1. Wstęp
2. Podstawy teoretyczne
3. Projekt?
4. Implementacja?
5. Badania
6. Podsumowanie

1.

wstęp, motywacja, cel i zakres

2.

background teoretyczny, teoria

problem klas mri

3.

projekt sys - opis, wymagania, jak wyglądają dane, wym funkcjonalne/nie funkcjonalne

aspekt systemowy

opis podejść, po kolei wszystkich ML i CNN, motywacja wyboru metody, szybkość, opis wycięcia VGG

wyniki tesowania, metryki

opis technologii - po 2 zdania biblioteki, opis zapisu modeli implementacji

4. podusmowanie

ROZDZIAŁ 2 – ZAGADNIENIA TEORETYCZNE

# Uczenie maszynowe

Dziedzina uczenia maszynowego, która powstała w latach 60-tych XX wieku, obecnie odgrywa bardzo ważną rolę w gwałtownie rozwijającym się sektorze analityki dużych zbiorów danych. Algorytmy uczenia wykorzystywane są przy wielu zadaniach przewidywania lub klasyfikowania danych.

Terminem uczenia maszynowego możemy określić gałąź informatyki oraz sztucznej inteligencji, zajmującej się utworzeniem modelu, który jest w stanie samodzielnie, pod wpływem otrzymywanych danych, uczyć się, czyli maksymalizować swoją skuteczność w wykonywaniu konkretnego zadania.

Istnieją różne rodzaje modeli uczenia maszynowego różniące się sposobem uczenia. Główną role odgrywa wpływ człowieka na dostarczane dane. Dwoma podstawowymi typami są:

* Uczenie nadzorowane – dane otrzymywane przez algorytm uczący posiadają etykiety, które są wykorzystywane w procesie uczenia, a także do określenia, przy pomocy różnych metryk, jak dobrze działa ostatecznie stworzony model
* Uczenie nienadzorowane – dane otrzymywane przez algorytm uczący nie posiadają etykiet. Zadaniem algorytmu jest zidentyfikowanie relacji i wzorców w przekazanych danych bez pomocy człowieka

Proces nadzorowanego uczenia, na którym skupia się ta praca, składa się z trzech głównych, powtarzających się części:

* Procesu decyzyjnego – Algorytm użyty w procesie uczenia na podstawie otrzymanych danych dokonuje predykcji.
* Funkcji błędu – Funkcja ta służy do oceny predykcji dokonanej w procesie decyzyjnym z prawdziwą wartością dostarczaną z danymi. Jest to miara dokładności modelu.
* Procesu optymalizacji – Metoda, której celem jest zaktualizowanie działania algorytmu używanego w procesie decyzyjnym, biorąca pod uwagę wartość funkcji błędu, w taki sposób, aby zminimalizować błąd w następnej iteracji.

# Głębokie uczenie maszynowe

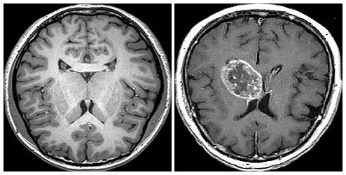
Terminy „Uczenie maszynowe” i „Głębokie uczenie” są często używane zamiennie, jednak oznaczają one inne podejścia. Samo głębokie uczenie jest poddziedziną uczenia maszynowego, a główną różnicą tych metod jest sposób uczenia się algorytmu. Głębokie uczenie maszynowe wykorzystuje wielowarstwowe sieci neuronowe, które cechują się bardzo duża elastycznością, a także umiejętnością wykorzystywania danych, które nie były wstępnie przetwarzane. Cel głębokiego uczenia jest taki sam jak uczenia maszynowego, czyli stworzenie sieci neuronowej, która dla otrzymywanych danych maksymalizuje swoją skuteczność

# Problem klasyfikacji

Jednym z podstawowych zadań zarówno dla uczenia maszynowego jak i uczenia głębokiego jest problem klasyfikacji danych. Polega ono na przyporządkowaniu danych wejściowych do jednej klasy, pochodzącej ze zbioru predefiniowanych klas. Głównym celem samej klasyfikacji jest stworzenie formalnego modelu, nazywanego klasyfikatorem, którego zadaniem jest klasyfikowanie nowo otrzymanych danych i przydzielenie ich do konkretnej klasy.

Zadanie klasyfikacji posiada wiele zastosowań np.: klasyfikowanie spamu e-mail, klasyfikowanie opinii na podstawie recenzji, wykrywanie obiektów na zdjęciach. Jednym z popularnych sektorów, w których wykorzystywane jest to zadanie jest sektor medyczny. Dane z systemów medycznych mogą zostać użyte do stworzenia klasyfikatora zdolnego do bardzo szybkiego klasyfikowania schorzeń u dotychczasowych jak i nowych pacjentów.

W tej pracy skupiono się tylko na problemie klasyfikacji binarnej, charakteryzującej się występowaniem jedynie dwóch klas. Dane mogą zostać zaklasyfikowane jako klasa pozytywna (1) lub klasa negatywna (0), tak jak zostało to zaprezentowane na Rysunku 1.

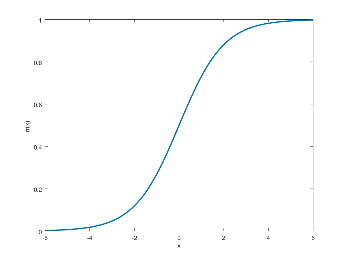


Rys. 1 Przykład klasyfikacji binarnej na podstawie skanów MRI. Po lewej stronie znajduje się skan pacjenta bez guza mózgu - klasa negatywna (0). Po prawej stronie skan pacjenta z guzem mózgu - klasa pozytywna (1).

# Regresja logistyczna

Znanym od dawna modelem wykorzystywanym w klasyfikacji jest regresja logistyczna (ang. *logistic regression*). Regresja logistyczna jest przykładem modelu zaliczanego do grupy uczenia nadzorowanego. Podejście to w swojej podstawowej wersji jest wykorzystywane tylko do klasyfikacji binarnej. Regresja logistyczna wykorzystuje funkcję logistyczną opisaną wzorem:

, oraz widoczną na Rysunku 69,



# Liniowa maszyna wektorów nośnych

Nowszym, często używanym modelem w zadaniach klasyfikacji, jest maszyna wektorów nośnych (ang. *support vector machines,* SVM). Zadaniem algorytmu uczenia SVM jest znalezienie takiej hiperpłaszczyzny, która jest w stanie oddzielić od siebie dane podawane w procesie uczenia. Zaletą SVM jest to, że nie wybiera on dowolnej płaszczyzny, a stara się maksymalizować margines między klasami, czyli maksymalizuje dystans pomiędzy hiperpłaszczyzną a najbliższymi przedstawicielami obydwu klas. Większy margines zapewnia lepszą separację klas, czyli zmniejsza generalizację. Po wyuczeniu modelu, nowe dane są klasyfikowane na podstawie ich położenia względem hiperpłaszczyzny. Liniowy SVM charakteryzuje się dużą skutecznością, gdy dane posiadają dużą ilość cech, co jest problemem w klasyfikacji obrazów.

# Analiza głównych składowych

Modele uczone na danych posiadających bardzo dużo wymiarów mogą borykać się z problem przetrenowania, czyli zbyt dużym dostosowaniem się do danych treningowych. Efektem tego jest słabsza jakość predykcji na nowych danych. Techniką, która może służyć do redukcji wymiarów danych jest analiza głównych składowych (ang. *principal component analysis*, PCA). Zastosowanie tego rozwiązania pozwala na pozostawienie cech, które posiadają największą wariancję, czyli wnoszą największą ilość informacji. Do zalet PCA można zaliczyć:

* Łatwość obliczenia – technika PCA opiera się na algebrze liniowej, która nie jest trudna obliczeniowo dla komputera
* Zmniejsza wymiarowość danych – zmniejszenie wymiarowości danych zapobiega przetrenowaniu modelu
* Przyspieszenie uczenia modeli – mniejsza ilość cech wpływa pozytywnie na szybkość uczenia algorytmów uczenia maszynowego

Zmniejszenie ilości cech może być przydatne w procesie uczenia modelu, jednakże mimo wszystko wiąże się z częściową utratą odrzuconych informacji. Dlatego ważne jest balansowanie pomiędzy redukcją wymiarów a zachowaniem informacji w danych.

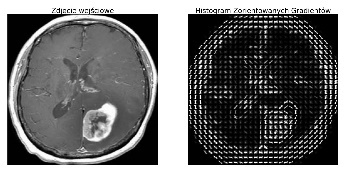
# Histogram zorientowanych gradientów

Zadania uczenia maszynowego opierające się na pracy z obrazami często wymagają wstępnego przetworzenia danych. Jedną z możliwości jest użycie narzędzi do deskrypcji cech, czyli wydobycia z obrazu przydatnych informacji oraz pominięcie nieistotnych z punktu widzenia uczenia maszynowego. Taką techniką jest histogram zorientowanych gradientów (ang. *histogram of oriented gradients*, HOG), który służy do wykrywania krawędzi i rogów na obrazie.

Głównymi punktami w procesie obliczania cech HOG jest:

* Obliczanie gradientu – oblicza się przestrzenne gradienty poziome i pionowe, które później wykorzystuje się do obliczenia wielkości i kąta gradientu.
* Dzielenie obrazu – obraz dzieli się na małe komórki np.: 8x8 pikseli, gdzie następnie w każdej z nich poszczególne piksele głosują na klasę histogramu w zależności od kąta gradientu.
* Deskrypcja cech – w ostatnim etapie sąsiednie komórki są grupowane w bloki, które poddawane są normalizacji, a następnie łączone by stworzyć deskryptor.

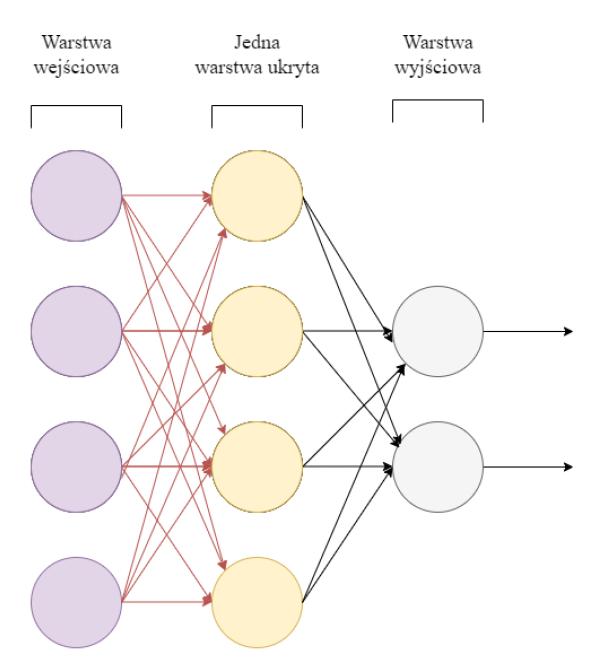
Efekt zastosowania techniki histogramu zorientowanych gradientów przedstawia Rysunek 69.



# Sieć neuronowa

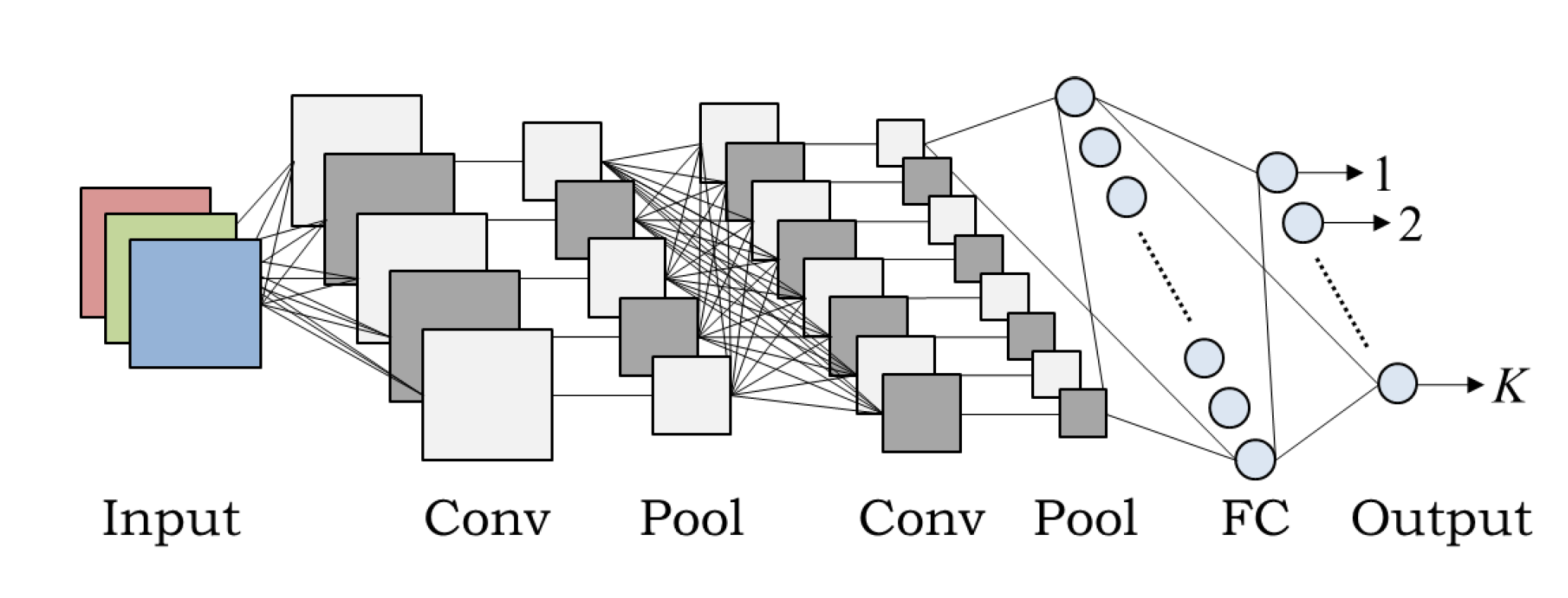
Tworzenie oraz uczenie sieci neuronowych (ang. *Neural Networks*, NN) staje się z roku na rok coraz bardziej popularną metodą uczenia maszynowego. Do ich największych zalet należy ogromna elastyczność, brak wymagania formalizacji rozwiązywanych zadań oraz zdolność do uczenia się na podstawie otrzymywanych danych i możliwość uogólniania zdobywanej wiedzy.

Architektura sieci neuronowych opiera się na neuronach uporządkowanych w warstwy, co zostało zaprezentowane na rysunku 69. Każda sieć neuronowa posiada warstwę wejściową( ang. *input layer*) i warstwę wyjściową (ang. *output layer*), które odpowiadają kolejno za przyjmowanie danych oraz zwracanie predykcji, a także warstwy ukryte (ang. *hidden layers*). Ilość warstw ukrytych mówi o stopniu głębokości sieci. Aby sieć określić głęboką siecią neuronową (ang. *Deep Neural Network*, DNN) powinna posiadać co najmniej cztery ukryte warstwy. Nadzorowane uczenie sieci neuronowej zakłada posiadanie zbioru danych wraz z etykietami, czyli oznaczonymi poprawnie klasami w przypadku zadania klasyfikacji. Sam proces uczenie przebiega w epokach. Pierwszym etapem każdej epoki jest wprowadzanie danych do wejścia sieci, następnie propagowanie ich przez wszystkie warstwy ukryte, a ostatecznie warstwa wyjściowa jako wynik zwraca predykcję modelu. W zadaniu klasyfikacji jest to prawdopodobieństwo przypisania danych wejściowych do jednej z klas. Po dokonaniu predykcji obliczany jest błąd sieci, określony funkcją kosztu. Aby minimalizować funkcję kosztu model wykorzystuje algorytm wstecznej propagacji, który oblicza pochodne cząstkowe funkcji kosztu względem parametrów sieci, a następnie modyfikuje je aby poprawić działanie sieci w następnej epoce.



# Konwolucyjne sieci neuronowe

Zadania związane z przetwarzaniem obrazów są częstym elementem rozwiązywanym przez głębokie uczenie. Powszechnie stosowaną klasą sztucznych sieci neuronowych używaną w takich problemach są konwolucyjne sieci neuronowe (ang. Convolutional Neural Network, CNN), różniące się od standardowych rozwiązań wykorzystywaniem operacji zwanej konwolucją. Dzięki temu cechują się zwiększoną skutecznością oraz szybkością uczenia względem zwykłej gęstej sieci neuronowej.



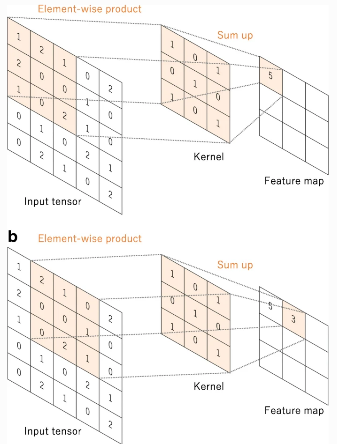
Podobnie do standardowej gęstej sieci neuronowej, CNN składa się z warstwy wejściowej, wyjściowej oraz licznych warstw ukrytych. Warstwy ukryte można podzielić na bloki, które składają się z następujących warstw:

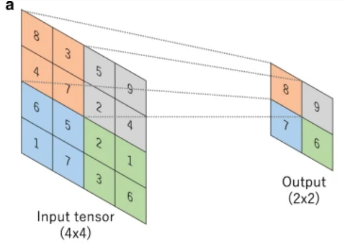
* jedna lub więcej warstw konwolucyjnych (ang. *convolution layers*) – zadaniem każdej z warstw jest zastosowanie operacji konwolucji pomiędzy otrzymanymi danymi z poprzedniej warstwy a jednym z filtrów przypisanych do niej samej. Proces ten został przedstawiony na Rysunku 69. Wagi filtrów są optymalizowane w procesie uczenia. Efektem tej warstwy jest mapa cech (ang. *feature map*), która jest podawana następnej warstwie.
* warstwa łączenia (ang. *pooling layer*) – zadaniem tej warstwy jest stopniowe zmniejszanie wymiarowości oraz uodpornienie na niewielkie zmiany sygnału wejściowego. Zmniejszenie wymiarowości, czyli zmniejszenie ilości uczonych parametrów, niesie ze sobą dodatkowo skrócenie czasu potrzebnego na uczenie i może pomóc z przeuczeniem. Operacja łączenia polega na wykorzystaniu okna przesuwnego o wybranej wielkości (zwykle nieduże np.: 2x2 piksele) oraz wybranym kroku okna i zastosowaniu funkcji maksimum, lub średniej. Proces ten pokazuje rysunek 69.

Każda z warstw konwolucyjnych jest zakończona nieliniową funkcją aktywacji. Obecnie najczęściej stosowaną funkcją jest ReLU (ang. *rectified linear unit*). Jej największą zaletą jest prostota implementacji. Brak obliczeń wykładniczych zmniejsza czas potrzebny przy uczeniu sieci. Funkcja ReLU jest opisana wzorem:

W procesie projektowania sieci konwolucyjnej określa się jaką ilość filtrów posiadać będą poszczególne warstwy konwolucyjne. Każdego z filtrów używa się do generowania map cech poprzez operację konwolucji, tak jak zostało to przedstawione na rysunku 69.

Ostatnie warstwy CNN zwykle stanowi kilka warstw gęstych, których zadaniem jest dokonanie samej klasyfikacji. W związku z tym, że warstwy konwolucyjne operują na danych wielowymiarowych, przed przekazaniem ich do warstw głębokich należy przekształcić je w jednowymiarowy wektor. W sieciach neuronowych zadanie to wykonuje specjalna warstwa spłaszczająca (ang. *flatten layer*).





# Ocena jakości klasyfikatora

W celu zweryfikowania jakości klasyfikatora binarnego użyto różnych metryk, aby zmniejszyć ryzyko błędnej oceny działania. Wszystkie wykorzystane metryki bazują na przedstawionej w Tabeli 123 macierzy pomyłek.

Tab. 1 Macierz pomyłek (ang. confusion matrix)

|  |  |  |  |
| --- | --- | --- | --- |
|  | | **Przewidziana** **wartość** | |
| **1** | **0** |
| **Rzeczywista**  **wartość** | **1** | **TP**  prawdziwie pozytywny | **FN**  fałszywie negatywny |
| **0** | **FP**  fałszywie pozytywny | **TN**  prawdziwie negatywny |

Pierwszą z metryk jest dokładność (ang. *accuracy*), mówiąca o stosunku poprawnie zaklasyfikowanych próbek względem całego zbioru.

Kolejną metryką jest precyzja (ang. *precision*), która jest określona następującym stosunkiem:

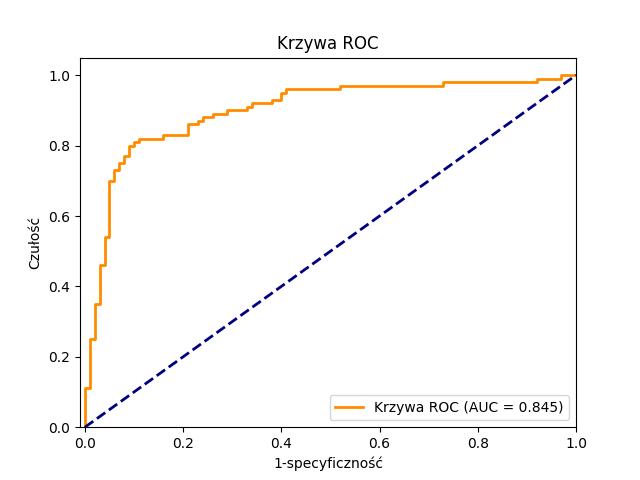
Przez wartość tej metryki rozumiemy odsetek prawdziwie pozytywnym przypadków z wszystkich, które klasyfikator określił jako pozytywne. Mała wartość precyzji sugeruje, że klasyfikator błędnie klasyfikuje negatywną klasę jako pozytywną.

Następną metryką jest czułość (ang. *recall*), czyli zdolność klasyfikatora do wykrywania przykładów pozytywnych

Kolejną metryką, która jest połączeniem dwóch poprzednich jest metryka F1. Jest ona średnią harmoniczną precyzji i czułości.

Ostatnią użytą metryką jest krzywa ROC (ang. *receiver operating characteristic curve).* Krzywa ta jest wykresem graficznym, który obrazuje zdolność klasyfikatora do oddzielania klasy pozytywnej od negatywnej. Metryka jest używana tylko w przypadku klasyfikatora binarnego. Tworzy się ją poprzez wykreślenie stosunku wskaźnika czułości opisanego wzorem (wzór recall), oraz 1-specyficzności (ang. specificity) opisanego wzorem

dla różnej wielkości progu klasyfikacji. Przykładowy wygląd krzywej ROC został przedstawiony na Rysunku 69.



Do przypisania wartości dla metryki ROC wykorzystuje się pole pod powierzchnią krzywej (AUC, ang. *area under curve).* Wartość AUC równa 0.5 oznacza brak umiejętność klasyfikatora do rozróżniania klas, a tym bardziej wartość AUC jest bliższa 1, tym bardziej klasyfikator poprawnie oddziela klasy pozytywne od negatywnych. Wartość poniżej 0.5 oznacza, że klasyfikator jest gorszy niż losowe klasyfikowanie do jednej z klas.

ROZDZIAŁ 3 – PROJEKT SYSTEMU

# Opis projektu

Głównym zadaniem aplikacji ma być szybka i łatwa możliwość sklasyfikowania wybranego przez użytkownika skanu MRI mózgu przez wybrany przez użytkownika, wytrenowany klasyfikator. Aplikacja ma działać lokalnie, tak samo jak lokalnie ma być dokonywana predykcja. Użytkownik ma mieć możliwość ręcznego wyboru oraz załadowania modelu, oraz dowolnego skanu. Użytkownik powinien otrzymać wynik klasyfikacji w formie tekstowej, informującej go czy przesłany skan został zaklasyfikowany do klasy pozytywnej lub negatywnej. Proces uczenia oraz testowania modeli ma odbywać się lokalnie na komputerze. Cały projekt ma zostać zaimplementowany w języku Pyton w wersji 3.10.8.

# Wymagania funkcjonalne aplikacji

**F1** Wybór modelu z lokalnego komputera

Użytkownik ma mieć możliwość wyboru i załadowania modelu z lokalnego komputera poprzez wybór ścieżki do modelu.

**F2** Wybór zdjęcia z lokalnego komputera

Użytkownik ma mieć możliwość wyboru i załadowania dowolnego zdjęcia z lokalnego komputera poprzez wybór ścieżki do zdjęcia.

**F3** Wyświetlanie załadowanego zdjęcia

Aplikacja powinna wyświetlić wybrane i załadowane przez użytkownika zdjęcie.

**F4** Klasyfikowanie wybranego zdjęcia

Aplikacja ma po wybraniu modelu oraz zdjęcia dokonać klasyfikacji załadowanego zdjęcia wykorzystując załadowany model.

**F5** Wyświetlanie wyniku klasyfikacji

Użytkownik po wykonaniu klasyfikacji ma zostać poinformowany w formie tekstowej o wyniku zwróconym przez model. Tekst ma informować użytkownika czy wybrane zdjęcie według modelu pochodzi z jednej z dwóch klas: „tumor”, „no tumor”.

# Wymagania niefunkcjonalne

**NF1** Interfejs graficzny ma być wykonany w języku angielskim

**NF2** Aplikacja ma informować tekstowo o swoim statusie

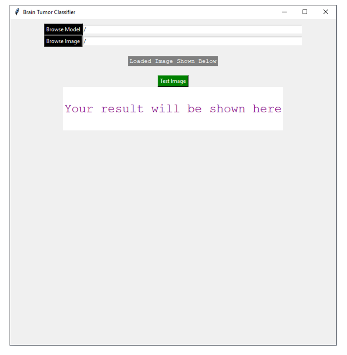
# Diagram przypadków użycia

Przedstawiony na Rysunku 69 diagram przypadków użycia, który został stworzony opierając się na wymienionych wymaganiach funkcjonalnych, ma na celu wizualne opisanie wszystkich funkcjonalności aplikacji.

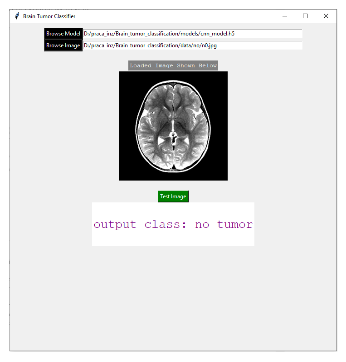


# Prototyp interfejsu graficznego

Aplikacja graficzna ma zostać wykonana przy pomocy dołączonej do języka Python biblioteki Tkinter. Cały interfejs będzie składać się z jednego widoku, w którym będą zawarte wszystkie funkcjonalności. Prototyp wyglądu, przed interakcją użytkownika został przedstawiony na Rysunku 911. Interfejs będzie składał się z trzech sekcji. W górnej sekcji znajdą się dwa przyciski odpowiedzialne kolejno za wybranie i wczytanie modelu, oraz za wybranie wczytanie zdjęcia. Dodatkowo obok przycisków zostanie pokazana ścieżka wybranego modelu i zdjęcia. W środkowej sekcji znajdować się będzie podgląd załadowanego zdjęcia, a także przycisk służący do jego klasyfikacji. Ostatnią część aplikacji zajmować będzie okno, w którym wypisywane będą wyniki klasyfikacji. Prototyp wyglądu po dokonaniu klasyfikacji przedstawia Rysunek 912.



Rys. 2 Prototyp wyglądu interfejsu przed interakcją użytkownika



Rys. 3 Prototyp interfejsu użytkownika po wykonaniu predykcji

ROZDZIAŁ 4 – BADANIA

# Cel, zakres i metodyka badania

Celem przeprowadzonych eksperymentów było wdrożenie a następnie porównanie kilku różnych podejść uczenia maszynowego i uczenia głębokiego w zadaniu klasyfikacji skanów MRI. Zbadany został także wpływ wielkości zbioru danych na wyniki tworzonych modeli. Każdy z modeli był uczony na oryginalnym zbiorze danych, a także na syntetycznie powiększonym zbiorze. Do porównania ze sobą wszystkich wyuczonych modeli wykorzystano dokładność klasyfikacji przeprowadzonej na zbiorze testowym.

Parametry modelu, który cechował się największą dokładnością, były następnie dostosowywane, aby jeszcze bardziej poprawić jego dokładność, a także inne przygotowane metryki.

# Zbiór danych

Do przeprowadzenia badań wykorzystany został zbiór Br35H. Zawiera on 3000 obrazów w formacie jpg przedstawiających skany MRI o różnych wymiarach. Zbiór jest podzielony równo po 1500 obrazów na dwie klasy: klasę pozytywną – skan przedstawiający guza mózgu oraz klasę negatywną – skan nie przedstawiający guza mózgu.

Aby przygotować zbiór do uczenia modeli, został on podzielony na 3 podzbiory: zbiór uczący, zbiór walidacyjny i zbiór testowy. Zbiór uczący jest używany w procesie trenowania modelu. To na podstawie tych danych model ma nauczyć się rozpoznawania wzorców. Do śledzenia procesu uczenia wykorzystuje się zbiór walidacyjny. Po każdej epoce procesu uczenia oblicza się skuteczność oraz wartość funkcji kosztu. Zbiór testowy służy do sprawdzenia jak działa wytrenowany model po zakończeniu procesu uczenia. Ważnym aspektem jest, aby dane zawarte w zbiorach nie powielały się. Oznacza to, że model nie powinien nigdy mieć wglądu w dane testowe podczas uczenia. Gwarantuje to, że podczas sprawdzania jakości modelu skutecznie zweryfikujemy jego możliwość do generalizacji danego problemu. Dokładne rozłożenie danych na zbiory zostało przedstawione w Tabeli 69.

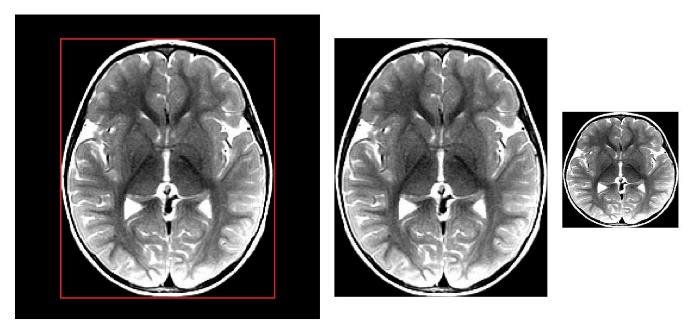
|  |  |  |  |
| --- | --- | --- | --- |
|  | Klasa pozytywna | Klasa negatywna | Suma |
| Zbiór uczący | 1200 | 1200 | 2400 |
| Zbiór walidacyjny | 100 | 100 | 200 |
| Zbiór testowy | 200 | 200 | 400 |

* 1. Wstępne przetwarzanie danych

Jednym ze sposobów na zwiększenie jakości modelu jest wstępne przetworzenie dostarczanych dla niego danych. W tej pracy użyto trzech kroków dla każdego załadowanego obrazu.

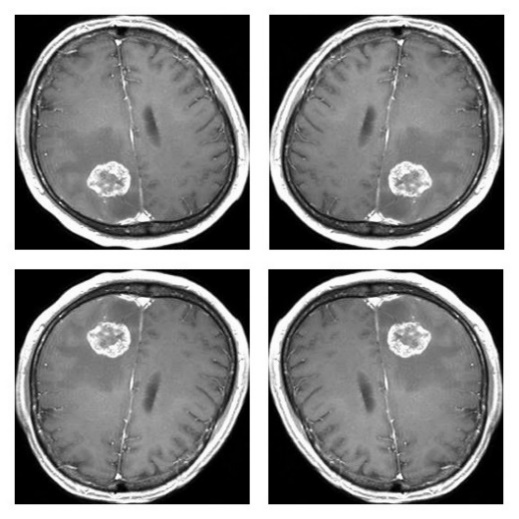
1. Konwersja obrazów do skali szarości – jest to prosta metoda pozwalająca na znaczne zmniejszenie wymiarów ładowanych obrazów.
2. Automatyczne przycięcie obrazów względem konturu – metoda została zaimplementowana dzięki bibliotece OpenCV. Jej zadaniem jest znalezienie najbardziej wysuniętych punktów wchodzących w kontur skanu mózgu, a następnie wycięcie niepotrzebnego tła. Dzięki temu zmniejsza się ilość danych, które nie wnoszą żadnych przydatnych dla procesu uczenia informacji.
3. Zmiana rozmiaru obrazu – aby ustandaryzować wprowadzane dane, ich wymiary są zmieniane do jednego standardu – 240 x 240 pikseli.

Cały proces przetwarzania zdjęcia został przedstawiony na Rysunku 2137.



* 1. Powiększenie zbioru danych

Aby wykorzystać możliwości uczenia maszynowego i sieci neuronowych do uczenia się wykorzystano technikę powiększania zbioru danych (ang. *Data augmentation*). Zastosowanie jej polega na sztucznym zwiększeniu ilości obrazów, używając do tego już istniejących. Syntetycznie wytworzone dane mogą mieć pozytywny wpływ na jakość uczonego modelu, a także na ograniczenie zbyt dużego dopasowania do danych. W tej pracy wykorzystano jedynie technikę odbić lustrzanych (ang. *mirroring*), polegającą na odbijaniu obrazu względem jednej osi lub obydwu na raz. Proces ten pokazuje Rysunek 2137.



Po zastosowaniu powiększenia zbioru danych ich ilość zwiększyła się czterokrotnie. Ostateczna ilość obrazów w poszczególnych zbiorach pokazuje Tabela 69.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Klasa pozytywna | Klasa negatywna | Suma |
| Zbiór uczący | 4800 | 4800 | 9600 |
| Zbiór walidacyjny | 400 | 400 | 800 |
| Zbiór testowy | 800 | 800 | 1600 |

# Opis podejść

Celem badania było zastosowanie różnych i znanych podejść i porównanie ich pod względem dokładności. Wybrane sposoby można podzielić na metody uczenia maszynowego i te wykorzystujące uczenie głębokich sieci neuronowych.

Badania związane z uczeniem maszynowym zakładały użycie znanych i powszechnie używanych klasyfikatorów liniowych, a także połączenie ich w jednym przypadku z techniką redukcji wymiarów danych a w drugim z techniką ekstrakcji cech. Przed rozpoczęciem procesu uczenia dane były poddane procesowi standaryzacji cech, czyli odjęcia średniej oraz skalowaniu do wariancji jednostkowej. Dla każdego podejścia używane były dwa klasyfikatory: regresja logistyczna oraz liniowa maszyna wektorów nośnych, z których ostatecznie wybierany był jeden, cechujący się większą dokładnością.

Pierwsze podejście zakładało wyłącznie wytrenowanie modelu klasyfikacji na ustandaryzowanych danych. Była to próba skorzystania z najprostszego rozwiązania, które może służyć za porównanie dla bardziej skomplikowanych metod.

Następne podejście wykorzystywało redukcję wymiarów. W badaniu użyto metody analizy głównych składowych (ang. *principal component analysis*, PCA) do redukcji ilości cech.

W ostatnim podejściu zastosowano technikę ekstrakcji cech nazwaną histogramem zorientowanych gradientów (ang. *histogram of oriented gradients*, HOG).

W drugiej części badań zaprojektowano oraz wytrenowano dwie sieci neuronowe rozwiązujące zadanie klasyfikacji binarnej obrazów. Pierwszą z sieci była

# Porównanie modeli

# Dostosowanie najlepszego modelu

# Wyniki testowania

# Użyte technologie i biblioteki

W tym podrozdziale zostały wymienione i opisane główne technologie i biblioteki potrzebne do wykonania projektu:

1. Python – język programowania wysokiego poziomu, który posiada wiele zróżnicowanych bibliotek umożliwiających wykorzystanie go do wielu zadań. Python jest jednym z popularniejszych języków wykorzystywanych w dziedzinie uczenia maszynowego i głębokiego uczenia.
2. Tkinter – wbudowana biblioteka języka Python umożliwiająca tworzenie prostego interfejsu graficznego. Może być używana do tworzenia interfejsów aplikacji dla większości systemów Unix, w tym macOS i Windows.
3. TensorFlow – popularna otwarto źródłowa biblioteka języka Python wykorzystywana w procesie uczenia maszynowego lub uczenia głębokich sieci neuronowych. Cechuje się łatwością implementacji łącząc ją z wysokopoziomowym API oraz szybkością działania w związku z użyciem języka C++ w samym silniku.
4. Keras – otwarto źródłowa biblioteka języka Python służąca jako interfejs biblioteki TensorFlow. Jej głównym celem jest zapewnienie łatwości obsługi dla użytkownika, przy zachowaniu dużej swobody działania.
5. Scikit-learn – otwarto źródłowa biblioteka języka Python posiadająca wiele algorytmów uczenia maszynowego. Biblioteka zawiera także dużo przydatnych narzędzi m.in. wstępnego przetwarzania danych
6. NumPy - otwarto źródłowa biblioteka pozwalająca na zaawansowaną obsługę wielowymiarowych tabel i macierz, a także na wykonywaniu na nich łatwych i szybkich obliczeń
7. Matplotlib – otwarto źródłowa biblioteka wykorzystywana do tworzenia wykresów oraz wizualizacji

Bibliografia

[nr] inicjał imienia, nazwisko autora, tytuł pracy, miejsce wydania pracy, wydawnictwo, rok wydania, ewentualnie numery stron.

[1] Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.

[2] Harris, C.R., Millman, K.J., van der Walt, S.J. et al. Array programming with NumPy. Nature 585, 357–362, 2020

[3] J. D. Hunter, "Matplotlib: A 2D Graphics Environment", Computing in Science & Engineering, vol. 9, no. 3, pp. 90-95, 2007

[4] H.P. Langtangen, Python Scripting For Computational Science, Texts in Computational Science and Engineering (Springer Berlin Heidelberg, 2009).

[5] Hoo, Zhe Hui, Jane Candlish, and Dawn Teare. "What is an ROC curve?." *Emergency Medicine Journal* 34.6 (2017): 357-359. ( ROC curve )

[6] Shorten, Connor, and Taghi M. Khoshgoftaar. "A survey on image data augmentation for deep learning." *Journal of big data* 6.1 (2019): 1-8. (augmentation)

[7] Jolliffe, Ian T., and Jorge Cadima. "Principal component analysis: a review and recent developments." Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 374.2065 (2016): 20150202. (PCA)

[8] Zhou, Wei, et al. "Histogram of oriented gradients feature extraction from raw Bayer pattern images." IEEE Transactions on Circuits and Systems II: Express Briefs 67.5 (2020): 946-950. (HOG)

[9] Hsu, Chih-Wei, Chih-Chung Chang, and Chih-Jen Lin. "A practical guide to support vector classification." (2003): 1396-1400. (SVM)

[10] Akinori Hidaka i Takio Kurita. „Consecutive Dimensionality Reduction by Canonical

Correlation Analysis for Visualization of Convolutional Neural Networks”. W:

t. 2017. Grud. 2017, s. 160–167. (CNN obrazek i opis)

[11] Rikiya Yamashita, Mizuho Nishio, Richard Kinh Gian Do i Kaori Togashi. „Convolutional neural networks: an overview and application in radiology”. W: Insights

into Imaging 9.4 (sierp. 2018), s. 611–629. issn: 1869-4101. (CNN obrazek i opis)