Ikhlaq Ahmad

Dr. Tatiana Erekhinskaya

CS 6320

May 6, 2024

Machine-Generated Text Detection: Another Step Towards Academic Integrity

Introduction:

Machine-generated text is not a new dimension in artificial intelligence. Historically, many such applications like chatbots Eliza have been a subject of experimentation with one goal in mind: Human-like language generation. Recent advances in the machine learning and artificial intelligence domains have made it possible to achieve human-level accuracy in automated language generation. With higher level of accuracy and nearly undifferentiability texts from language models, it is now a serious concern for originality and integrity of work, especially in academia.

Goal:

The goal of this project is to develop a model capable of distinguishing whether a given text is authored by a human or generated by an artificial intelligence (AI) system. Text authorship detection has various practical applications, including content filtering, plagiarism detection, and identifying fake news. In this project, we employ a Convolutional Neural Network (CNN) architecture for text classification. CNNs have proven to be effective in natural language processing tasks, especially in tasks like text classification, sentiment analysis, and language translation.

Dataset:

For training and evaluation, we use a dataset from Kaggle.com named AI_Human.csv containing texts authored by humans and texts generated by AI systems. The dataset is pre-processed to ensure uniformity in text length and format. Each text sample is labeled as either 0 for Human or 1 for AI. The dataset contains around 480000 rows of predetermined essays.

	A	B
1	text	generated
2	Cars. Cars have been around since they became famous in the 1900s, when Henry Ford created and built the first ModelT. Cars have played a major role in our every day	0
3	Transportation is a large necessity in most countries worldwide. With no doubt, cars, buses, and other means of transportation make going from place to place easier and	0
4	"America's love affair with it's vehicles seems to be cooling" says Elisabeth rosenthal. To understand rosenthal's perspective, it is easier to suggest that America's car	0
5	How often do you ride in a car? Do you drive a one or any other motor vehicle to work? The store? To the mall? Have you ever thought about how many people in the world	0
6	Cars are a wonderful thing. They are perhaps one of the worlds greatest advancements and technologies. Cars get us from point a to point i. That is exactly what we want	0
7	The electrol college system is an unfair system, people don't have the right to select their own president, they dont have the right to select a president. Because, when	0
8	Dear state senator, It is the utmost respect that I ask for the method for presidential election be changed from the electoral college to something more suitable for the	0
9	Fellow citizens, cars have become a major role in our daily lives. They have their many excellent uses, however there are advantages of limiting that usage. To name a few	0
10	"It's official: The electoral college is unfair, outdated, and irrational" Plumer, Source 2. Many do not like the electoral college for these reasons and many others such as it	0
11	The Electoral College has been kept for centuries, established by the founding fathers and established in the Constitution. I think that the process should be maintained	0
12	Dear senator, Retain the Electoral College. The Electoral College consists of 538 electors and a majority of 270 electors is is required to elect the President. Each state	0
13	"When I had a car I was always tense. I'm much happier this way." Heidrun Walter source 1. For many years, people have depended on cars to take them places and allow	0
14	Voting for a president is one of the most important decisions you can make. This person you vote for will be changing the country for better, or for worse. Electoral colleges	0
15	dear senator, I have come to a conclusion on how I think the president and his candidates should be voted for. In one hand we've got the Electoral College, and the	0
16	Everyone wants to go to the park with their children or grandchildren, or maybe just take a nice walk. With the way we live now everything will be too polluted because	0
17	There are many advantages to limiting car usage in our community. Other countries such as France, Germany, and Colombia are home to cities that are working towards	0
18	dear state senator, i believe that we should abolish the electoral college and just leave up to popular vote. The electoral college should be abolished because it is unfair,	0
19	Transportation has become one of the largest emissions throughout the world, and many do not seem to wonder how beneficial limiting car usage can be. Not only does	0
20	We live in a world where the only way to get from place to place is to use a car. If we were to limit the amount of car usage all around the world we would be living in a	0
21	The long list of benefits that comes from limited car usage are mostly based on where you live and how much you do your part. Laws and restrictions have been arriving	0
22	The system of the Electoral College is a widely argued debate as to if it should be continued or if it should be gotten rid of. The electoral system shouldn't be used as a	0
23	Dear Senator, Many people might agree theat we shouldn't have an Electoral College at all, since most of thee time it doesn't seem to work at all since of instead just	0
24	In the old world, people didn have car's, they did everything on foot. They got food, ran errands, and just to get out of the house with the family. So how come now adays	0
25	.	0
26	Dear Florida State Senator, Although many could argue that the Electoral College is a fair system of deciding the president, far too many people consider it	0
27	Thousands of people around the world use cars. It's an allaround good mode of transportation. Unfortunately, the fumes that cars emit are harmful for the environment.	0
28	Although numerous amounts of people want to banish the Electoral College, there are more rights than wrongs about it. The Electoral College is in the Constitution for an	0
29	There are many different things that you could do to help out the ecosystem, and reducing your car use is certainly one of them. A study conducted in Europe shows that	0

35403	Title: The Face on Mars: A Mysterious Discovery	1
35404	Title: Understanding the Facial Action Coding System: Decoding Human Emotions	1
35405	Driverless cars, also known as autonomous vehicles, are a rapidly advancing technology that has the potential to revolutionize the way we travel. These self-driving cars	1
35406	Title: A Cowboy Who Rode the Waves	1
35407	Title: Car-free Cities: A Sustainable and Healthy Alternative for Urban Living	1
35408	The Electoral College is a complex and sometimes controversial part of the United States' presidential election process. It was designed as a compromise between	1
35409	Title: Exploring Venus	1
35410	Title: The Mysterious Face on Mars: A Middle School Student's Perspective	1
35411	The Facial Action Coding System (FACS) is an essential tool for understanding and analyzing human facial expressions. Developed by Dr. Paul Ekman, a prominent expert	1
35412	Driverless cars, also known as self-driving or autonomous vehicles, are a revolutionary concept in the world of transportation. These cars use advanced technologies	1
35413	Have you ever wondered what it would be like to surf? It's a thrilling, adrenaline-pumping experience where the ocean meets the shore, and you become one with the	1
35414	Car-free cities are an innovative concept gaining traction in the world today, as they aim to promote environmentally sustainable and healthy lifestyles while reducing	1
35415	The Electoral College is a system used in the United States to elect the President and Vice President. It consists of 538 electors who represent the 50 states and the	1
35416	Title: Exploring Venus	1
35417	Title: The Face on Mars	1
35418	The Facial Action Coding System (FACS) is a method for analyzing and categorizing the movements and expressions of the human face. It was developed by Dr. Paul	1
35419	Driverless cars have gained immense popularity in recent years due to the promise of increased safety, reduced traffic, and improved mobility. These vehicles, also	1
35420	A Cowboy Who Rode the Waves	1
35421	Title: Car-Free Cities: A Future Sustainable Solution	1
35422	The electoral college is a unique system used in American presidential elections to choose the president and the vice president. It was designed with the intention to	1
35423	Title: Exploring Venus: A Journey Through the Mysteries of Earth's Twin Planet	1
35424	In 1976, the Viking 1 spacecraft sent back pictures of a rock formation on Mars. This rock formation, located in a region known as Cydonia Mensae, resembled a human	1
35425	The Facial Action Coding System (FACS) is a method of categorizing and identifying human facial movements. Developed by Dr. Paul Ekman and Dr. Wallace Friesen,	1
35426	Title: The Future of Transportation: A Look into Driverless Cars	1
35427	Title: A Cowboy Who Rode the Waves	1
35428	Car-free cities, also known as "car-lite" or "car-free" areas, have gained immense popularity over the years due to their various benefits and the growing concern for the	1
35429	The electoral college is a system used by the United States to elect the President and Vice President. It has been a subject of debate and controversy for many years, with	1
35430	Title: Exploring Venus	1
35431	The Face on Mars is a fascinating topic for many middle and high school students who are interested in space exploration, astronomy, or just plain curiosity. This	1
35432	Title: Facial Action Coding System - Analyzing Human Emotions through Facial Expressions	1
35433	Driverless cars, also known as autonomous vehicles or self-driving cars, have been a hot topic of discussion in the transportation industry for quite some time. These	1
35434	Title: A Cowboy Who Rode the Waves	1
35435	Title: Embracing Car-Free Cities: The Benefits and Challenges Ahead	1
35436	The electoral college is a unique method used in the United States to elect the President and Vice President. It consists of 538 electors, with each state getting a specific	1
35437	Title: Exploring Venus	1
35438	Title: The Face on Mars	1

Model Architecture:

The Convolutional Neural Network (CNN) architecture chosen for this text authorship detection task is designed to effectively capture and learn hierarchical features from textual data. The architecture comprises several layers arranged in a sequential manner:

- **Embedding Layer:** The input text data was first passed through an embedding layer, which converted the text into dense vector representations called word embeddings. These embeddings captured semantic relationships between words and enable the model to learn contextual information.
- **Convolutional Layers:** Convolutional layers consists of multiple filters or kernels that slide across the input embeddings to extract features. Each filter performs a convolution

operation, capturing different patterns or features present in the text. The use of multiple filters with varying sizes allows the model to learn both local and global features.

- **Activation Function:** After each convolution operation, an activation function (such as ReLU - Rectified Linear Unit) is applied elementwise to introduce non-linearity into the model, enabling it to learn complex relationships within the data.
- **Max Pooling Layers:** Following the convolutional layers, max pooling layers were employed to down sample the feature maps obtained from the convolutional operations. Max pooling retains the most significant information from each feature map while reducing the dimensionality of the data, which helps in preventing overfitting and improving computational efficiency.
- **Flatten Layer:** The output feature maps from the max pooling layers are flattened into a one-dimensional vector, which serves as the input to the subsequent fully connected layers.
- **Dense Layers:** Fully connected dense layers are added to the model to perform classification based on the learned features. These layers enable the model to learn complex decision boundaries and mappings between the input features and output labels.
- **Output Layer:** The final layer of the network is a single neuron with a sigmoid activation function, which produces a probability score indicating the likelihood of the input text being authored by a human or generated by an AI system. The output value close to 1 signifies high confidence in Human authorship, while a value close to 0 indicates high confidence in AI authorship.

The overall architecture of the CNN model is designed to efficiently process textual data, automatically learn relevant features, and make accurate predictions regarding the authorship of

the input text. By leveraging the hierarchical nature of convolutional operations, the model can effectively capture patterns and relationships within the text, enabling it to discriminate between human and AI-generated content.

Training

Training the Convolutional Neural Network (CNN) model for text authorship detection involves several key steps aimed at optimizing the model's parameters and improving its performance on the task. The training procedure can be outlined as follows:

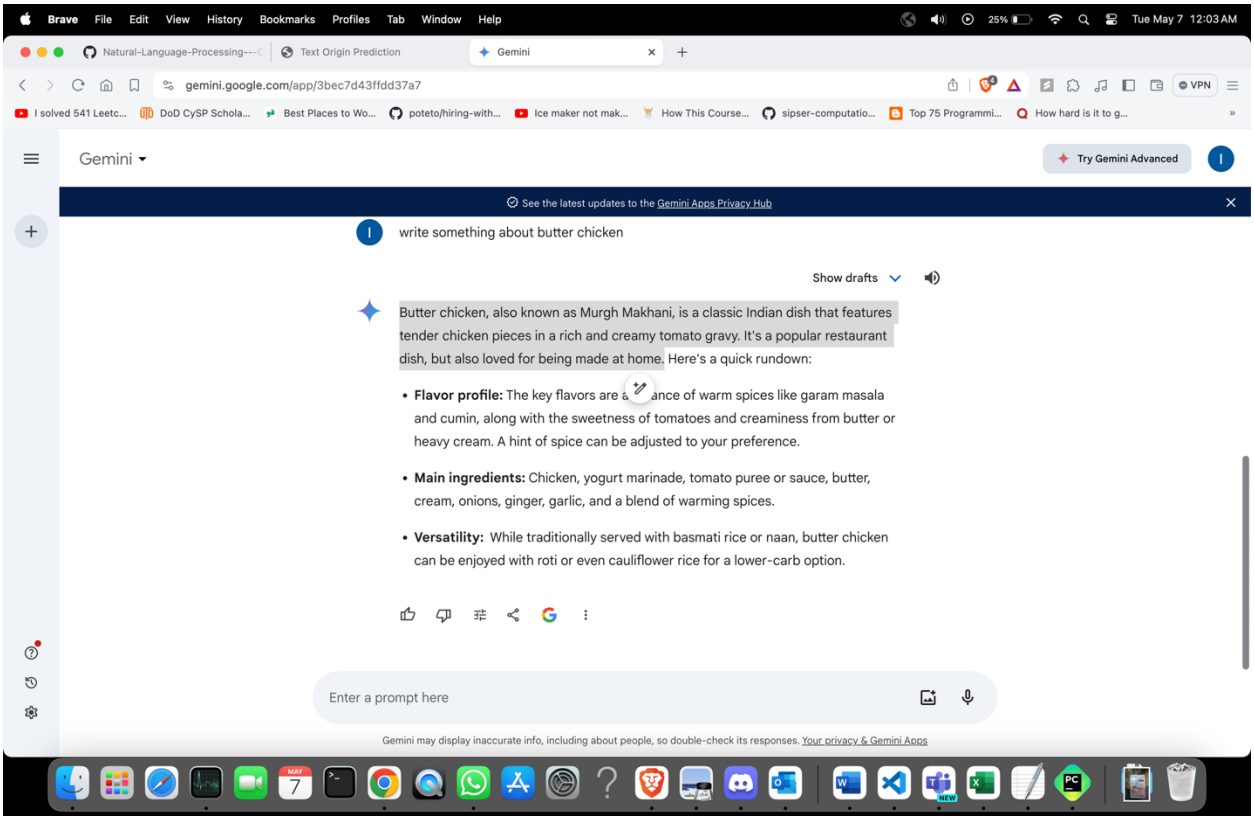
- **Data Preprocessing:** Before training the model, the dataset was preprocessed to ensure uniformity and compatibility with the model architecture. This preprocessing included tokenization, padding, and vectorization of the text data.
- **Data Splitting:** The dataset was divided into three subsets: training, validation, and test sets. The training set was used to train the model, the validation set was used to tune hyperparameters and monitor performance during training, and the test set was used to evaluate the final performance of the trained model.
- **Tokenization:** Text data was tokenized to convert words into numerical representations. This involves creating a vocabulary of unique words present in the dataset and mapping each word to an integer index.
- **Model Compilation:** The model was compiled with appropriate hyperparameters, including the choice of loss function, optimizer, and evaluation metrics. For binary classification tasks like text authorship detection, binary cross-entropy loss and Adam optimizer are commonly used choices.
- **Training:** The model is trained on the training data using the fit () method. During training, the model iteratively adjusts its parameters (weights and biases) based on the

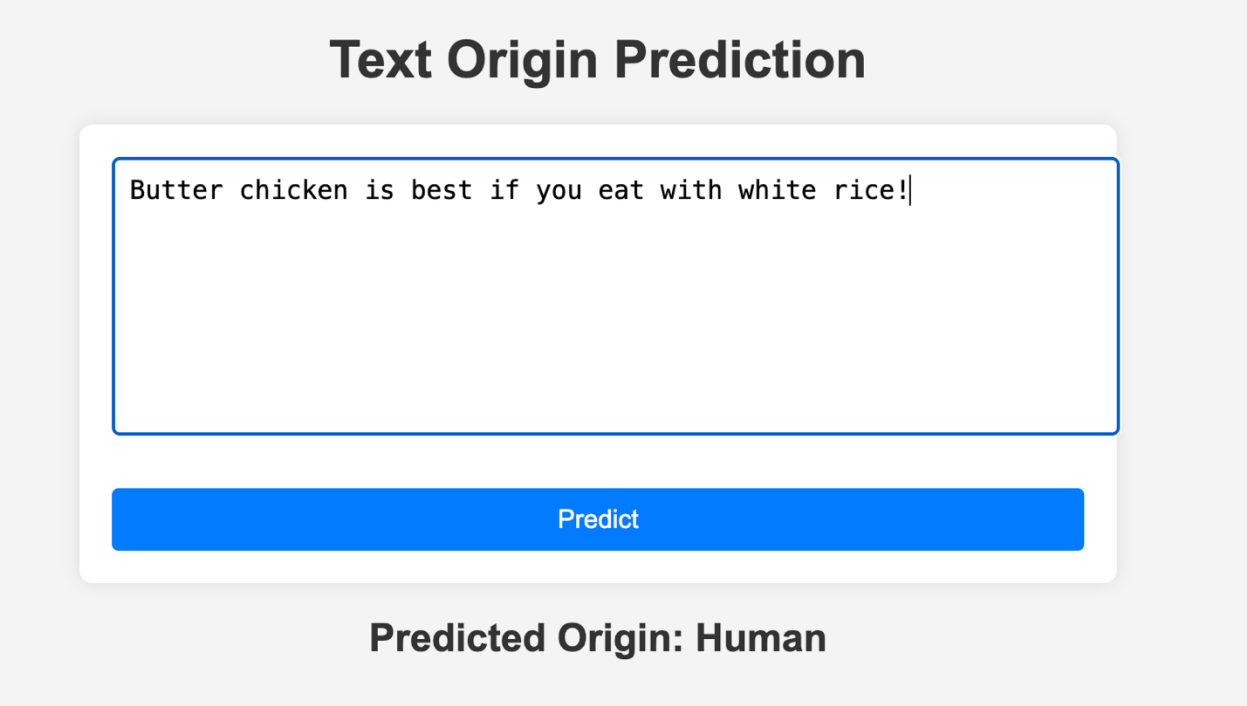
gradients of the loss function with respect to these parameters. The training process involves multiple epochs, where each epoch corresponds to one pass through the entire training dataset.

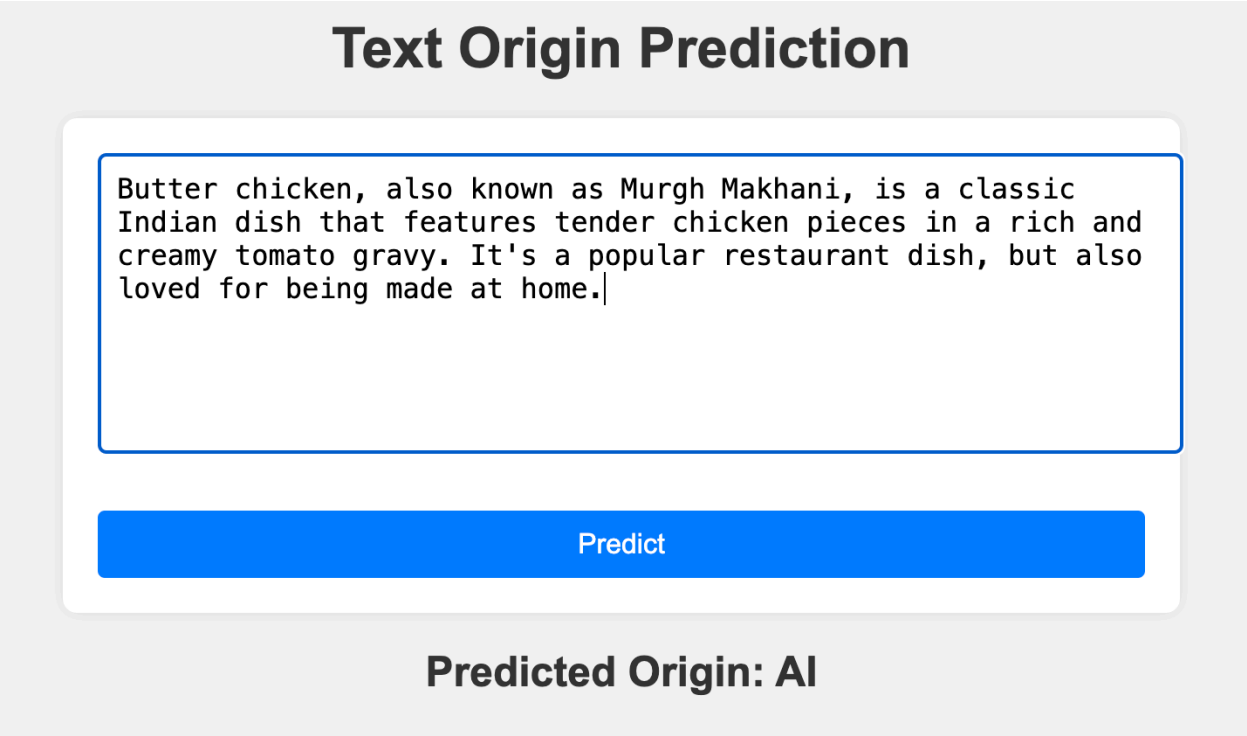
- **Validation:** Model performance is monitored on the validation set after each epoch to assess its generalization ability and prevent overfitting.
- **Early Stopping:** To prevent overfitting, early stopping can be employed, where training was halted if the performance on the validation set does not improve for a certain number of epochs.
- **Hyperparameter Tuning:** Hyperparameters such as learning rate, batch size, and number of layers can significantly impact the model's performance. These hyperparameters are tuned using techniques like grid search or random search to find the optimal configuration.
- **Evaluation:** Once training is complete, the final performance of the trained model is evaluated on the test set, which provides an unbiased estimate of its performance on unseen data.
- **Model Saving:** Finally, if the trained model meets the desired performance criteria, it can be saved to disk for future use or deployment in production environments.

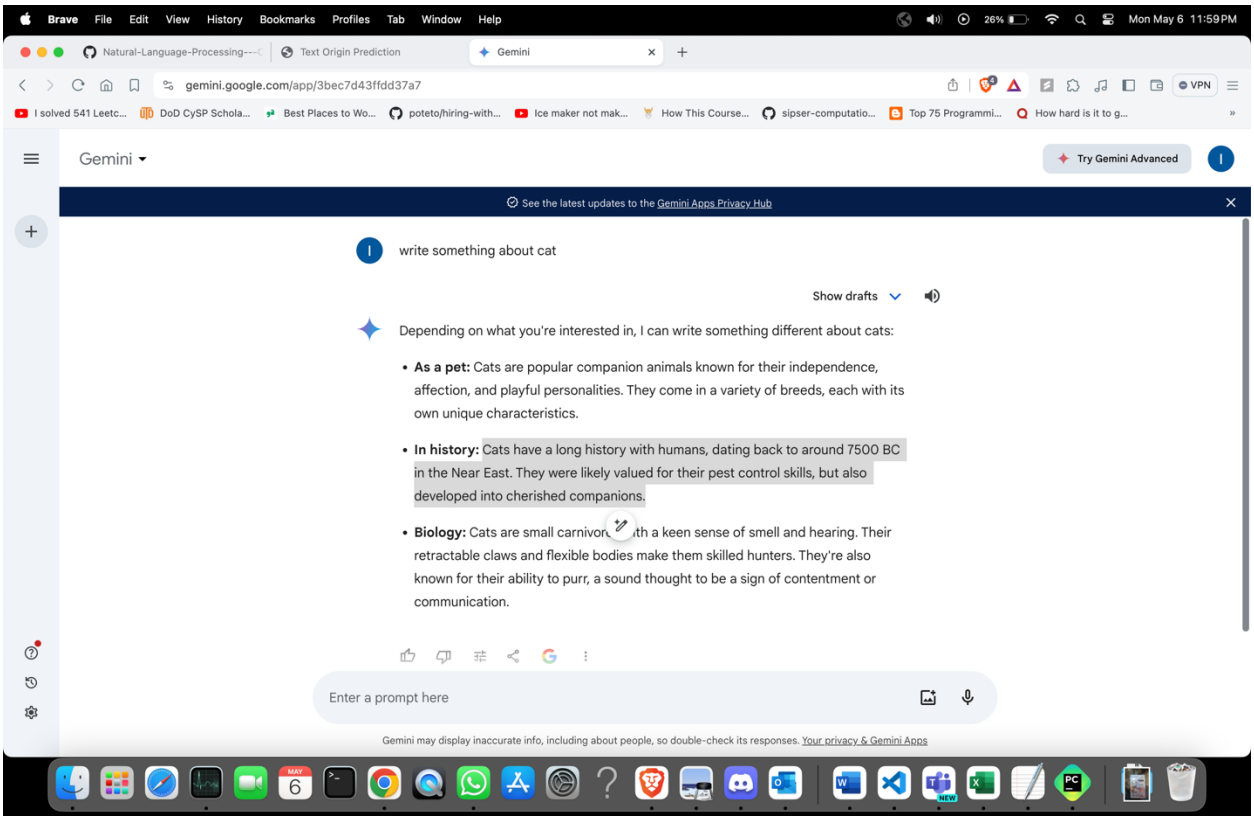
By following these steps, we can train a CNN model for text authorship detection that achieves high accuracy and generalization ability on unseen data. Iterative refinement of the model architecture and hyperparameters may be necessary to achieve optimal performance.

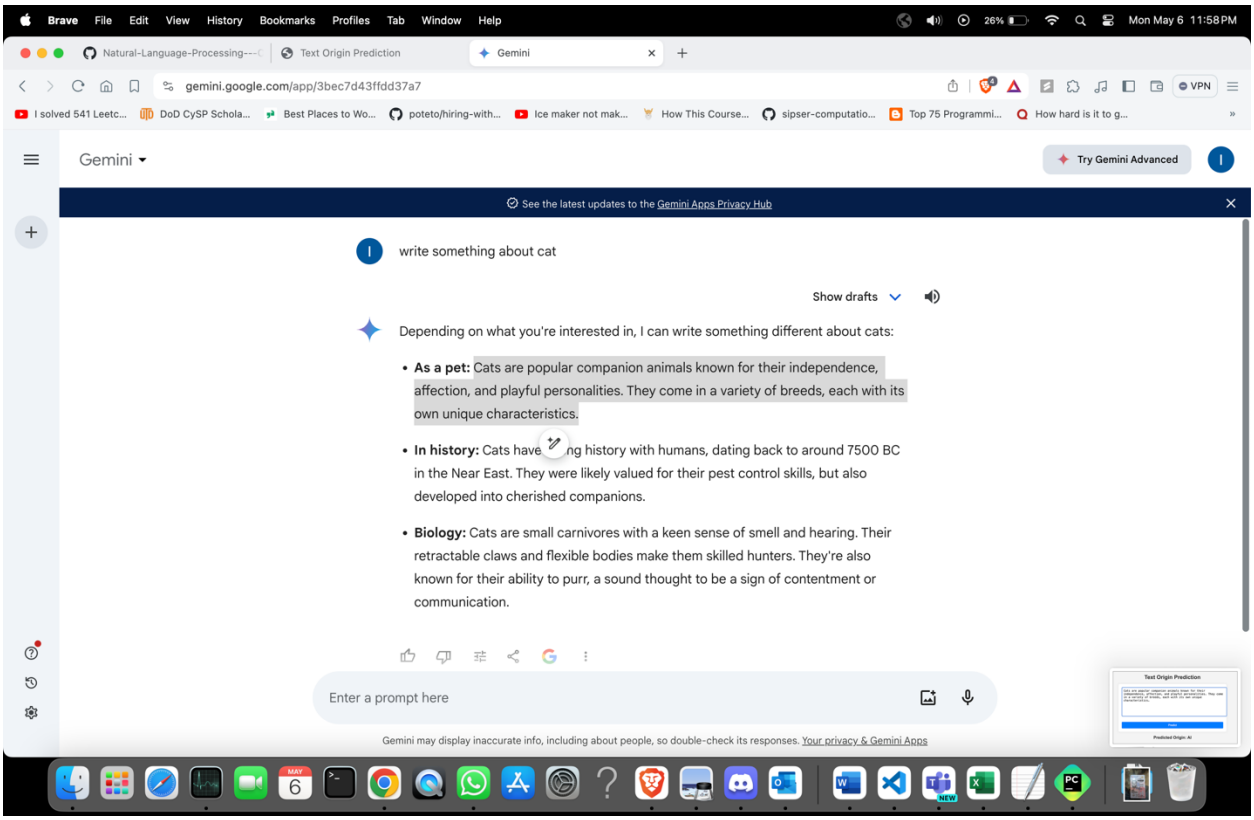
Sample Run

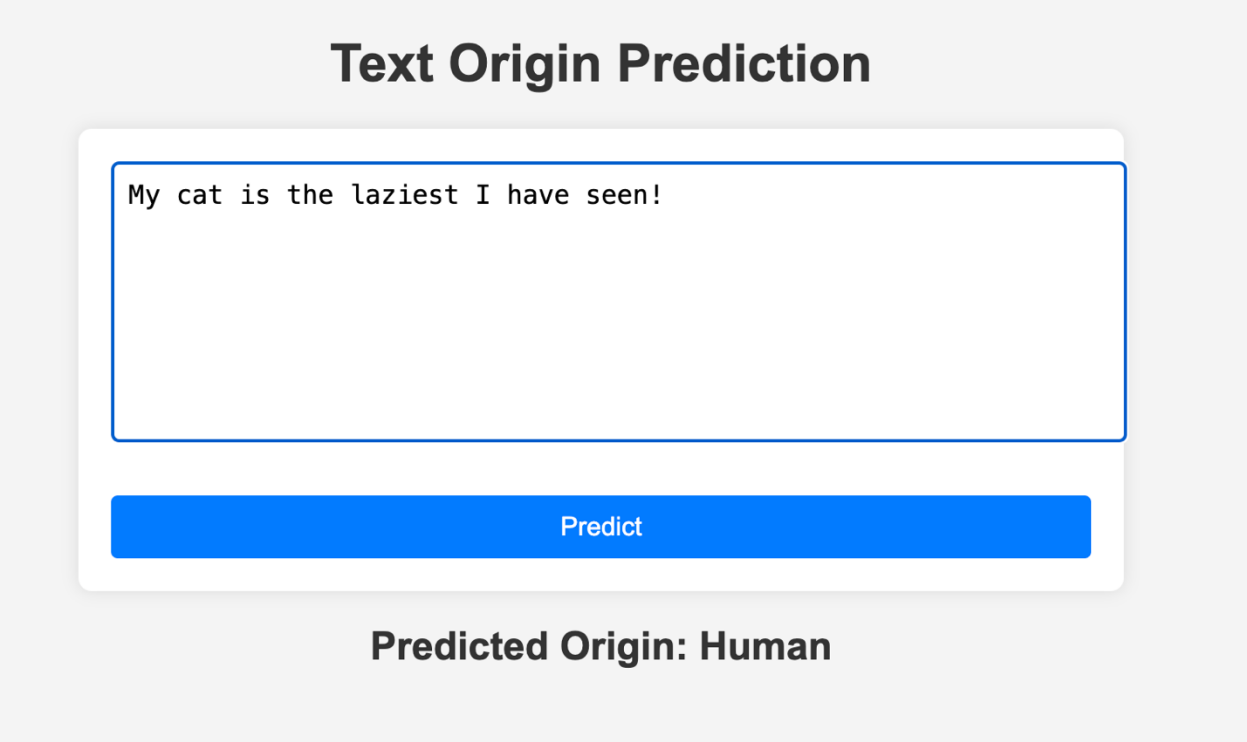

Brave File Edit View History Bookmarks Profiles Tab Window Help
Tue May 7 12:03 AM
gemini.google.com/app/3bec7d43ffdd37a7
Gemini
Try Gemini Advanced
See the latest updates to the Gemini Apps Privacy Hub
write something about butter chicken
Show drafts
Butter chicken, also known as Murgh Makhani, is a classic Indian dish that features tender chicken pieces in a rich and creamy tomato gravy. It's a popular restaurant dish, but also loved for being made at home. Here's a quick rundown:
• Flavor profile: The key flavors are a ance of warm spices like garam masala and cumin, along with the sweetness of tomatoes and creaminess from butter or heavy cream. A hint of spice can be adjusted to your preference.
• Main ingredients: Chicken, yogurt marinade, tomato puree or sauce, butter, cream, onions, ginger, garlic, and a blend of warming spices.
• Versatility: While traditionally served with basmati rice or naan, butter chicken can be enjoyed with roti or even cauliflower rice for a lower-carb option.
Enter a prompt here
Gemini may display inaccurate info, including about people, so double-check its responses. Your privacy & Gemini Apps

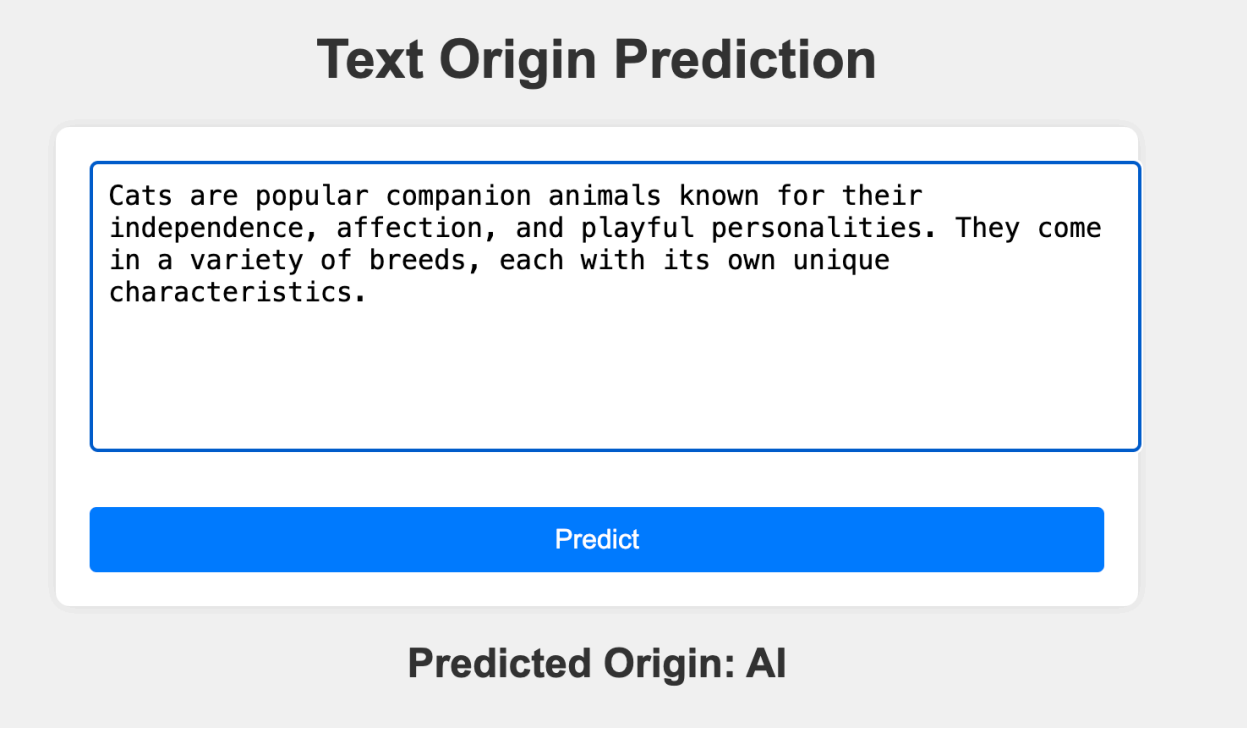

Text Origin Prediction
Butter chicken is best if you eat with white rice!
Predict
Predicted Origin: Human

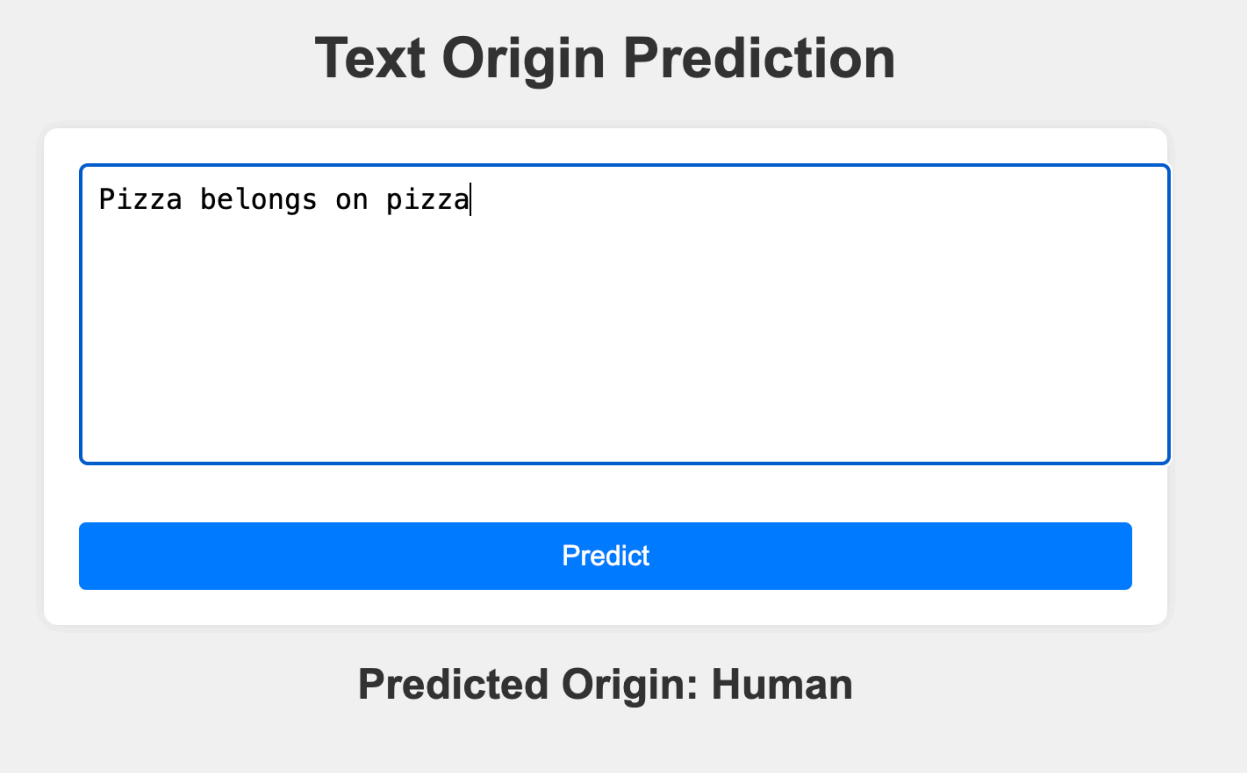

Text Origin Prediction
Butter chicken, also known as Murgh Makhani, is a classic Indian dish that features tender chicken pieces in a rich and creamy tomato gravy. It's a popular restaurant dish, but also loved for being made at home.
Predict
Predicted Origin: AI


Brave File Edit View History Bookmarks Profiles Tab Window Help
Mon May 6 11:59 PM
gemini.google.com/app/3bec7d43ffdd37a7
Gemini
Try Gemini Advanced
See the latest updates to the Gemini Apps Privacy Hub
write something about cat
Show drafts
Depending on what you're interested in, I can write something different about cats:
As a pet: Cats are popular companion animals known for their independence, affection, and playful personalities. They come in a variety of breeds, each with its own unique characteristics.
In history: Cats have a long history with humans, dating back to around 7500 BC in the Near East. They were likely valued for their pest control skills, but also developed into cherished companions.
Biology: Cats are small carnivores with a keen sense of smell and hearing. Their retractable claws and flexible bodies make them skilled hunters. They're also known for their ability to purr, a sound thought to be a sign of contentment or communication.
Enter a prompt here
Gemini may display inaccurate info, including about people, so double-check its responses. Your privacy & Gemini Apps


Brave File Edit View History Bookmarks Profiles Tab Window Help
Mon May 6 11:58 PM
gemini.google.com/app/3bec7d43ffdd37a7
Gemini
Try Gemini Advanced
See the latest updates to the Gemini Apps Privacy Hub
write something about cat
Show drafts
Depending on what you're interested in, I can write something different about cats:
As a pet: Cats are popular companion animals known for their independence, affection, and playful personalities. They come in a variety of breeds, each with its own unique characteristics.
In history: Cats have a long history with humans, dating back to around 7500 BC in the Near East. They were likely valued for their pest control skills, but also developed into cherished companions.
Biology: Cats are small carnivores with a keen sense of smell and hearing. Their retractable claws and flexible bodies make them skilled hunters. They're also known for their ability to purr, a sound thought to be a sign of contentment or communication.
Enter a prompt here
Gemini may display inaccurate info, including about people, so double-check its responses. Your privacy & Gemini Apps


Text Origin Prediction
My cat is the laziest I have seen!
Predict
Predicted Origin: Human


Text Origin Prediction
Cats are popular companion animals known for their independence, affection, and playful personalities. They come in a variety of breeds, each with its own unique characteristics.
Predict
Predicted Origin: AI



Limitations:

I used epoch = 4 which is very low compared to epoch = 10 at least. Due to the large dataset and limited computing power, it wasn't possible to compute using more layers. Additional computing power and larger dataset can enhance the quality of the model and accuracy of predictions.