



AKADEMIA GÓRNICZO-HUTNICZA IM. STANISŁAWA STASZICA W KRAKOWIE

WYDZIAŁ INFORMATYKI, ELEKTRONIKI I TELEKOMUNIKACJI

Sieć rozpoznająca monety

Autor:	<i>Arkadiusz Kaczmarczyk</i>
Kierunek studiów:	<i>Elektronika i Telekomunikacja</i>
Przedmiot:	<i>Systemy na bazie algorytmów sztucznej inteligencji</i>

Kraków, 2021

1. Wstęp

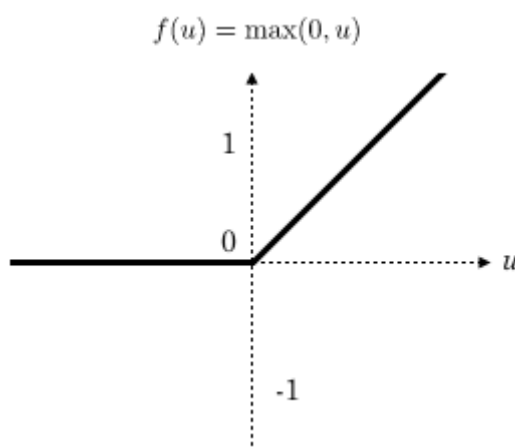
Celem niniejszego projektu było stworzenie algorytmu bazującego na sztucznej inteligencji, zdolnego do rozpoznawania monety znajdującej się na zdjęciu. Jako, iż liczba monet jest bardzo duża, autor pracy ograniczył się do rozpoznawania tylko polskich monet o nominale: 1 grosz, 2 grosze, 5 groszy, 10 groszy, 20 groszy, 50 groszy, 1 złoty, 2 złote i 5 złote. Sieć zdolna jest do rozpoznawania tylko monet obiegowych (bez monet kolekcjonerskich, okolicznościowych i wycofanych z oficjalnego obiegu).

2. Realizacja

W celu rozpoznawania monet autor zaprojektował sieć o strukturze:

LINEAR -> RELU -> LINEAR -> RELU -> LINEAR -> SOFTMAX.

Warstwa pierwsza i druga składa się z dwóch podwarstw: warstwy LINEAR – liniowa funkcja postaci $y = ax + b$, gdzie „a” i „b” to obliczane współczynniki sieci, a „y” i „x” to dane dostarczane do sieci. Następnie dane poddawane są działaniu warstwy RELU, a więc nieliniowej funkcji aktywacji **ReLU (Rectified Linear Unit)**. Jej wygląd przedstawiony został na rysunku poniżej.



Funkcja ReLU

Warstwa trzecia również składa się z dwóch podwarstw: warstwy liniowej LINEAR oraz specjalnej warstwy SOFTMAX. Jest to funkcja która na podstawie wejściowego wektora danych wylicza prawdopodobieństwo dla kilku węzłów wyjściowych.

Ze względu na ograniczenia sprzętowe i czas obliczeń, sieć do nauki i rozpoznawania wykorzystuje zdjęcia rozmiaru 64 x 64 piksele. Jest to małe zdjęcie jednakże, jest ono w formacie RGB co daje $64 \times 64 \times 3 = 12288$ danych wejściowych. Z racji tak dużej liczby danych zastosowana została stosunkowo mała liczba węzłów w sieci – 1 warstwa zawiera 40 węzłów, 2 warstwa zawiera 20 węzłów, a trzecia, końcowa warstwa, zawiera 10 węzłów

wyściowych (9 rodzajów monet oraz kategoria „inne”). Do zaprojektowania sieci wykorzystany został język *Python* oraz framework *tensorflow*.

3. Wyniki

Dzięki zastosowaniu opisanej powyżej sieci, autorowi udało się uzyskać następujące wyniki:

```
Cost after epoch 0: 2.359269
Cost after epoch 100: 1.458900
Cost after epoch 200: 0.991891
Cost after epoch 300: 0.675036
Cost after epoch 400: 0.454649
Cost after epoch 500: 0.255536
Cost after epoch 600: 0.135462
Cost after epoch 700: 0.069518
Cost after epoch 800: 0.034447
Cost after epoch 900: 0.017537
Cost after epoch 1000: 0.009589
Cost after epoch 1100: 0.006593
Cost after epoch 1200: 0.004426
Cost after epoch 1300: 0.003655
Cost after epoch 1400: 0.002629
Parameters have been trained!
Train Accuracy: 1.0
Test Accuracy: 0.5769231
```

Dane treningowe zawierały 1771 zdjęć z 10 kategorii (9 rodzajów monet oraz kategoria „inne”). Sieć wytrenowana została w 1500 cyklach treningowych. Jeden cykl treningowy, polega na załadowaniu wszystkich zdjęć do sieci i wykonania obliczeń, a następnie obliczeniu funkcji błędu i odpowiedniej zmiany parametrów, w celu zminimalizowania funkcji błędu. Jak widzimy zgodnie z założeniami funkcja „cost function” maleje wraz z liczbą cykli treningowych. Wytrenowana sieć uzyskuje 100% trafności podczas treningu i 57,7% skuteczności podczas przeprowadzonego testu. Test przeprowadzony został na 130 zdjęciach.

4. Modyfikacje i ulepszenia

Stworzona przez autora sieć nie jest idealna i można w niej zastosować wiele modyfikacji polepszających jej możliwości i zastosowania.

Pierwszym ważnym czynnikiem polepszającym zdolności algorytmu jest zgromadzenie większej bazy danych zawierającej zdjęcia monet. Autorowi udało się zgromadzić tylko 1771 zdjęć treningowych i 130 zdjęć testowych, co jak powszechnie wiadomo jest liczbą bardzo małą. Jednakże zgromadzenie większej ilości danych jest mozolne i czasochłonne.

Jako że autor dysponuje ograniczonymi zasobami sprzętowymi, program przetwarza tylko zdjęcia o rozdzielczości 64 x 64 piksele. Jak widać istnieje tu dużo miejsca do poprawy (im

większe zdjęcie tym więcej widać detali). Wraz z zastosowaniem lepszego sprzętu można stworzyć bardziej skomplikowany algorytm – sieć zawierającą więcej warstw i więcej węzłów.

Ostatnią możliwą dużą modyfikacją jest rozszerzenie zestawu rozpoznawanych monet. Oprócz dodania monet z innych państw, można rozpoznawać także monety kolekcjonerskie i nieobiegowe.