

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ  
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
**«БЕЛГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ НАЦИОНАЛЬНЫЙ  
ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ»**  
(Н И У «Б е л Г У»)

ИНСТИТУТ ИНЖЕНЕРНЫХ И ЦИФРОВЫХ ТЕХНОЛОГИЙ  
КАФЕДРА МАТЕРИАЛОВЕДЕНИЯ И НАНОТЕХНОЛОГИЙ

**ИСПОЛЬЗОВАНИЕ ГЕНЕТИЧЕСКИХ АЛГОРИТМОВ И МАШИННОГО  
ОБУЧЕНИЯ ДЛЯ ПРЕДСКАЗАНИЯ ПЛАСТИЧНОСТИ  
ВЫСОКОЭНТРОПИЙНЫХ СПЛАВОВ**

Выпускная квалификационная работа  
обучающегося по направлению подготовки

28.03.03 Наноматериалы

очной формы обучения,

группы 12001714

Вережак Михаила Викторовича

# Цели и задачи работы

Целью работы является разработка модели машинного обучения с использованием генетического алгоритма для предсказания пластических свойств высокоэнтропийных сплавов.

Задачи данной работы:

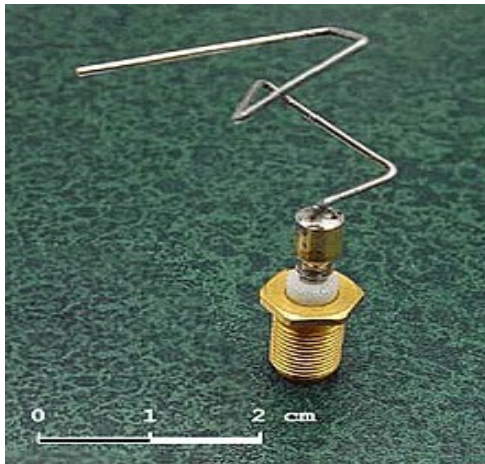
- 1).Создание базы данных сплавов на основе литературных данных
- 2).Выбор признаков для высокоэнтропийных сплавов с помощью генетического алгоритма, которые могут успешно использоваться для предсказания пластичности данного класса материалов.
- 3).Подбор подходящих моделей машинного обучения и оценка их качества работы
- 4).Тестирование модели по предсказанию пластичности на выплавленных сплавах

# Материалы и методики

- Выбор обучающего алгоритма машинного обучения
- Реализация генетического алгоритма
- Создание и выбор набора признаков
- Сбор базы данных из источников
- Испытание на одноосное сжатие
- Подготовка образцов для испытаний на сжатие

# Генетические алгоритмы

Генетический алгоритм — стохастический алгоритм поиска, применяемый для решения задач оптимизации. Идея генетических алгоритмов заимствована у живой природы и состоит в организации эволюционного процесса, конечной целью которого является получение оптимального решения в сложной комбинаторной задаче.



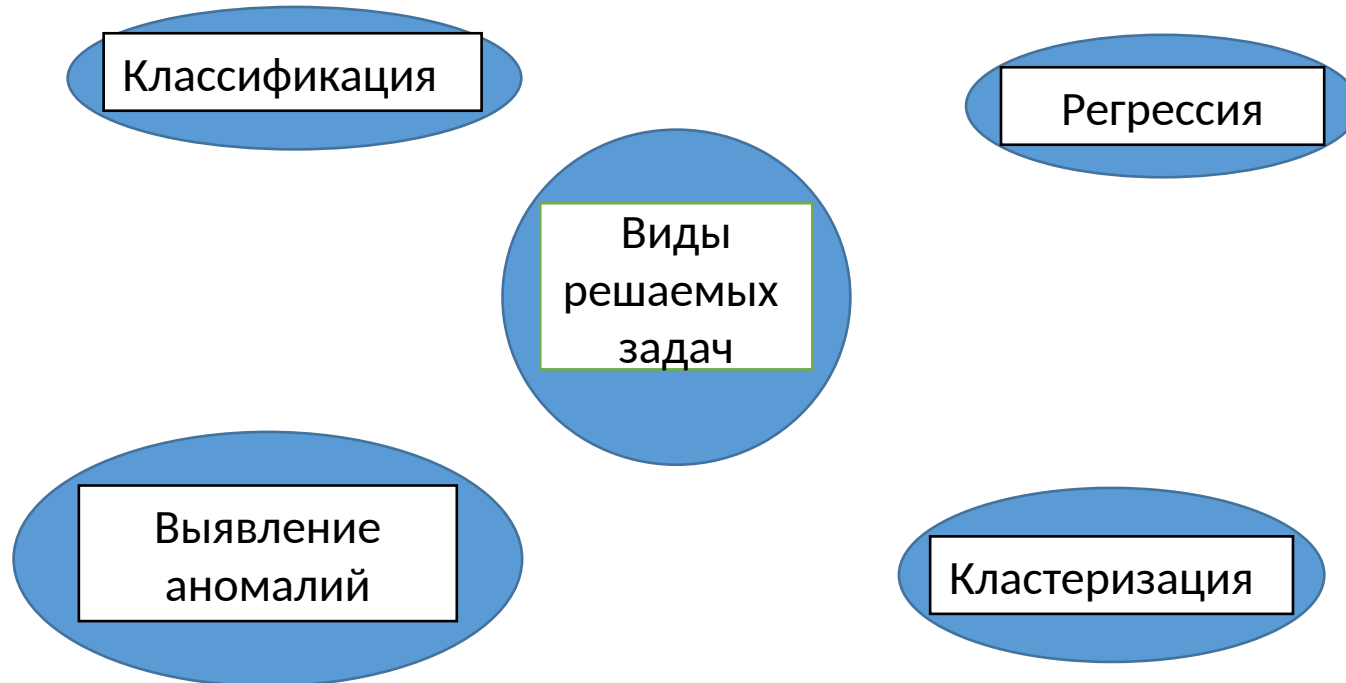
Антенна космического аппарата NASA. Сложная форма была найдена эволюционной компьютерной программой.



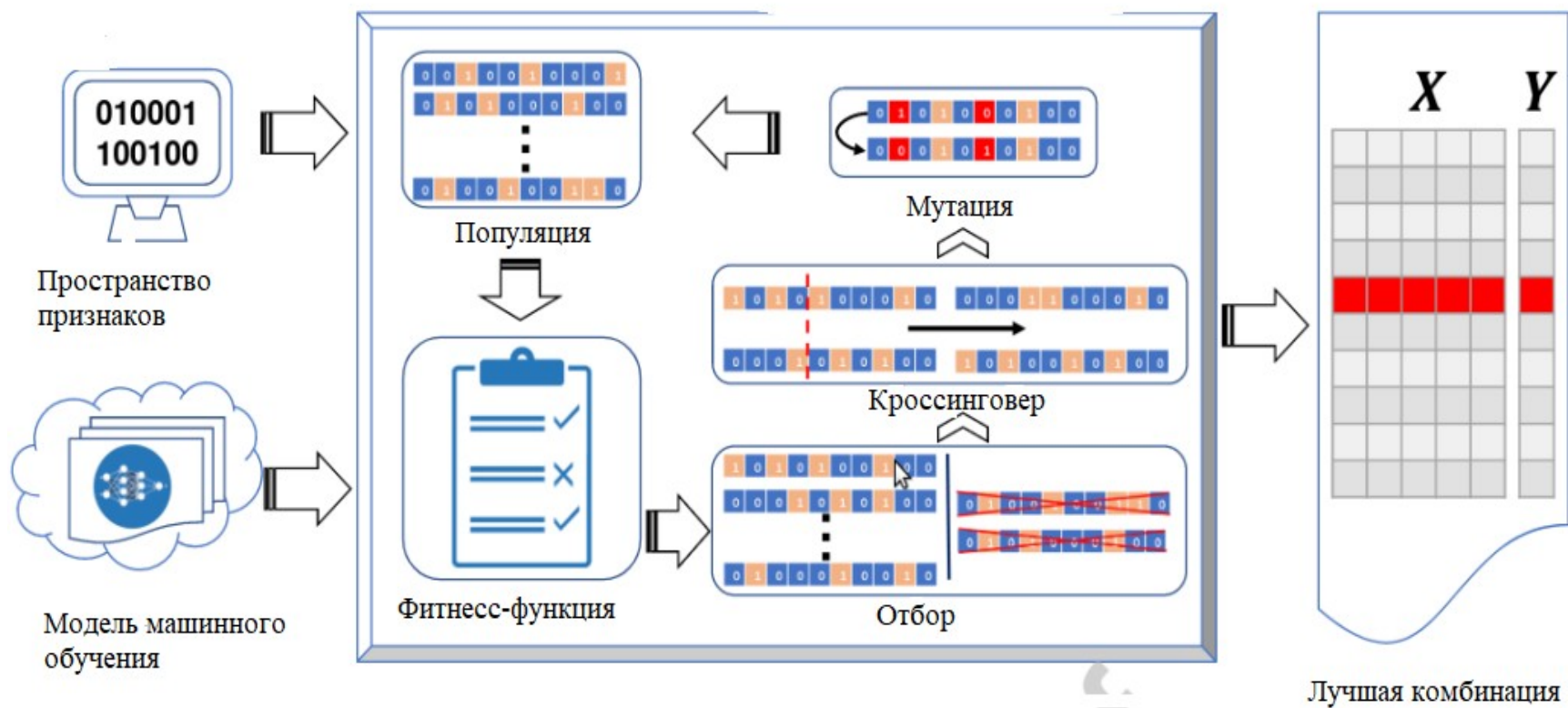
Алгоритм, предсказывающий свойства и структуру материалов по химическому составу.

# Машинное обучение

Машинное обучение – это область научного знания, имеющая дело с алгоритмами, способными обучаться. Необходимость использования методов машинного обучения объясняется тем, что для многих сложных задач очень сложно разработать "явный" алгоритм их решения, однако часто можно научить компьютер обучиться решению этих задач.



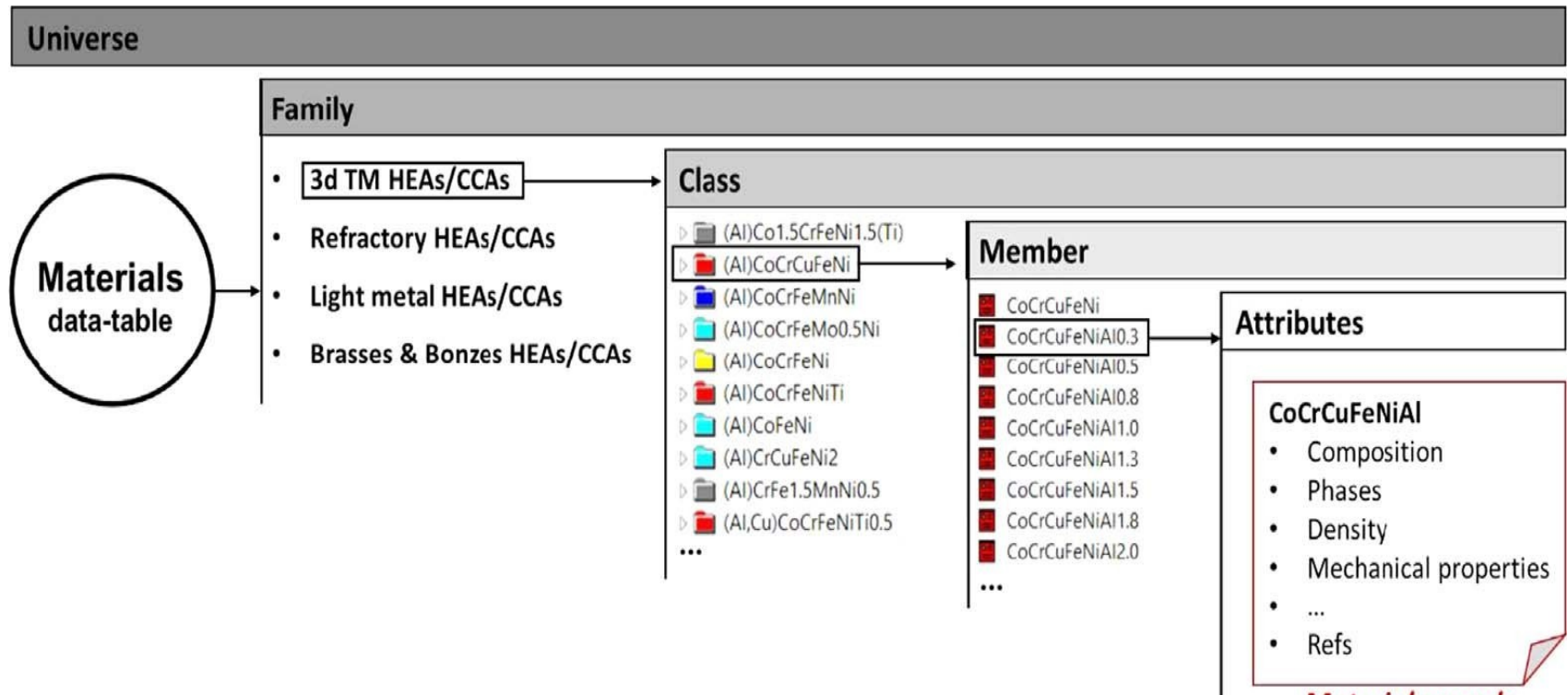
# Комбинация генетического алгоритма и подходов машинного обучения



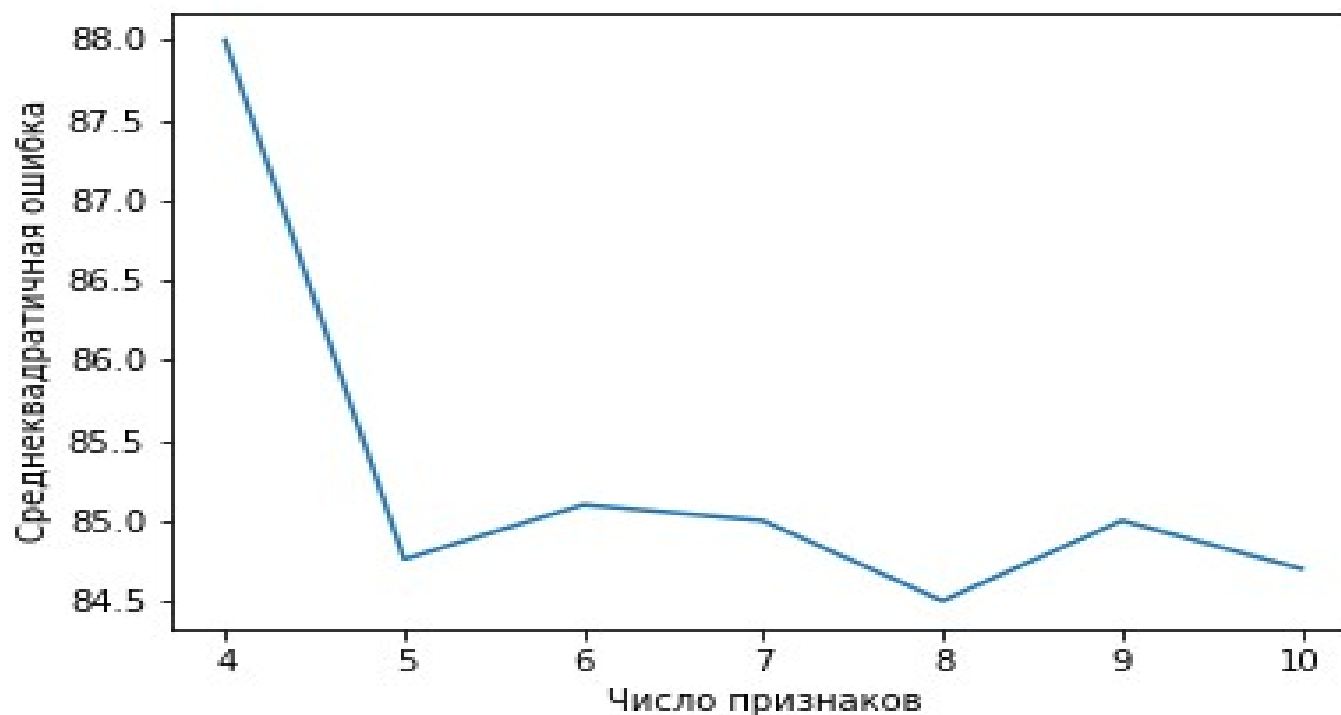
Схема, используемая для решения поставленной задачи по предсказанию пластичности высокоэнтропийных сплавов.

# Создание обучающего набора данных, сбор признаков и выбор алгоритмов машинного обучения

В качестве алгоритмов машинного обучения выбрали: метод опорных векторов, линейную и байесовская регрессия, метод к ближайших соседей. Мы свели признаки материалов, которые потенциально влияют на пластичность высокоэнтропийных сплавов (термодинамические и кинетические свойства), такие как энтальпия, энтропия и многие другие. Сплавы были отобраны из баз данных по высокоэнтропийным сплавам (Кембриджская база данных). Набор представляет собой таблицу, в которой указан химический состав и пластичность.

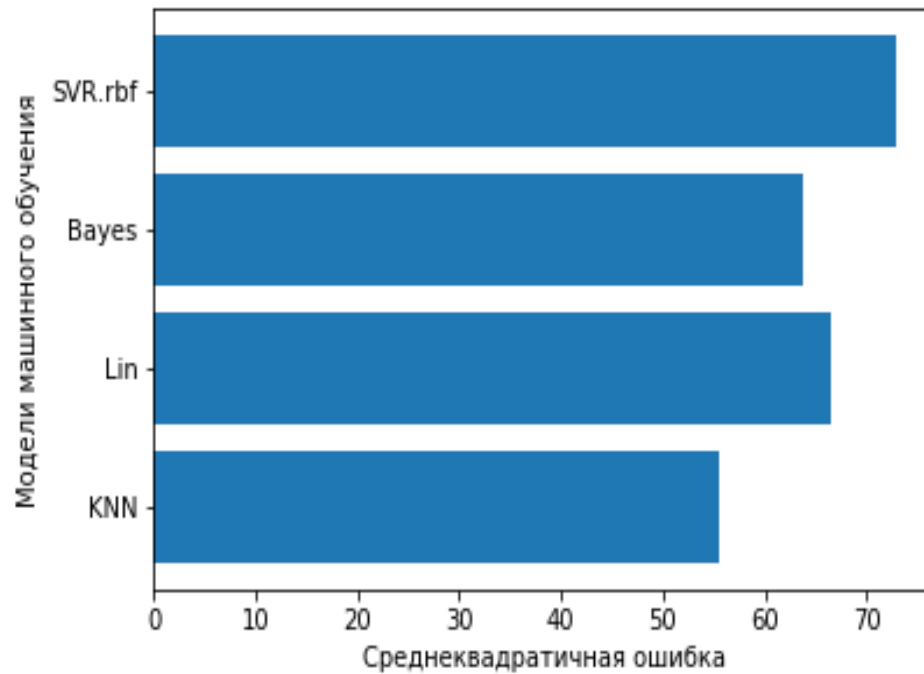


# Результаты работы генетического алгоритма



Из приведённых графиков видно, что оптимальным числом признаков для решения представленной задачи является 8 признаков и параметры подобраны правильно.

# Точность алгоритмов машинного обучения



Гистограмма точности моделей машинного обучения

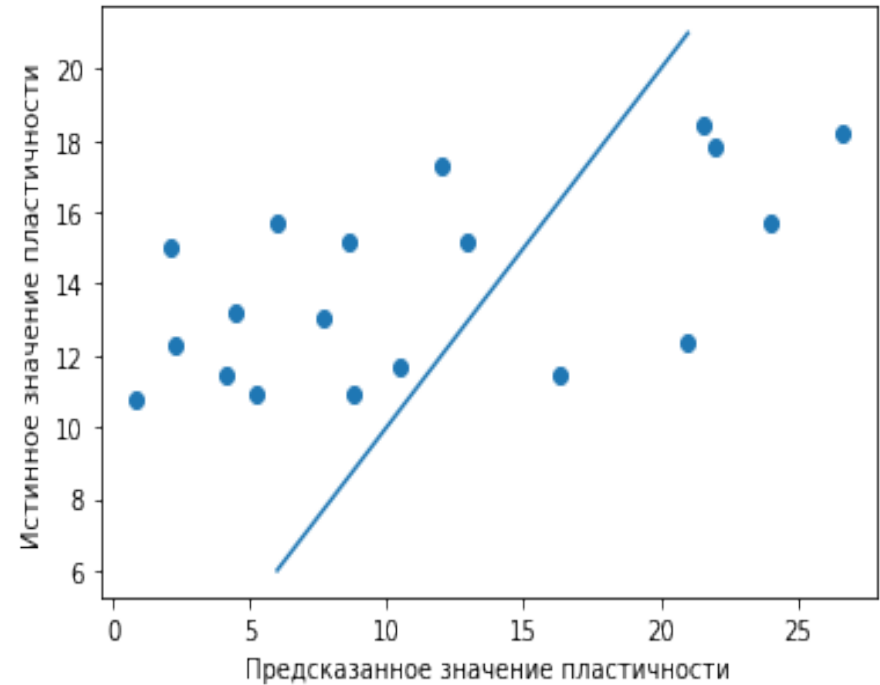


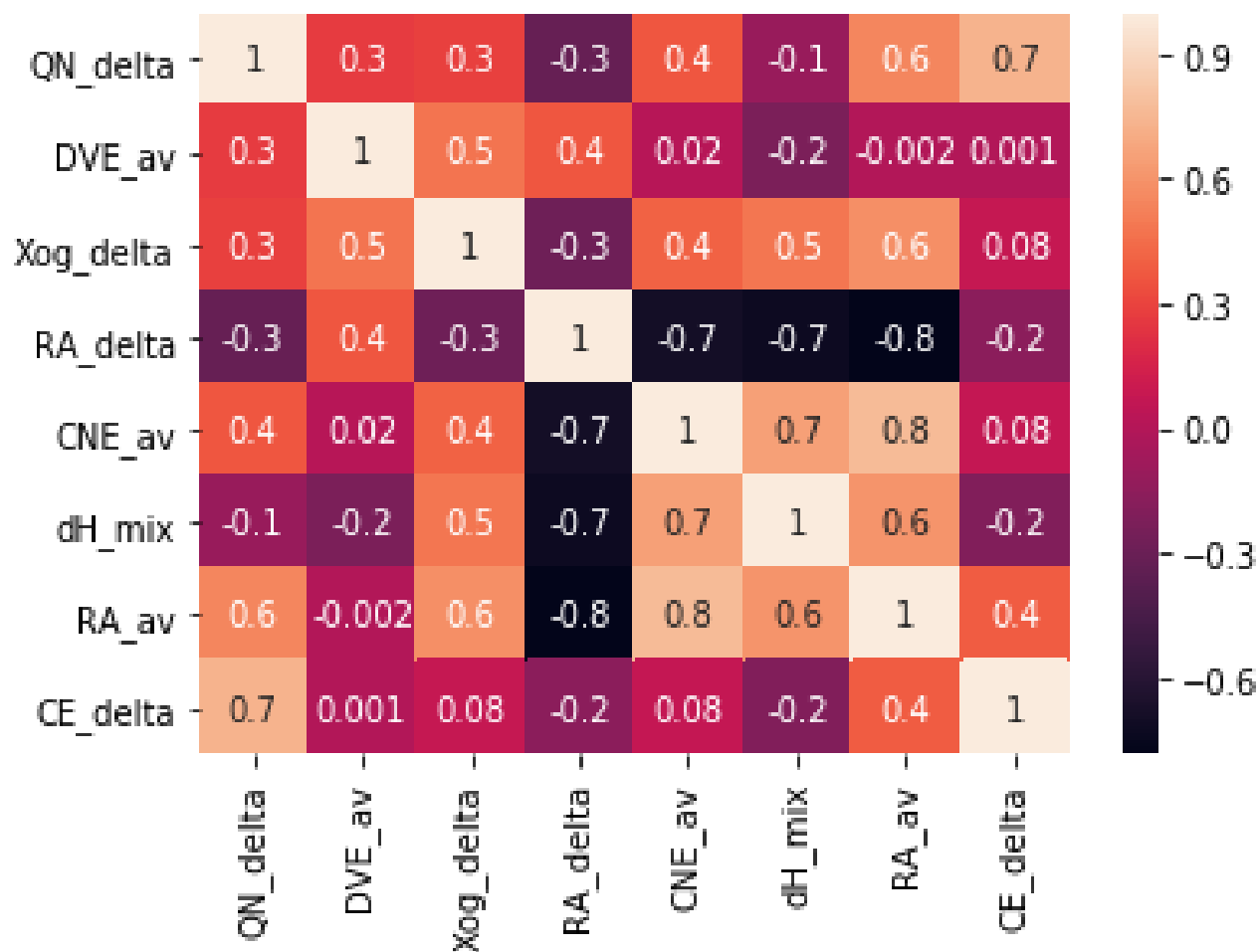
График разброса для истинных и предсказанных значений пластичности

# Результаты выбора признаков и их анализ

Делая статистический анализ полученных признаков можно сделать вывод, что во 3 моделях машинного обучения присутствует дескриптор электроотрицательности по Оганову, в 4-ёх моделях – присутствует энтальпия смешения и радиус атомов.

| Модель обучения        | Признаки               |                                                                            |        | Ошибка |
|------------------------|------------------------|----------------------------------------------------------------------------|--------|--------|
| Линейная регрессия     | Модель обучения        | Признаки                                                                   | Ошибка | 8,15   |
|                        | Линейная регрессия     | $\delta(X_{og}), C, X_{abs}, dH_{mix}, X_{MB}, EV, \delta(CNE), \delta RA$ | 8,15   |        |
|                        | Метод опорных векторов | $C, WF, X_{og}, D, VA, SIE, AN, dH_{mix}$                                  | 8,54   |        |
|                        | Байесовская регрессия  | $\delta X_{og}, \delta(EM), EA, AEN, \delta RA, VA, \delta(RAM), dH_{mix}$ | 8      |        |
|                        | К ближайших соседей    | $\delta QN, DVE, \delta X_{og}, \delta RA, CNE, dH_{mix}, RA, \delta CE$   | 7,31   |        |
| Метод опорных векторов | Модель обучения        | Признаки                                                                   | Ошибка | 8,54   |
|                        | Линейная регрессия     | $\delta(X_{og}), C, X_{abs}, dH_{mix}, X_{MB}, EV, \delta(CNE), \delta RA$ | 8,15   |        |
|                        | Метод опорных векторов | $C, WF, X_{og}, D, VA, SIE, AN, dH_{mix}$                                  | 8,54   |        |
|                        | Байесовская регрессия  | $\delta X_{og}, \delta(EM), EA, AEN, \delta RA, VA, \delta(RAM), dH_{mix}$ | 8      |        |
|                        | К ближайших соседей    | $\delta QN, DVE, \delta X_{og}, \delta RA, CNE, dH_{mix}, RA, \delta CE$   | 7,31   |        |
| Байесовская регрессия  | Модель обучения        | Признаки                                                                   | Ошибка | 8      |
|                        | Линейная регрессия     | $\delta(X_{og}), C, X_{abs}, dH_{mix}, X_{MB}, EV, \delta(CNE), \delta RA$ | 8,15   |        |
|                        | Метод опорных векторов | $C, WF, X_{og}, D, VA, SIE, AN, dH_{mix}$                                  | 8,54   |        |
|                        | Байесовская регрессия  | $\delta X_{og}, \delta(EM), EA, AEN, \delta RA, VA, \delta(RAM), dH_{mix}$ | 8      |        |
|                        | К ближайших соседей    | $\delta QN, DVE, \delta X_{og}, \delta RA, CNE, dH_{mix}, RA, \delta CE$   | 7,31   |        |
| К ближайших соседей    | Модель обучения        | Признаки                                                                   | Ошибка | 7,31   |
|                        | Линейная регрессия     | $\delta(X_{og}), C, X_{abs}, dH_{mix}, X_{MB}, EV, \delta(CNE), \delta RA$ | 8,15   |        |
|                        | Метод опорных векторов | $C, WF, X_{og}, D, VA, SIE, AN, dH_{mix}$                                  | 8,54   |        |
|                        | Байесовская регрессия  | $\delta X_{og}, \delta(EM), EA, AEN, \delta RA, VA, \delta(RAM), dH_{mix}$ | 8      |        |
|                        | К ближайших соседей    | $\delta QN, DVE, \delta X_{og}, \delta RA, CNE, dH_{mix}, RA, \delta CE$   | 7,31   |        |

## Матрицы корреляций признаков для метода ближайших соседей



# Тестовые сплавы

Химический состав тестовых сплавов и их пластичность

| Alloys | Al | Cr | Nb | Ti | V  | Zr | $\delta_{\text{пр}},\%$ | $\delta,\%$ |
|--------|----|----|----|----|----|----|-------------------------|-------------|
| A1     | 13 | 0  | 10 | 44 | 28 | 5  | -                       | 7           |
| A2     | 12 | 2  | 19 | 44 | 16 | 7  | -                       | 6           |
| A3     | 15 | 11 | 6  | 37 | 20 | 11 | 1                       | 8           |
| A4     | 11 | 7  | 25 | 50 | 14 | 11 | -                       | 3           |
| A5     | 14 | 13 | 23 | 22 | 26 | 2  | 12                      | 12          |
| A6     | 12 | 1  | 22 | 31 | 20 | 14 | 10                      | 7           |
| A7     | 14 | 3  | 21 | 34 | 4  | 24 | 1                       | 8           |
| A8     | 10 | 0  | 20 | 36 | 13 | 21 | 1                       | 9           |

# Выводы

Разработана модель машинного обучения с использованием генетического алгоритма для предсказания пластических свойств высокоэнтропийных сплавов.

Показано, что при 8 признаках генетический алгоритм демонстрирует наилучшую точность для решения задачи по предсказанию пластичности высокоэнтропийных сплавов. В качестве моделей машинного обучения использовали метод опорных векторов, метод ближайших соседей, байесовский регрессор и линейная регрессия. Оптимальным вариантом выбора модели является метод случайных соседей, который дал минимальную среднеквадратичную ошибку.

Задачу считаю выполненной, но для дальнейшего улучшения точности предсказания моделей необходим более детальный анализ. В частности, необходимо учитывать микроструктуру материала, а также наличие дефектов, таких как дислокации, вакансии и т.п.

Спасибо за внимание !