kaggle

Late Submission ···

LLM - Detect AI Generated Text

Identify which essay was written by a large language model

Overview | Data | Code | Models | Discussion | Leaderboard | Rules | Team | Submissions

Overview

In recent years, large language models (LLMs) have become increasingly sophisticated, capable of generating text that is difficult to distinguish from human-written text. In this competition, we hope to foster open research and transparency on AI detection techniques applicable in the real world.

This competition challenges participants to develop a machine learning model that can accurately detect whether an essay was written by a student or an LLM. The competition dataset comprises a mix of student-written essays and essays generated by a variety of LLMs.

Competition Host
The Learning Agency Lab

Prizes & Awards
$110,000
Awards Points & Medals

Participation
5,264 Competitors
4,358 Teams

7th Place Solution in Efficiency Prize
Yellowleaf

Detect LLM Text - Efficiency LB

Notebook   Input   Output   Logs   Comments (58)

EfficiencyRank	TeamName	PrivateScore	ScoringTime	EfficiencyScore	DateSubmitted
1	Kelvin	0.919670	1092	-1.851545	Sat Dec 2 14:23:13 2023
2	Zaakcii Ru	0.917835	1389	-1.838617	Sat Jan 6 18:08:29 2024
3	Beat TFIDF!	0.902804	538	-1.834069	Sat Jan 20 01:44:40 2024
4	SuperFei	0.916737	1673	-1.827600	Mon Jan 22 12:13:19 2024
5	Neko	0.899194	614	-1.824324	Mon Jan 15 03:58:57 2024
6	Cheer	0.912350	1669	-1.818730	Mon Jan 22 15:39:06 2024
7	yellowleaf	0.891007	368	-1.815134	Sat Jan 20 17:43:58 2024

1. Problem description

2. Data

3. Solution

4. Results

5. Important findings

Problem description

- Determine whether a given essay was student written
- or LLM(Large Larguage Model)generated

The mainstream approach in competition is to seek or generate diverse datasets of student writing and large language model-generated data for data augmentation. Many Kagglers have generously shared their own datasets.

 Our approach also involves searching for external open-source data, and we have utilized the following external datasets

Solution

Removing duplicate text and data filtering. This part is the same as the public work.

Tokenization using Byte-pair Encoding. This part is the same as the public work.

Generate TF-IDF vectors for model training. We adjusting parameters such as min_df, max_df, and max_features to effectively reduce the feature dimensionality and minimize the impact of noisy data.

Train the MultinomialNB and stochastic gradient descent. We adjusted the parameter alpha which means the additive smoothing for MultinomialNB and the train steps for SGD.

Ensemble MNB and SGD. We have attempted to combine Multinomial Naive Bayes (MNB) and Stochastic Gradient Descent (SGD) using different weights.

Details of the submission

Version	Public Score	Private Score
1	0.962507	0.861747
2	0.962465	0.853448
3	0.967997	0.891007
4	0.967267	0.898650
5	0.967885	0.890983

kaggle

Reference

1.https://www.kaggle.com/code/batprem/llm-daigt-excluded-prompts?scriptVersionId=158926419

2.https://www.kaggle.com/datasets/thedrcat/daigt-v2-train-dataset

3.https://www.kaggle.com/datasets/alejopaullier/argugpt

4.https://www.kaggle.com/datasets/kagglemini/train-00000-of-00001-f9daec1515e5c4b9

5.https://www.kaggle.com/datasets/pbwic036/commonlit-data

6.https://www.kaggle.com/datasets/wcqyfly/argu-train

7.https://www.kaggle.com/competitions/llm-detect-ai-generated-text/discussion/468908

8.https://www.kaggle.com/competitions/llm-detect-ai-generated-text/discussion/455711

Question and Answer

kaggle