

# model

December 6, 2023

```
[1]: import numpy as np
```

```
[2]: from datasets import load_dataset
dataset = load_dataset("squad_v2")
```

```
[49]: rand_indices = []
for i in range(15):
    rand_indices.append(np.random.randint(0, len(dataset["train"])))
```

```
[3]: dft = dataset['train']
```

```
[4]: lengths_context = [len(x) for x in dft['context']]
lengths_question = [len(y) for y in dft['question']]
lengths_answer = [len(x['text'][0]) if x['text'] else 0 for x in
    ↪dft['answers']]
```

```
[5]: avg_context = sum(lengths_context) / len(lengths_context)
min_context = min(lengths_context)
max_context = max(lengths_context)
print(f'Context length: avg: {avg_context}, min: {min_context}, max:
    ↪{max_context}')
```

Context length: avg: 754.5662873410631, min: 151, max: 3706

```
[6]: avg_question = sum(lengths_question) / len(lengths_question)
min_question = min(lengths_question)
max_question = max(lengths_question)
print("Question length: avg: {}, min: {}, max: {}".format(avg_question,
    ↪min_question, max_question))
```

Question length: avg: 58.50773870272178, min: 1, max: 25651

```
[7]: avg_answer = sum(lengths_answer) / len(lengths_answer)
min_answer = min(lengths_answer)
max_answer = max(lengths_answer)
print("Answer length: avg: {}, min: {}, max: {}".format(avg_answer, min_answer,
    ↪max_answer))
```

Answer length: avg: 13.423760157766711, min: 0, max: 239

```
[56]: from transformers import AutoTokenizer, TFRobertaForQuestionAnswering
import tensorflow as tf
tokenizer = AutoTokenizer.from_pretrained("deepset/roberta-base-squad2")
model = TFRobertaForQuestionAnswering.from_pretrained("deepset/
↳roberta-base-squad2")
```

Some weights of the PyTorch model were not used when initializing the TF 2.0 model TFRobertaForQuestionAnswering: ['roberta.embeddings.position\_ids']  
- This IS expected if you are initializing TFRobertaForQuestionAnswering from a PyTorch model trained on another task or with another architecture (e.g. initializing a TFBertForSequenceClassification model from a BertForPreTraining model).

- This IS NOT expected if you are initializing TFRobertaForQuestionAnswering from a PyTorch model that you expect to be exactly identical (e.g. initializing a TFBertForSequenceClassification model from a BertForSequenceClassification model).

All the weights of TFRobertaForQuestionAnswering were initialized from the PyTorch model.

If your task is similar to the task the model of the checkpoint was trained on, you can already use TFRobertaForQuestionAnswering for predictions without further training.

```
[43]: def roberta(question, text, model, tokenizer):
    inputs = tokenizer(question, text, add_special_tokens=True,
↳return_tensors="tf")
    output = model(**inputs)
    answer_start_index = int(tf.math.argmax(output.start_logits, axis=1)[0])
    answer_end_index = int(tf.math.argmax(output.end_logits, axis=-1)[0])
    predict_answer_tokens = inputs.input_ids[0, answer_start_index :
↳answer_end_index + 1]
    return tokenizer.decode(predict_answer_tokens)
```

```
[60]: for i in rand_indices:
    print(f'Question: {dft["question"][i]}')
    print(f'Context: {dft["context"][i]}')
    print(f"Answer using Roberta Model:
↳{roberta(dft['question'][i], dft['context'][i], model, tokenizer)}")
    print(f'Ground Truth Answer: {dft["answers"][i]["text"]}')
    print('-----')
    print('-----')
    print('\n')
```

Question: The conversion for what system took place in the railway station?  
Context: In the Soviet Union, in the 1970s, a comparison was made between systems electrified at 3 kV DC and 25 kV AC (50 Hz). The results showed that percentage losses in the overhead wires (catenary and contact wires) was over 3

times greater for 3 kV DC than for 25 kV AC. But when the conversion losses were all taken into account and added to overhead wire losses (including cooling blower energy) the 25 kV AC lost a somewhat higher percent of energy than for 3 kV DC. Thus in spite of the much higher losses in the catenary, the 3 kV DC was a little more energy efficient than AC in providing energy from the USSR power grid to the terminals of the traction motors (all DC at that time). While both systems use energy in converting higher voltage AC from the USSR's power grid to lower voltage DC, the conversions for the DC system all took place (at higher efficiency) in the railway substation, while most of the conversion for the AC system took place inside the locomotive (at lower efficiency). Consider also that it takes energy to constantly move this mobile conversion hardware over the rails while the stationary hardware in the railway substation doesn't incur this energy cost. For more details see: Wiki: Soviet Union DC vs. AC.

Answer using Roberta Model:

Ground Truth Answer: []

-----  
-----

Question: What does the directional selection of beneficial mutations lead to?

Context: Most mutations within genes are neutral, having no effect on the organism's phenotype (silent mutations). Some mutations do not change the amino acid sequence because multiple codons encode the same amino acid (synonymous mutations). Other mutations can be neutral if they lead to amino acid sequence changes, but the protein still functions similarly with the new amino acid (e.g. conservative mutations). Many mutations, however, are deleterious or even lethal, and are removed from populations by natural selection. Genetic disorders are the result of deleterious mutations and can be due to spontaneous mutation in the affected individual, or can be inherited. Finally, a small fraction of mutations are beneficial, improving the organism's fitness and are extremely important for evolution, since their directional selection leads to adaptive evolution.:7.6

Answer using Roberta Model: adaptive evolution

Ground Truth Answer: ['adaptive evolution']

-----  
-----

Question: What was charted in 1939?

Context: During the difficult 1930s of the Great Depression, government at all levels was integral to creating jobs. The city provided recreational and educational programs, and hired people for public works projects. In 1932, Raleigh Memorial Auditorium was dedicated. The North Carolina Symphony, founded the same year, performed in its new home. From 1934 to 1937, the federal Civilian Conservation Corps constructed the area now known as William B. Umstead State Park. In 1939, the State General Assembly chartered the Raleigh-Durham Aeronautical Authority to build a larger airport between Raleigh and Durham, with the first flight occurring in 1943.

Answer using Roberta Model: Raleigh-Durham Aeronautical Authority  
Ground Truth Answer: ['Raleigh-Durham Aeronautical Authority to build a larger airport between Raleigh and Durham,']  
-----  
-----

Question: When was the Chremonidean War?

Context: Independent city states were unable to compete with Hellenistic kingdoms and were usually forced to ally themselves to one of them for defense, giving honors to Hellenistic rulers in return for protection. One example is Athens, which had been decisively defeated by Antipater in the Lamian war (323-322) and had its port in the Piraeus garrisoned by Macedonian troops who supported a conservative oligarchy. After Demetrius Poliorcetes captured Athens in 307 and restored the democracy, the Athenians honored him and his father Antigonos by placing gold statues of them on the agora and granting them the title of king. Athens later allied itself to Ptolemaic Egypt to throw off Macedonian rule, eventually setting up a religious cult for the Ptolemaic kings and naming one of the city phyles in honor of Ptolemy for his aid against Macedon. In spite of the Ptolemaic monies and fleets backing their endeavors, Athens and Sparta were defeated by Antigonos II during the Chremonidean War (267-61). Athens was then occupied by Macedonian troops, and run by Macedonian officials.

Answer using Roberta Model: 267-61  
Ground Truth Answer: ['267-61']  
-----  
-----

Question: Which Armenian church recognizes the Pope?

Context: While the Armenian Apostolic Church remains the most prominent church in the Armenian community throughout the world, Armenians (especially in the diaspora) subscribe to any number of other Christian denominations. These include the Armenian Catholic Church (which follows its own liturgy but recognizes the Roman Catholic Pope), the Armenian Evangelical Church, which started as a reformation in the Mother church but later broke away, and the Armenian Brotherhood Church, which was born in the Armenian Evangelical Church, but later broke apart from it. There are other numerous Armenian churches belonging to Protestant denominations of all kinds.

Answer using Roberta Model: Armenian Catholic Church  
Ground Truth Answer: ['Armenian Catholic Church']  
-----  
-----

Question: What characteristic did Ibn Sina find out about infectious diseases?

Context: Ibn Sina (Avicenna) is regarded as the most influential philosopher of Islam. He pioneered the science of experimental medicine and was the first

physician to conduct clinical trials. His two most notable works in medicine are the Kitāb al-shifāʾ ("Book of Healing") and The Canon of Medicine, both of which were used as standard medicinal texts in both the Muslim world and in Europe well into the 17th century. Amongst his many contributions are the discovery of the contagious nature of infectious diseases, and the introduction of clinical pharmacology.

Answer using Roberta Model: contagious nature

Ground Truth Answer: ['contagious']

-----  
-----

Question: What type of rich foods did people refrain from eating during Lent?

Context: Traditionally the feast also applied to sexual desires, which were supposed to be suppressed during the following fasting. Before Lent began, all rich food and drink were consumed in what became a giant celebration that involved the whole community, and is thought to be the origin of Carnival. The Lenten period of the Liturgical calendar, the six weeks directly before Easter, was originally marked by fasting and other pious or penitential practices.

During Lent, no parties or celebrations were held, and people refrained from eating rich foods, such as meat, dairy, fat and sugar.

Answer using Roberta Model: meat, dairy, fat and sugar

Ground Truth Answer: ['meat, dairy, fat and sugar']

-----  
-----

Question: What predecessor's a Digimon similar to?

Context: Digimon started out as digital pets called "Digital Monsters", similar in style and concept to the Tamagotchi. It was planned by WiZ and released by Bandai on June 26, 1997. The toy began as the simple concept of a Tamagotchi mainly for boys. The V-Pet is similar to its predecessors, with the exceptions of being more difficult and being able to fight other Digimon v-pets. Every owner would start with a Baby Digimon, train it, evolve it, take care of it, and then have battles with other Digimon owners to see who was stronger. The Digimon pet had several evolution capabilities and abilities too, so many owners had different Digimon. In December, the second generation of Digital Monster was released, followed by a third edition in 1998.

Answer using Roberta Model: <s>

Ground Truth Answer: []

-----  
-----

Question: How many Mexicans were killed in the battle of Sacramento?

Context: On February 8, 1847, Doniphan continued his march with 924 men mostly from Missouri; he accompanied a train of 315 wagons of a large commercial caravan heading to the state capital. Meanwhile, the Mexican forces in the state

had time to prepare a defense against the Americans. About 20 miles (32 km) north of the capital where two mountain ranges join from east to west is the only pass into the capital; known as Sacramento Pass, this point is now part of present-day Chihuahua City. The Battle of Sacramento was the most important battle fought in the state of Chihuahua because it was the sole defense for the state capital. The battle ended quickly because of some devastating defensive errors from the Mexican forces and the ingenious strategic moves by the American forces. After their loss at the Battle of Sacramento, the remaining Mexican soldiers retreated south, leaving the city to American occupation. Almost 300 Mexicans were killed in the battle, as well as almost 300 wounded. The Americans also confiscated large amounts of Mexican supplies and took 400 Mexican soldiers prisoners of war. American forces maintained an occupation of the state capital for the rest of the Mexican-American War.

Answer using Roberta Model: Almost 300

Ground Truth Answer: ['Almost 300']

-----

-----

Question: In terms of international trophies, whos is the most successful club?  
Context: Football is the most popular national sport of Egypt. The Cairo Derby is one of the fiercest derbies in Africa, and the BBC picked it as one of the 7 toughest derbies in the world. Al Ahly is the most successful club of the 20th century in the African continent according to CAF, closely followed by their rivals Zamalek SC. Al Ahly was named in 2000 by the Confederation of African Football as the "African Club of the Century". With twenty titles, Al Ahly is currently the world's most successful club in terms of international trophies, surpassing Italy's A.C. Milan and Argentina's Boca Juniors, both having eighteen.

Answer using Roberta Model: Al Ahly

Ground Truth Answer: ['Al Ahly']

-----

-----

Question: When an isolated system is given less degrees of freedom, what happens to total energy?

Context: This principle is vitally important to understanding the behaviour of a quantity closely related to energy, called entropy. Entropy is a measure of evenness of a distribution of energy between parts of a system. When an isolated system is given more degrees of freedom (i.e., given new available energy states that are the same as existing states), then total energy spreads over all available degrees equally without distinction between "new" and "old" degrees. This mathematical result is called the second law of thermodynamics.

Answer using Roberta Model: <s>

Ground Truth Answer: []

-----

-----

Question: Temperatures higher than 300 degrees Celsius were shown to do what?  
Context: People can be exposed to asphalt in the workplace by breathing in fumes or skin absorption. The National Institute for Occupational Safety and Health (NIOSH) has set a Recommended exposure limit (REL) of 5 mg/m<sup>3</sup> over a 15-minute period. Asphalt is basically an inert material that must be heated or diluted to a point where it becomes workable for the production of materials for paving, roofing, and other applications. In examining the potential health hazards associated with asphalt, the International Agency for Research on Cancer (IARC) determined that it is the application parameters, predominantly temperature, that effect occupational exposure and the potential bioavailable carcinogenic hazard/risk of the asphalt emissions. In particular, temperatures greater than 199 °C (390 °F), were shown to produce a greater exposure risk than when asphalt was heated to lower temperatures, such as those typically used in asphalt pavement mix production and placement.

Answer using Roberta Model: <s>

Ground Truth Answer: []

-----  
-----

Question: What other countries are referred to as middle powers?  
Context: China, France, Russia, the United Kingdom and the United States are often referred to as great powers by academics due to "their political and economic dominance of the global arena". These five nations are the only states to have permanent seats with veto power on the UN Security Council. They are also the only recognized "Nuclear Weapons States" under the Nuclear Non-Proliferation Treaty, and maintain military expenditures which are among the largest in the world. However, there is no unanimous agreement among authorities as to the current status of these powers or what precisely defines a great power. For example, sources have at times referred to China, France, Russia and the United Kingdom as middle powers.

Answer using Roberta Model: <s>

Ground Truth Answer: ['China, France, Russia and the United Kingdom']

-----  
-----

Question: Who did not connect with the soldier?  
Context: Apparently the sailor did not connect with the soldier, as Mahan believed he was innovating the term Middle East. It was, however, already there to be seen.

Answer using Roberta Model: the sailor

Ground Truth Answer: ['the sailor']

-----  
-----

Question: What organization recommended that a lactose-free diet should always be continued?

Context: WHO recommends a child with diarrhea continue to be fed. Continued feeding speeds the recovery of normal intestinal function. In contrast, children whose food is restricted have diarrhea of longer duration and recover intestinal function more slowly. A child should also continue to be breastfed. The WHO states "Food should never be withheld and the child's usual foods should not be diluted. Breastfeeding should always be continued." And in the specific example of cholera, CDC also makes the same recommendation. In young children who are not breast-fed and live in the developed world, a lactose-free diet may be useful to speed recovery.

Answer using Roberta Model: <s>

Ground Truth Answer: []

-----  
-----

```
[58]: from openai import OpenAI
def connect_openai(prompt):
    client = OpenAI(api_key =
        ↪'sk-ESEIBz3OR4sTz957m6nAT3B1bkFJqym7jkHGYkEb5Wxx53L0')
    responses = client.chat.completions.create(model = 'gpt-3.5-turbo',
        ↪messages = [{'role':'user', 'content': prompt}])
    return responses.choices[0].message.content
```

```
[59]: print(connect_openai("I will be providing you with context and question. Please
        ↪provide me with the answer. For Example - Context: The answer to the
        ↪question is 42. Question: What is the answer to the question? Answer: 42"))
```

Context: The capital city of France is Paris.

Question: What is the capital city of France?

Answer: Paris

```
[62]: for i in rand_indices:
    print(f"Question: {dft['question'][i]}")
    print(f"Context: {dft['context'][i]}\n")
    print(connect_openai(f"Context:{dft['context'][i]} Question:
        ↪{dft['question'][i]}. Give me the a to the point answer"))
    print(f'Ground Truth Answer: {dft["answers"][i]["text"]}')
    print('-----')
    print('-----')
    print('\n')
```

Question: The conversion for what system took place in the railway station?

Context: In the Soviet Union, in the 1970s, a comparison was made between systems electrified at 3 kV DC and 25 kV AC (50 Hz). The results showed that



percentage losses in the overhead wires (catenary and contact wires) was over 3 times greater for 3 kV DC than for 25 kV AC. But when the conversion losses were all taken into account and added to overhead wire losses (including cooling blower energy) the 25 kV AC lost a somewhat higher percent of energy than for 3 kV DC. Thus in spite of the much higher losses in the catenary, the 3 kV DC was a little more energy efficient than AC in providing energy from the USSR power grid to the terminals of the traction motors (all DC at that time). While both systems use energy in converting higher voltage AC from the USSR's power grid to lower voltage DC, the conversions for the DC system all took place (at higher efficiency) in the railway substation, while most of the conversion for the AC system took place inside the locomotive (at lower efficiency). Consider also that it takes energy to constantly move this mobile conversion hardware over the rails while the stationary hardware in the railway substation doesn't incur this energy cost. For more details see: Wiki: Soviet Union DC vs. AC.

The conversion for the DC system took place in the railway substation.

Ground Truth Answer: []

-----  
-----

Question: What does the directional selection of beneficial mutations lead to?

Context: Most mutations within genes are neutral, having no effect on the organism's phenotype (silent mutations). Some mutations do not change the amino acid sequence because multiple codons encode the same amino acid (synonymous mutations). Other mutations can be neutral if they lead to amino acid sequence changes, but the protein still functions similarly with the new amino acid (e.g. conservative mutations). Many mutations, however, are deleterious or even lethal, and are removed from populations by natural selection. Genetic disorders are the result of deleterious mutations and can be due to spontaneous mutation in the affected individual, or can be inherited. Finally, a small fraction of mutations are beneficial, improving the organism's fitness and are extremely important for evolution, since their directional selection leads to adaptive evolution.:7.6

The directional selection of beneficial mutations leads to adaptive evolution.

Ground Truth Answer: ['adaptive evolution']

-----  
-----

Question: What was charted in 1939?

Context: During the difficult 1930s of the Great Depression, government at all levels was integral to creating jobs. The city provided recreational and educational programs, and hired people for public works projects. In 1932, Raleigh Memorial Auditorium was dedicated. The North Carolina Symphony, founded the same year, performed in its new home. From 1934 to 1937, the federal Civilian Conservation Corps constructed the area now known as William B. Umstead

State Park. In 1939, the State General Assembly chartered the Raleigh-Durham Aeronautical Authority to build a larger airport between Raleigh and Durham, with the first flight occurring in 1943.

The Raleigh-Durham Aeronautical Authority, to build a larger airport between Raleigh and Durham, was chartered in 1939.

Ground Truth Answer: ['Raleigh-Durham Aeronautical Authority to build a larger airport between Raleigh and Durham,']

-----

Question: When was the Chremonidean War?

Context: Independent city states were unable to compete with Hellenistic kingdoms and were usually forced to ally themselves to one of them for defense, giving honors to Hellenistic rulers in return for protection. One example is Athens, which had been decisively defeated by Antipater in the Lamian war (323-322) and had its port in the Piraeus garrisoned by Macedonian troops who supported a conservative oligarchy. After Demetrius Poliorcetes captured Athens in 307 and restored the democracy, the Athenians honored him and his father Antigonos by placing gold statues of them on the agora and granting them the title of king. Athens later allied itself to Ptolemaic Egypt to throw off Macedonian rule, eventually setting up a religious cult for the Ptolemaic kings and naming one of the city phyles in honor of Ptolemy for his aid against Macedon. In spite of the Ptolemaic monies and fleets backing their endeavors, Athens and Sparta were defeated by Antigonos II during the Chremonidean War (267-61). Athens was then occupied by Macedonian troops, and run by Macedonian officials.

The Chremonidean War occurred from 267-261 BCE.

Ground Truth Answer: ['267-61']

-----

Question: Which Armenian church recognizes the Pope?

Context: While the Armenian Apostolic Church remains the most prominent church in the Armenian community throughout the world, Armenians (especially in the diaspora) subscribe to any number of other Christian denominations. These include the Armenian Catholic Church (which follows its own liturgy but recognizes the Roman Catholic Pope), the Armenian Evangelical Church, which started as a reformation in the Mother church but later broke away, and the Armenian Brotherhood Church, which was born in the Armenian Evangelical Church, but later broke apart from it. There are other numerous Armenian churches belonging to Protestant denominations of all kinds.

The Armenian Catholic Church recognizes the Pope.

Ground Truth Answer: ['Armenian Catholic Church']

-----  
-----  
  
Question: What characteristic did Ibn Sina find out about infectious diseases?  
Context: Ibn Sina (Avicenna) is regarded as the most influential philosopher of Islam. He pioneered the science of experimental medicine and was the first physician to conduct clinical trials. His two most notable works in medicine are the Kitāb al-shifāʾ ("Book of Healing") and The Canon of Medicine, both of which were used as standard medicinal texts in both the Muslim world and in Europe well into the 17th century. Amongst his many contributions are the discovery of the contagious nature of infectious diseases, and the introduction of clinical pharmacology.

Ibn Sina discovered that infectious diseases are contagious in nature.  
Ground Truth Answer: ['contagious']  
-----  
-----

Question: What type of rich foods did people refrain from eating during Lent?  
Context: Traditionally the feast also applied to sexual desires, which were supposed to be suppressed during the following fasting. Before Lent began, all rich food and drink were consumed in what became a giant celebration that involved the whole community, and is thought to be the origin of Carnival. The Lenten period of the Liturgical calendar, the six weeks directly before Easter, was originally marked by fasting and other pious or penitential practices. During Lent, no parties or celebrations were held, and people refrained from eating rich foods, such as meat, dairy, fat and sugar.

During Lent, people refrained from eating rich foods such as meat, dairy, fat, and sugar.  
Ground Truth Answer: ['meat, dairy, fat and sugar']  
-----  
-----

Question: What predecessor's a Digimon similar to?  
Context: Digimon started out as digital pets called "Digital Monsters", similar in style and concept to the Tamagotchi. It was planned by WiZ and released by Bandai on June 26, 1997. The toy began as the simple concept of a Tamagotchi mainly for boys. The V-Pet is similar to its predecessors, with the exceptions of being more difficult and being able to fight other Digimon v-pets. Every owner would start with a Baby Digimon, train it, evolve it, take care of it, and then have battles with other Digimon owners to see who was stronger. The Digimon pet had several evolution capabilities and abilities too, so many owners had different Digimon. In December, the second generation of Digital Monster was released, followed by a third edition in 1998.

A Digimon is similar to the Tamagotchi.

Ground Truth Answer: []

-----  
-----  
  
Question: How many Mexicans were killed in the battle of Sacramento?

Context: On February 8, 1847, Doniphan continued his march with 924 men mostly from Missouri; he accompanied a train of 315 wagons of a large commercial caravan heading to the state capital. Meanwhile, the Mexican forces in the state had time to prepare a defense against the Americans. About 20 miles (32 km) north of the capital where two mountain ranges join from east to west is the only pass into the capital; known as Sacramento Pass, this point is now part of present-day Chihuahua City. The Battle of Sacramento was the most important battle fought in the state of Chihuahua because it was the sole defense for the state capital. The battle ended quickly because of some devastating defensive errors from the Mexican forces and the ingenious strategic moves by the American forces. After their loss at the Battle of Sacramento, the remaining Mexican soldiers retreated south, leaving the city to American occupation. Almost 300 Mexicans were killed in the battle, as well as almost 300 wounded. The Americans also confiscated large amounts of Mexican supplies and took 400 Mexican soldiers prisoners of war. American forces maintained an occupation of the state capital for the rest of the Mexican-American War.

Almost 300 Mexicans were killed in the Battle of Sacramento.

Ground Truth Answer: ['Almost 300']

-----  
-----  
  
Question: In terms of international trophies, whos is the most successful club?

Context: Football is the most popular national sport of Egypt. The Cairo Derby is one of the fiercest derbies in Africa, and the BBC picked it as one of the 7 toughest derbies in the world. Al Ahly is the most successful club of the 20th century in the African continent according to CAF, closely followed by their rivals Zamalek SC. Al Ahly was named in 2000 by the Confederation of African Football as the "African Club of the Century". With twenty titles, Al Ahly is currently the world's most successful club in terms of international trophies, surpassing Italy's A.C. Milan and Argentina's Boca Juniors, both having eighteen.

Al Ahly is currently the most successful club in terms of international trophies.

Ground Truth Answer: ['Al Ahly']

Question: When an isolated system is given less degrees of freedom, what happens to total energy?

Context: This principle is vitally important to understanding the behaviour of a quantity closely related to energy, called entropy. Entropy is a measure of evenness of a distribution of energy between parts of a system. When an isolated system is given more degrees of freedom (i.e., given new available energy states that are the same as existing states), then total energy spreads over all available degrees equally without distinction between "new" and "old" degrees. This mathematical result is called the second law of thermodynamics.

When an isolated system is given less degrees of freedom, the total energy becomes more concentrated and less evenly distributed between the parts of the system.

Ground Truth Answer: []

-----  
-----

Question: Temperatures higher than 300 degrees Celsius were shown to do what?

Context: People can be exposed to asphalt in the workplace by breathing in fumes or skin absorption. The National Institute for Occupational Safety and Health (NIOSH) has set a Recommended exposure limit (REL) of 5 mg/m<sup>3</sup> over a 15-minute period. Asphalt is basically an inert material that must be heated or diluted to a point where it becomes workable for the production of materials for paving, roofing, and other applications. In examining the potential health hazards associated with asphalt, the International Agency for Research on Cancer (IARC) determined that it is the application parameters, predominantly temperature, that effect occupational exposure and the potential bioavailable carcinogenic hazard/risk of the asphalt emissions. In particular, temperatures greater than 199 °C (390 °F), were shown to produce a greater exposure risk than when asphalt was heated to lower temperatures, such as those typically used in asphalt pavement mix production and placement.

Temperatures higher than 300 degrees Celsius were shown to increase the occupational exposure risk and the potential bioavailable carcinogenic hazard/risk of asphalt emissions.

Ground Truth Answer: []

-----  
-----

Question: What other countries are referred to as middle powers?

Context: China, France, Russia, the United Kingdom and the United States are often referred to as great powers by academics due to "their political and economic dominance of the global arena". These five nations are the only states to have permanent seats with veto power on the UN Security Council. They are also the only recognized "Nuclear Weapons States" under the Nuclear Non-

Proliferation Treaty, and maintain military expenditures which are among the largest in the world. However, there is no unanimous agreement among authorities as to the current status of these powers or what precisely defines a great power. For example, sources have at times referred to China, France, Russia and the United Kingdom as middle powers.

Some other countries often referred to as middle powers include Canada, Australia, Germany, Japan, and Brazil.

Ground Truth Answer: ['China, France, Russia and the United Kingdom']

-----  
-----

Question: Who did not connect with the soldier?

Context: Apparently the sailor did not connect with the soldier, as Mahan believed he was innovating the term Middle East. It was, however, already there to be seen.

The sailor did not connect with the soldier.

Ground Truth Answer: ['the sailor']

-----  
-----

Question: What organization recommended that a lactose-free diet should always be continued?

Context: WHO recommends a child with diarrhea continue to be fed. Continued feeding speeds the recovery of normal intestinal function. In contrast, children whose food is restricted have diarrhea of longer duration and recover intestinal function more slowly. A child should also continue to be breastfed. The WHO states "Food should never be withheld and the child's usual foods should not be diluted. Breastfeeding should always be continued." And in the specific example of cholera, CDC also makes the same recommendation. In young children who are not breast-fed and live in the developed world, a lactose-free diet may be useful to speed recovery.

There is no specific organization mentioned in the given context that recommends a lactose-free diet to be always continued.

Ground Truth Answer: []

-----  
-----

```
[1]: claude = []  
with open('claude2_results.txt', 'r') as f:  
    for line in f:
```

```
claude.append(line.strip())
```

```
[3]: claude
```

```
[3]: ['The conversion from high voltage AC to lower voltage DC for the 3 kV DC
electrified railway system took place in the railway station.',
'Adaptive evolution.',
'The Raleigh-Durham Aeronautical Authority was chartered in 1939 to build a
larger airport between Raleigh and Durham.',
'So the Chremonidean War occurred between 267 BC and 261 BC.',
'the Armenian church that recognizes the Pope is the Armenian Catholic
Church.',
'Their contagious nature.',
'People refrained from eating rich foods such as meat, dairy, fat and sugar
during Lent.',
'A Digimon is similar to the Tamagotchi.',
'Almost 300 Mexicans were killed in the battle of Sacramento.',
'Al Ahly is the most successful club in terms of international trophies.',
'When an isolated system is given less degrees of freedom, total energy is
concentrated into fewer available states.',
'Temperatures higher than 300 degrees Celsius were shown to produce a greater
exposure risk than when asphalt was heated to lower temperatures.',
'France, Russia, and the United Kingdom have at times been referred to as
middle powers.',
'The sailor did not connect with the soldier.',
'WHO (the World Health Organization) recommended that breastfeeding should
always be continued.']
```

As we can see the Claude 2 model makes a mistake in thinking WHO recommended the health procedure. There were some places where the model (Claude) is able to handle context based queries continuously and perfectly. All three models perform really good while Chat GPT is performing to the point by giving full line answers and more human like replies compared to Roberta

```
[ ]:
```