

**WYDZIAŁ ZARZĄDZANIA**

Katedra Informatyki Biznesowej i Inżynierii Zarządzania

Projekt dyplomowy

*Aplikacja mobilna do nauki matematyki dla szkoły podstawowej z wykorzystaniem machine learningu do rozpoznawania odpowiedzi pisanych ręcznie*

*A mobile math learning app for elementary school using machine learning to recognize handwritten answers*

Autor: *Monika Helena Etrych*

Kierunek studiów: Informatyka i Ekonometria

Opiekun pracy: *dr Beata Basiura*

Kraków, 2023 r.

Spis treści

[Wstęp 3](#_Toc123635373)

[1. Rozpoznawanie cyfr i ich interpretacja 4](#_Toc123635374)

[1.1. Interpretacja obrazów przez komputer 4](#_Toc123635375)

[1.2. Metody uczenia maszynowego do rozpoznawania cyfr 4](#_Toc123635376)

[1.3. Obróbka obrazów z cyframi 4](#_Toc123635377)

[1.4. Weryfikacja modelu do rozpoznawania cyfr 4](#_Toc123635378)

[2. Model aplikacji do nauki uczenia matematyki 4](#_Toc123635379)

[2.1. Założenia modelu aplikacji 4](#_Toc123635380)

[2.2. Wybrane technologie zastosowane do budowy aplikacji 4](#_Toc123635381)

[2.3. Wybrane metody uczenia maszynowego zastosowane w aplikacji 4](#_Toc123635382)

[2.4. Korzystanie z modelu ML poprzez API HTTP 4](#_Toc123635383)

[3. Opis implementacji i działania aplikacji 4](#_Toc123635384)

[3.1. Ekran startowy i podstawowe funkcjonalności 4](#_Toc123635385)

[3.2. Schematy, opisy poszczególnych funkcjonalności, opis kodów 4](#_Toc123635386)

[3.3. Przykłady zastosowania aplikacji w konkretnych zadaniach 4](#_Toc123635387)

[Zakończenie 5](#_Toc123635388)

[Bibliografia 6](#_Toc123635389)

[Spis tabel, wykresów i rysunków 7](#_Toc123635390)

[Aneks (kod, tabelka z danymi, treść ankiety) 8](#_Toc123635391)

# Wstęp

1-2 stron, cel pracy, ważny problem, może być hipoteza badawcza, napisać na koniec!

Problem rozpoznawania cyfr – problematyczne, bo każdy ma inny charakter, pochyłość, duże małe. Różnice kulturowe. Pisanie cyfry od góry do dołu lub od dołu.

# Rozpoznawanie cyfr i ich interpretacja

## Interpretacja obrazów przez komputer

Obrazy reprezentowane przez komputer składają z pikseli - kwadratów, które posiadają kilka wartości określających kolor. Wyróżnia się 4 podstawowe reprezentacje: binarną, monochromatyczną (ang. grayscale), RGB i RGBA.

Reprezentacja binarna oraz monochromatyczna składają się tylko z jednej wartości dla danego pixela. W przypadku reprezentacji binarnej wartość jest zerojedynkowa, gdzie 0 oznacza, że pixel ma kolor czarny, a 1 to kolor biały. W pozostałych reprezentacjach wartość jest z zakresu od 0 do 255. W reprezentacji monochromatycznej 0 oznacza kolor czarny, 255 - biały, natomiast wartości pośrednie, to odcienie szarości.

RGB zawiera trzy wartości określające kolor: czerwony, zielony i niebieski. RGBA posiada dodatkowo jeszcze jedną wartość określającą poziom przezroczystości. [1]

Co jeszcze:

CMY i CMYK?

Grafika rastowa i wektorowa?

Formaty?

TRYBY

Może obrazek obrazka w różnych stylach?

Wyróżnia się 3 podstawowe tryby koloru:

* tryb [czarno-biały](https://pl.wikipedia.org/wiki/Achromatyzm_(grafika)) – występują w nim wyłącznie białe i czarne [piksele](https://pl.wikipedia.org/wiki/Piksel),
* tryb [monochromatyczny](https://pl.wikipedia.org/wiki/Monochromatyzm_(grafika)) (tzw. skala szarości) – oprócz koloru czarnego i białego występuje cała gama jasności pośrednich (szarości),
* tryb [kolorowy](https://pl.wikipedia.org/wiki/Barwa) – piksele mają przyporządkowane kolory z określonego modelu [przestrzeni barw](https://pl.wikipedia.org/wiki/Przestrze%C5%84_barw), np. [RGB](https://pl.wikipedia.org/wiki/RGB), [CMYK](https://pl.wikipedia.org/wiki/CMYK) i inne.

## Metody uczenia maszynowego do rozpoznawania cyfr

Metody uczenia maszynowego są coraz częściej wykorzystywane do rozpoznawania cyfr w różnych dziedzinach, takich jak bankowość, czy usługi pocztowe. Uczenie maszynowe pozwala na automatyczne i precyzyjne rozpoznawanie cyfr. Wiele papierowych formularzy w sposób szybki i dokładny może zostać przeniesiona do komputera. Natomiast rozpoznawanie kodów pocztowych znacząco przyspiesza proces sortowania listów i przesyłek [3].

Do rozpoznawania cyfr wykorzystuje się algorytmy uczenia nadzorowanego. W tym podejściu, dane uczące zawierają cyfry wraz z odpowiadającymi im etykietami, czyli poprawnymi wartościami cyfr. Algorytm uczony na takich danych jest w stanie nauczyć się rozpoznawać cyfry na podstawie cech wizualnych, takich jak kształt i proporcje. [1][5]

Wśród popularnych algorytmów uczenia maszynowego stosowanych do rozpoznawania cyfr znajdują się między innymi konwolucyjne sieci neuronowe (CNN), wielowarstwowe sieci neuronowe (MLP) oraz algorytm maszyn wektorów nośnych (SVM). Zastosowanie tych algorytmów wraz z odpowiednio przygotowanymi danymi uczącymi, pozwala na uzyskanie skuteczności rozpoznawania cyfr na poziomie nawet 99% [4][5].

Algorytm maszyn wektorów nośnych (SVM) polega na maksymalizacji odległości pomiędzy hiperpłaszczyzną, a marginesem przynależności do poszczególnych klas, poprzez zmianę położenia hiperpłaszczyzny. Istnieje wiele możliwych hiperpłaszczyzn, a jej kształt zależy od złożoności problemu - liczby klas do zaklasyfikowania. Kiedy liczba klas wynosi 2, to hiperpłaszczyzna jest prostą, a dla 3 - płaszczyzną 2D itd. [6]

Wymienione metody zostały porównane pod kątem czasu wykonania oraz dokładności dla danych treningowych i testowych. Najwyższy wynik dokładności dla danych treningowych osiągnął SVM - 99.98%, a dla testowych CNN - 99.31% [5].

Co jeszcze:

- jeszcze kilka odwołań do badań

- może opisać inne metody jak działają

- CNN na pewno opisać później w 2 rozdziale

1. Support Vector Machine – SVM - Maszyna wektorów nośnych
2. Gaussian Naive Bayes – GNB -
3. Decision Trees
4. Random Forest
5. K Nearest Neighbours - KNN
6. Stochastic Gradient Descent - SGD
7. Convolutional Neural Network
8. CNN, MLP

## Obróbka obrazów z cyframi

W procesie rozpoznawania cyfr przez model, kluczowe znaczenie ma odpowiednie przekształcenie obrazów przed ich przekazaniem do modelu. Wstępne przetwarzanie obrazów ma na celu przygotowanie danych do dalszej analizy. Ponadto zmniejsza czas potrzebny do wytrenowania modelu i zwiększa szybkość otrzymania wyniku. Na przykład kiedy wejściowy obraz jest stosunkowo duży, zmniejszenie rozmiaru znacząco zmniejsza czas trenowania, bez znacznej utraty dokładności modelu.

Sieci neuronowe wymagają, aby dane wejściowe miały ten sam rozmiar, ponieważ sieci te korzystają z macierzy wag, które mają ustaloną liczbę kolumn i wierszy. Przetwarzanie obrazów przez ustawienie stałych wymiarów jest jednym ze sposobów na dostosowanie rozmiaru obrazów do wymogów sieci neuronowej.

Zmiany rozmiaru obrazu można dokonać na dwa sposoby. Pierwszy z nich polega na zmniejszeniu obrazu przez wycinanie obszaru, który zawiera obraz, a drugi sposób polega na zwiększeniu rozmiaru poprzez dodanie pustych pikseli.

Model głębokiego uczenia może ulec przeuczeniu, czyli zbytniemu dopasowaniu do danych trenujących, co skutkuje słabymi wynikami na nowych danych. Aby zapobiec temu problemowi, konieczna jest odpowiednia liczba danych trenujących. Jednym ze sposobów na zwiększenie liczby danych jest zastosowanie augmentacji, czyli przekształcanie istniejących obrazów poprzez np. losowe zmiany rotacji lub jasności. Augmentacja jest łatwym sposobem na zwiększenie różnorodności danych trenujących i poprawę jakości modelu.

Te same zabiegi przetwarzania danych mogą być stosowane zarówno w ramach wstępnego przetwarzania danych, jak i augmentacji. Jednakże, wstępne przetwarzanie danych jest stosowane zarówno na danych treningowych, jak i testowych, podczas gdy augmentacja jest stosowana wyłącznie na danych treningowych po wstępnym przetworzeniu. [2]

Co jeszcze:

~~Do najczęstszych przekształceń należą: centrowanie, obcinanie brzegów, zamiana trybu kolorów na greyscale.~~

Preprocessing: Resize, Orientation

Augmentation: Grayscale, Random Flips, Random Rotations, Random Brightness and Exposure, Random Noise

Co użyte w aplikacji

## Weryfikacja modelu do rozpoznawania cyfr

Interpretacja wyników, doczytać o tych różnych wskaźnikach, accuracy, i te ile dobrych dobrze klasyfikuje ile złych źle itp.

Correctly Classified

Incorrectly classified

Kappa statistic -co to ? Kappa statistics value ranges from 0 to 1.  
Value 0 means totally disagreement and 1  
means full agreement.

Mean absolute error

Root Mean squared error

Relative absolute error

Root relative absolute error

dokładność, czułość, specyficzność oraz krzywe ROC i AUC.

# Model aplikacji do nauki uczenia matematyki

## Założenia modelu aplikacji

## Wybrane technologie zastosowane do budowy aplikacji

## Wybrane metody uczenia maszynowego zastosowane w aplikacji

## Korzystanie z modelu ML poprzez API HTTP

# 

# Opis implementacji i działania aplikacji

## Ekran startowy i podstawowe funkcjonalności

## Schematy, opisy poszczególnych funkcjonalności, opis kodów

## Przykłady zastosowania aplikacji w konkretnych zadaniach

# Zakończenie

Wyniki, rezultaty, udowodnienie tezy, co na przyszłość rozwinąć

# Bibliografia

[1] [Candocia](https://www.linkedin.com/in/max-candocia-a270b574/) M.*, A Simple Explanation of how Computers Recognize Images,* [Analysis for Many Audiences,](https://maxcandocia.com/)  2016, url: <https://maxcandocia.com/article/2016/Apr/06/how-computers-recognize-images/>, dostęp: 23.02.2023r.

[2] Nelson J*., What is Image Preprocessing and Augmentation?,* Roboflow, 2020, url: <https://blog.roboflow.com/why-preprocess-augment/>, dostęp: 28.02.2023r.

[3] S. M. Shamim, Mohammad Badrul Alam Miah, Angona Sarker, Masud Rana, Abdullah Al Jobair, 2018, *Handwritten Digit Recognition Using Machine Learning,* Algorithms Indonesian Journal of Science & Technology (1) vol. 3, str.: 29-39 ISSN 2528-1410, ISSN 2527-8045

# [4] Javed M., *The Best Machine Learning Algorithm for Handwritten Digits Recognition,* 2020,

url:<https://towardsdatascience.com/the-best-machine-learning-algorithm-for-handwritten-digits-recognition-2c6089ad8f09>, dostęp: 1.03.2023

[5] Dixit R., Kushwah R., Pashine S., *Handwritten Digit Recognition using Machine and  
Deep Learning Algorithms,* Artykuł Naukowy 2021, url: <https://arxiv.org/pdf/2106.12614.pdf>, dostęp: 1.03.2023

# [6] Gandhi R., *Support Vector Machine — Introduction to Machine Learning Algorithms*, 2018, url: <https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47>, dostęp: 1.03.2023

GNB - <https://www.youtube.com/watch?v=H3EjCKtlVog>

Decision trees - <https://towardsdatascience.com/decision-trees-in-machine-learning-641b9c4e8052>

Random forest - <https://towardsdatascience.com/understanding-random-forest-58381e0602d2>

KNN - <https://medium.com/swlh/k-nearest-neighbor-ca2593d7a3c4>

<https://learn.g2.com/k-nearest-neighbor>

SGD - <https://towardsdatascience.com/gradient-descent-811efcc9f1d5>

# Spis tabel, wykresów i rysunków

# Aneks (kod, tabelka z danymi, treść ankiety)

(NOTATKI)

## Spis ludności w USA

## Rozpoznawanie japońskich znaczków xD

Podobnym zadaniem jest rozpoznawanie przez algorytm znaków specyficznych dla różnych języków np. języka chińskiego czy arabskiego.

## Akcelerometry

Data augmentation – ciekawe :D