### Uniwersytet Warmińsko-Mazurski w Olsztynie Wydział Matematyki i Informatyki

### Andrzej Strzeszewski

Kierunek: Informatyka

Uczenie zespołowe, zastosowanie i porównania efektywności wybranych algorytmów w dostępnych zbiorach danych.

Praca magisterska wykonana w Katedra Metod Matematycznych Informatyki pod kierunkiem dr. Andrzeja Jankowskiego

# University of Warmia and Mazury in Olsztyn Faculty of Mathematics and Computer Science

### Andrzej Strzeszewski

Field of Study: Computer Science

## Ensemble learning, application and performance comparisons of selected algorithms on available datasets

Master's Thesis is performed in the Chair of Mathematical Methods of Computer Science under supervision of Andrzej Jankowski, PhD

## Streszczenie

Praca zawiera opisy algorytmów uczenia maszynowego. Celem pracy jest połączenie tych algorytmów metodami uczenia zespołowego oraz sprawdzenie czy dane rozwiązanie jest dokładniejsze. Dla rozwiązań liczone będą błędy na podstawie różnych mierników. Takich jak: Mean Absolute Error, Mean Absolute Percentage Error. Rozważania zakończyłem AB-testami, w których z odpowiednim poziomem ufności stwierdziłem czy uczenie zespołowe daje większą efektywność niż pojedyncze algorytmy.

## Abstract

Ensemble learning, application and performance comparisons of selected algorithms on available datasets

dfsf

# Spis treści

Streszczenie	1
Abstract	2
Wstęp	4
Rozdział 1. Wymagania aplikacji	5
Rozdział 2. Użyte technologie	6
Rozdział 3. Źródła danych	7
Rozdział 4. Wizualizacja danych	8
Rozdział 5. Trenowanie modeli	9
Rozdział 6. Opis wykorzystanych algorytmów	10
6.1. Mierzenie dokładności algorytmów decyzyjnych 6.2. Regresja liniowa 6.3. knn 6.4. Sieć neuronowa 6.5. Drzewo decyzyjne – regresyjne 6.6. Regresyjny las losowy (ang. Regression forest) 6.7. Drzewo decyzyjne wzmocnione (ang. Boosted decision tree)  Rozdział 7. Implementacja i omówienie kodu 7.1. Algorytmy 7.2. Łączenie algorytmów 7.2.1. Bagging 7.2.2. Boosting	10 10 10 10 10 10 10 11 11 11 11 11
Rozdział 8. Sprawdzanie efektywności uczenie zaespołowego- AB testy	12
Rozdział 9. Podsumowanie	13
Spis rysunków	14
Spis tabel	15
Rozdział 10. Spis algorytmów	16
Rozdział 11. Indeks stosowanych oznaczeń	17
Rozdział 12. Używane źródła danych i dokumentów z Internetu	18
Bibliografia	19

# Wstęp

sdfsdfs

# Wymagania aplikacji

sgsgfsdg

# Użyte technologie

fgxf

# Źródła danych

dasdsa

# Wizualizacja danych

dasdsa

## Trenowanie modeli

dasdsa

## Opis wykorzystanych algorytmów

(tutaj jeszcze nie wiem jakich użyję więc wrzucam jak najwięcej, żeby mieć potem w czym wybierać)

#### 6.1. Mierzenie dokładności algorytmów decyzyjnych

tutaj coś o błędach

#### 6.2. Regresja liniowa

dfghdfgh

#### 6.3. knn

dfghdfgh

#### 6.4. Sieć neuronowa

dfghdfgh

#### 6.5. Drzewo decyzyjne – regresyjne

dfghdfgh

#### 6.6. Regresyjny las losowy (ang. Regression forest)

fghfgh

#### 6.7. Drzewo decyzyjne wzmocnione (ang. Boosted decision tree)

ghjghj hjkgk

## Implementacja i omówienie kodu

### 7.1. Algorytmy

fudtuy

## 7.2. Łączenie algorytmów

fudtuy

### **7.2.1.** Bagging

dsa

#### 7.2.2. Boosting

## Sprawdzanie efektywności uczenie zaespołowego- AB testy

 $\operatorname{cgjh}$ 

## Podsumowanie

gh

# Spis rysunków

dsa

# Spis tabel

sdadsa

# Spis algorytmów

sdadsa

## Indeks stosowanych oznaczeń

sdadsa

## Używane źródła danych i dokumentów z Internetu

 $Geographic\ coordinate\ system\ \texttt{https://en.wikipedia.org/wiki/Geographic\_coordinate\_system,\ dostep\ online:\ 20.01.2022.}$ 

## Bibliografia

[1] Geographic coordinate system https://en.wikipedia.org/wiki/Geographic\_coordinate\_system, dostęp online: 20.01.2022.