



Модерирование онлайн-дискуссий

Руководитель: Резник Сергей Александрович

Студент: Курочкин Владислав Юрьевич

О проекте (о чем? почему полезен? есть ли аналоги? статус проекта на начало семестра)

Глобальная цель – автоматизация модерирования информационно-коммуникативной среды (определение спама, обсуждений не по теме, блокировка нежелательных пользователей, поиск похожих тем – ориентирование человека, формирование списка часто задаваемых вопросов и т.д.).

В данное время невозможна полная автоматизация модерирования информационно-коммуникативной среды в силу сложности задачи.

На данный момент удалось обучить несколько моделей на нескольких датасетах и сравнить их.

Постановка задачи

Требуется:

- Найти или создать датасет путем парсинга какого-либо сайта. Датасет должен иметь информацию о принадлежности комментариев/сообщений определенным темам (классам) или информацию о том, являются ли комментарии/сообщения спамом;
- Провести обучение наивного байесовского классификатора и логистической регрессии на полученном датасете, сравнить модели.

Что было сделано? SMS-spam

- Наивный байесовский классификатор;
- Логистическая регрессия.

	v1	v2
0	ham	Go until jurong point, crazy.. Available only ...
1	ham	Ok lar... Joking wif u oni...
2	spam	Free entry in 2 a wkly comp to win FA Cup fina...
3	ham	U dun say so early hor... U c already then say...
4	ham	Nah I don't think he goes to usf, he lives aro...

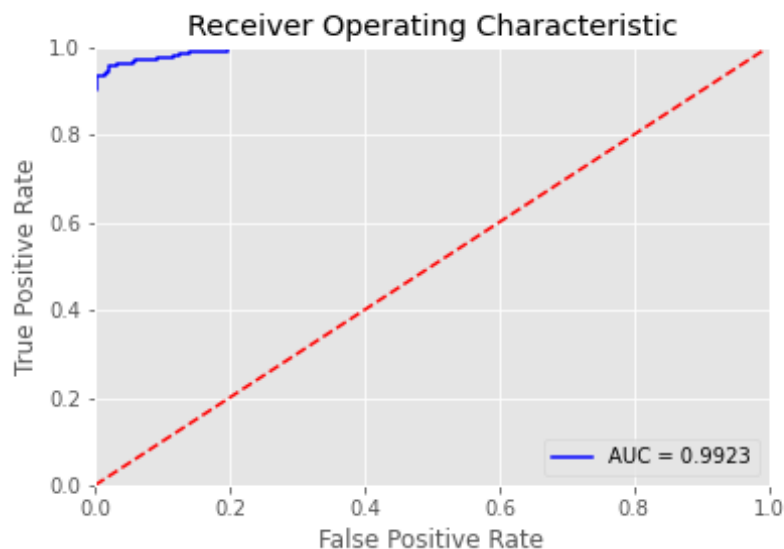
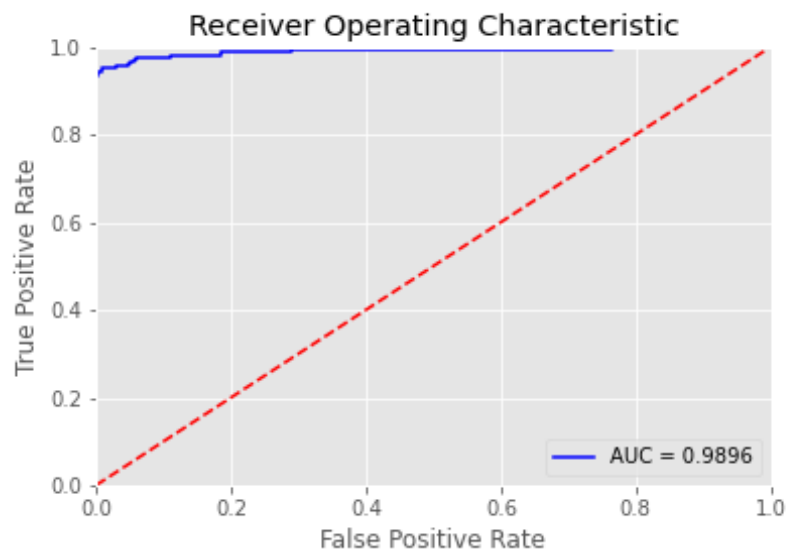
Что было сделано? SMS-spam

Победители: BernoulliNB (CountVectorizer) и LogisticRegression (CountVectorizer).

	0	1		0	1		0	1		0	1
0	3250	0	0	1574	1	0	3250	0	0	1572	3
1	2	481	1	17	247	1	0	483	1	25	239
train, 0.561			test, 0.561			train, 0.5			test, 0.5		

Что было сделано? SMS-spam

Победители: BernoulliNB (CountVectorizer) и LogisticRegression (CountVectorizer).



Что было сделано? SMS-spam

Победители: BernoulliNB (CountVectorizer) и LogisticRegression (CountVectorizer).

	precision	recall	f1-score	support
0	0.999385	1.000000	0.999692	3250.0
1	1.000000	0.995859	0.997925	483.0
accuracy	0.999464	0.999464	0.999464	3733.0
macro avg	0.999692	0.997930	0.998809	3733.0
weighted avg	0.999465	0.999464	0.999464	3733.0

	precision	recall	f1-score	support
0	1.0	1.0	1.0	3250.0
1	1.0	1.0	1.0	483.0
accuracy	1.0	1.0	1.0	3733.0
macro avg	1.0	1.0	1.0	3733.0
weighted avg	1.0	1.0	1.0	3733.0

	precision	recall	f1-score	support
0	0.989315	0.999365	0.994315	1575.0
1	0.995968	0.935606	0.964844	264.0
accuracy	0.990212	0.990212	0.990212	1839.0
macro avg	0.992641	0.967486	0.979579	1839.0
weighted avg	0.990270	0.990212	0.990084	1839.0

	precision	recall	f1-score	support
0	0.984346	0.998095	0.991173	1575.0
1	0.987603	0.905303	0.944664	264.0
accuracy	0.984774	0.984774	0.984774	1839.0
macro avg	0.985974	0.951699	0.967918	1839.0
weighted avg	0.984813	0.984774	0.984496	1839.0



Что было сделано? Spam email

- Наивный байесовский классификатор;
- Логистическая регрессия;
- Дерево решений;
- Случайный лес.

Что было сделано? Spam email

Победитель: RandomForestClassifier.

	0	1
0	1833	20
1	84	1145

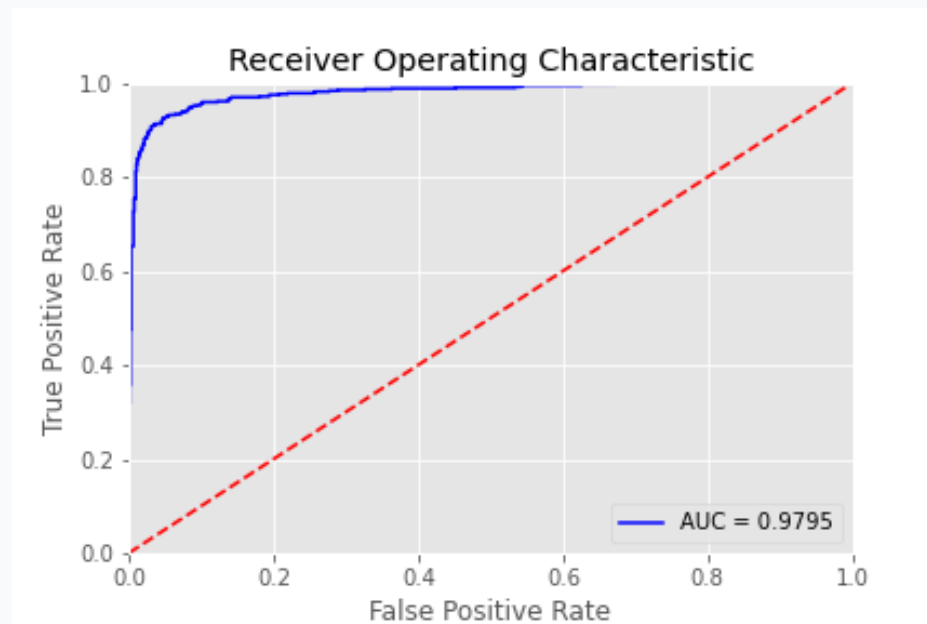
	0	1
0	908	27
1	59	525

	precision	recall	f1-score	support
0	0.956182	0.989207	0.972414	1853.0
1	0.982833	0.931652	0.956558	1229.0
accuracy	0.966256	0.966256	0.966256	3082.0
macro avg	0.969507	0.960429	0.964486	3082.0
weighted avg	0.966809	0.966256	0.966091	3082.0

	precision	recall	f1-score	support
0	0.938987	0.971123	0.954784	935.0
1	0.951087	0.898973	0.924296	584.0
accuracy	0.943384	0.943384	0.943384	1519.0
macro avg	0.945037	0.935048	0.939540	1519.0
weighted avg	0.943639	0.943384	0.943063	1519.0

Что было сделано? Spam email

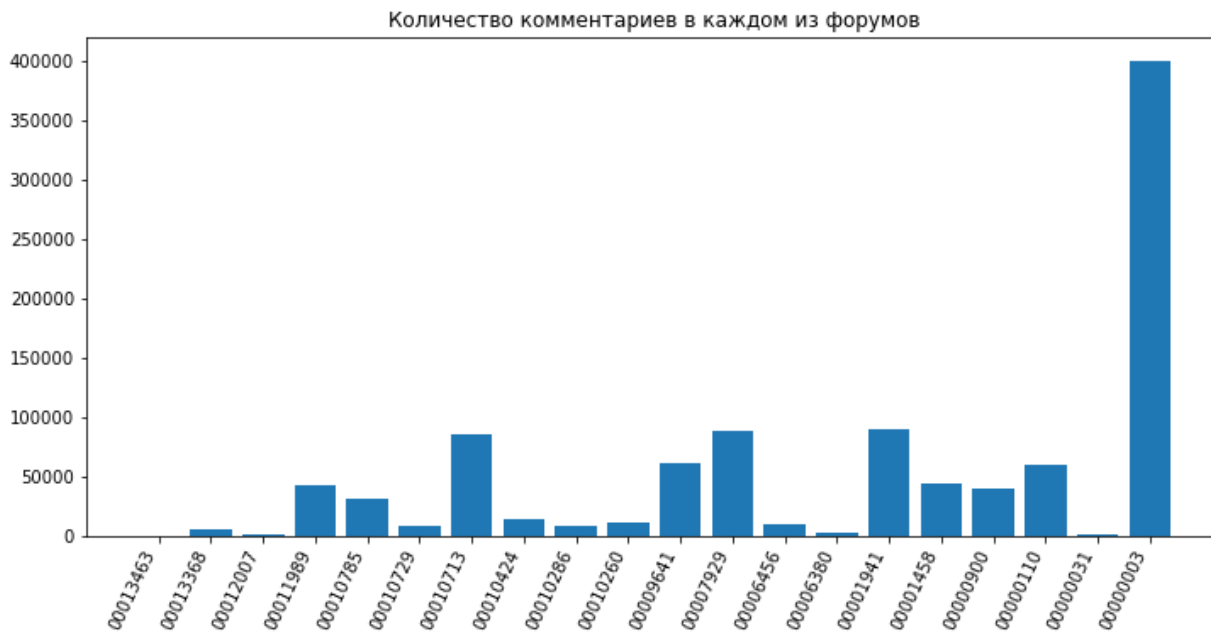
Победитель: RandomForestClassifier.



Что было сделано? dxdu

	Точность на train	Точность на test
0	68.183%	56.430%

Победитель: MultinomialNB (CountVectorizer, stop_words).



Проблемы в ходе выполнения проекта

- Большой объем несбалансированных данных с сайта dxdu;
- Долгое обучение ансамблей на данных, связанных с NLP;
- Данные, полученные путем парсинга, являются очень сырыми.

Заключение (чему научились и какие технологии были использованы? Перспективы проекта)

- Наивный байесовский классификатор и логистическая регрессия – крутые ребята😊
- Не забываем и про деревья!
- Надо пробовать нейронные сети и методы снижения размерности.



Ссылки

- Github: <https://github.com/Vladislav0504/project1.git>;
- SMS-spam: <https://www.kaggle.com/uciml/sms-spam-collection-dataset>;
- Spam email: <https://archive.ics.uci.edu/ml/datasets/spambase>;
- Сайт dxdy: <https://dxdy.ru/ekonomika-i-finansovaya-matematika-f5.html>.



Спасибо за внимание!