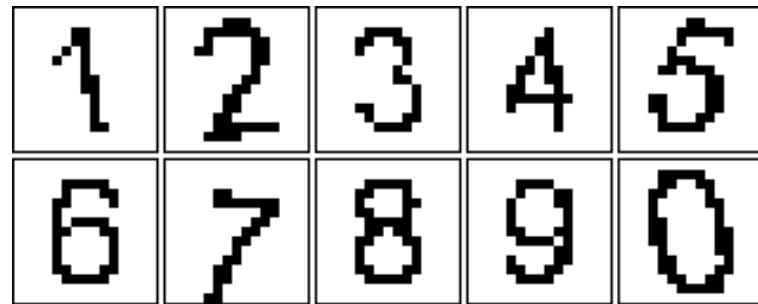
Двоичные массивы

Точность классификации

99%



Способность к обобщению:
инвариантность распознавания к повороту и масштабированию



Классификация текстов

Антигены

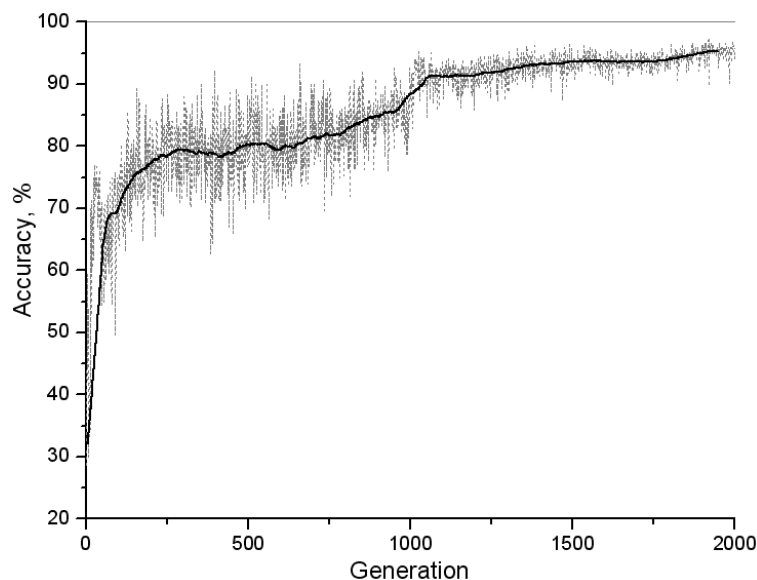
Части исходного кода на одном из
6 языков программирования

Антитела

Строки переменной длины

Точность классификации

96%



Характерные элементы антител:

Assembler @, mov, lf

Basic \$

C fp, {#13, }, <, \n

Fortran #10c

Lisp ;;,),))#13

Pascal en, in#13, :=

Заключение

Особенности системы:

- › Эффективное обучение
- › Высокая степень обобщения
- › Децентрализованная память
- › Коллективный характер классификации
- › Устойчивость к удалению антител

Дальнейшие исследования:

- › Тестирование на различных задачах
- › Переменная скорость обучения
- › “Мягкое” связывание
- › Пост-обработка популяции антител

Веб-страница проекта доступна по адресу:

inversed.ru/AIS.htm