

# Fake Reviews Detection

# Towards Undeceived: Fake Reviews Detection (2022)

[Link](#)

- Набор данных с сайта Amazon с 1996 по 2018 г.
- 24 категории (книги, электроника, фильмы и т.д.)
- Каждый товар имеет не менее 5 ОТЗЫВОВ.
- Для обучения объединили имя рецензента, заголовок рецензии и текст рецензии.

TABLE I  
DATA DESCRIPTION

Variable	Description
reviewerID	ID of the reviewer
verified	If the customer's account has been verified
asin	Amazon product ID
reviewerName	name of the reviewer
reviewText	text of the review
overall	rating of the product
summary	summary of the review
unixReviewTime	time of the review(unix time)
reviewTime	time of the review(raw)
image	image link for the review(if any)

TABLE III  
LIST OF DISCLAIMER EXPRESSIONS

Model Name	Recall	Precision	F1-score
BERT	0.81	0.80	0.79
Logistic Regression	0.97	0.97	0.98
Random Forest	0.99	0.97	0.99
Multinomial Naïve Bayes	0.68	0.65	0.80
Bernoulli Naïve Bayes	0.89	0.88	0.89

# Fake Reviews Classification using Deep Learning (2023)

[Link](#)

- Для обработки текста не было предпринято ничего кроме токенизации, поскольку удаление стоп-слов, лемматизация или стемминг потенциально могут удалить характеристики важные для классификаторов.
- Были выделены 5 нетекстовых признаков: рейтинг рецензента, дата публикации, день недели, количество слов, предложений.
- Отдельно использовались еще 7 «поведенческих» признаков:
  - рейтинг рецензента,
  - средний рейтинг для отелей и ресторанов,
  - разница между отдельной оценкой рецензента и средней оценкой рецензента (отклонение рецензента),
  - разница между средним рейтингом рецензента минус рейтинг рецензента для конкретного отеля (отклонение отеля),
  - дата регистрации рецензента,
  - дата публикации рецензии,
  - сравнение местоположения рецензента и отеля/ресторана, который он оценивает (True-False)

# Fake Reviews Classification using Deep Learning (2023)

[Link](#)

TABLE 3 CLASSIFICATION RESULTS OF TEXTUAL MODEL

Domain	Precision	Recall	F1	Accuracy
Restaurants	83.189	83.194	83.279	83.279
Hotels	85.180	79.989	77.625	77.620

Case 1: CNN обученная на тексте

TABLE 4 NON-TEXTUAL RESULTS

Domain	Classifier	Precision	Recall	F1	Accuracy
Restaurants	Log Regression	63	64	61.5	63.22
	SVM	<b>65</b>	<b>64.5</b>	<b>64.5</b>	<b>65.40</b>
	MLP	63	62	63	63.14
Hotels	Log Regression	58.5	58.5	58	58.37
	SVM	<b>64.5</b>	<b>64</b>	<b>64</b>	<b>65.46</b>
	MLP	64	58.512	56.522	59.211

Case 2: Модели обученные только на нетекстовых признаках

TABLE 5 ONLY BEHAVIORAL RESULTS

Domain	Classifier	Precision	Recall	F1	Accuracy
Restaurants	Log Regression	77	75	75	75.35
	SVM MLP	82	79.5	80	79.85
	Log Regression	<b>88.50</b>	<b>87.5</b>	<b>87.5</b>	<b>87.55</b>
Hotels	SVM	73.5	71	70.5	70.83
	MLP	71.5	73.5	73	73.39
		<b>77</b>	<b>76</b>	<b>76</b>	<b>76.28</b>

Case 3: Модели обученные только на поведенческих признаках

TABLE 6 NON-TEXTUAL AND BEHAVIORAL RESULTS

Domain	Classifier	Precision	Recall	F1	Accuracy
Restaurants	Log Regression	76.125	76.5	76.5	77.215
	SVM	77	78	77	76.868
	MLP	<b>89</b>	<b>89</b>	<b>89</b>	<b>88.214</b>
Hotels	Log Regression	78.5	78.5	78	78.20
	SVM	82.5	80.5	80.5	80.44
	MLP	<b>87</b>	<b>86</b>	<b>84.5</b>	<b>85.57</b>

Case 4: Модели обученные на нетекстовых и поведенческих признаках

# Fake Reviews Classification using Deep Learning (2023)

[Link](#)

TABLE 7 RESULTS OF CASE 5

Domain	Precision	Recall	F1	Accuracy
Restaurants	85.321	84.224	84.225	84.167
Hotels	78.122	75.623	75.451	74.198

Case 5: CNN обученная на тексте и нетекстовых признаках

TABLE 9 RESULTS OF CASE 7

Domain	Precision	Recall	F1	Accuracy
Restaurants	88.129	88.310	87.386	87.286
Hotels	82.044	82.236	82.523	82.623

Case 7: CNN обученная на всех признаках

TABLE 8 RESULTS OF CASE 6

Domain	Precision	Recall	F1	Accuracy
Restaurants	86	86.525	84.365	84.404
Hotels	81.512	80.521	77.205	79.160

Case 6: CNN обученная на тексте и поведенческих признаках

TABLE 10 RESULTS OF ENSEMBLE CLASSIFIER

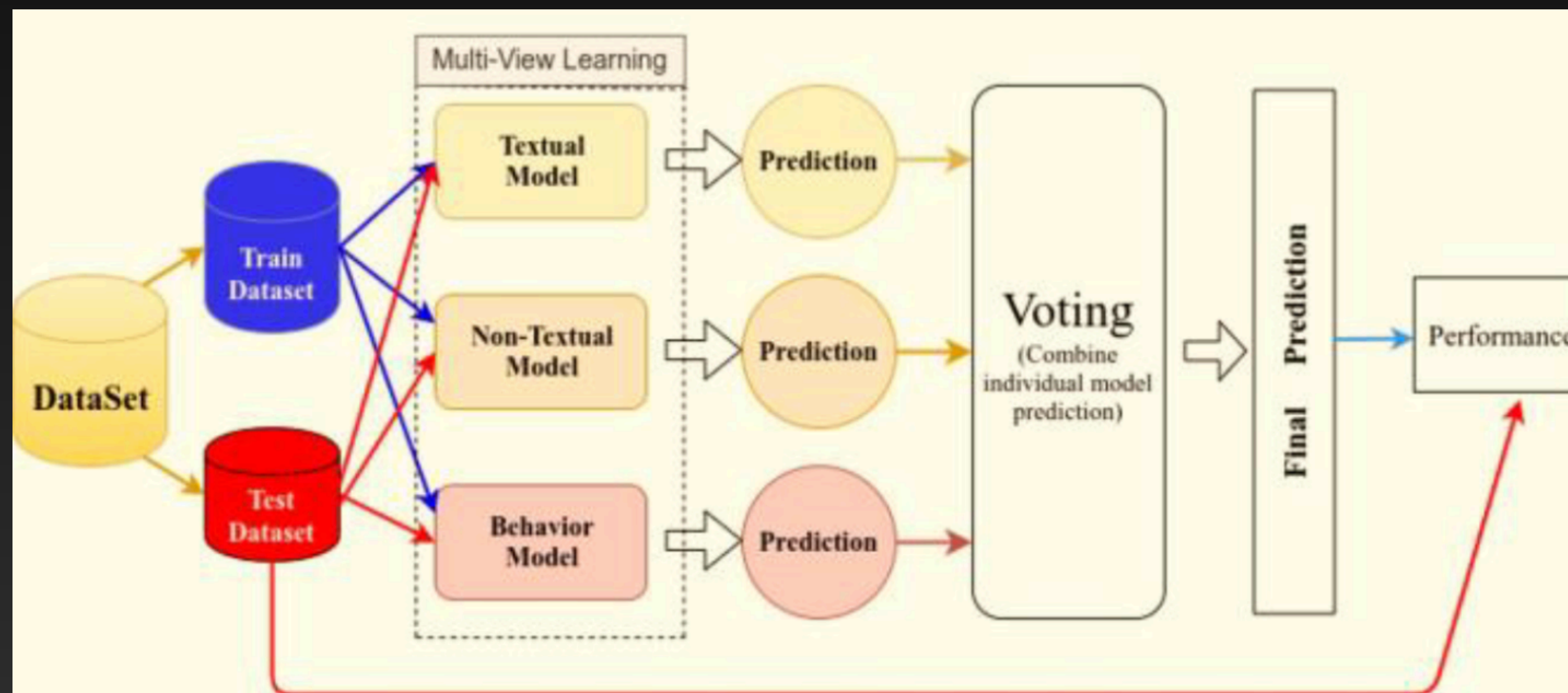
Domain	Precision	Recall	F1	Accuracy
Restaurants	93.317	92.314	92.034	92.742
Hotels	91.366	91.467	91.167	91.266

Case 8: Ансамбль трех моделей, каждая из которых обучалась на своем классе признаков



# Fake Reviews Classification using Deep Learning (2023)

[Link](#)



Case 8: Ансамбль трех моделей

В качестве моделей взяли модели с лучшими показателями для каждого класса:

- Текст - CNN
- Нетекстовые признаки - SVM
- Поведенческие признаки - Log Reg

# Detecting Fake Reviews: Just a Matter of Data (2023)

[Link](#)

Использовался набор данных, содержащий отзывы о приложениях в магазине приложений Apple для iOS.

В качестве модели использовался случайный лес.

Данные были поделены на три части:

- количество отзывов, предоставленных каждым рецензентом
- процент отзывов на одну звезду
- время жизни аккаунта рецензента
- среднее время между всеми отзывами, оставленными каждым рецензентом
- общее количество отзывов, полученных для всех версий каждого приложения
- процентное соотношение полученных отзывов на каждую звезду
- текст
- длина текста в символах

# Detecting Fake Reviews: Just a Matter of Data (2023)

[Link](#)

**Table 1. Feature set combinations.**

FS-Combination 1	TF-IDF
FS-Combination 2	TF-IDF + further review features from Step B (review features)
FS-Combination 3	TF-IDF + review features + product features
FS-Combination 4	TF-IDF + review features + product features + reviewer features

**Table 3. Evaluation metrics.**

Used FS-Combination	Evaluation metrics			
	Accuracy	Recall	F1	AUC
FS-Combination 1	0.7917	0.7411	0.7806	0.7917
FS-Combination 2	0.8092	0.8663	0.8195	0.8092
FS-Combination 3	0.9216	0.9209	0.9215	0.9216
FS-Combination 4	0.9438	0.9430	0.9437	0.9438

