

**AGH**



# Analiza fałszywych wiadomości

Analysis of fake news

Realizacja: Piotr Kardaś

Opiekun: dr inż. Anna Zygmunt

2021



## Lizbona. Sąd: 152 pacjentów, a nie 17 tys. zmarło z powodu COVID-19

Dodano: wczoraj 21:23



### NAJNOWSZE

20:55 Restauracja c  
Obstrzenia i

20:31 "Inność nie m  
został... miss



STOWARZYSZENIERKW.ORG

To oficjalne - Senat USA ogłosił dziś: koronawirus to kłamstwo; media ukrywają prawdę - Stowarzyszenie RKW

11

4 udostępnienia



### Analiza stanu wiedzy

analiza stylu, treści, propagacji, postawy, reputacji i cech wizualnych  
opis najpopularniejszych algorytmów uczenia maszynowego oraz metryk  
opis modeli epidemiologicznych oraz współczynnika  $R_0$

### Budowa klasyfikatora opartego na technikach NLP

analiza cech tekstu - bigramy, nacechowanie, obiektywność i inne  
obliczenie wartości metryk na 3 różnych zbiorach danych  
porównanie wyników z innymi pracami wykorzystującymi te same zbiory danych

### Porównanie propagacji fałszywych wiadomości i propagacji epidemii

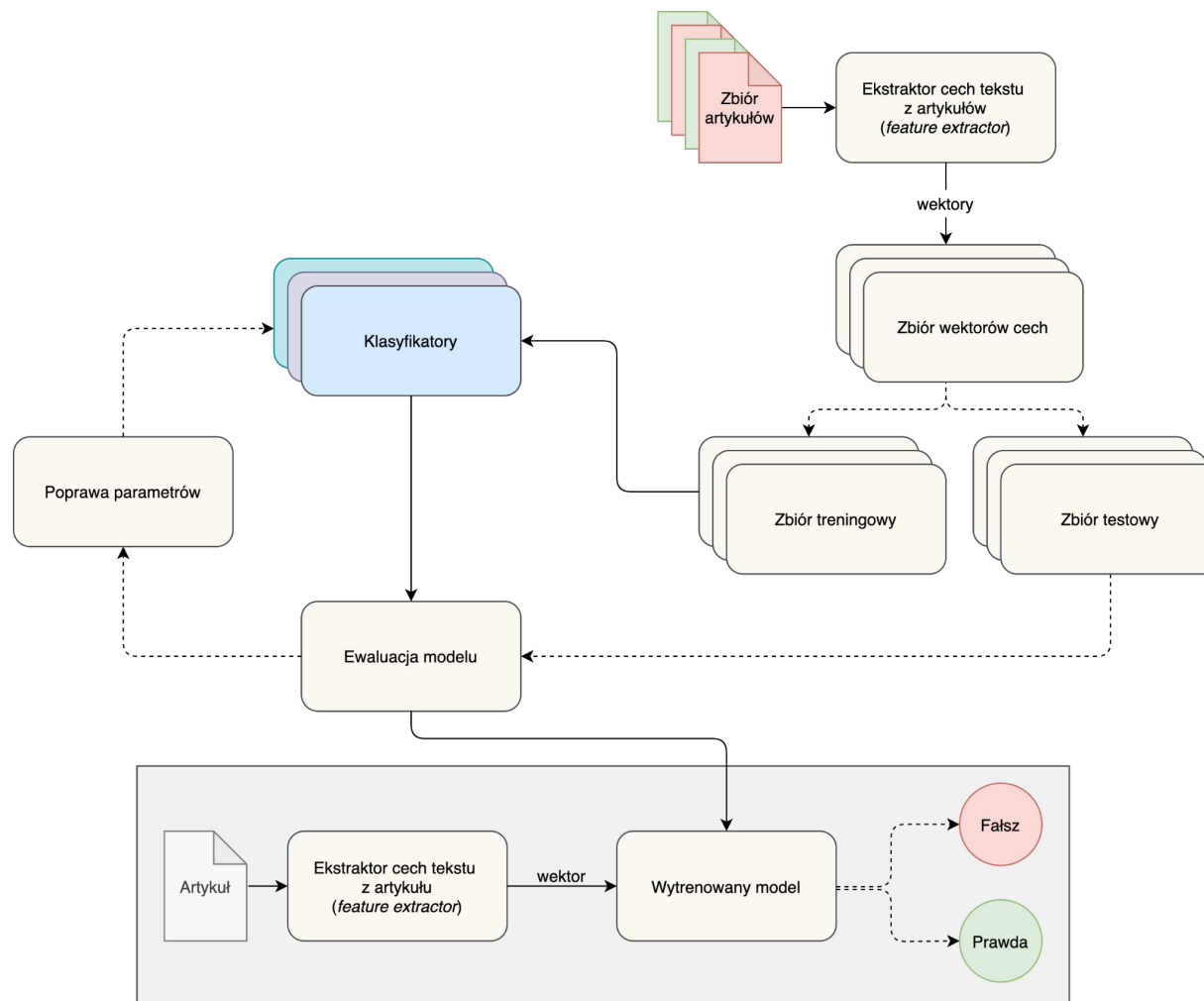
badanie współczynnika  $R_0$   
analiza propagacji w przestrzeni i czasie  
porównanie z epidemiami



|   |
|---|
| Liczba pomyłek (błędów ortograficznych)                           |
| Liczba unikalnych słów  |
| Liczba zdań   |
| Średnia liczba czasowników, rzeczowników i przymiotników w zdaniu |
| Długość artykułu  |
| Sentyment artykułu  |
| Subiektywność artykułu  |
| 10 najpopularniejszych bigramów                                   |

*Istnieje możliwość dodawania kolejnych cech.*

# Klasyfikacja fałszywych wiadomości





Passive-Aggressive Classifier

SDG Classifier

Gradient Boosting Classifier

Logistic Regression

Random Forest Classifier

Extra Trees Classifier

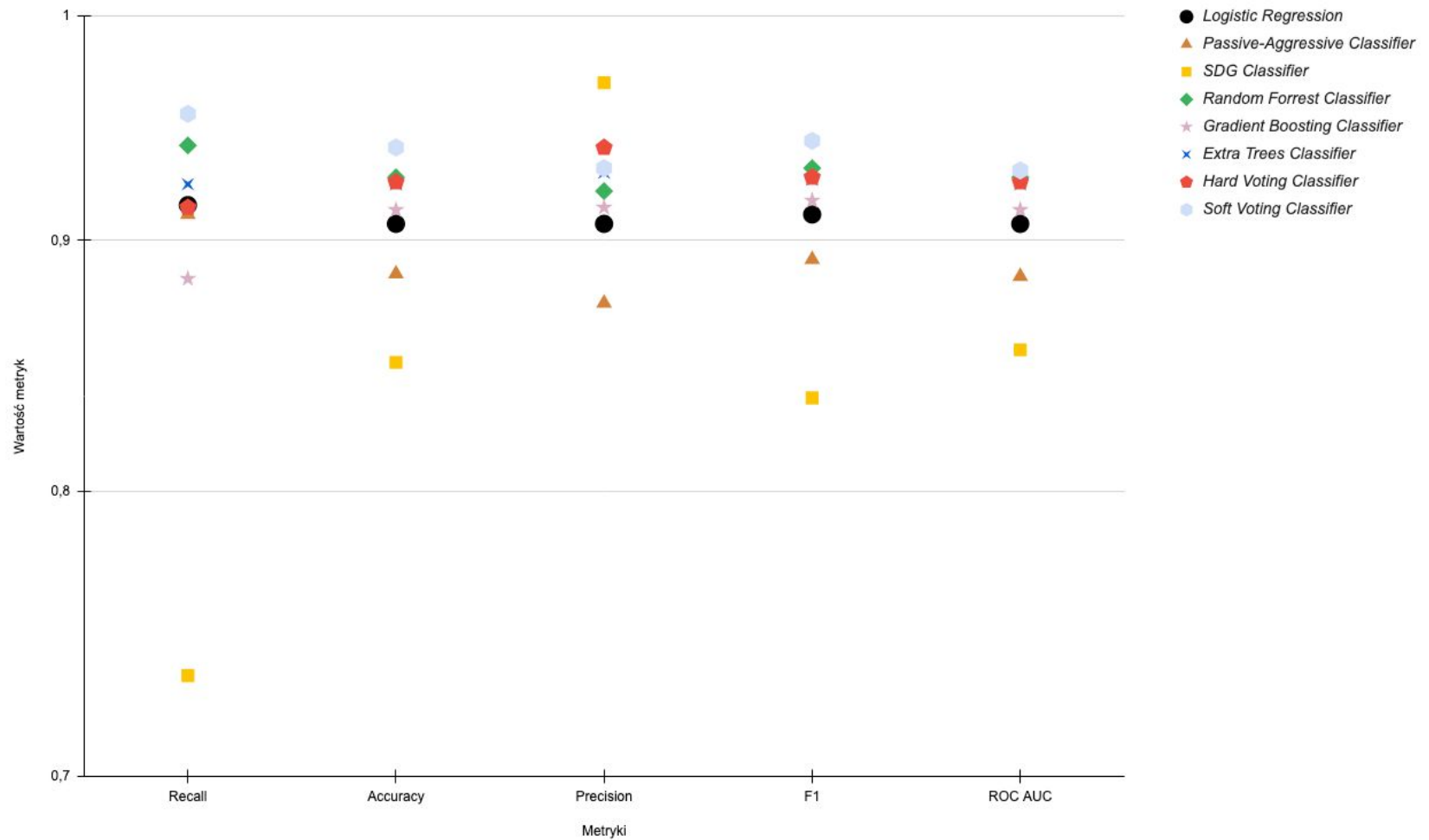
Hard Voting Classifier

*Logistic Regression  
Random Forest Classifier  
Gradient Boosting Classifier  
Extra Trees Classifier*

Soft Voting Classifier

*Logistic Regression  
Random Forest Classifier  
Gradient Boosting Classifier  
Extra Trees Classifier*

# Wynik klasyfikacji

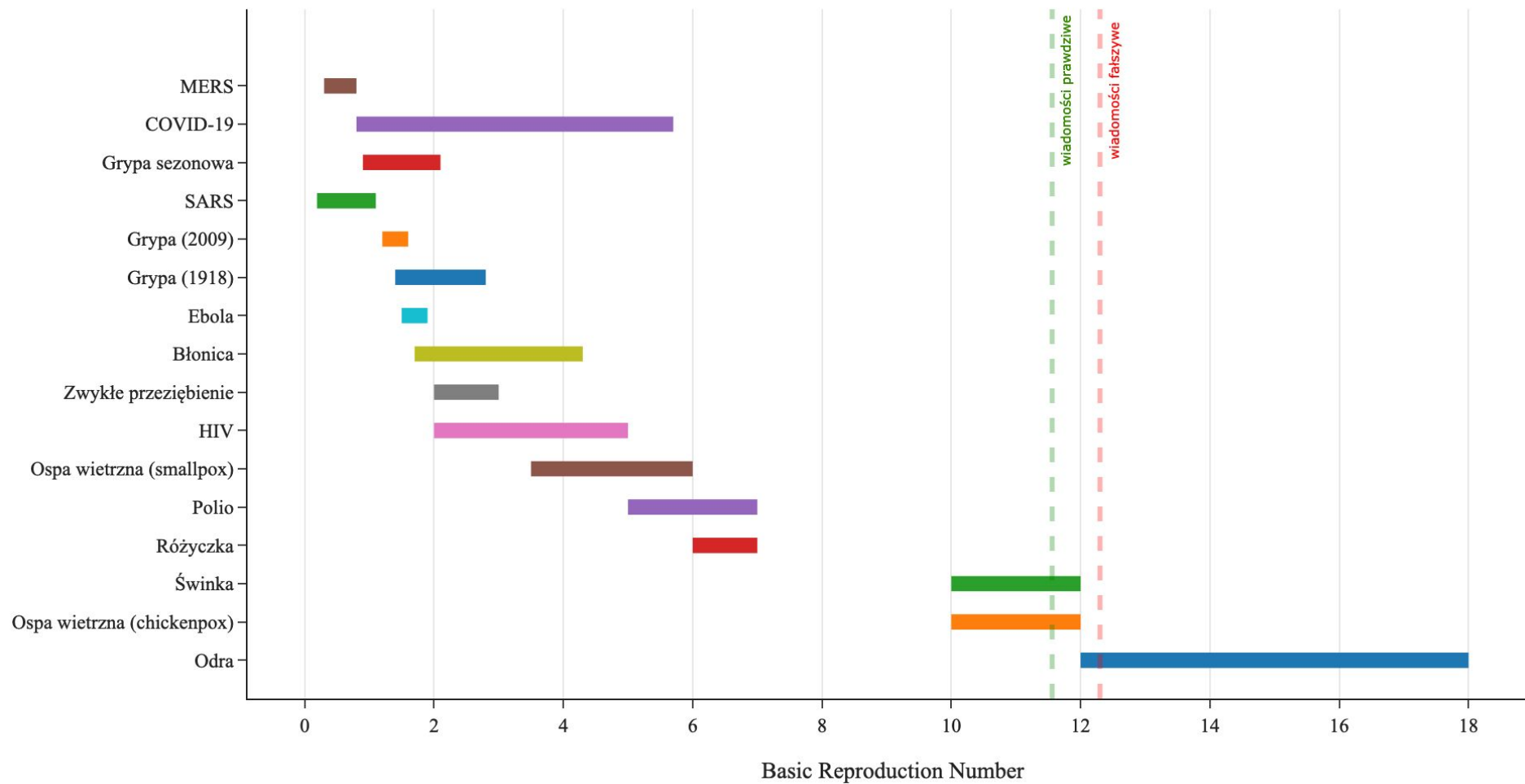




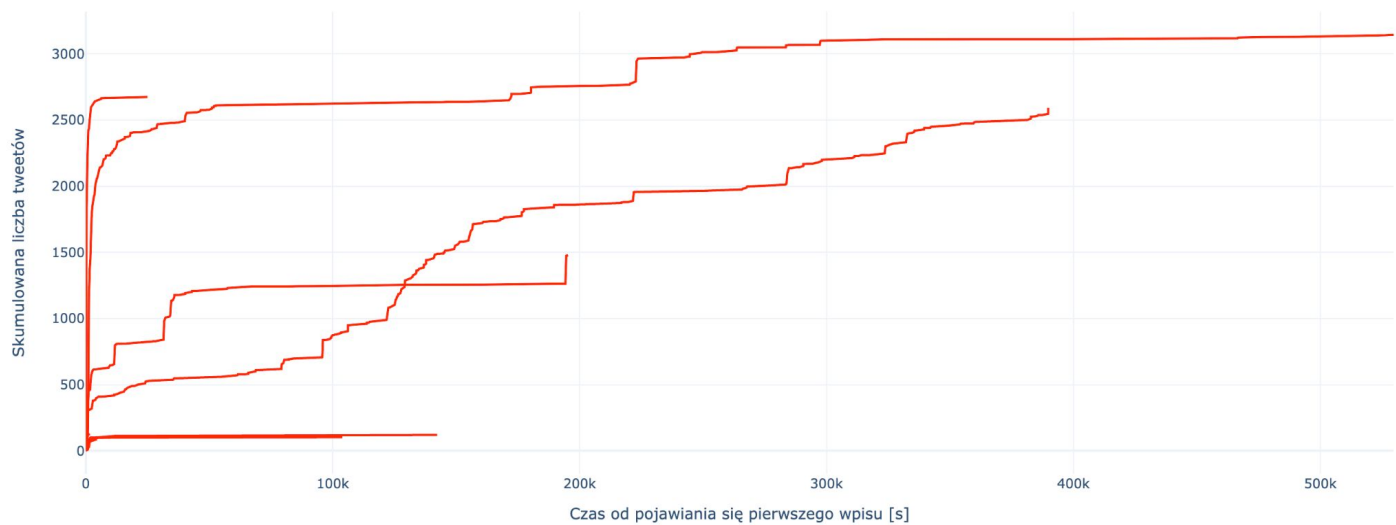
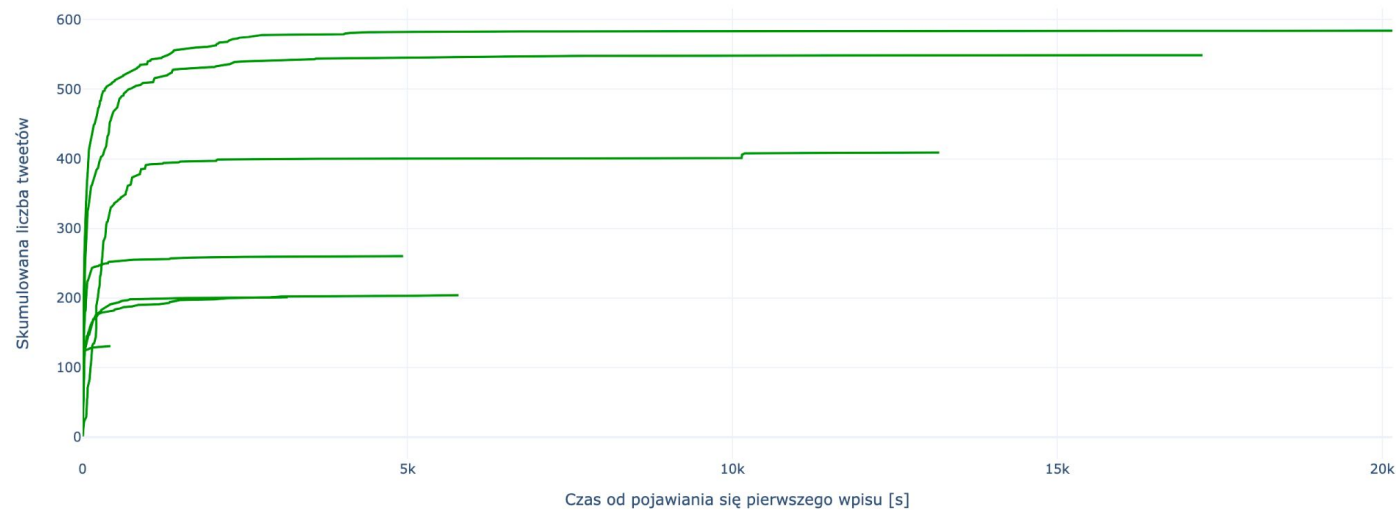
- metody agregujące kilka klasyfikatorów w jeden dają bardzo dobre rezultaty
- udało się uzyskać klasyfikator o zbliżonej jakości (wskazania metryk bliskie lub niższe o kilka procent względem istniejących rozwiązań)
- trudność w porównywaniu rozwiązań, autorzy wykorzystują różne metryki



# Analiza współczynnika R0



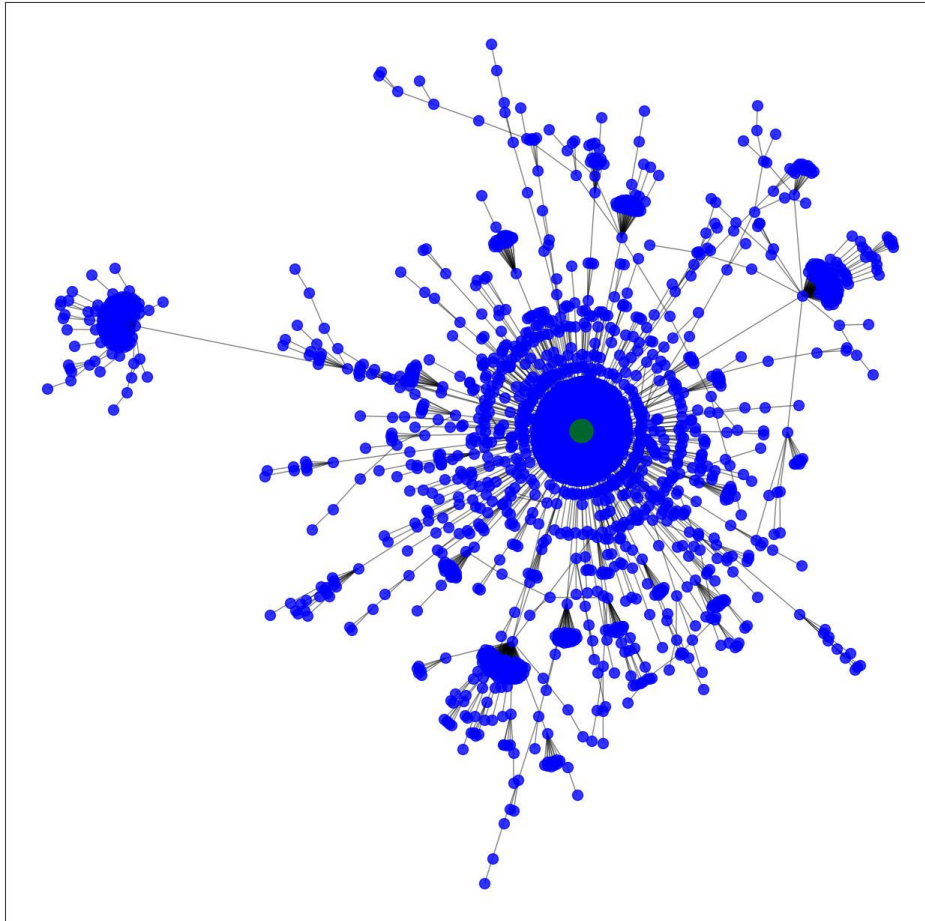
# Analiza propagacji w czasie



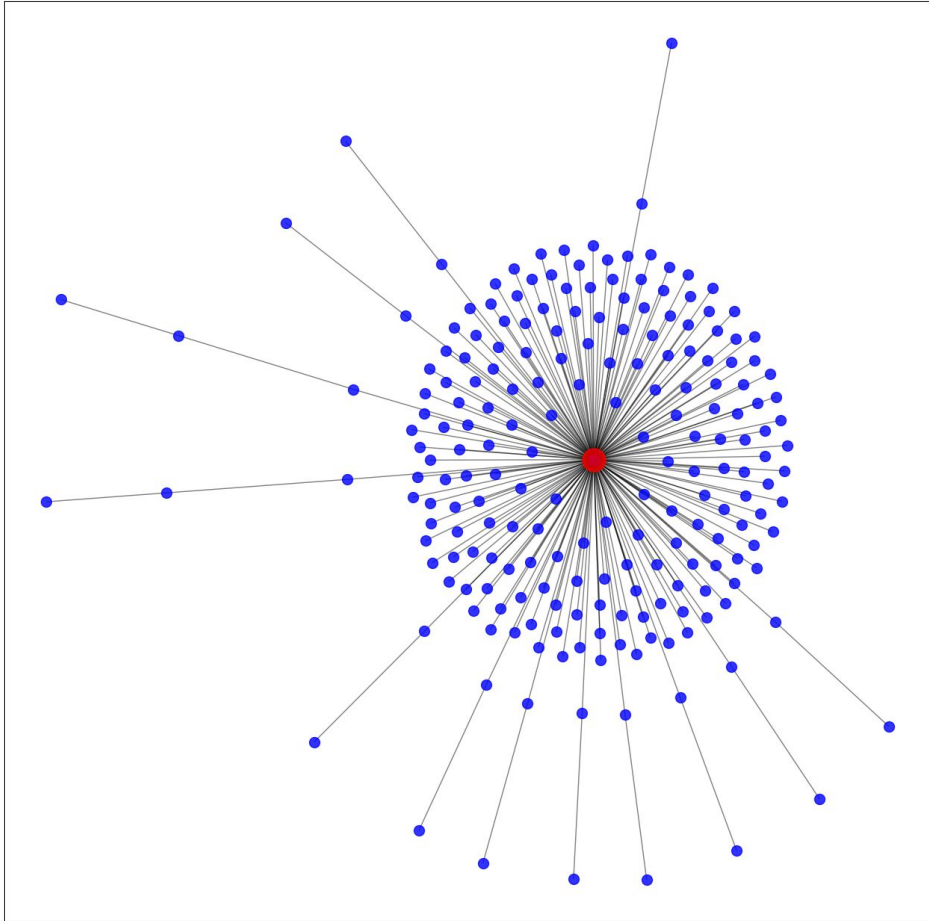
# Analiza propagacji w przestrzeni



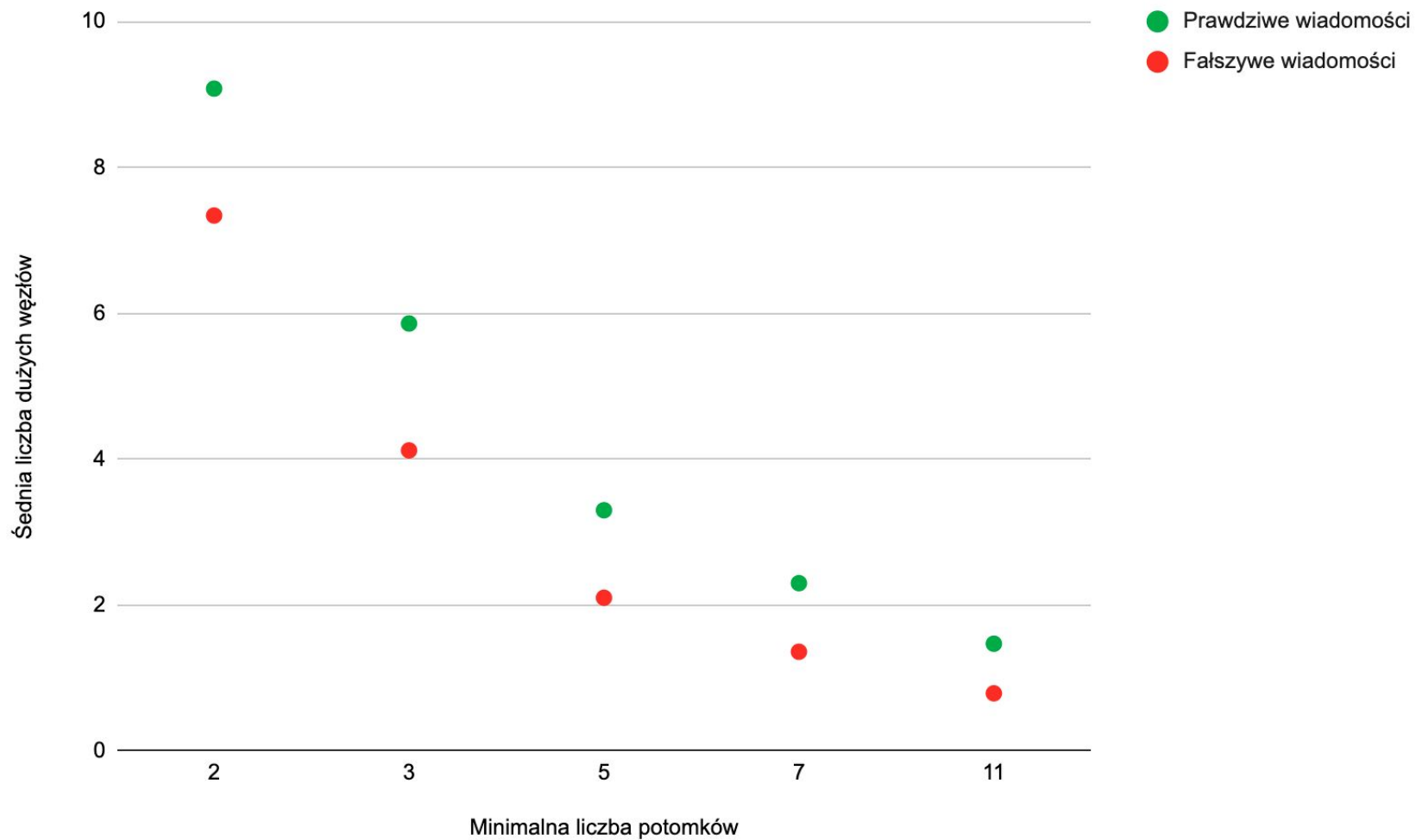
true: 'the white house is lit up rainbow. URL'



false: 'el chapo puts \$100 million bounty on donald trump...dead or alive! URL'



## Analiza propagacji w przestrzeni





- fałszywe wiadomości rozprzestrzeniają się niczym odra, a prawdziwe rozpowszechniają się podobnie do ospy i świnki
- fałszywe wiadomości rozpowszechniają się inaczej od wiadomości prawdziwych



- metody agregujące kilka klasyfikatorów w jeden pozwoliły na uzyskanie rozwiązania o podobnej jakości lub o kilka procent gorszych od istniejących rozwiązań
- fałszywe wiadomości propagują się inaczej od prawdziwych wiadomości
- istnieją pewne podobieństwa w propagacji epidemii oraz propagacji fałszywych informacji
- dalszy rozwój części klasyfikacyjnej powinien zostać skupiony na wykorzystaniu sieci neuronowych oraz propozycji dodatkowych cech tekstu
- w kwestii propagacji wiadomości na zgromadzeniu dodatkowych zbiorów danych oraz użyciu bardziej zaawansowanych mechanizmów matematycznych do porównania propagacji z rzeczywistymi epidemiami



Dziękuję za uwagę