

# Rozpoznawanie języka tekstu na podstawie częstotliwości liter

Krzysztof Kutt, Michał Nowak

**Abstrakt—TODO: Tutaj leci abstrakt**

**Słowa kluczowe—sieć neuronowa, język, rozpoznawanie.**

## I. WPROWADZENIE

**P**RZEDSTAWIONA w tym artykule aplikacja powstała w ramach przedmiotu Sztuczne Sieci Neuronowe, semestr letni 2011/2012, prowadzący: mgr inż. Tomasz Orzechowski.

### A. Sieci neuronowe

Sztuczne Sieci Neuronowe to modele matematyczne, stworzone na wzór naturalnych połączeń neuronów znajdujących się w ludzkim mózgu. Każdy sztuczny neuron zbiera sygnały wejścia, dokonuje na nich odpowiedniej transformacji i zwraca odpowiednią wartość na wyjściu. Sztuczna sieć neuronowa może składać się z tysięcy takich elementarnych neuronów. W zależności od budowy sieci oraz od wykorzystywanej przez neurony funkcji aktywacji, wyróżniamy różne rodzaje sieci. Tutaj można jeszcze skorzystać z książki Tadeusiewicza [1], strona 12.

### B. Rodzaje sieci

Skorzystać z książki Tadeusiewicza [1] - główne rodzaje sieci to rozdziały książki - napisać zdanie-dwa o każdym z nich

### C. Zastosowania sieci

Wiki i książka tadeusiewicza.

### D. Rozpoznawanie języka

Coś ogólnie o naszym problemie. Do czego to może być potrzebne, kto już walczył z tym problemem i jak (np google translator ma opcję rozpoznawania języka). Przedstawione opcje to nie musi być tylko częstotliwość liter - pewnie są jakieś inne (choćby sprawdzanie, czy występują znaki specjalne danego języka :P)

### E. Określenie języka przy pomocy sieci neuronowej

Czyli wprowadzenie teoretyczne do naszego projektu. Uznaliśmy, że ten problem można rozwiązać przy pomocy sieci neuronowej. Uznaliśmy, że fajnie będzie to zrobić przy pomocy prostej sieci, złożonej z trzech warstw, z neuronami z aktywacją sigmoidalną (i dlaczego akurat taką :P)

K. Kutt i M. Nowak, Katedra Automatyki, Wydział Elektrotechniki, Automatyki, Informatyki i Elektroniki, Akademia Górniczo-Hutnicza, Kraków, Polska, e-mail: kkutt@student.agh.edu.pl, mnowak@student.agh.edu.pl

## II. IMPLEMENTACJA

Tutaj idzie sekcja o tym jak podzieliliśmy na moduły, jaką sieć wybraliśmy, jaką bibliotekę (i dlaczego) Do środkowej warstwy wzięliśmy 10 neuronów - tak z dupy trochę - może to potestować? (to może być też w testach w sumie)

## III. TESTY

Przeprowadzone testy. Wyniki. Informacje o częstotliwości liter, ale też o skuteczności sieci. Wykresy? Co nam wyszło, dla których języków działa, które okazały się podobne, etc

## IV. WNIOSKI

Projekt jest fajny, bądź nie i jak go można dalej rozwijać, albo gdzie go można zastosować

## LITERATURA CYTOWANA

- [1] R. Tadeusiewicz, *Sieci neuronowe*, wyd. 2, Warszawa, Polska: Akademicka Oficyna Wydawnicza RM, 1993.
- [2] H. Kopka and P. W. Daly, *A Guide to L<sup>A</sup>T<sub>E</sub>X*, 3rd ed. Harlow, England: Addison-Wesley, 1999.