

Sztuczna inteligencja i inżynieria wiedzy laboratorium

Ćwiczenie 1. Algorytmy genetyczne

1. Wstęp

Problem kwadratowego przydziału zasobów QAP jest ważnym problemem alokacji zasobów spotykanych w przemyśle. Polega on na wskazaniu lokalizacji dla N fabryk, tak aby zminimalizować koszt transportu pomiędzy nimi. Do rozwiązania tego problemu można wykorzystać algorytm genetyczny, który jest metaheurystyką, która naśladuje ewolucję naturalną metodą ciśnienia selekcyjnego i doboru naturalnego.

2. Sposób rozwiązania:

a. Opis osobnika

Na samym początku musiałem się zastanowić czym jest osobnik. Zdefiniowałem go jako tablica liczb bez powtórzeń, każda liczba zawiera informacje o tym w której lokalizacji ma powstać dana fabryka. Rozmiar tablicy jest zależny od ilości lokalizacji/fabryk.

b. Ocena osobnika

Aby ocenić osobnika musiałem mieć do dyspozycji 2 macierze, przepływu i odległości. Ocena przebiega poprzez dodawanie do siebie pomnożonych przez siebie wybranych wartości tych dwóch macierzy. Wynik, który uzyskujemy po wykonaniu tej operacji jest wartością, która definiuje jak dobry jest osobnik. Im ta wartość jest mniejsza tym lepiej.

c. Opis metod ewolucyjnych

W przebiegu algorytmu zastosowałem trzy metody ewolucyjne: selekcję, krzyżowanie i mutację w podanej kolejności. Szczegółowe ich działania wygląda następująco:

- **Selekcja :**
 - **Turniejowa** – wybieram losowo podaną liczbę osobników z populacji, następnie sprawdzam który z nich ma najniższą wartość oceny (czy który jest najlepszy) i ten osobnik zostaje wybrany.
 - **Ruletkowa** – im osobnik ma mniejsza wartość oceny (czyli im jest lepszy) tym większą szanse ma na wylosowanie. Następnie następuje losowanie i osobnik zostaje wybrany.
- **Krzyżowanie** – przed wykonaniem tej metody musimy uwzględnić prawdopodobieństwa jej wystąpienia. Krzyżowanie możemy zacząć gdy mamy dwóch osobników wybranych metodą selekcji. Następnie losujemy wartość większą od zera i mniejszą lub równą ilości genów które posiada osobnik. Wartość ta dzieli rodziców na 2 części, tworząc z nich dziecko. Następnym etapem jest naprawa osobnika ponieważ nie może on mieć dwóch takich samych genów.
- **Mutacja** – wykonując tą metodę musimy uwzględnić prawdopodobieństwo jej wykonania dla każdego genu w osobniku. Sam proces polega na zamianie genu z losowo wybranym.

d. Opis rozwiązania problemu

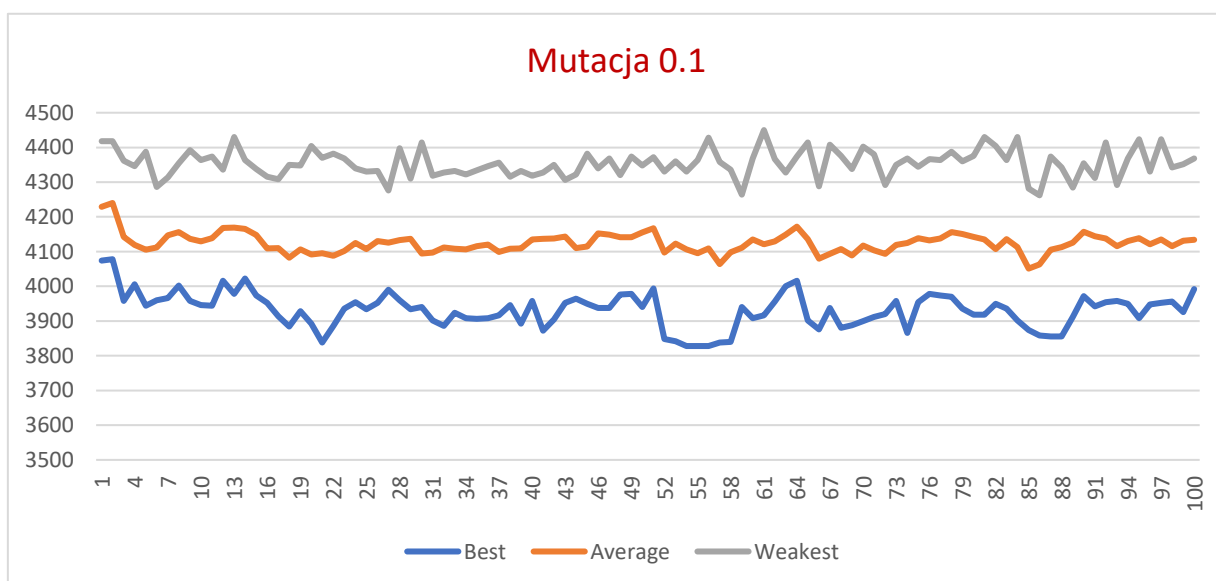
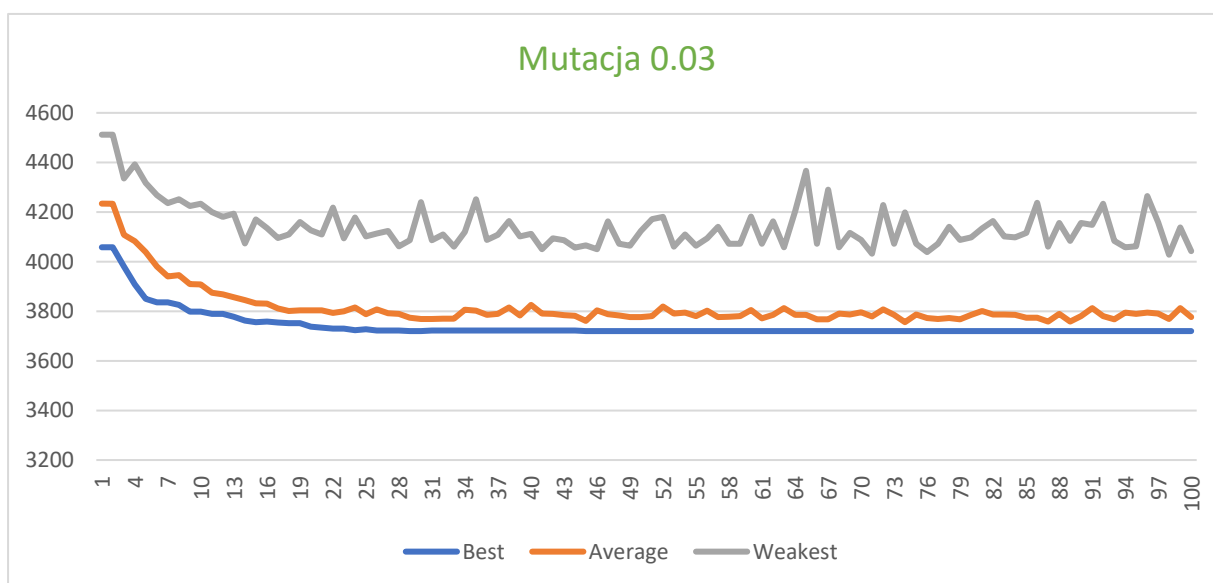
- I. Wczytanie pliku z macierzami i utworzenie ich w programie
- II. Inicjalizacja populacji o określonej liczbie losowych osobników
- III. Obliczenie wartości oceny dla każdego osobnika
- IV. Zapis do pliku najlepszego, najgorszego i średnią oceny osobników w generacji
- V. Rozpoczęcie pętli określającej ilość generacji
- VI. Rozpoczęcie pętli która utworzy osobników następnej generacji, która wykonuje następujące kroki:
 - i. Selekcja dwóch osobników.
 - ii. Krzyżowanie tych osobników uwzględniając prawdopodobieństwo
 - iii. Mutacja powstałego osobnika uwzględniając prawdopodobieństwo na poziomie genu
 - iv. Dodanie osobnika do nowej generacji
- VII. Zapis do pliku najlepszego, najgorszego i średnią oceny osobników w generacji
- VIII. Nowa populacja staje się starą populacją

3. Wyniki badań

1. Badanie Mutacji - HAD16.dat

PopSize:100 Gen:100 Px:0.7 Tour:5

Prawdopodobieństwo Mutacji (Pm)	Średnia Średniej z Najlepszych
0.01	3779,63
0.02	3764,627
0.03	3754,884
0.05	3774,57
0.1	3912,963



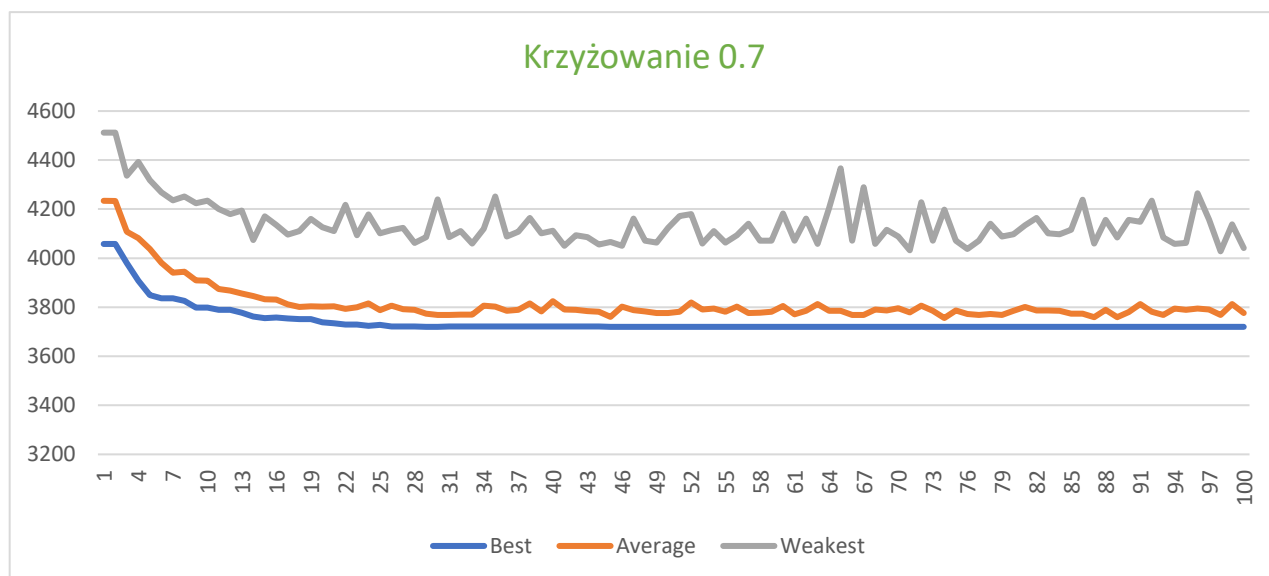
Podsumowanie i wnioski:

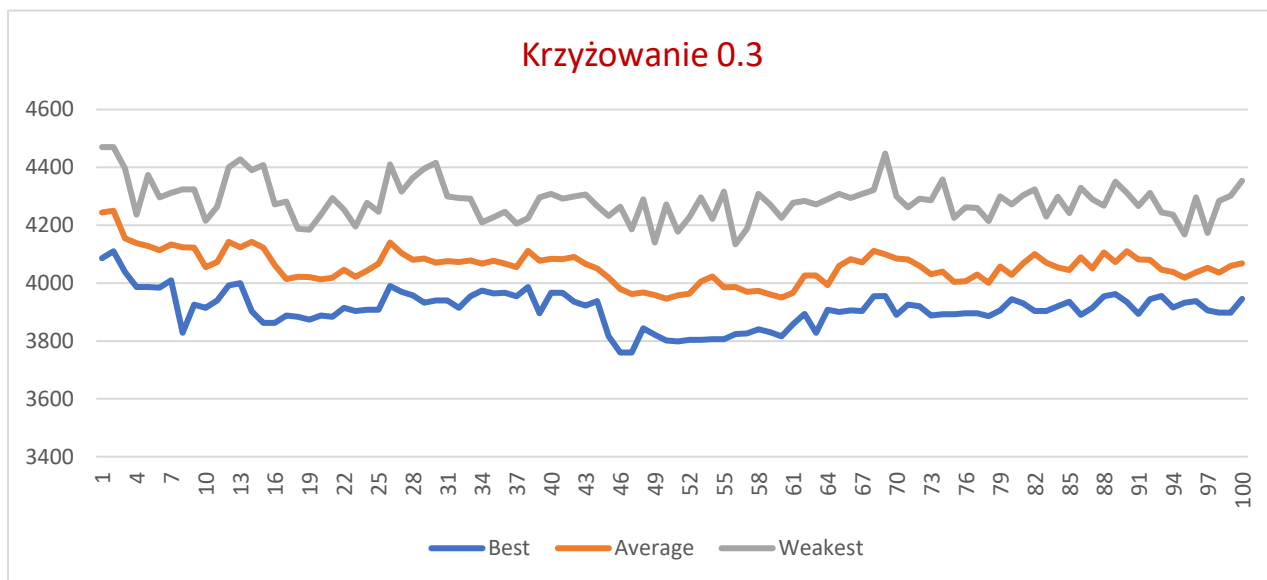
Im większe prawdopodobieństwo mutacji tym więcej osobników jest zmieniana w danym pokoleniu. Mutacja nie zawsze powodowała poprawienie jakości osobników, często zdarzało się, że nowe osobniki było gorsze niż ich niezmutowane wersje. Zwiększanie prawdopodobieństwa mutacji od pewnego momentu nie dawało korzystnych wyników. Według moich badań najlepsze prawdopodobieństwo to 3%, zwiększając je od tego momentu pogorszymy jakość naszych wyników.

2. Badanie Krzyżowania – HAD16.dat

PopSize:100 Gen:100 Pm:0.03 Tour:5

Prawdopodobieństwo Krzyżowania (Px)	Średnia Średniej z Najlepszych
1	3775,173
0.8	3770,307
0.7	3754,884
0.6	3757
0.3	3909,323





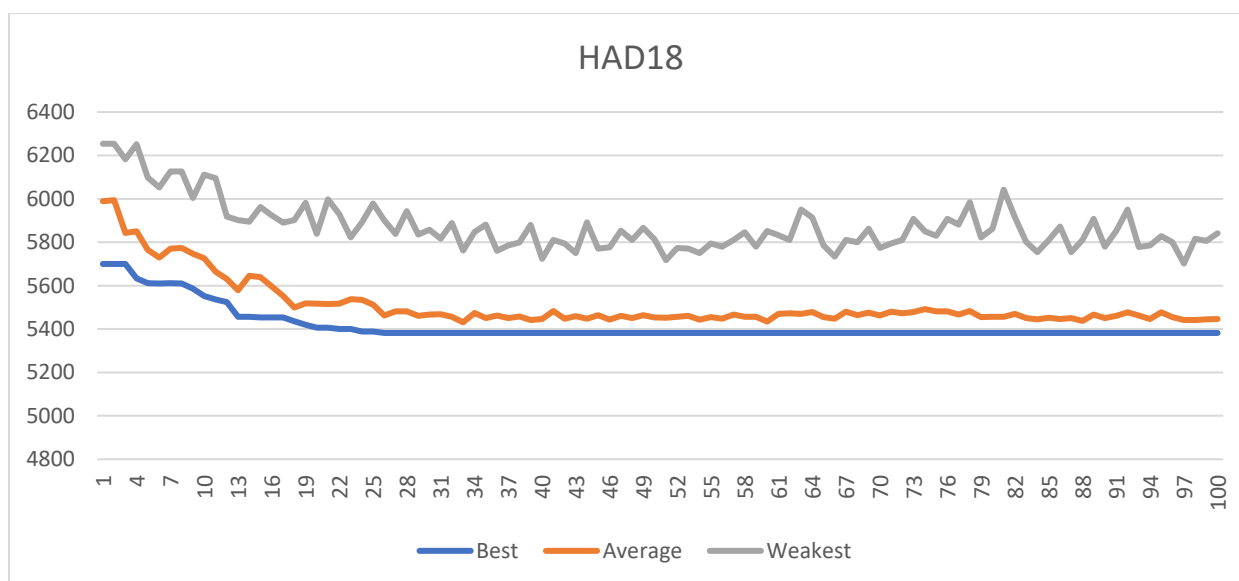
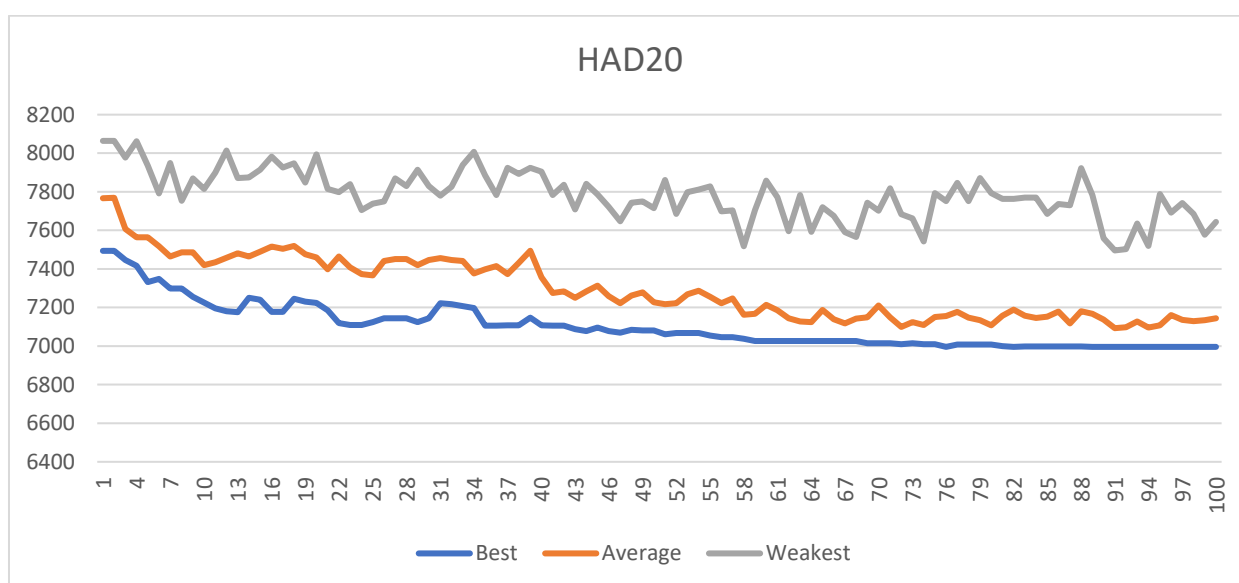
Podsumowanie i wnioski:

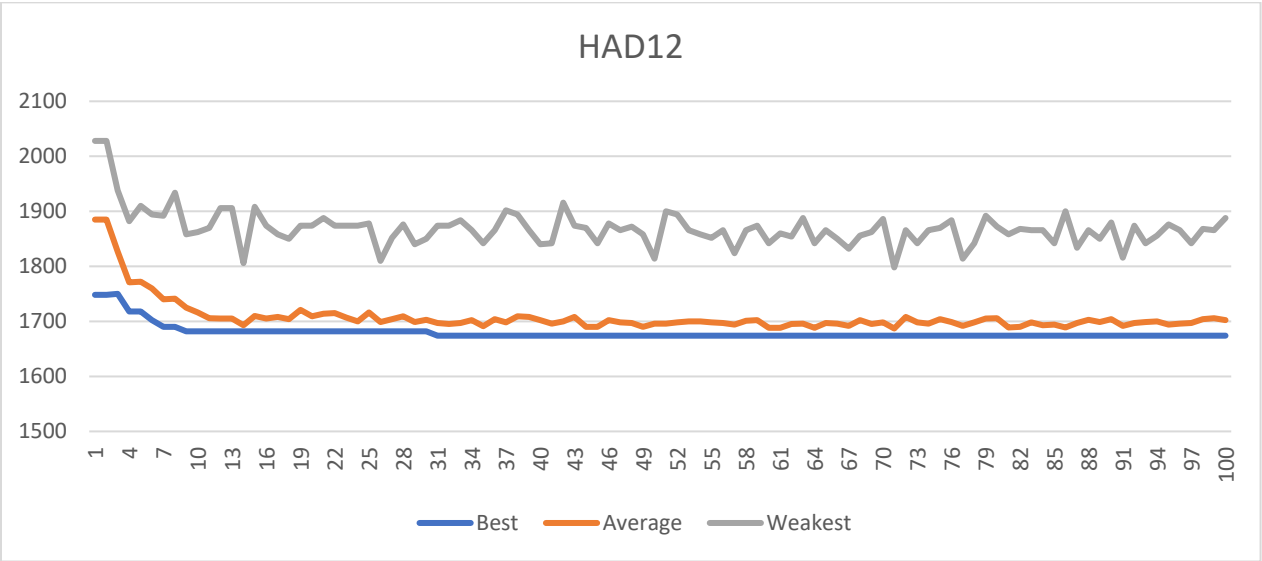
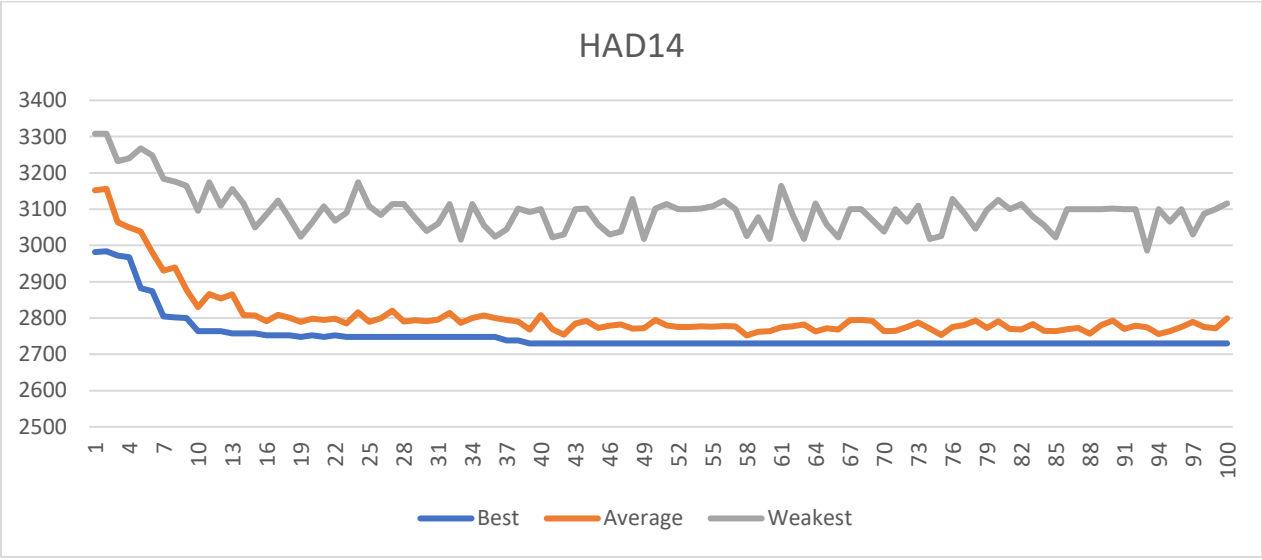
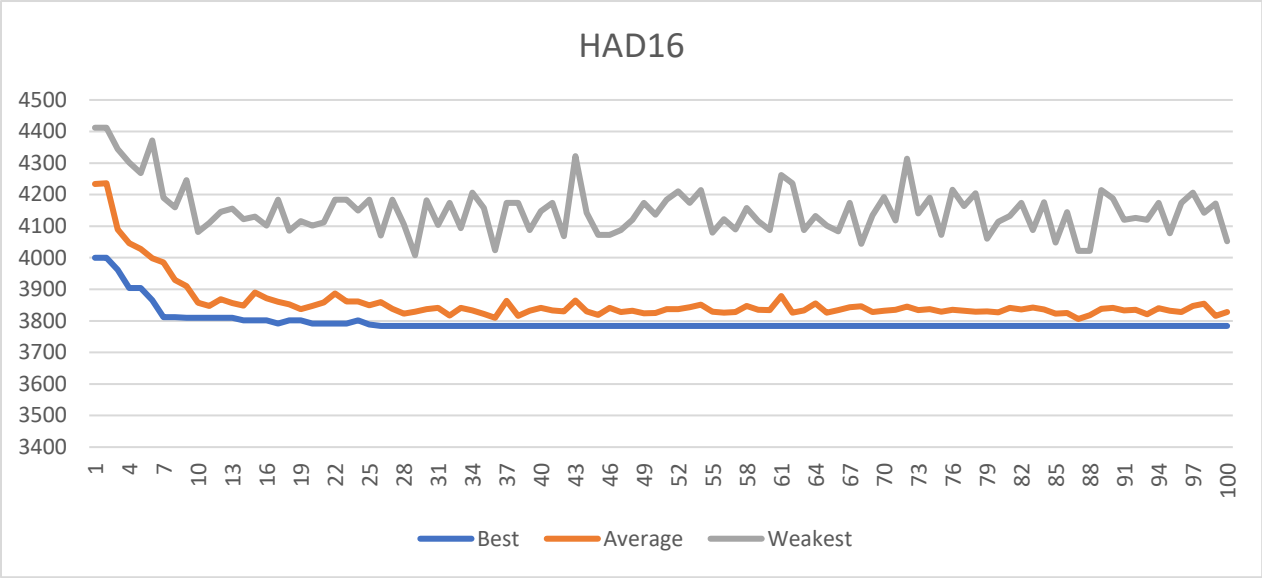
Z moich obserwacji wynika że prawdopodobieństwo krzyżowania, które daje najlepsze rezultaty to 0.7. Gdy zwiększamy tą wartość to zbyt dużo osobników którzy mają najlepsze wyniki zostaje poddanych krzyżowaniu co za tym idzie końcowy wynik nie jest tak dobry. Natomiast zmniejszając tą wartość rozwój kolejnych pokoleń nie daje satysfakcjonujących wyników, co widać na wykresie gdzie prawdopodobieństwo krzyżowania jest ustawione na 0.3.

3. Badania na 5 plikach testowych

PopSize:100 Gen:100 Pm:0.03 Px:0.7 Tour:5

Nazwa Pliku	Średnia	Odchylenie	OPT
HAD12	1675,697	8,252655	1652
HAD14	2756,187	6,576044	2724
HAD16	3785,023	17,51657	3720
HAD18	5429,6	18,10071	5358
HAD20	7058,563	28,47923	6922





Podsumowanie i wnioski:

W każdym z badanych plików wynik końcowy był zbliżony do idealnego, ale nie udało się go osiągnąć. Ze względu na wielkość macierzy możemy zaobserwować że oś najlepszych zaczyna się stabilizować w innych miejscach. Dla HAD12 wystarczyło 31 generacji, aby otrzymać najlepszy wynik. Natomiast dla HAD20 stało się to dopiero przy 82 generacji.

4. Porównanie algorytmu z nieewolucyjną metodą

Jako nieewolucyjną metodę podejścia do problemu wybrałem losowe tworzenie populacji i sprawdzanie wartości oceny dla każdego z nich. Metoda nieewolucyjna daje przypadkowe rezultaty, które trudno przewidzieć i uznać za optymalne. Wyniki będą porzucane, nie będzie można dostrzec żadnej tendencji. Jest to metoda na „chybił trafił”, w której przypadek decyduje o utworzeniu danego osobnika.