

POLITECHNIKA WARSZAWSKA
WYDZIAŁ MATEMATYKI I NAUK INFORMACYJNYCH



SIECI NEURONOWE

**Rozpoznawanie i klasyfikacja pisanych
cyfr przy użyciu modeli
matematycznych**

Autorzy:

Anna ZAWADZKA
Piotr WASZKIEWICZ

8 listopada 2016

1 Opis problemu badawczego

Problem badawczy przedstawiony na stronie <https://www.kaggle.com/c/digit-recognizer> polega na rozpoznawaniu i klasyfikacji ręcznie pisanych cyfr poprzez przetwarzanie i analizę obrazów przedstawiających odpowiednie symbole. Zbiory danych zostaną zaczerpnięte z publicznej bazy danych MNIST[1].

00112233445566778899
00112233445566778899

2 Planowane do wykorzystania techniki

Projekt zakłada realizację zadania poprzez zbadanie różnych metod, ze szczególnym uwzględnieniem różnych modeli sieci neuronowych. Na chwilę obecną planowana jest konstrukcja sieci klasyfikującej typu *feed-forward* zgodnie z opisem na stronie biblioteki PyBrain[8]. Wykorzystane zostaną również jedne z najpopularniejszych obecnie klasyfikatorów: maszyny wektorów podpierających (SVM)[3], Lasy Losowe[2], kNN[4] oraz model regresji wielomianowej.

3 Cel badań

Celem badań jest porównanie jakości klasyfikacji dla różnych modeli matematycznych i wskazanie najskuteczniejszego klasyfikatora pod względem czasu uczenia, wydajności i jakości udzielanych odpowiedzi. Oprócz tego badania mają na celu rozszerzenie istniejącego wektora cech o nowe, unikalne wartości które polepszą jakość klasyfikacji. Przykładem takich cech może być liczba przecięć w napisanym symbolu, liczba zakończeń lub procent powierzchni zajmowanej przez narysowany symbol. W trakcie obliczeń podjęta zostanie próba odrzucenia tych cech które przeszkadzają lub pogarszają działanie modeli. Przeprowadzone badania obejmą również wybór optymalnych parametrów dla poszczególnych klasyfikatorów metodą GridSearch[7].

4 Opis danych

Zbiory danych treningowych oraz testowych pochodzą z publicznej bazy danych MNIST[1]. Każdy element ze zbioru treningowego jest opisany 785 wartościami. Pierwsza liczba określa zakodowaną cyfrę (wartość z przedziału $[0, 9]$), kolejne 784 wartości są z przedziału $[0, 255]$ i opisują kolory pikseli zeskanowanej cyfry w skali szarości dla obrazka o wymiarach 28x28 pikseli. Zbiór testowy w przeciwieństwie do treningowego nie zawiera informacji o reprezentowanej klasie. Zbiór treningowy i testowy zawierają odpowiednio 60,000 i 10,000 elementów.

5 Sposób weryfikacji rezultatów

Podczas ewaluacji otrzymywanych rozwiązań minimalizowana będzie funkcja błędu opisana wzorem

$$f(M, d) = e(M, d) + t(M, d)$$

gdzie M oznacza model, d zbiór testowy, $e()$ współczynnik *Error rate*, czyli miarę określającą stosunek źle zaklasyfikowanych elementów do wszystkich obiektów w zbiorze, oraz $t()$ funkcję czasu liczoną jako liczbę sekund potrzebną na realizację obliczeń.

Literatura

- [1] LeCun, Y., Cortes, C., and Burges, C., *The MNIST database of handwritten digits*, in: <http://yann.lecun.com/exdb/mnist>.
- [2] Breiman, L., *Random Forests*. Machine Learning 45 (1), 2001
- [3] Cortes, C., Vapnik, V., *Support-vector networks*. Machine Learning 20 (3), 1995.
- [4] Altman N. S., *An introduction to kernel and nearest-neighbor nonparametric regression*. The American Statistician 46 (3), 1992.
- [5] Scholkopf, B., Williamson, R., Smola, A., Shawe-Taylor, J., Platt, J., *Support Vector Method for Novelty Detection*, Advances in Neural Information Processing Systems 12, 1992.
- [6] Wang, Y., Casasent, D., *A Support Vector Hierarchical Method for multi-class classification and rejection*, Proc. of Int. Joint Conf. on Neural Networks, 2009.
- [7] http://scikit-learn.org/stable/modules/grid_search.html.
- [8] <http://pybrain.org/docs/tutorial/fnn.html>