

Streszczenie

Jan Sołtysik

16 czerwca 2020

Celem pracy mojej pracy licencjackiej pt. "Zastosowanie algorytmu genetycznego do optymalizacji struktury sieci neuronowej" jest zaimplementowanie **algorytmu genetycznego (AG)** umożliwiającego znalezienie optymalnej struktury **sztucznej sieci neuronowej (SSN)**. Sztuczne sieci neuronowe i uczenie głębokie są ostatnimi laty tematem numer jeden w świecie informatyki. Pomimo tego że nie jest to nowa dyscyplina, jej początki sięgają roku 1943. Jednak dopiero rozwój technologii oraz łatwy dostęp do ogromnych zbiorów danych pozwoliły pokazać SSN prawdziwy potencjał. Budowanie modeli SSN jest jednak procesem skomplikowanym, wymaga znajomości analizy matematycznej, statystyki, algebry itd. które potrzebne są do zrozumienia algorytmów uczenia maszynowego. Współczesne biblioteki implementujące SSN znacznie upraszczają ich używanie, lecz wciąż to użytkownik musi wybrać multum hiperparametrów sieci takich jak: ilość i rodzaje warstw, funkcje aktywacyjne itd. Aby zautomatyzować proces wyboru hiperparametrów zmodyfikuję AG aby przystosować go do pracy z SSN.

Do zrealizowania powyższego celu wykorzystałem język programowania **python** wraz z popularnymi bibliotekami:

- **numpy** - implementacja operacji numerycznych
- **tensorflow** - umożliwia budowanie SSN warstwami odpowiednio specyfikując hiperparametry sieci. Proces uczenia SSN został też uproszczony w wersji 2.0 biblioteki ponieważ sprowadza się do wywołania pojedynczej metody **fit**.

Struktura pracy jest następująca. We wstępie opisuję cel pracy, jej inspirację oraz krótko opisuję co jest zawarte w każdym z rozdziałów. W następnym rozdziale zawarte są potrzebne definicje i wzory potrzebne do zrozumienia elementarnego AG oraz sposobu działania SSN. Rozdział trzeci zawiera modyfikację AG która umożliwia znalezienie optymalnej struktury SSN. W rozdziale czwartym zawarte są szczegóły dotyczące implementacji opisanego algorytmu. Ostatni rozdział zawiera optymalne sieci znalezione przez stworzony w ramach pracy algorytm. Algorytm został przetestowany na dwóch zbiorach danych zarówno dla problemu klasyfikacji jak i regresji: **Fashion-Mnist** - zadanie rozpoznania kategorii ubrania podając jego zdjęcie, oraz **Boston Housing** - przewidywanie ceny domu w zależności od wartości zadanych dla jego parametrów.