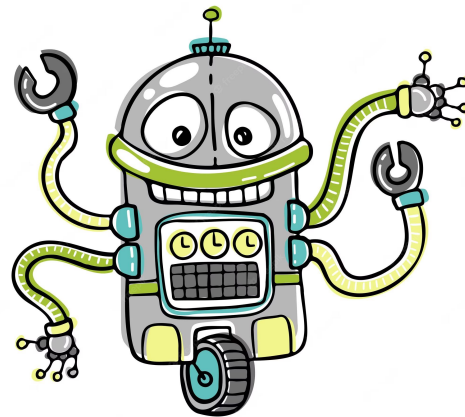


Fake News Detection

Wstęp do uczenia maszynowego

Alicja Charuza, Mateusz Gałęziwski



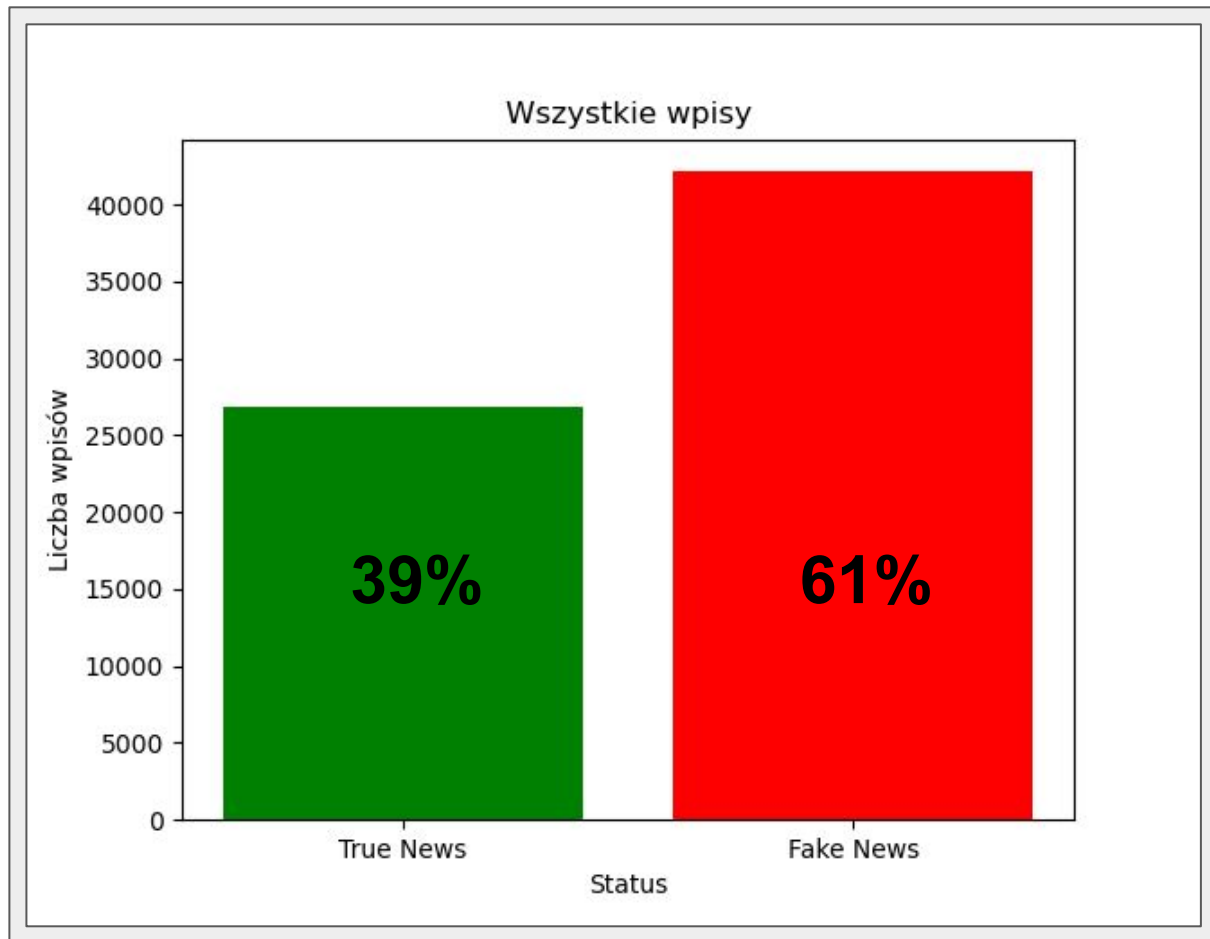
Źródło danych: <https://www.kaggle.com/datasets/mohamadaflahkhan/fake-news-dataset-combined-different-sources>

Po co klasyfikować fake newsy?

Komu mógłby się przydać nasz model?

O zbiorze danych:

- Dane zawierają **tytuł**, **tekst właściwy** oraz **etykietę fake/true**
- Ilość wszystkich zarejestrowanych wpisów: **69045**
- Wpisy pochodzą z różnych platform



Krok 1: detekcja języków

W jakich językach występują wpisy w zbiorze?

Co możemy zauważyć?

- Łączna ilość wpisów po angielsku: **67 419**
- Prawie wszystkie wpisy w innym języku są **fałszywe!**

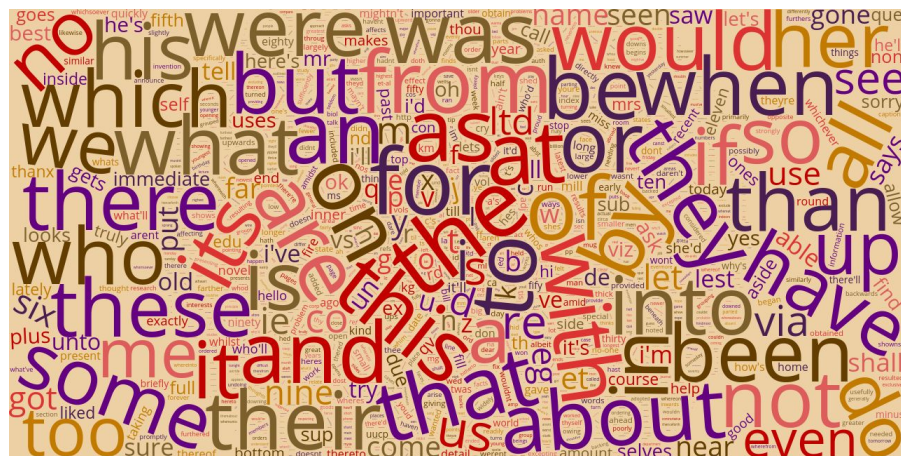
language	Ground	Label
af	fake	3
ar	fake	22
ca	fake	3
cy	fake	3
da	fake	2
de	fake	133
el	fake	4
en	fake	40538
	true	26881
es	fake	175
et	fake	1
fi	fake	5
fr	fake	53
hr	fake	3
hu	fake	1
id	fake	2
it	fake	11
lt	fake	1
nl	fake	4
no	fake	4
pl	fake	4
pt	fake	15
ro	fake	5
ru	fake	203
sl	fake	1
so	fake	8
sv	fake	1
sw	fake	14
tl	fake	2
tr	fake	10
unknown	fake	925
	true	5
vi	fake	2
zh-cn	fake	1

Przygotowanie danych tekstowych w j. angielskim

usunięcie :

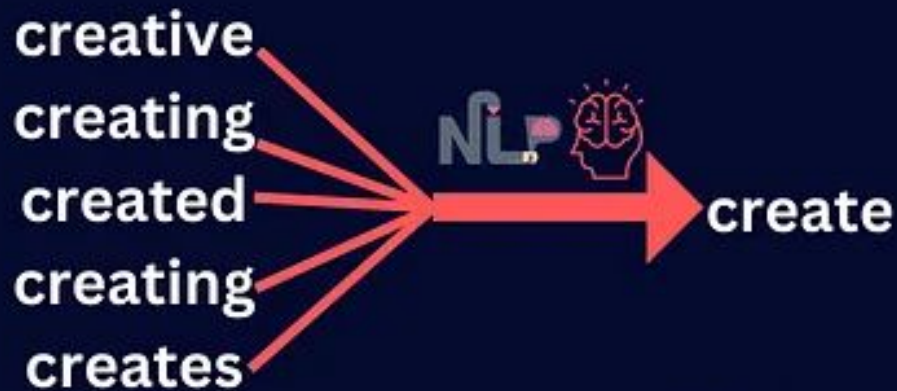
- interpunkcji
- cyfr
- pojedynczych znaków tj. ' (^|).\$
- stopwords w języku angielskim
- zamiana wszystkich liter na małe

stopwords



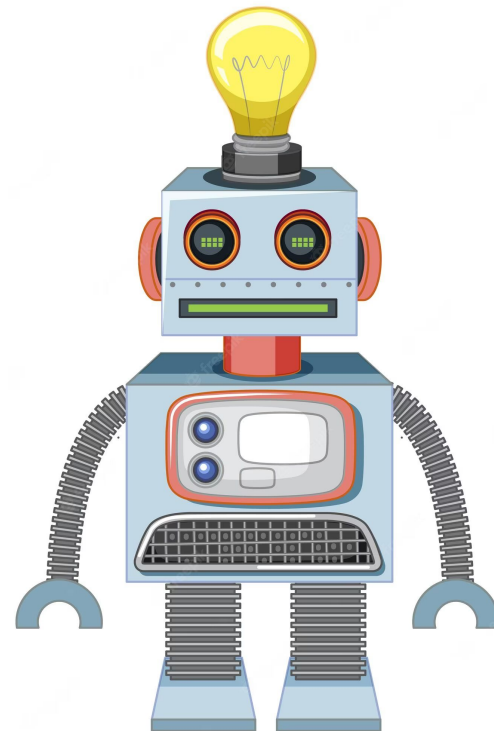
Stemizacja słów

Stemming in NLP



Rozważane modele

1. Regresja logistyczna
2. Drzewo decyzyjne
3. Las losowy
4. Liniowy Support Vector Classifier
5. MLPClassifier
6. Voting modeli (1, 2, 3)



Łatwo wyjaśnialne modele

Logistic Regression:

Train results:

F1 score: 0.9671120640147018
Accuracy: 0.9739633850630456
Recall: 0.9595077110081288
Precision: 0.9748379129360913
ROC AUC: 0.9961894397662193

Test results:

F1 score: 0.9454706927175844
Accuracy: 0.9565525049532975
Recall: 0.936653176139363
Precision: 0.9544558006096467
ROC AUC: 0.9891074094628323

Decision Tree with parameters:

Train results:

F1 score: 0.9532666274278988
Accuracy: 0.9639003394762367
Recall: 0.922813948188103
Precision: 0.9857977601038792
ROC AUC: 0.9790426964647868

Test results:

F1 score: 0.9247779590357077
Accuracy: 0.941268044155109
Recall: 0.8977652648249164
Precision: 0.9534666417492057
ROC AUC: 0.9443657733630982

Random Forest with parameters:

Train results:

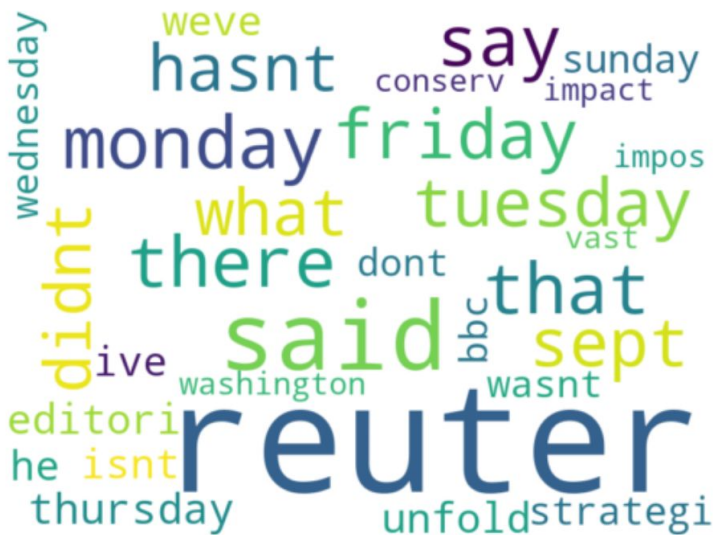
F1 score: 0.9971047619047619
Accuracy: 0.9976964112512124
Recall: 0.9942262402187951
Precision: 1.0
ROC AUC: 0.9999850848127622

Test results:

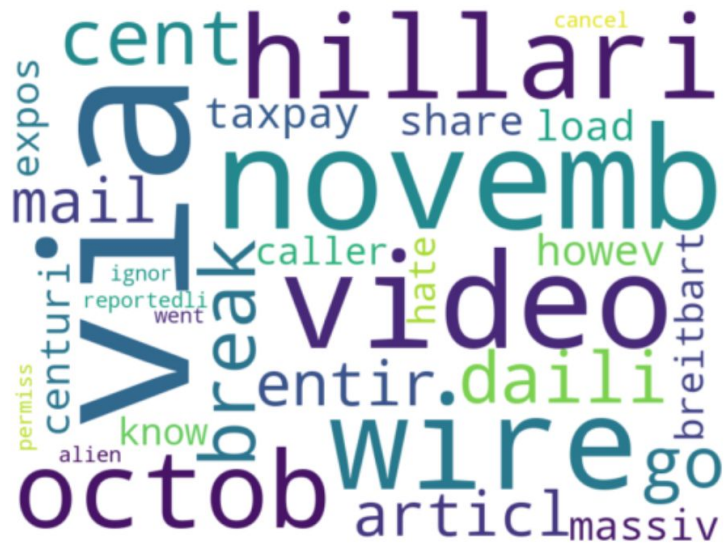
F1 score: 0.9319765309864321
Accuracy: 0.947495046702519
Recall: 0.8944219602322717
Precision: 0.972822966507177
ROC AUC: 0.986741036003936

1. Regresja logistyczna

Słowa najbardziej wpływające na
klasyfikację wpisu jako **prawdziwy news**

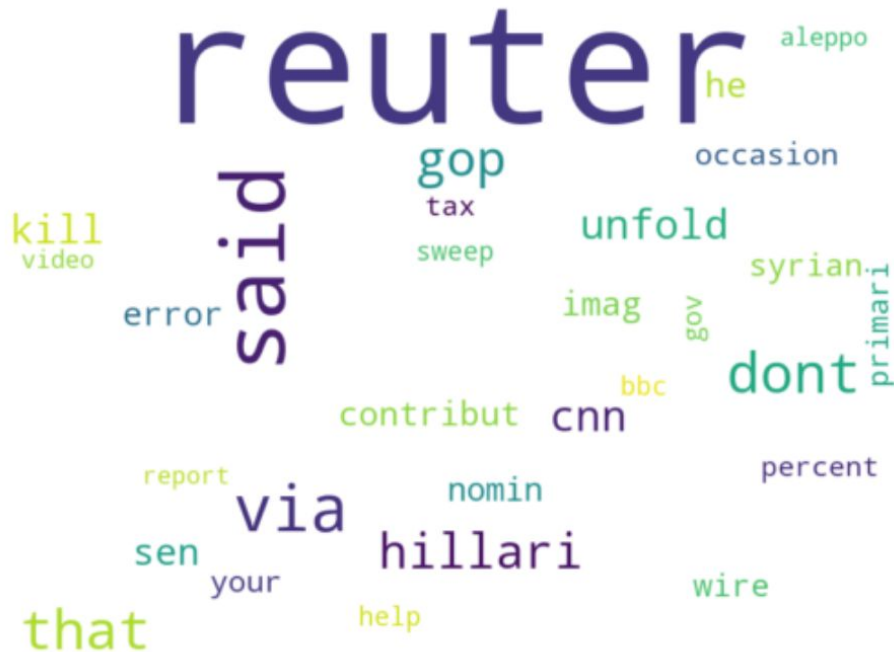


Słowa najbardziej wpływające na klasyfikację wpisu jako **fałszywy news**



2. Drzewo decyzyjne

Słowa najbardziej wpływające na klasyfikację wpisu



3. Las losowy

Słowa najbardziej wpływające na klasyfikację wpisu



Bardziej złożone modele vs. soft voting

Linear SVM:

Train results:

F1 score: 0.9670026443873836
Accuracy: 0.9739027643064986
Recall: 0.9584441236800122
Precision: 0.9757153905645785
ROC AUC: 0.9960747732779475

Test results:

F1 score: 0.945418929748019
Accuracy: 0.9566232663458817
Recall: 0.9341896885447827
Precision: 0.9569214131218458
ROC AUC: 0.9888077364415244

MLP Classifier:

Train results:

F1 score: 0.9997721230535511
Accuracy: 0.9998181377303589
Recall: 0.9999240294765631
Precision: 0.9996202627781575
ROC AUC: 0.9999999272056108

Test results:

F1 score: 0.9504404306432958
Accuracy: 0.9605859043305972
Recall: 0.9398205173323949
Precision: 0.9613030957523399
ROC AUC: 0.9914050553541202

Soft voting:

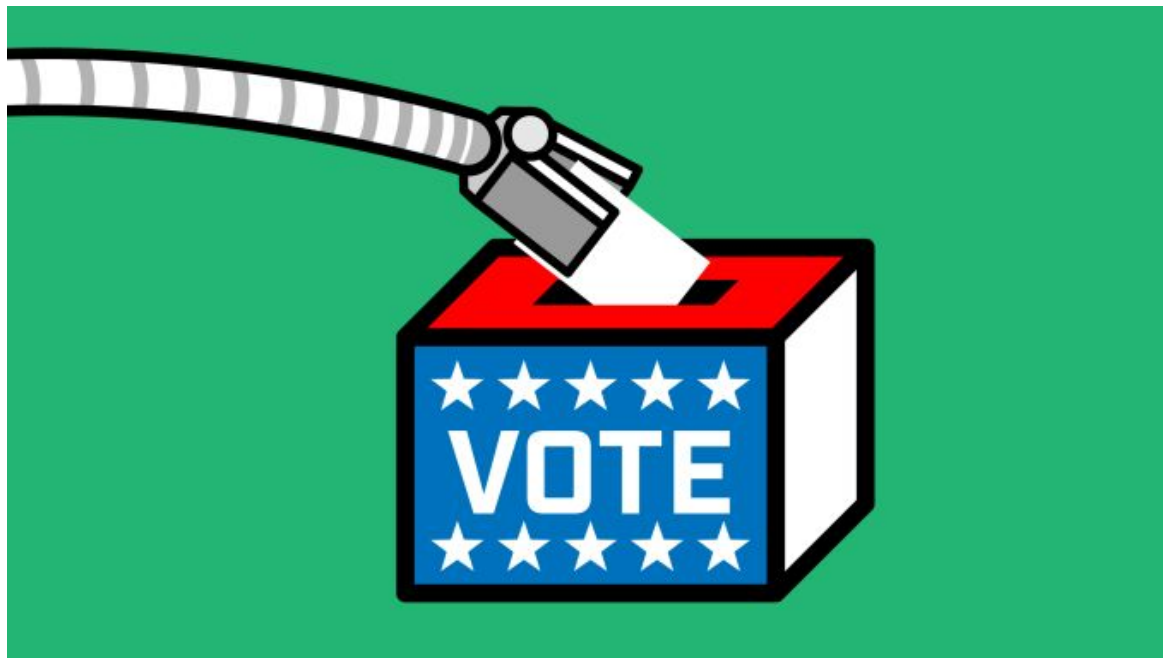
Train results:

F1 score: 0.9808231812766612
Accuracy: 0.9848751212415131
Recall: 0.9694598495783636
Precision: 0.9924560584849899
ROC AUC: 0.9993800791494928

Test results:

F1 score: 0.9504132231404958
Accuracy: 0.9609397112935183
Recall: 0.9308463839521379
Precision: 0.9708203340062397
ROC AUC: 0.9919432963411713

Voting

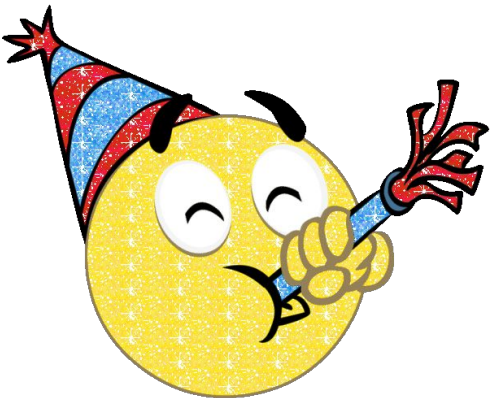


Źródło: <https://techcrunch.com/2019/05/06/microsoft-aims-to-modernize-and-secure-voting-with-electionguard/>

Modele wchodzące w skład modelu wdrożeniowego

1. Regresja logistyczna
2. Drzewo decyzyjne
3. Las losowy
- ~~4. Linear Support Vector Classifier~~
- ~~5. MLPClassifier~~





96.09%



newsów model zaklasyfikował dobrze

Co można zrobić inaczej?

Wyniki modelu ostatecznego po usunięciu
kolumn reuter, cnn, bbc i nazw miesięcy oraz
dni tygodnia

Wyniki modeli bez wcześniej wymienionych kolumn

Logistic Regression:

Train results:

F1 score: 0.9500113955785155
Accuracy: 0.9601067054686553
Recall: 0.9454861636171178
Precision: 0.9545801526717558
ROC AUC: 0.9924548195658993

Test results:

F1 score: 0.9214292142921428
Accuracy: 0.938247152861286
Recall: 0.9126403993581743
Precision: 0.930388949472919
ROC AUC: 0.982201656276749

Decision Tree with parameters:

Train results:

F1 score: 0.8962441855995079
Accuracy: 0.918182369346429
Recall: 0.8813700287312869
Precision: 0.9116289982012982
ROC AUC: 0.9559824553958323

Test results:

F1 score: 0.8353208921256258
Accuracy: 0.8720379146919431
Recall: 0.8179711178463184
Precision: 0.8534226190476191
ROC AUC: 0.8942430827633289

Random Forest with parameters:

Train results:

F1 score: 0.9975746551462786
Accuracy: 0.9980599005699042
Recall: 0.9951610464237108
Precision: 1.0
ROC AUC: 0.9999791963347656

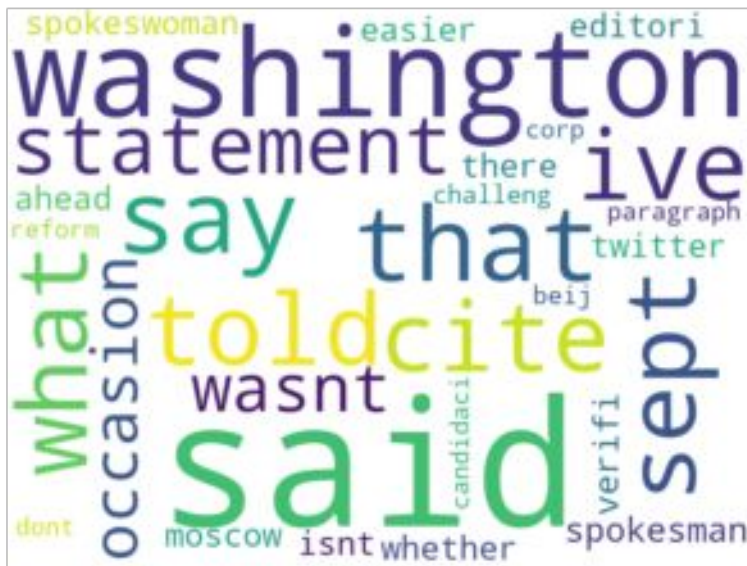
Test results:

F1 score: 0.9128469980632666
Accuracy: 0.9331541345405673
Recall: 0.882331966482439
Precision: 0.9455483377913642
ROC AUC: 0.9810902293018089

Są gorsze... ale może to lepiej?

Regresja Logistyczna

Słowa najbardziej wpływające na klasyfikację wpisu jako **prawdziwy news**



Słowa najbardziej wpływające na klasyfikację wpisu jako **falszywy news**



Drzewo Decyzyjne

Widzimy, że drzewo decyzyjne teraz polega na większej liczbie zmiennej, co uczyni model bardziej stabilnym.



Las Losowy



Wyniki Soft Votingu porównane z wynikami przed poprawkami

przed:

Soft voting:

Train results:

F1 score: 0.9808231812766612

Accuracy: 0.9848751212415131

Recall: 0.9694598495783636

Precision: 0.9924560584849899

ROC AUC: 0.9993800791494928

Test results:

F1 score: 0.9504132231404958

Accuracy: 0.9609397112935183

Recall: 0.9308463839521379

Precision: 0.9708203340062397

ROC AUC: 0.9919432963411713

po:

Soft voting:

Train results:

F1 score: 0.9808434978778725

Accuracy: 0.9848126591487814

Recall: 0.9697565401481929

Precision: 0.9921868956447745

ROC AUC: 0.9993500973240415

Test results:

F1 score: 0.9504207830223198

Accuracy: 0.9616608898634789

Recall: 0.9261900517026208

Precision: 0.9759534097313545

ROC AUC: 0.9922461121013969

Jest lepiej,
mimo że
poszczególne
modele wyszły
gorzej.





96.16%



newsów model po poprawkach zaklasyfikował dobrze

Czas na pytania



Dziękujemy za uwagę :)

