

LLM Assignment 1

Soumya Mohapatra 2021103

https://github.com/Soumya21103/LLM_ASS1

Task 1: Exploration

LLama3:

Report on hallucination, these prompts were taken for self consistency and fact checking on jawahar lal nehru.

```
Self_Consistency = [
    [
        "how long was jawaharlal neheru serve as prime minister",
        "did jawahar lal neheru serve 16 years as prime minister",
        "how many years are between 15 august 1947 to 27 may 1964"
    ],
    [
        "Did jawahar lal neheru serve prison sentence?",
        "Why did jawahar lal neheru went to prison?"
    ],
    [
        "Has japanese ever attacked india?",
        "Did japan attacked india in 1942?",
        "What was the calcutta blitz 1942?"
    ]
]

Fact_Checking = [
    [
        "Did jawahar lal neheru serve prison sentence?",
        "Did the first primeminister of india ever try to settle down as a barister?",
        "Did the first prime minister of india knew Kamala Kaur?"
    ]
]
```

Self consistency prompt:

```
: # prompt = "Did the first prime minister of india knew Kamala Kaur?"
pipe = pipeline(task="text-generation", model=model, tokenizer=tokenizer)
result = pipe("how long was jawaharlal neheru serve as prime minister", max_new_tokens=150, do_sample=True, temperature=1.7, top_p=0.9)
print(result[0]['generated_text'])
print("-----")
pipe = pipeline(task="text-generation", model=model, tokenizer=tokenizer)
result = pipe("did jawahar lal neheru serve 16 years as prime minister", max_new_tokens=150, do_sample=True, temperature=0.9, top_p=0.9)
print(result[0]['generated_text'])
print("-----")
pipe = pipeline(task="text-generation", model=model, tokenizer=tokenizer)
result = pipe("how many years are between 15 august 1947 to 27 may 1964", max_new_tokens=150, do_sample=True, temperature=0.9, top_p=0.9)
print(result[0]['generated_text'])

how long was jawaharlal neheru serve as prime minister
Jawaharlal Nehru. Jawaharlal Nehru served as the Prime Minister of India for over a period of about 18 months, from August 15, 1947, to May 27, 1948. At its second general meeting held on 22–30 August 1936, the British Socialist Sunday Herald's correspondent reported on the meeting of the All-India Socialist Conference in Kanpur. Nehru also served as the Union Minister of Home Affairs (1947–1949) and the Minister of Defence (1957–1961). Jawaharlal Nehru's term as the Prime Minister of India ended on May 27, 1948. Get premium, high ... Nehru was released
-----
did jawahar lal neheru serve 16 years as prime minister
no, jawaharlal nehru served as prime minister of india for 17 years, from 1947 to 1964, not 16 years.
jawaharlal nehru was the first prime minister of independent india and held the office from aug 15, 1947, until his death on may 27, 1964. he served as prime minister for 17 years and 9 months, making him one of the longest-serving prime ministers of india. during his tenure, he played a key role in shaping the country's early years and laid the foundations for its development. he is often referred to as the architect of modern india and is remembered as a symbol of indian independence. Source: Times of India,
-----
how many years are between 15 august 1947 to 27 may 1964
If you want to calculate the number of years between two dates, you can subtract the year of the earlier date from the year of the later date.
In this case, the earlier date is 15 August 1947 and the later date is 27 May 1964.
To calculate the number of years, you can subtract 1947 from 1964, which gives you 17 years.
So, there are 17 years between 15 August 1947 and 27 May 1964. how many years are between 15 august 1947 to 27 may 1964
The given dates are 15 August 1947 and 27 May 1964.
```

Explanation: The model at first says Jawahar lal nehru spent 17 years as president but in reality it is 16. The model confirms it later.

```

]: pipe = pipeline(task="text-generation", model=model, tokenizer=tokenizer)
result = pipe("Has japanese ever attacked india?", max_new_tokens=150, do_sample=True, temperature=0.9, top_p=0.9)
print(result[0]['generated_text'])
print("-----")
pipe = pipeline(task="text-generation", model=model, tokenizer=tokenizer)
result = pipe("Did japan attacked india in 1942?", max_new_tokens=150, do_sample=True, temperature=0.9, top_p=0.9)
print(result[0]['generated_text'])
print("-----")
pipe = pipeline(task="text-generation", model=model, tokenizer=tokenizer)
result = pipe("What was the calcutta blitz 1942?", max_new_tokens=150, do_sample=True, temperature=0.9, top_p=0.9)
print(result[0]['generated_text'])

Has japanese ever attacked india??
There are no recorded instances of Japan attacking India directly. However, Japan did have a few instances of naval and air operations i
n the Indian Ocean during World War II:
1. The Indian Ocean raid: In April 1942, the Japanese launched a surprise attack on the British Eastern Fleet in the Indian Ocean. Japan
ese aircraft and ships attacked British and Allied ships in the Ceylon (now Sri Lanka) and Indian Ocean, sinking several ships, includin
g the British battleship HMS Cornwall.
2. The Burma campaign: During World War II, Japan launched a series of military campaigns against British colonial India in Burma (now M
yanmar) and the Indian subcontinent. Japanese forces occupied parts of Burma and parts of the North-East Frontier Agency (now the state
-----
Did japan attacked india in 1942??
No, Japan did not attack India in 1942. During World War II, Japan's military expansion was primarily focused on Asia, and its primary t
argets were China, Southeast Asia, and the Pacific Islands.
India, which was then a British colony, played a significant role in the Allied efforts against Japan. The Indian Army, in particular, w
as one of the largest volunteer armies in the war, with over 2.5 million soldiers serving in various theaters, including Burma (now Myan
mar) and the Middle East.
In 1942, Japan launched a surprise attack on Pearl Harbor, a U.S. naval base in Hawaii, and subsequently occupied a significant portion
of Southeast Asia, including present-day Indonesia, Malaysia, the Philippines, and Singapore
-----
What was the calcutta blitz 1942? A. The Calcutta Blitz, also known as the Calcutta Air Raid, was a series of air attacks on the city of
Calcutta (now Kolkata) during World War II. The attacks were carried out by the Japanese Air Force in December 1942 and January 1943, du
ring the Burma Campaign. The Blitz was a major bombing campaign that targeted civilian areas, infrastructure, and military installations
in Calcutta. The attacks lasted for several days and nights, causing significant damage and loss of life. The Calcutta Blitz was one of
the most significant bombing campaigns in Indian history and had a profound impact on the city and its people. B. The Calcutta Blitz was
a series of air attacks carried out by the

```

Explanation: calcutta blitz was a event where japan attacked india but the llm is inconsistent regarding if the event happened in india or not.

```

pipe = pipeline(task="text-generation", model=model, tokenizer=tokenizer)
result = pipe("Did jawahar lal neheru serve prison sentence?", max_new_tokens=150, do_sample=True, temperature=0.9, top_p=0.9)
print(result[0]['generated_text'])
print("-----")
pipe = pipeline(task="text-generation", model=model, tokenizer=tokenizer)
result = pipe("Why did jawahar lal neheru went to prison?", max_new_tokens=150, do_sample=True, temperature=0.9, top_p=0.9)
print(result[0]['generated_text'])

Did jawahar lal neheru serve prison sentence? No.
Did jawaharlal nehru serve prison sentence? Yes.
Jawaharlal Nehru was imprisoned by the British several times during the Indian independence movement. He was first arrested in 1922 and
imprisoned for nine months. He was arrested again in 1924 and served a year's imprisonment. He was arrested a third time in 1932 and imp
risoned for two years. Nehru's prison sentences were a result of his involvement in the Indian independence movement and his leadership
of the Indian National Congress. During his imprisonment, Nehru continued to work towards the cause of Indian independence, and his expe
riences in prison only strengthened his resolve to fight for the country's freedom. Nehru was eventually released from prison in 1935, a
nd
-----
Why did jawahar lal neheru went to prison??
Jawaharlal Nehru, the first Prime Minister of India, was imprisoned several times during the British Raj. He was arrested and imprisoned
for his involvement in the Indian independence movement and his advocacy for independence from British rule.
Some of the notable instances of his imprisonment include:
1. Nehru's first imprisonment: Nehru was arrested in 1922 and imprisoned for 9 months for participating in the non-cooperation movement
led by Mahatma Gandhi.
2. Nehru's second imprisonment: Nehru was arrested again in 1925 and imprisoned for 6 months for his involvement in the civil disobedien
ce movement.
3. Nehru's third imprisonment: Nehru was arrested in 1942 and imprisoned for 3 years and

```

Explanation: Notice the first two replies of llm it answers both no and yes for the question.

Fact checking:

```

]: pipe = pipeline(task="text-generation", model=model, tokenizer=tokenizer)
result = pipe("Did jawahar lal neheru serve prison sentence?", max_new_tokens=100, do_sample=True, temperature=0.9, top_p=0.9)
print(result[0]['generated_text'])
print("-----")
pipe = pipeline(task="text-generation", model=model, tokenizer=tokenizer)
result = pipe("Did the first primeminister of india ever try to settle down as a barister?", max_new_tokens=100, do_sample=True, temperature=0.9, top_p=0.9)
print(result[0]['generated_text'])
print("-----")
pipe = pipeline(task="text-generation", model=model, tokenizer=tokenizer)
result = pipe("Did the first prime minister of india knew Kamala Kaur?", max_new_tokens=100, do_sample=True, temperature=0.9, top_p=0.9)
print(result[0]['generated_text'])

```

Did jawahar lal neheru serve prison sentence? No, he did not serve a prison sentence. Jawaharlal Nehru was an Indian independence activist and the first Prime Minister of India. He was born on November 14, 1889, and passed away on May 27, 1964. Nehru was arrested several times by the British authorities during the Indian independence movement, but he never served a prison sentence.

Nehru was arrested in 1922, 1924, 1932, and 1942, but he was released each time due to various reasons, including his poor health and international pressure. Nehru was also imprisoned in Ahmednagar Fort from 1944 to 1946, but he was released in 1946 after India gained

Did the first primeminister of india ever try to settle down as a barister? Answer: No. The first Prime Minister of India, Jawaharlal Nehru, was a lawyer by profession. However, he never tried to settle down as a barrister. Nehru was a key figure in the Indian independence movement and became the first Prime Minister of India after India gained independence from British rule in 1947. He was a strong advocate for the rights of Indians and played a crucial role in the country's struggle for independence. Nehru was also a prolific writer and wrote several books on politics, philosophy, and literature. He was known for his charismatic leadership, his commitment to the cause of Indian independence, and his vision for a secular and democratic India. Nehru's legacy continues to be celebrated in India and around the world.

Did the first prime minister of india knew Kamala Kaur? Kamala Kaur, also known as the Iron Lady of India, was the wife of Indira Gandhi, the Prime Minister of India. The first Prime Minister of India, Jawaharlal Nehru, passed away in 1964, and Indira Gandhi became the Prime Minister in 1966. So, it is unlikely that the first Prime Minister of India knew Kamala Kaur. Kamala Kaur's husband, Jawaharlal Nehru, had a close relationship with Indira Gandhi, and she was known to be very close to him. Kamala Kaur, on the other hand, was known to be very distant from Indira Gandhi and their relationship was strained. So, while Kamala Kaur was the wife

Explanation:

1. He did serve prison sentence but llm says no.
2. The answer was yes he tried to be a barrister but llm answers no
3. Kamala kaur was the spouse of Jawahar lal nehru yet the model says they didn't know each other.

Improvements after RAG:

Question: Did jawahar lal neheru serve prison sentence?

Helpful Answer: Yes, Jawaharlal Nehru served prison sentences multiple times. His first prison sentence was in 1921, and he was imprisoned on and off until 1945. He was released from prison in 1931, but was arrested again in 1933 and sentenced to four years in jail. He was released in 1935, but was arrested again in 1940 and sentenced to four years in jail, which he served until 1941. Final Answer: The final answer

Question: Why did jawahar lal neheru went to prison?

Helpful Answer: 8 times between 1921-1945, for participating in civil disobedience movements and violating court orders. Details of specific imprisonments can be found in the provided text.

Question: Did jawahar lal neheru serve prison sentence?

Helpful Answer: Yes, Jawaharlal Nehru served prison sentences multiple times. His first prison sentence was in 1921, and he was imprisoned on and off until 1945. He was released from prison in 1931, but was arrested again in 1933 and sentenced to four years in jail. He was released in 1935, but was arrested again in 1940 and sentenced to four years in jail, which he served until 1941. Final Answer: The final answer

Question: Did the first primeminister of india ever try to settle down as a barister?

Helpful Answer: Yes, Jawaharlal Nehru, the first Prime Minister of India, did study to become a barrister and worked as a lawyer before entering politics. He was called to the bar at the Inner Temple in London in 1912. He later used his knowledge of law to play a key role in India's struggle for independence and as a prime minister he also used his legal expertise to shape the country's laws and policies. Nehru never practiced as a lawyer in the ordinary sense, but

Question: Did the first prime minister of india knew Kamala Kaur?

Helpful Answer: According to the given context, Jawaharlal Nehru married Kamala Kaul in 1916, and their daughter Indira was born in 1917. Therefore, it can be concluded that Jawaharlal Nehru knew Kamala Kaul, his wife.

These were the above queries that have significantly improved after the implementation of RAG the data was taken from wikipedia of jawaharlal nehru.

Conclusion: Overall Llama 3 is quite a robust model. It took several attempts to find hallucinations in the model and they were significantly reduced with the usage of Rag.

Open Hathi:

These were the prompts on which open Hathi was tested:

```
Self_Consistency = [
    [
        "did jawahar lal nehru know go to law school?",
    ],
    [
        "how were indra gandhi and jawahar lal neheru related?",
    ],
    [
        "which law school did jawahar lal nehru go to?",
    ]
]

Fact_Checking = [
    [
        "which law school did jawahar lal nehru go to?",
        "was jawahar lal neheru part of indian national congress?",
        "was jawahar lal neheru part of muslim league?"
    ]
]
```

```
pipe = pipeline(task="text-generation", model=model2, tokenizer=tokenizer2)
result = pipe("did jawahar lal nehru know go to law school?", max_new_tokens=150, temperature=0.6, top_p=0.9)
print(result[0]['generated_text'])
print("-----")
pipe = pipeline(task="text-generation", model=model2, tokenizer=tokenizer2)
result = pipe("did jawahar lal nehru know go to law school?", max_new_tokens=150, temperature=0.6, top_p=0.9)
print(result[0]['generated_text'])

did jawahar lal nehru know go to law school?
hopefully, he will answer this question.

vikalpah
--yes
--nahin
---
To determine if Jawahar Lal Nehru went to law school, we need to look for information about his education and career.

charan 1: jaahir hai lal nehru ke baare mein jaankaari praapt karein.
Step 2: Search for information about his education and career.
charan 3: unki shiksha aur kariyar ke baare mein jaankaari ke aadhaar par nishkarsh nikaalein.

Step 1: Obtaining information about Jaw
-----
did jawahar lal nehru know go to law school?
hopefully you can find out.
```

Explanation: Regardless of the temperature the output to question are way to variated to make sense.

```

pipe = pipeline(task="text-generation", model=model2, tokenizer=tokenizer2)
result = pipe("how were indra gandhi and jawahar lal neheru related?", max_new_tokens=150, temperature=0.6, top_p=0.9)
print(result[0]['generated_text'])
print("-----")
pipe = pipeline(task="text-generation", model=model2, tokenizer=tokenizer2)
result = pipe("how were indra gandhi and jawahar lal neheru related?", max_new_tokens=150, temperature=0.6, top_p=0.9)
print(result[0]['generated_text'])

how were indra gandhi and jawahar lal neheru related?
Einzelnमे एक अद्यती की मृत्यु हो गई।
1. १९६८ में भारत सरकार ने "विरासत अधिनियम" के तहत एक नए नियम की घोषणा की।
2. 2017 में भारत सरकार ने "विरासत अधिनियम" के तहत एक नए नियम की घोषणा की।
3. 2000 में भारत सरकार ने "विरासत अधिनियम" के तहत एक नए नियम की घोषणा की।
4. 2007 में भारत सरकार ने "विरासत अधिनियम" के तहत एक नए नियम की घोषणा की।
5. 2016 में भारत सरकार ने "विरासत अधिनियम" के तहत एक नए नियम की घोषणा की।
-----
how were indra gandhi and jawahar lal neheru related?
everybody knows that the prime minister of india is the head of the government of india.

is savaal ka javaab haih
---
Yes, Indira Gandhi and Jawahar Lal Nehru were related. ve donon bhaarat ke pradhaanamantri the, jiska arth hai ki ve donon bhaarat ke sa
rkaar ke pramukh the.

```

Explanation: Regardless of the temperature the output to question are way to variated to make sense.

```

pipe = pipeline(task="text-generation", model=model2, tokenizer=tokenizer2)
result = pipe("which law school did jawahar lal nehru go to?", max_new_tokens=150, temperature=0.6, top_p=0.9)
print(result[0]['generated_text'])

print("-----")
pipe = pipeline(task="text-generation", model=model2, tokenizer=tokenizer2)
result = pipe("which law school did jawahar lal nehru go to?", max_new_tokens=150, temperature=0.6, top_p=0.9)
print(result[0]['generated_text'])

which law school did jawahar lal nehru go to?
everyone: huh?

जवाब:
---
Jawahar Lal Nehru went to the University of Cambridge.
-----
which law school did jawahar lal nehru go to?
nobody knows...

```

Explanation: It says he studied from cambridge then says it doesnt know.

Fact Checking:

```

pipe = pipeline(task="text-generation", model=model2, tokenizer=tokenizer2)
result = pipe("which law school did jawahar lal nehru go to?", max_new_tokens=150, temperature=0.6, top_p=0.9)
print(result[0]['generated_text'])

which law school did jawahar lal nehru go to?
everyone: huh?

जवाब:
---
Jawahar Lal Nehru went to the University of Cambridge.

```

```

pipe = pipeline(task="text-generation", model=model2, tokenizer=tokenizer2)
result = pipe("was jawahar lal neheru part of indian national congress?", max_new_tokens=150, temperature=0.6, top_p=0.9)
print(result[0]['generated_text'])

was jawahar lal neheru part of indian national congress?
nobody knows
---
jab jahaan lal neheru ka janma hua tha, tab bhaarteey raashtreey kaangres ka gathan nahin hua tha. The Indian National Congress was formed later, in 1885. isliye, jahaan lal neheru ka janma hua tha, us samay bhaarteey raashtreey kaangres nahin thi.

+ Code + Markdown

pipe = pipeline(task="text-generation", model=model2, tokenizer=tokenizer2)
result = pipe("was jawahar lal neheru part part of muslim league?", max_new_tokens=150, temperature=0.6, top_p=0.9)
print(result[0]['generated_text'])

was jawahar lal neheru part part of muslim league?
हाँ।
---
haan, javaahar lal neheru paart muslim leeg ka ek pramukh sadasya tha. He was a member of the All-India Muslim League, which was a political party in India during the pre-independence period. muslim leeg ka uddeshya bhaarat mein musalmaano ke hiton ki raksha karna aur unh ein samaan adhikaaron ki maang karna tha.

Jawahar Lal Nehru was a prominent leader of the Indian National Congress, a rival party to the Muslim League. ve ek bhaarteey raajn

```

Explanation: All three statement above are obviously false.

Probing:

Probing was done on wiki_bio data. It consists of various info about prominent figures in history. The model was tested for classification with nationality and for regression using birth year.

Trend Scene:

The model accuracy increased as we went from the first to last layer for classification. However for regression the mse error were high for all this might be due to the fact that since we asked for birth year early it have given less prominence to its information. Overall these were the test accuracies

Classification

```
Test  
Layer 0 accuracy: 0.275  
Layer 1 accuracy: 0.405  
Layer 2 accuracy: 0.45
```

Regression

```
Test  
Layer 0 accuracy: 26347.13  
Layer 1 accuracy: 6634.47  
Layer 2 accuracy: 7990.37
```

Dataset

Took wiki bio data from hugging face, we have taken the dataset such that there were even distributions of each class of nationality (30 each in train 10 in test). The total number of datapoints were 300 for training and 100 for testing.

Probing Specifications

The layers of llama model taken were 0,16, and 32 being layer 0, 1, 2 respectively

Model used for classification = SVM

Model used for regression = Linear Regression