




Algorytmy genetyczne – przeszukiwanie dużych przestrzeni rozwiązań

- [Wprowadzenie](#)
- [Przeczytaj](#)
- [Audiobook](#)
- [Sprawdź się](#)
- [Dla nauczyciela](#)



Algorytmy genetyczne – przeszukiwanie dużych przestrzeni rozwiązań

Źródło: CHUTERSNAP, domena publiczna.

Główną koncepcją algorytmów genetycznych jest efektywne przeszukiwanie przestrzeni rozwiązań do danego problemu. Dzięki swojej losowości i operacjom genetycznym są w stanie stosunkowo szybko znaleźć optymalne rozwiązanie.

Algorytmy genetyczne są najczęściej stosowane do rozwiązywania problemów nieliniowych i trudnych do opisanego w sensie matematycznym. Jednym z nich jest problem grupowania, którym zajmiemy się w tym e-materiale.

Pozostałe materiały z serii:

- [Algorytmy genetyczne,](#)
- [Algorytmy genetyczne – populacja,](#)
- [Algorytmy genetyczne – projektowanie,](#)
- [Algorytmy genetyczne – zastosowanie.](#)

Twoje cele

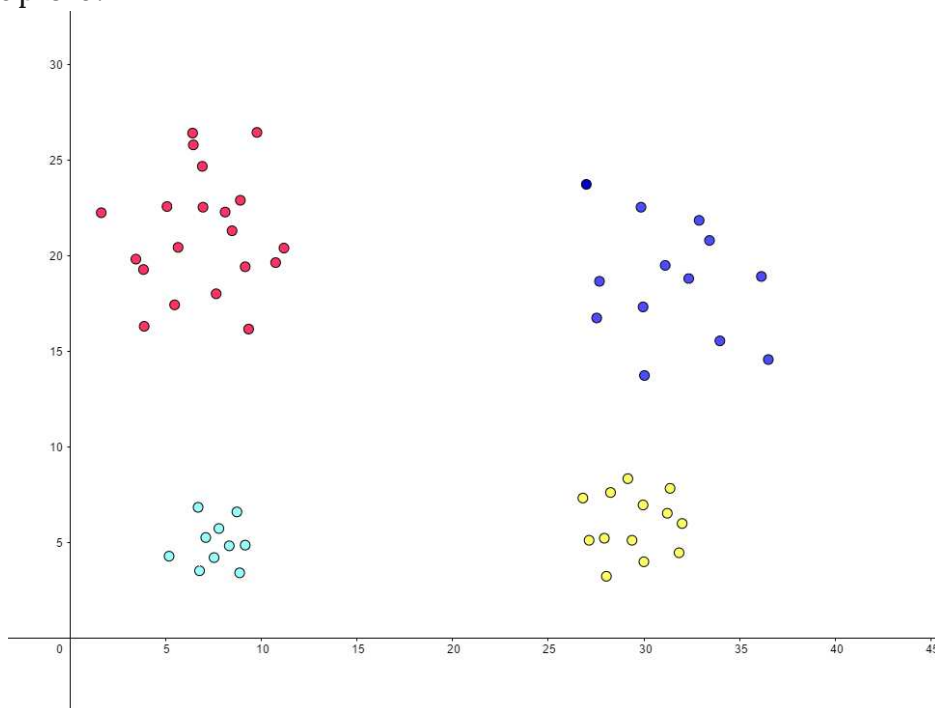
- Dowiesz się, na czym polega problem grupowania.
- Zapoznasz się z koncepcją wykorzystania algorytmów genetycznych w grupowaniu.
- Poznasz przykładowe problemy grupowania.

Przeczytaj

Czym jest problem grupowania?

Problem grupowania należy do kategorii problemów posiadających bardzo dużą przestrzeń rozwiązań. Jego istotą jest pogrupowanie danych przekazanych na wejściu względem pewnych wspólnych cech.

Problem ten można łatwo przedstawić w formie punktów w dwuwymiarowym układzie współrzędnych. Na wejściu programu podajemy współrzędne każdego z punktów, a następnie algorytm powinien wskazać ich skupiska. Przy czym każdy punkt musi należeć do jakiegoś skupiska.



Źródło: Contentplus.pl Sp. z o.o., licencja: CC BY-SA 3.0.

Standardowe algorytmy do zagadnień grupowania wymagają od użytkownika informacji wstępnej, takiej jak:

- liczba grup,
- promienie grup,
- gęstość punktów w grupach.

Czasami ciężko jednak jest podać taką informację. Dodatkowo tradycyjne algorytmy mają tendencję do wpadania w ekstrema lokalne. Tych wad pozbawione są algorytmy genetyczne. Można je dodatkowo ulepszyć poprzez wybór bardzo dobrej populacji początkowej jednym z tradycyjnych algorytmów.

Algorytm k-średnich

Algorytm k-średnich jest tradycyjnym algorytmem wykorzystywanym do rozwiązywania problemów grupowania. Swoje działanie opiera on na **centroidach** utworzonych na podstawie średniej współrzędnych punktów w grupie.

Jest to algorytm iteracyjno-optymalizacyjny. W kolejnych iteracjach przenosi on obiekty pomiędzy skupieniami, dążąc do zwiększenia podobieństw między obiektami w danej grupie.

Wadą tego algorytmu jest to, że na samym początku wymaga informacji o liczbie grup w analizowanym zbiorze danych. Dodatkowo jest on wrażliwy na odstające elementy oraz **szum**.

Zastosowanie algorytmu genetycznego

Nasz algorytm genetyczny będzie zajmował się wyszukiwaniem środków każdego ze skupisk punktów. Genem będzie para liczb rzeczywistych, które reprezentują współrzędne punktu na płaszczyźnie. Chromosom zatem będzie zbiorem potencjalnych środków skupisk.

1 ; 8	2 ; 4	7 ; 3
-------	-------	-------

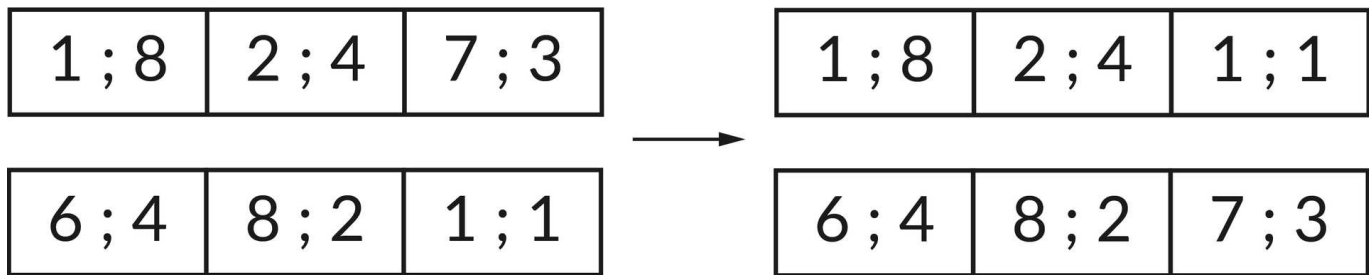
Źródło: Contentplus.pl Sp. z o.o., licencja: CC BY-SA 3.0.

Pierwszą populację możemy wybrać zupełnie losowo albo spośród podanych na wejściu punktów. Wadą takiego podejścia jest to, że dojście do optymalnego rozwiązania może zająć dużą ilość czasu.

Dobrym pomysłem jest na samym początku wykorzystanie tradycyjnego algorytmu znajdowania grup. Dzięki temu uzyskamy w miarę przybliżone rozwiązanie, które następnie możemy zoptymalizować naszym algorytmem genetycznym.

Operacja krzyżowania

Operacja krzyżowania w naszym algorytmie genetycznym będzie dosyć standardowa. Możemy tutaj stosować zarówno krzyżowanie jednopunktowe, jak i wielopunktowe. Nie ma to aż tak dużego wpływu na jego funkcjonowanie.



Źródło: Contentplus.pl Sp. z o.o., licencja: CC BY-SA 3.0.

Operacja mutacji

Tak samo jak w przypadku algorytmów genetycznych projektujących obwody elektryczne, będziemy tutaj wykorzystywać trzy rodzaje mutacji. Pierwsza z nich będzie polegała na dodawaniu kolejnego genu do danego chromosomu, a zatem kolejnego potencjalnego środka skupiska.

Drugi rodzaj mutacji jest przeciwieństwem pierwszego i będzie on polegał na usunięciu losowego punktu skupiska z danego chromosomu. Natomiast ostatni rodzaj polega na losowej zmianie wartości współrzędnych zawartych w danym genie.

Technika równoległych populacji

Aby ulepszyć działanie naszego algorytmu genetycznego, możemy wprowadzić pewną modyfikację jego działania. Zamiast pracować na jednej, możemy równoległe ewoluować kilka niezwiązanych ze sobą populacji. Następnie co parę iteracji naszego algorytmu stworzymy nowe populacje poprzez krzyżowanie najlepszych osobników ze wszystkich grup między sobą. Pozwala to na lepsze i szybsze przeszukiwanie przestrzeni rozwiązań.

Użycie wniosków wyciągniętych ze startowej populacji

Jeżeli zdecydujemy się na użycie tradycyjnego algorytmu do wyznaczenia początkowej populacji, to możemy wyciągnąć z niej parę przydatnych wniosków, które usprawnią dalszy proces ewolucyjny.

Przykładowo, dzięki punktom wyznaczonym przez tradycyjny algorytm jesteśmy w stanie określić minimalną oraz maksymalną ilość skupisk w naszych danych wejściowych. Możemy tę informację uwzględnić w trakcie przeprowadzania mutacji i utrzymywać liczbę genów w chromosomach w danych granicach.

Przykładowe problemy grupowania

Z problemem grupowania można się często spotkać w branży marketingowej. Często grupuje się klientów o podobnych zachowaniach oraz zainteresowaniach. Dzięki temu

można wyświetlać dla nich odpowiednie reklamy, modyfikować ceny w sklepach internetowych oraz proponować im promocje na produkty, które mogą ich zainteresować.

Jeżeli prowadzimy własny serwis internetowy, to bardzo często zbieramy dużą ilość danych na temat naszych odwiedzających. Ich klasyfikacja pozwala znaleźć przydatne zależności, które mogą nam pomóc w dostosowaniu treści znajdujących się na naszej stronie.

Grupowanie jest również wykorzystywane do podziału obrazu na regiony pod względem własności, takich jak kolor, tekstura, czy intensywność. Taki uproszczony obraz jest prostszy do analizy przez algorytmy rozpoznawania obrazu. Jest ono również wykorzystywane do klasyfikacji dokumentów tekstowych w wyszukiwarkach internetowych.

Słownik

szum informacyjny

nadmiar informacji utrudniający wyodrębnienie informacji prawdziwych i istotnych
(źródło: *Słownik języka polskiego PWN*)

centroid

punkt leżący wewnątrz skupiska obiektów; reprezentuje jego środek

Audiobook

Polecenie 1

Przesłuchaj audiobook, a następnie zastanów się, jak może przebiegać rozwój algorytmów genetycznych w następnych latach.

Audiobook można wysłuchać pod adresem: <https://zpe.gov.pl/b/PjNJlFzX>

Spróbujmy przybliżyć historię algorytmów genetycznych, aby zrozumieć, jak narodził się i rozwijał pomysł zastosowania biologicznych procesów w rozwiązywaniu problemów z różnych dziedzin życia.

Wszystko rozpoczęło się w 1950 roku, kiedy opublikowano przełomowy artykuł Alana Turinga pod tytułem „Maszyny liczące a inteligencja”. Opisał on w nim nie tylko koncepcję słynnego testu, ale również rozważania dotyczące działania maszyny w nim uczestniczącej.

Wbudowany w niej proces uczenia się miał przebiegać w sposób przypominający ewolucję. Można to nazwać pierwszym krokiem w przeniesieniu genetyki do świata techniki. Było to także ważne wydarzenie z punktu widzenia rozwoju sztucznej inteligencji.

W następnych latach wydanych zostało jeszcze kilka publikacji naukowych o symulacji ewolucji, jednakże dopiero po roku 1960 nastąpiło większe zainteresowanie tym tematem. Duży wpływ wywarł na to australijski genetyk Alex Fraser. Wraz z Donaldem G Burnellem opisali w swojej książce dokładne symulacje zawierające wiele zasad i elementów powszechnie stosowanych dzisiaj w algorytmach genetycznych.

W podobnym czasie działał również Hans-Joachim Bremermann, który jako jeden z pierwszych uczonych podjął próbę rozwiązania problemów optymalizacji z użyciem symulacji ewolucji. Uczony ten był także prekursorem tworzenia teorii algorytmów ewolucyjnych.

Na przełomie lat sześćdziesiątych oraz siedemdziesiątych XX wieku pojawiły się również kolejne duże osiągnięcia w omawianej dziedzinie. Zaczęto rozwiązywać pierwsze bardziej skomplikowane problemy inżynierskie z użyciem algorytmów ewolucyjnych.

Znaczącą rolę w popularyzacji algorytmów genetycznych w następnych latach odegrał John Henry Holland. Opracował on między innymi twierdzenie o przewidywaniu jakości następnych generacji, nazwane potem od jego nazwiska.

Prawdziwy przełom nastąpił jednak w roku 1985, kiedy w Pittsburghu zorganizowano pierwszą międzynarodową konferencję dotyczącą algorytmów genetycznych. Grupa profesorów i badaczy zainteresowanych tematem była tak wielka, że już 4 lata później pojawiła się potrzeba stworzenia większej organizacji zajmującej się tym działem.

W ten sposób utworzono Międzynarodowe Stowarzyszenie Algorytmów Genetycznych (International Society for Genetic Algorithms). Była to inicjatywa, która mocno przyczyniła się do rozwoju algorytmów genetycznych w następnych latach.

Warto przy okazji wspomnieć również o polskich uczonych, którzy zajmują się bądź zajmowali się omawianym tematem. Należą do nich między innymi Zbigniew Michalewicz, autor książki „Algorytmy genetyczne + struktury danych = programy ewolucyjne”, a także Andrzej Buller i Marek Kasperski.

Współcześnie algorytmy genetyczne są już szeroko stosowane i raczej rzadko myślimy o tym, ile pracy i czasu wymagało ich rozwinięcie aż do obecnej postaci. Warto jednak pamiętać o stojącej za nimi historii. Znając bowiem genezę i pierwotne założenia danej koncepcji, możemy ją zazwyczaj lepiej zrozumieć.

Sprawdź się

Pokaż ćwiczenia:   

Ćwiczenie 1



Wskaż, czy algorytm k-średnich jest wrażliwy na odstające elementy oraz szum.

nie

tak

Ćwiczenie 2



Wskaż, na czym polega technika równoległych populacji.

Na równoległym ewoluowaniu kilku niezwiązanych ze sobą populacji, przy czym po każdej iteracji algorytmu krzyżujemy najlepszych osobników ze wszystkich grup między sobą.

Na równoległym ewoluowaniu kilku niezwiązanych ze sobą populacji, przy czym co parę iteracji algorytmu krzyżujemy najlepszych osobników ze wszystkich grup między sobą.

Na równoległym ewoluowaniu kilku niezwiązanych ze sobą populacji, przy czym na koniec działania algorytmu wybieramy najlepsze rozwiązanie spośród wszystkich populacji.

Na równoległym ewoluowaniu kilku niezwiązanych ze sobą populacji, przy czym co parę iteracji nie wykonujemy operacji mutacji.

Ćwiczenie 3



Wskaż, w jaki sposób jesteśmy w stanie ulepszyć algorytm genetyczny rozwiązujący problem grupowania.

- Możemy zrezygnować z operacji mutacji, aby przyspieszyć działanie algorytmu.
- Możemy zwiększyć liczbę iteracji algorytmu, aby go przyspieszyć.
- Możemy zastosować jeden z tradycyjnych algorytmów wykorzystywanych do rozwiązania problemu grupowania, aby wygenerować startową populację algorytmu genetycznego.
- Możemy zmniejszyć liczbę iteracji algorytmu, aby go przyspieszyć i zmniejszyć ryzyko utknięcia w ekstremum lokalnym.

Ćwiczenie 4



Zaznacz wszystkie właściwości algorytmu k-średnich.

- Swoje działanie opiera na centroidach.
- Jest to algorytm rekurencyjny.
- W swoim działaniu korzysta z operacji genetycznych.
- Jest to algorytm iteracyjno-optymalizacyjny.
- Jest algorytmem wykorzystywanym do rozwiązywania problemów grupowania.

Ćwiczenie 5



Wskaż, czy jeżeli użyjemy tradycyjnego algorytmu do wyznaczenia początkowej populacji, to możemy wyciągnąć z niej parę przydatnych wniosków.

- Nie, jedyną zaletą takiego podejścia jest lepsza populacja startowa algorytmu.
- Tak, jesteśmy w stanie określić minimalną oraz maksymalną ilość skupisk w naszych danych wejściowych.

Ćwiczenie 6



Wstaw brakujące wyrażenia tak, aby treść poniższego tekstu była prawdziwa.

wykorzystywane do rozwiązywania problemów grupowania posiadają parę wad.

Przed wszystkim zazwyczaj wymagają one od użytkownika pewnego rodzaju informacji

wstępnej. Może to być lub ich .

Zwykle trudno jest nam podać taką

informację. Dodatkowo ich kolejną wadą jest tendencja do utykania w ekstremach .

Tych wad algorytmy genetyczne.

obwód

objętość grup

nie są pozbawione

liczba grup

lokalnych

globalnych

Algorytmy genetyczne

są pozbawione

promień

Tradycyjne algorytmy

Ćwiczenie 7



Połącz każde pojęcie z właściwą definicją.

szum

nadmiar informacji utrudniający
wyodrębnienie informacji prawdziwych
i istotnych

problem grupowania

jego istotą jest uporządkowanie danych
przekazanych na wejściu względem
pewnych wspólnych cech

centroid

punkt leżący wewnątrz skupiska
obiektów; reprezentuje jego środek

Ćwiczenie 8



Wskaż, ile różnych rodzajów mutacji wykorzystujemy w algorytmie genetycznym rozwiązującym problem grupowania.

trzy

jeden

cztery

dwa

Dla nauczyciela

Autor: Bartosz Zadrozny

Przedmiot: Informatyka

Temat: Algorytmy genetyczne – przeszukiwanie dużych przestrzeni rozwiązań

Grupa docelowa:

Szkoła ponadpodstawowa, liceum ogólnokształcące, technikum, zakres rozszerzony

Podstawa programowa:

Cele kształcenia – wymagania ogólne

I. Rozumienie, analizowanie i rozwiązywanie problemów na bazie logicznego i abstrakcyjnego myślenia, myślenia algorytmicznego i sposobów reprezentowania informacji.

Treści nauczania – wymagania szczegółowe

I. Rozumienie, analizowanie i rozwiązywanie problemów.

Zakres rozszerzony. Uczeń spełnia wymagania określone dla zakresu podstawowego, a ponadto:

8) dyskutuje na temat roli myślenia komputacyjnego i jego metod, takich jak: abstrakcja, reprezentacja danych, dekompozycja problemu, redukcja, myślenie rekurencyjne, podejście heurystyczne w rozwiązywaniu problemów z różnych dziedzin.

III. Posługiwanie się komputerem, urządzeniami cyfrowymi i sieciami komputerowymi.

Zakres rozszerzony. Uczeń spełnia wymagania określone dla zakresu podstawowego, a ponadto:

2) dokonuje kompresji informacji, objaśnia różnice między kompresją stratną i bezstratną tekstów, obrazów, dźwięków, filmów;

Kształtowane kompetencje kluczowe:

- kompetencje cyfrowe;
- kompetencje osobiste, społeczne i w zakresie umiejętności uczenia się;
- kompetencje matematyczne oraz kompetencje w zakresie nauk przyrodniczych, technologii i inżynierii.

Cele operacyjne (językiem ucznia):

- Dowiesz się, na czym polega problem grupowania.
- Zapoznasz się z koncepcją wykorzystania algorytmów genetycznych w grupowaniu.
- Poznasz przykładowe problemy grupowania.

Strategie nauczania:

- konstruktywizm;
- konektywizm.

Metody i techniki nauczania:

- dyskusja;
- rozmowa nauczająca z wykorzystaniem multimediu i ćwiczeń interaktywnych;
- burza mózgów;
- mapa myśli.

Formy pracy:

- praca indywidualna;
- praca w parach;
- praca w grupach;
- praca całego zespołu klasowego.

Środki dydaktyczne:

- komputery z głośnikami, słuchawkami i dostępem do internetu;
- zasoby multimedialne zawarte w e-materiale;
- tablica interaktywna/tablica, pisak/kreda.

Przebieg lekcji

Przed lekcją:

1. **Przygotowanie do zajęć.** Nauczyciel loguje się na platformie i udostępnia e-materiał: „Algorytmy genetyczne – przeszukiwanie dużych przestrzeni rozwiązań”. Nauczyciel prosi uczniów o zapoznanie się z treściami w sekcji „Przeczytaj”.

Faza wstępna:

1. Uczniowie metodą burzy mózgów przypominają sobie najważniejsze informacje dotyczące algorytmów genetycznych. Chętna lub wybrana osoba zapisuje je na tablicy za pomocą mapy myśli.
2. Nauczyciel wyświetla uczniom temat zajęć oraz cele. Prosi, by na ich podstawie uczniowie sformułowali kryteria sukcesu.

Faza realizacyjna:

1. **Praca z tekstem.** Uczniowie analizują treści z sekcji „Przeczytaj” wyświetlone na tablicy.
2. **Praca z multimediami.** Uczniowie wspólnie słuchają treści zawartych w sekcji „Audiobook”. Nauczyciel inicjuje dyskusję na temat tego, jak może przebiec rozwój algorytmów genetycznych w następnych latach.
3. **Ćwiczenie umiejętności.** Uczniowie wykonują indywidualnie ćwiczenia nr 1-7, a następnie dzielą się wynikami swojej pracy z kolegą lub koleżanką.

Faza podsumowująca:

1. Nauczyciel wyświetla na tablicy temat lekcji i cele zawarte w sekcji „Wprowadzenie”. W kontekście ich realizacji podsumowuje przebieg zajęć, a także wskazuje mocne i słabe strony pracy uczniów.

Praca domowa:

1. Uczniowie wykonują ćwiczenie 8 z sekcji „Sprawdź się”.

Wskazówki metodyczne:

- Treści w sekcji „Audiobook” można wykorzystać na lekcji jako podsumowanie i utrwalenie wiedzy uczniów.