### GitHub Link: https://github.com/Samyak005/Multi-Hop-QG

### Qualitative analysis excel sheet link

https://docs.google.com/spreadsheets/d/1G769VevZx8HD7RenAX7CBNbWHffH

XORnn9UQ0WKShMw/edit#gid=0

### Demo Video

Link to a demo video of FastAPI deployment with all the models we have trained.

## Introduction

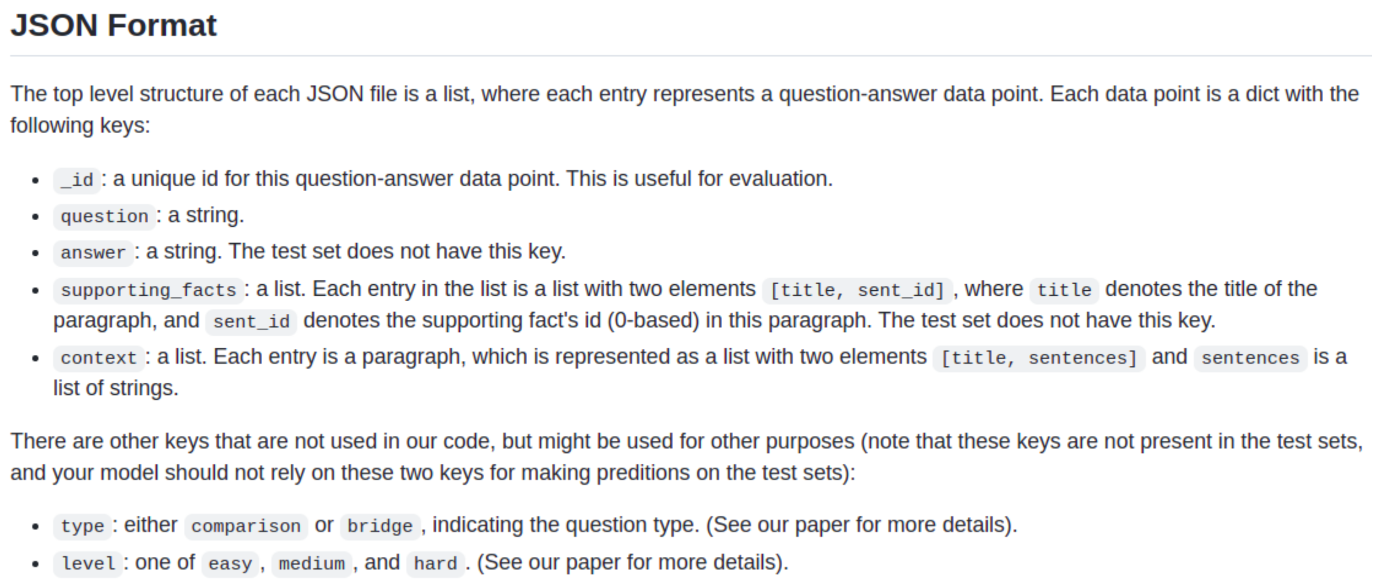
The task of Multi-hop Question Generation aims at generating questions that require finding a common concept and reasoning over multiple supporting documents. This task is quite apt on the HotpotQA dataset, which presents a diverse set of questions not constrained to any pre-existing knowledge base. Although the intended use of the dataset is for Question-Answering tasks, we can structure the inputs appropriately and use the same for the task of Question Generation.

The questions presented by the HotpotQA dataset can be primarily classified into two types:

1. Bridge Questions
These questions are framed connecting the descriptors of a common entity across two distinct supporting documents.
2. Comparison Questions
These questions are framed by comparing instances based on a common entity across two distinct supporting documents.

Each data point in the dataset has four components:

1. Answer
2. Question
3. Context Documents
4. Selected sentences from all the context documents which form the principal supporting facts for the Question Generation task.



### JSON Format

The top level structure of each JSON file is a list, where each entry represents a question-answer data point. Each data point is a dict with the following keys:

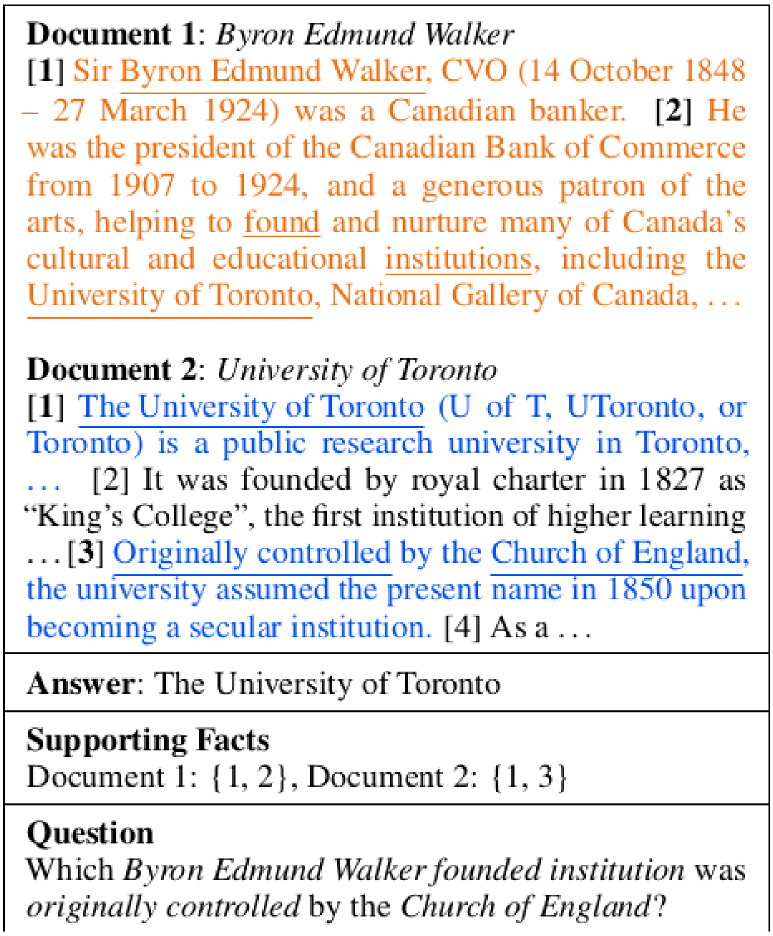
- `_id` : a unique id for this question-answer data point. This is useful for evaluation.
- `question` : a string.
- `answer` : a string. The test set does not have this key.
- `supporting_facts` : a list. Each entry in the list is a list with two elements `[title, sent_id]` , where `title` denotes the title of the paragraph, and `sent_id` denotes the supporting fact's id (0-based) in this paragraph. The test set does not have this key.
- `context` : a list. Each entry is a paragraph, which is represented as a list with two elements `[title, sentences]` and `sentences` is a list of strings.

There are other keys that are not used in our code, but might be used for other purposes (note that these keys are not present in the test sets, and your model should not rely on these two keys for making preditions on the test sets):

- `type` : either `comparison` or `bridge` , indicating the question type. (See our paper for more details).
- `level` : one of `easy` , `medium` , and `hard` . (See our paper for more details).

## Previous Work

Previously, to solve the problem of Multi-Hop question generation, Graph convolutional network and reinforcement learning based approaches have been used. Sachan et.al. use transformer model to tackle this problem.



**Document 1**: *Byron Edmund Walker*
[**1**] Sir Byron Edmund Walker, CVO (14 October 1848 – 27 March 1924) was a Canadian banker. [**2**] He was the president of the Canadian Bank of Commerce from 1907 to 1924, and a generous patron of the arts, helping to found and nurture many of Canada's cultural and educational institutions, including the University of Toronto, National Gallery of Canada, . . .

**Document 2**: *University of Toronto*
[**1**] The University of Toronto (U of T, UToronto, or Toronto) is a public research university in Toronto, . . . [2] It was founded by royal charter in 1827 as "King's College", the first institution of higher learning . . . [**3**] Originally controlled by the Church of England, the university assumed the present name in 1850 upon becoming a secular institution. [4] As a . . .

**Answer**: The University of Toronto

**Supporting Facts**
Document 1: {1, 2}, Document 2: {1, 3}

**Question**
Which *Byron Edmund Walker founded institution* was *originally controlled* by the *Church of England*?

## Transformer-based models

### Data Preprocessing

Data was converted from JSON format to CSV format. For the input, answer and context tokens were appended before them respectively.

CSV Format :

| text | question |
|---|---|
| <answer> answer <context> context | question |

We use HotpotQA's distractor setting that contains 2 gold and 8 distractor paragraphs for a question. We limit the context size to the 2 gold paragraphs, as the distractor paragraphs are irrelevant to the generation task.

### Performance with supporting facts as input

We first consider a simplified version of the task when only the supporting facts are used during training and testing. In other words, in this setting, we remove all sentences of the context documents that have not been annotated as supporting facts. This is an overly simplified setting since supporting fact annotations are not always available at test time.

### Performance with full context as input

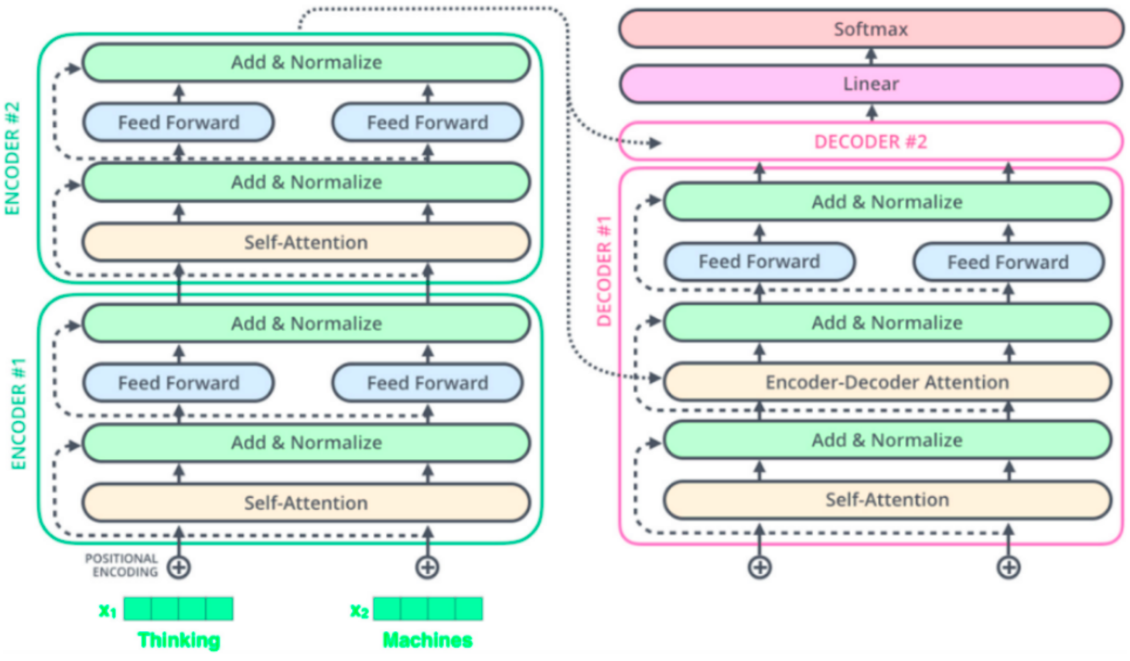
In a more realistic setting when the supporting facts are not available at test time, the model needs to process the full context. As the average document context is three times the size of the supporting facts in HotpotQA, this setting is potentially much more challenging.

# T5

T5 is a seq2seq transformer model. It is a "unified framework that converts every language problem into a text-to-text format". The T5 model has an encoder-decoder based transformer architecture which is
best suited for the text-to-text approach. The number of parameters is kept same as BERT (which is an encoder only model) by sharing them across decoder and encoder without a significant drop in performance.

The encoder consists of a stack of identical layers. Every layer is composed of two sub-layers. The first sub-layer of each encoder layer is a multi-head self-attention mechanism. The second sub-layer on the other hands is a fully connected position-wise feed-forward network. Residual connections are employed around these sub-layers, each followed by the normalization layer. Similar to the encoder, decoder also consists of a stack of identical layers. In the decoder, a third sub-layer is also inserted, in addition to the two sub-layers already present in the encoder layer. This third sub-layer performs multi-head attention on the output received by the encoder stack. Here also residual connections are employed around these sub-layers, like that of encoder, each followed by the normalization layer.

T5-base was used from the hugging face transformers library.



Architecture of T5

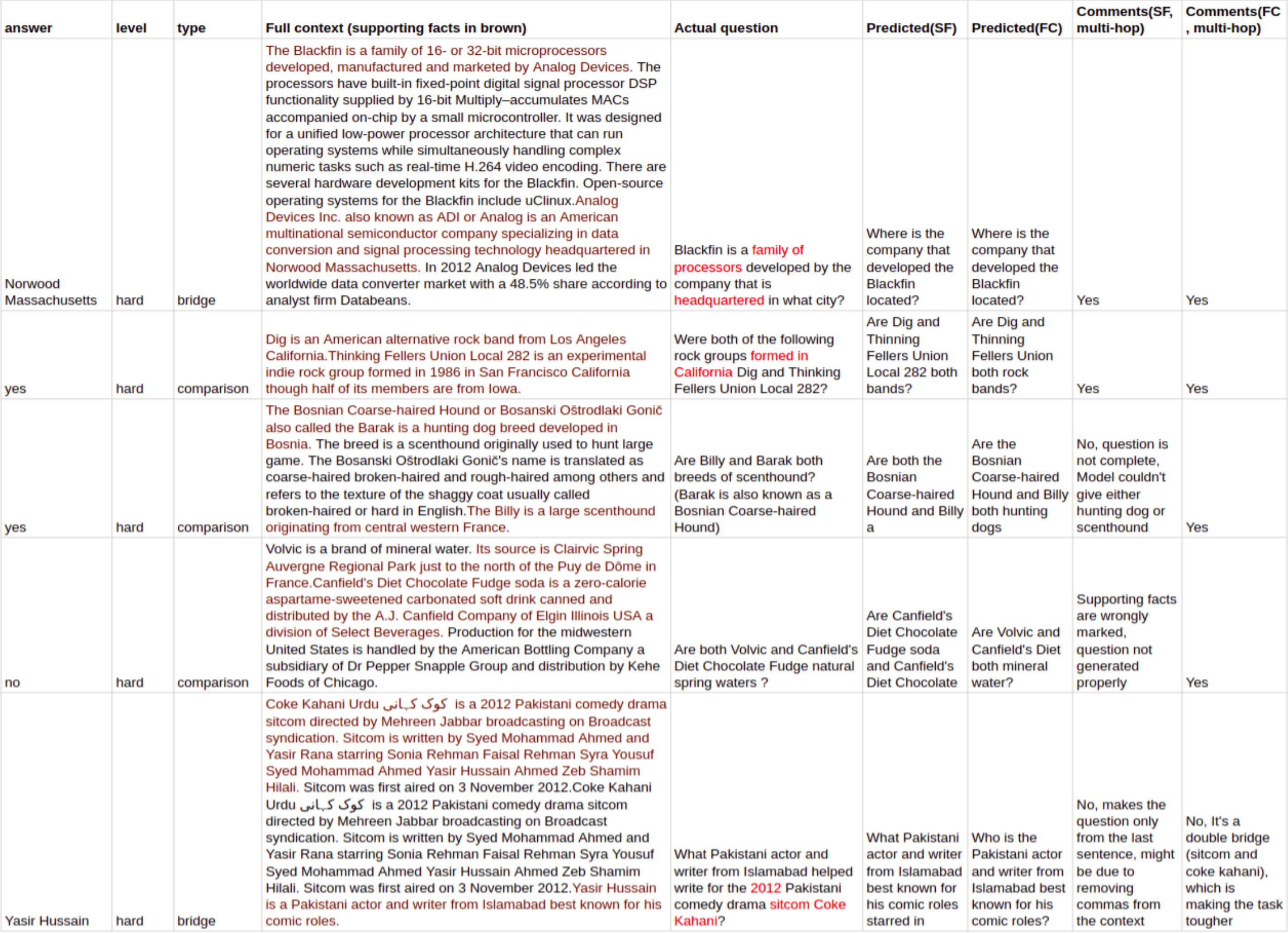
Maximum sequence length was decided after visualizing the sequence lengths using a histogram.

Training Details:
Trained on a single GPU
Epochs: 20
LR: 0.001
Batch-Size: 1

Sequence Length: 175 (supporting facts), 600 (full context)

## Qualitative analysis

(FC - full context, SF - supporting facts)



| answer | level | type | Full context (supporting facts in brown) | Actual question | Predicted(SF) | Predicted(FC) | Comments(SF, multi-hop) | Comments(FC , multi-hop) |
|---|---|---|---|---|---|---|---|---|
| Norwood Massachusetts | hard | bridge | The Blackfin is a family of 16- or 32-bit microprocessors developed, manufactured and marketed by Analog Devices. The processors have built-in fixed-point digital signal processor DSP functionality supplied by 16-bit Multiply–accumulates MACs accompanied on-chip by a small microcontroller. It was designed for a unified low-power processor architecture that can run operating systems while simultaneously handling complex numeric tasks such as real-time H.264 video encoding. There are several hardware development kits for the Blackfin. Open-source operating systems for the Blackfin include uClinux.Analog Devices Inc. also known as ADI or Analog is an American multinational semiconductor company specializing in data conversion and signal processing technology headquartered in Norwood Massachusetts. In 2012 Analog Devices led the worldwide data converter market with a 48.5% share according to analyst firm Databeans. | Blackfin is a family of processors developed by the company that is headquartered in what city? | Where is the company that developed the Blackfin located? | Where is the company that developed the Blackfin located? | Yes | Yes |
| yes | hard | comparison | Dig is an American alternative rock band from Los Angeles California.Thinking Fellers Union Local 282 is an experimental indie rock group formed in 1986 in San Francisco California though half of its members are from Iowa. | Were both of the following rock groups formed in California Dig and Thinking Fellers Union Local 282? | Are Dig and Thinning Fellers Union Local 282 both bands? | Are Dig and Thinning Fellers Union both rock bands? | Yes | Yes |
| yes | hard | comparison | The Bosnian Coarse-haired Hound or Bosanski Oštrodlaki Gonič also called the Barak is a hunting dog breed developed in Bosnia. The breed is a scenthound originally used to hunt large game. The Bosanski Oštrodlaki Gonič's name is translated as coarse-haired broken-haired and rough-haired among others and refers to the texture of the shaggy coat usually called broken-haired or hard in English.The Billy is a large scenthound originating from central western France. | Are Billy and Barak both breeds of scenthound? (Barak is also known as a Bosnian Coarse-haired Hound) | Are both the Bosnian Coarse-haired Hound and Billy a | Are the Bosnian Coarse-haired Hound and Billy both hunting dogs | No, question is not complete, Model couldn't give either hunting dog or scenthound | Yes |
| no | hard | comparison | Volvic is a brand of mineral water. Its source is Clairvic Spring Auvergne Regional Park just to the north of the Puy de Dôme in France.Canfield's Diet Chocolate Fudge soda is a zero-calorie aspartame-sweetened carbonated soft drink canned and distributed by the A.J. Canfield Company of Elgin Illinois USA a division of Select Beverages. Production for the midwestern United States is handled by the American Bottling Company a subsidiary of Dr Pepper Snapple Group and distribution by Kehe Foods of Chicago. | Are both Volvic and Canfield's Diet Chocolate Fudge natural spring waters ? | Are Canfield's Diet Chocolate Fudge soda and Canfield's Diet Chocolate | Are Volvic and Canfield's Diet both mineral water? | Supporting facts are wrongly marked, question not generated properly | Yes |
| Yasir Hussain | hard | bridge | Coke Kahani Urdu کوک کہانی is a 2012 Pakistani comedy drama sitcom directed by Mehreen Jabbar broadcasting on Broadcast syndication. Sitcom is written by Syed Mohammad Ahmed and Yasir Rana starring Sonia Rehman Faisal Rehman Syra Yousuf Syed Mohammad Ahmed Yasir Hussain Ahmed Zeb Shamim Hilali. Sitcom was first aired on 3 November 2012.Coke Kahani Urdu کوک کہانی is a 2012 Pakistani comedy drama sitcom directed by Mehreen Jabbar broadcasting on Broadcast syndication. Sitcom is written by Syed Mohammad Ahmed and Yasir Rana starring Sonia Rehman Faisal Rehman Syra Yousuf Syed Mohammad Ahmed Yasir Hussain Ahmed Zeb Shamim Hilali. Sitcom was first aired on 3 November 2012.Yasir Hussain is a Pakistani actor and writer from Islamabad best known for his comic roles. | What Pakistani actor and writer from Islamabad helped write for the 2012 Pakistani comedy drama sitcom Coke Kahani? | What Pakistani actor and writer from Islamabad best known for his comic roles starred in | Who is the Pakistani actor and writer from Islamabad best known for his comic roles? | No, makes the question only from the last sentence, might be due to removing commas from the context | No, It's a double bridge (sitcom and coke kahani), which is making the task tougher |

Analysis of 5 hard level examples was done from the test dataset(comments for poor predictions were made about what could have gone wrong)

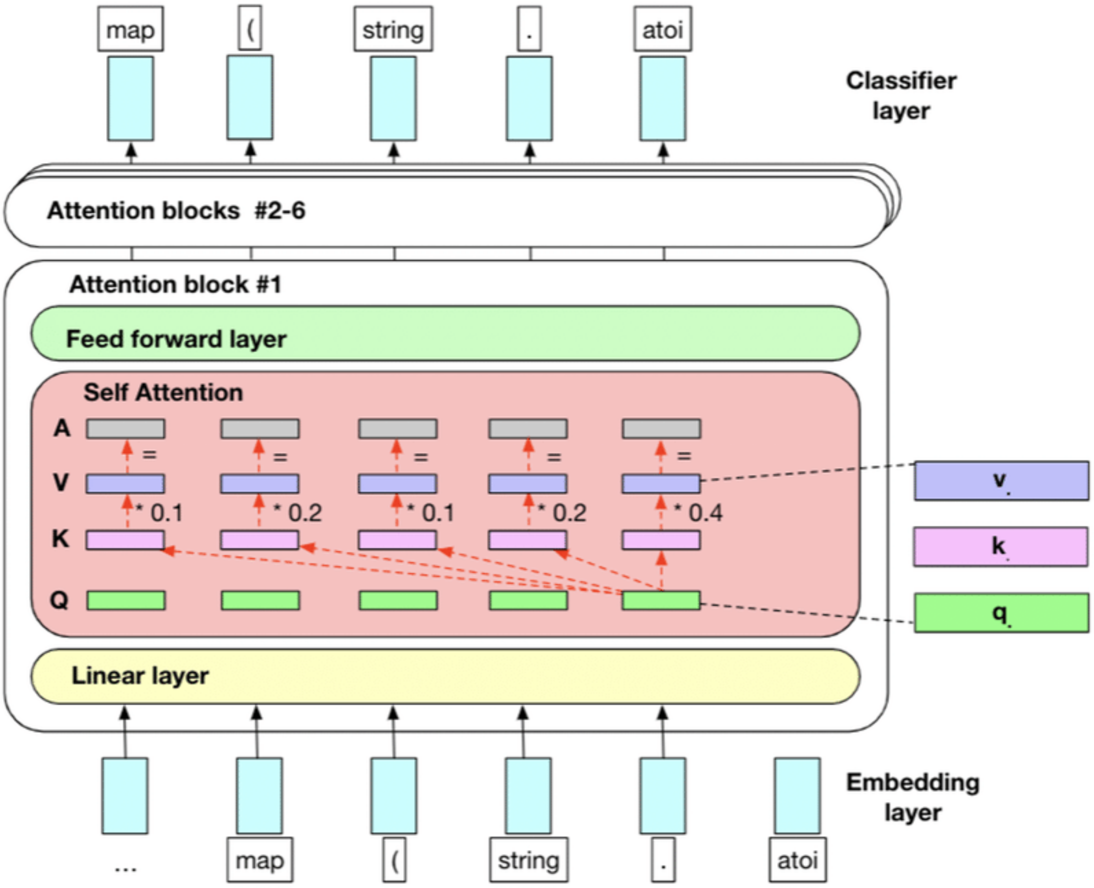
## GPT2

GPT2, the transformer-based language model, was used from the hugging-face transformers library. We first trained on the complete distractor dataset till 10000 steps, then tested on 3000 samples. The second time, we trained till 20 epochs, to get a loss of 0.08.

The GPT2 Architecture

The GPT2 architecture implements a deep neural network, specifically a transformer, which uses attention in place of previous recurrence- and convolution-based architectures. Attention mechanisms allow the model to selectively focus on segments of input text it predicts to be the most relevant. This model allows for greatly increased parallelization, and outperforms previous benchmarks for RNN/CNN/LSTM-based models.
The architecture is very similar to the decoder-only transformer.

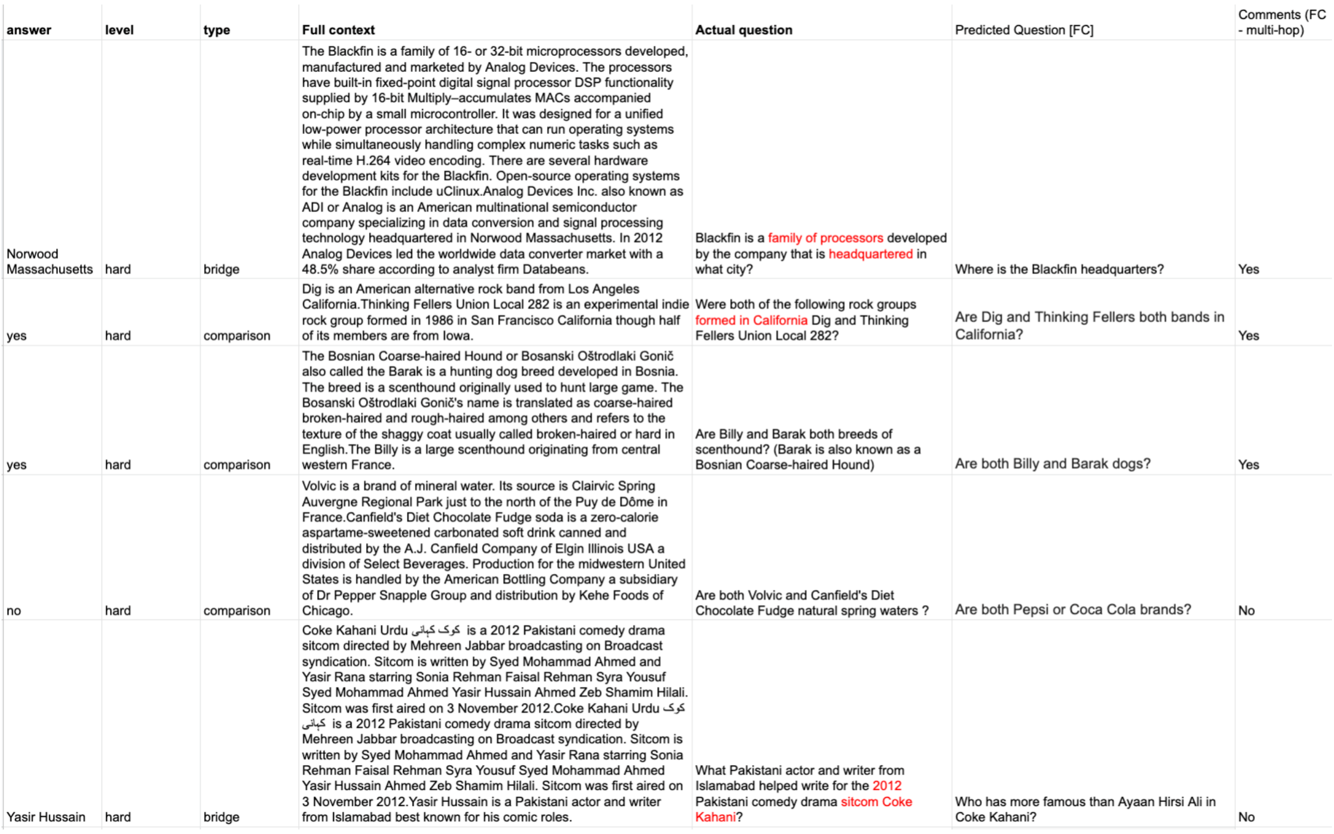
GPT-2 consists of solely stacked decoder blocks from the transformer architecture. In the standard transformer architecture, the decoder is fed a word embedding concatenated with a context vector, both generated by the encoder. Furthermore, in the standard transformer architecture self-attention is applied to the entire surrounding context, e.g. all of the other words in the sentence. In GPT-2 masked self-attention is used instead: the decoder is only allowed (via obfuscation masking of the remaining word positions) to glean information from the prior words in the sentence (plus the word itself).



Training Details:

Trained on a single GPU
Epochs:20
LR:0.001
Batch-Size:1
Loss: 0.08
Average Loss:0.32

## Qualitative Analysis

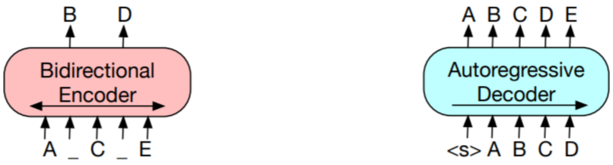


| answer | level | type | Full context | Actual question | Predicted Question [FC] | Comments (FC - multi-hop) |
|---|---|---|---|---|---|---|
| Norwood Massachusetts | hard | bridge | The Blackfin is a family of 16- or 32-bit microprocessors developed, manufactured and marketed by Analog Devices. The processors have built-in fixed-point digital signal processor DSP functionality supplied by 16-bit Multiply–accumulates MACs accompanied on-chip by a small microcontroller. It was designed for a unified low-power processor architecture that can run operating systems while simultaneously handling complex numeric tasks such as real-time H.264 video encoding. There are several hardware development kits for the Blackfin. Open-source operating systems for the Blackfin include uClinux.Analog Devices Inc. also known as ADI or Analog is an American multinational semiconductor company specializing in data conversion and signal processing technology headquartered in Norwood Massachusetts. In 2012 Analog Devices led the worldwide data converter market with a 48.5% share according to analyst firm Databeans. | Blackfin is a family of processors developed by the company that is headquartered in what city? | Where is the Blackfin headquarters? | Yes |
| yes | hard | comparison | Dig is an American alternative rock band from Los Angeles California.Thinking Fellers Union Local 282 is an experimental indie rock group formed in 1986 in San Francisco California though half of its members are from Iowa. | Were both of the following rock groups formed in California Dig and Thinking Fellers Union Local 282? | Are Dig and Thinking Fellers both bands in California? | Yes |
| yes | hard | comparison | The Bosnian Coarse-haired Hound or Bosanski Oštrodlaki Gonič also called the Barak is a hunting dog breed developed in Bosnia. The breed is a scenthound originally used to hunt large game. The Bosanski Oštrodlaki Gonič's name is translated as coarse-haired broken-haired and rough-haired among others and refers to the texture of the shaggy coat usually called broken-haired or hard in English.The Billy is a large scenthound originating from central western France. | Are Billy and Barak both breeds of scenthound? (Barak is also known as a Bosnian Coarse-haired Hound) | Are both Billy and Barak dogs? | Yes |
| no | hard | comparison | Volvic is a brand of mineral water. Its source is Clairvic Spring Auvergne Regional Park just to the north of the Puy de Dôme in France.Canfield's Diet Chocolate Fudge soda is a zero-calorie aspartame-sweetened carbonated soft drink canned and distributed by the A.J. Canfield Company of Elgin Illinois USA a division of Select Beverages. Production for the midwestern United States is handled by the American Bottling Company a subsidiary of Dr Pepper Snapple Group and distribution by Kehe Foods of Chicago. | Are both Volvic and Canfield's Diet Chocolate Fudge natural spring waters ? | Are both Pepsi or Coca Cola brands? | No |
| Yasir Hussain | hard | bridge | Coke Kahani Urdu کوک کہانی is a 2012 Pakistani comedy drama sitcom directed by Mehreen Jabbar broadcasting on Broadcast syndication. Sitcom is written by Syed Mohammad Ahmed and Yasir Rana starring Sonia Rehman Faisal Rehman Syra Yousuf Syed Mohammad Ahmed Yasir Hussain Ahmed Zeb Shamim Hilali. Sitcom was first aired on 3 November 2012.Coke Kahani Urdu کوک کہانی is a 2012 Pakistani comedy drama sitcom directed by Mehreen Jabbar broadcasting on Broadcast syndication. Sitcom is written by Syed Mohammad Ahmed and Yasir Rana starring Sonia Rehman Faisal Rehman Syra Yousuf Syed Mohammad Ahmed Yasir Hussain Ahmed Zeb Shamim Hilali. Sitcom was first aired on 3 November 2012.Yasir Hussain is a Pakistani actor and writer from Islamabad best known for his comic roles. | What Pakistani actor and writer from Islamabad helped write for the 2012 Pakistani comedy drama sitcom Coke Kahani? | Who has more famous than Ayaan Hirsi Ali in Coke Kahani? | No |

## BART

Bart is a sequence-to-sequence transformer-based language model. It has a bidirectional encoder(like BERT) and a left-to-right decoder(like GPT). The pretraining task involves randomly shuffling the order of the original sentences and a novel in-filling scheme, where spans of text are replaced with a single mask token.

The following image from the original paper that introduced Bart clearly shows the top-level architecture.



(a) BERT: Random tokens are replaced with masks, and the document is encoded bidirectionally. Missing tokens are predicted independently, so BERT cannot easily be used for generation.

(b) GPT: Tokens are predicted auto-regressively, meaning GPT can be used for generation. However words can only condition on leftward context, so it cannot learn bidirectional interactions.

BART is a denoising autoencoder for pretraining sequence-to-sequence models. BART is trained by (1) corrupting text with an arbitrary noising function, and (2) learning a model to reconstruct the original text.

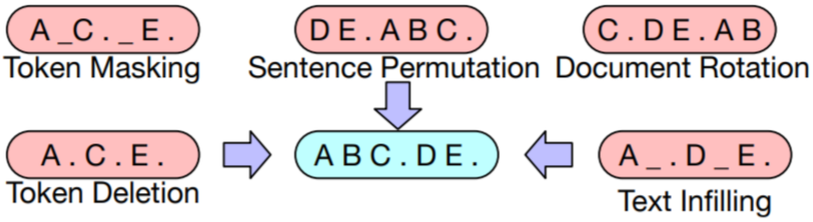
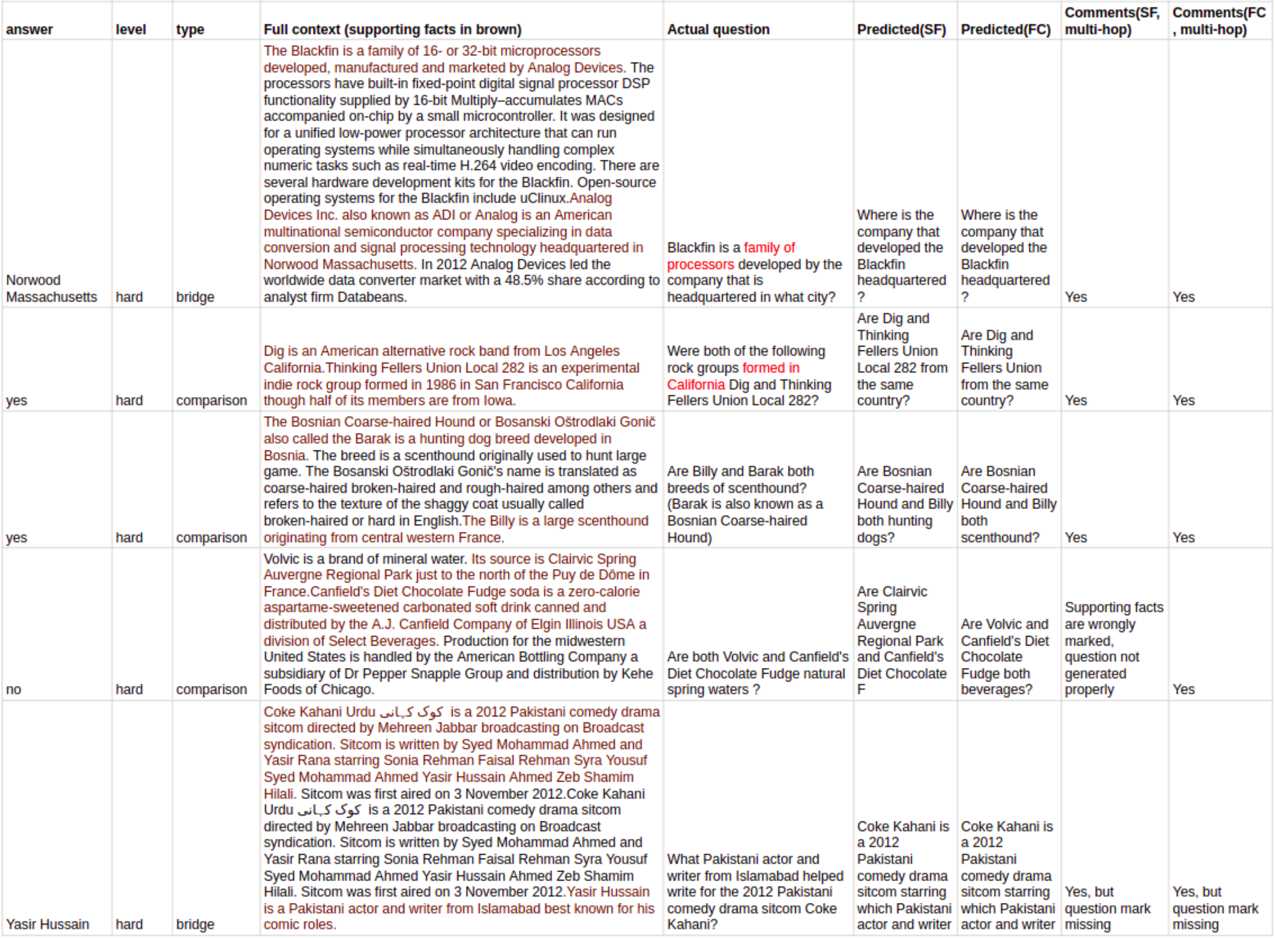


Figure 2: Transformations for noising the input that we experiment with. These transformations can be composed.

All of the images are taken from the original paper "BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension".

Training Details:
Trained on a single GPU
Epochs: 20
LR: 0.001
Batch-Size: 1
Sequence Length: 175 (supporting facts), 600 (full context)

# Qualitative Analysis



| answer | level | type | Full context (supporting facts in brown) | Actual question | Predicted(SF) | Predicted(FC) | Comments(SF, multi-hop) | Comments(FC , multi-hop) |
|---|---|---|---|---|---|---|---|---|
| Norwood Massachusetts | hard | bridge | The Blackfin is a family of 16- or 32-bit microprocessors developed, manufactured and marketed by Analog Devices. The processors have built-in fixed-point digital signal processor DSP functionality supplied by 16-bit Multiply–accumulates MACs accompanied on-chip by a small microcontroller. It was designed for a unified low-power processor architecture that can run operating systems while simultaneously handling complex numeric tasks such as real-time H.264 video encoding. There are several hardware development kits for the Blackfin. Open-source operating systems for the Blackfin include uClinux.Analog Devices Inc. also known as ADI or Analog is an American multinational semiconductor company specializing in data conversion and signal processing technology headquartered in Norwood Massachusetts. In 2012 Analog Devices led the worldwide data converter market with a 48.5% share according to analyst firm Databeans. | Blackfin is a family of processors developed by the company that is headquartered in what city? | Where is the company that developed the Blackfin headquartered ? | Where is the company that developed the Blackfin headquartered ? | Yes | Yes |
| yes | hard | comparison | Dig is an American alternative rock band from Los Angeles California.Thinking Fellers Union Local 282 is an experimental indie rock group formed in 1986 in San Francisco California though half of its members are from Iowa. | Were both of the following rock groups formed in California Dig and Thinking Fellers Union Local 282? | Are Dig and Thinking Fellers Union Local 282 from the same country? | Are Dig and Thinking Fellers Union from the same country? | Yes | Yes |
| yes | hard | comparison | The Bosnian Coarse-haired Hound or Bosanski Oštrodlaki Gonič also called the Barak is a hunting dog breed developed in Bosnia. The breed is a scenthound originally used to hunt large game. The Bosanski Oštrodlaki Gonič's name is translated as coarse-haired broken-haired and rough-haired among others and refers to the texture of the shaggy coat usually called broken-haired or hard in English.The Billy is a large scenthound originating from central western France. | Are Billy and Barak both breeds of scenthound? (Barak is also known as a Bosnian Coarse-haired Hound) | Are Bosnian Coarse-haired Hound and Billy both hunting dogs? | Are Bosnian Coarse-haired Hound and Billy both scenthound? | Yes | Yes |
| no | hard | comparison | Volvic is a brand of mineral water. Its source is Clairvic Spring Auvergne Regional Park just to the north of the Puy de Dôme in France.Canfield's Diet Chocolate Fudge soda is a zero-calorie aspartame-sweetened carbonated soft drink canned and distributed by the A.J. Canfield Company of Elgin Illinois USA a division of Select Beverages. Production for the midwestern United States is handled by the American Bottling Company a subsidiary of Dr Pepper Snapple Group and distribution by Kehe Foods of Chicago. | Are both Volvic and Canfield's Diet Chocolate Fudge natural spring waters ? | Are Clairvic Spring Auvergne Regional Park and Canfield's Diet Chocolate F | Are Volvic and Canfield's Diet Chocolate Fudge both beverages? | Supporting facts are wrongly marked, question not generated properly | Yes |
| Yasir Hussain | hard | bridge | Coke Kahani Urdu کوک کہانی is a 2012 Pakistani comedy drama sitcom directed by Mehreen Jabbar broadcasting on Broadcast syndication. Sitcom is written by Syed Mohammad Ahmed and Yasir Rana starring Sonia Rehman Faisal Rehman Syra Yousuf Syed Mohammad Ahmed Yasir Hussain Ahmed Zeb Shamim Hilali. Sitcom was first aired on 3 November 2012.Coke Kahani Urdu کوک کہانی is a 2012 Pakistani comedy drama sitcom directed by Mehreen Jabbar broadcasting on Broadcast syndication. Sitcom is written by Syed Mohammad Ahmed and Yasir Rana starring Sonia Rehman Faisal Rehman Syra Yousuf Syed Mohammad Ahmed Yasir Hussain Ahmed Zeb Shamim Hilali. Sitcom was first aired on 3 November 2012.Yasir Hussain is a Pakistani actor and writer from Islamabad best known for his comic roles. | What Pakistani actor and writer from Islamabad helped write for the 2012 Pakistani comedy drama sitcom Coke Kahani? | Coke Kahani is a 2012 Pakistani comedy drama sitcom starring which Pakistani actor and writer | Coke Kahani is a 2012 Pakistani comedy drama sitcom starring which Pakistani actor and writer | Yes, but question mark missing | Yes, but question mark missing |

# Evaluation Results

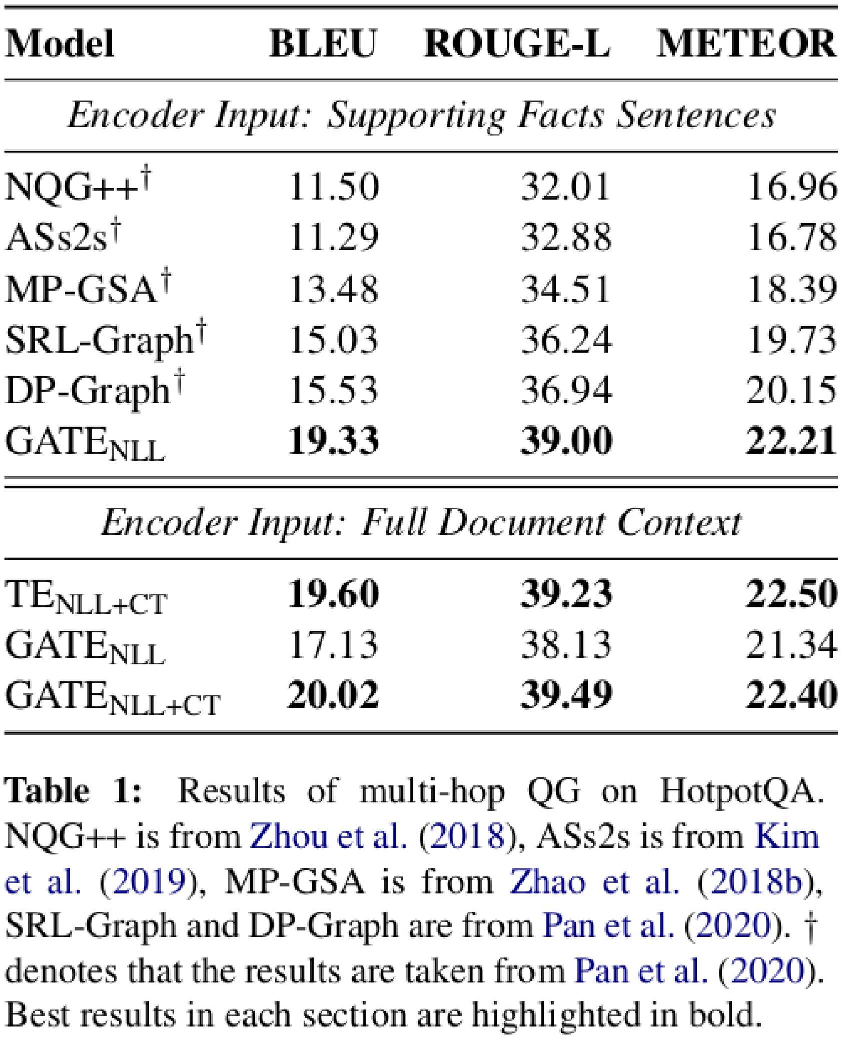
We have evaluated the results on the following metrics

1. BLEU (1 to 4, taking mean with equal weight)
2. METEOR
3. ROUGE_L

| Model | BLEU | ROUGE-L | METEOR |
|---|---|---|---|
| T5 with supporting facts | 22.5 | 32.9 | 19.1 |
| T5 with full context | 22.3 | 32.5 | 19.1 |

| | | | |
|---|---|---|---|
| GPT2 with full context | 6.3 | 16.2 | 9.4 |
| BART with supporting facts | 24.0 | 34.0 | 19.9 |
| BART with full context | 24.2 | 34.0 | 20.1 |

**Results from the Sachan et.al. paper**



| Model | BLEU | ROUGE-L | METEOR |
|---|---|---|---|
| *Encoder Input: Supporting Facts Sentences* | | | |
| NQG++$^\dagger$ | 11.50 | 32.01 | 16.96 |
| ASs2s$^\dagger$ | 11.29 | 32.88 | 16.78 |
| MP-GSA$^\dagger$ | 13.48 | 34.51 | 18.39 |
| SRL-Graph$^\dagger$ | 15.03 | 36.24 | 19.73 |
| DP-Graph$^\dagger$ | 15.53 | 36.94 | 20.15 |
| $\text{GATE}_{\text{NLL}}$ | **19.33** | **39.00** | **22.21** |
| *Encoder Input: Full Document Context* | | | |
| $\text{TE}_{\text{NLL+CT}}$ | **19.60** | **39.23** | **22.50** |
| $\text{GATE}_{\text{NLL}}$ | 17.13 | 38.13 | 21.34 |
| $\text{GATE}_{\text{NLL+CT}}$ | **20.02** | **39.49** | **22.40** |

**Table 1:** Results of multi-hop QG on HotpotQA. NQG++ is from Zhou et al. (2018), ASs2s is from Kim et al. (2019), MP-GSA is from Zhao et al. (2018b), SRL-Graph and DP-Graph are from Pan et al. (2020). † denotes that the results are taken from Pan et al. (2020). Best results in each section are highlighted in bold.

# Model deployment

We have implemented a FastAPI backend for deploying the model. The following section describes how to use it.

# Deploying model using FastAPI

## Installing requirements

```
# Installing FastAPI related libraries
pip install -r fastapi_requirements.txt
# Install pytorch and transformers 4.9.1 library
pip install transformers==4.9.1
```

## Trained models

We have trained three models, two of them trained with both supporting facts and full context.

Models details and corresponding names are as follows.

1. `t5_supp` - T5 with supporting facts.
2. `t5_full` - T5 with full context.
3. `gpt2` - GPT2 with full context.
4. `bart_supp` - BART with supporting facts.
5. `bart_full`- BART with full context.

The code for deployment assumes that the models except `gpt2` are in the folder `../trained_models`.

## Deploying backend

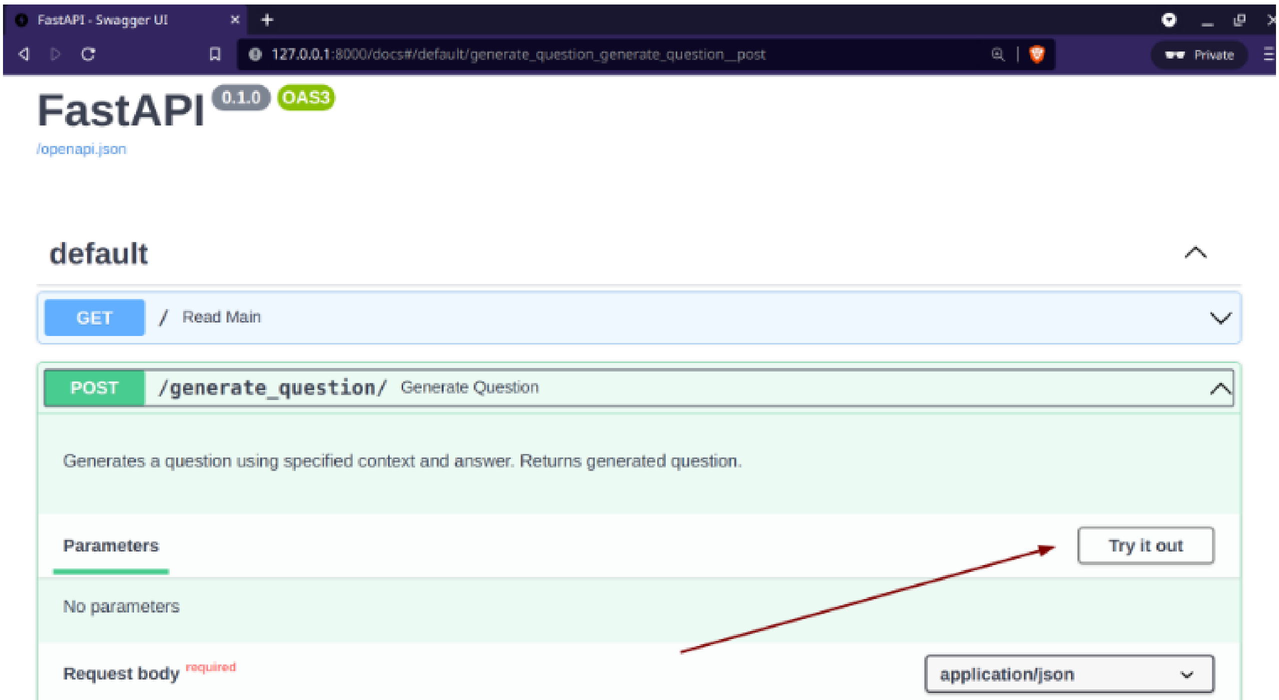
Put the pretrained models on the appropriate directories or modify the path files in the code.

### Starting backend

```
uvicorn main:app --reload # This will start FastAPI backend
```

## How to use FastAPI backend to generate questions

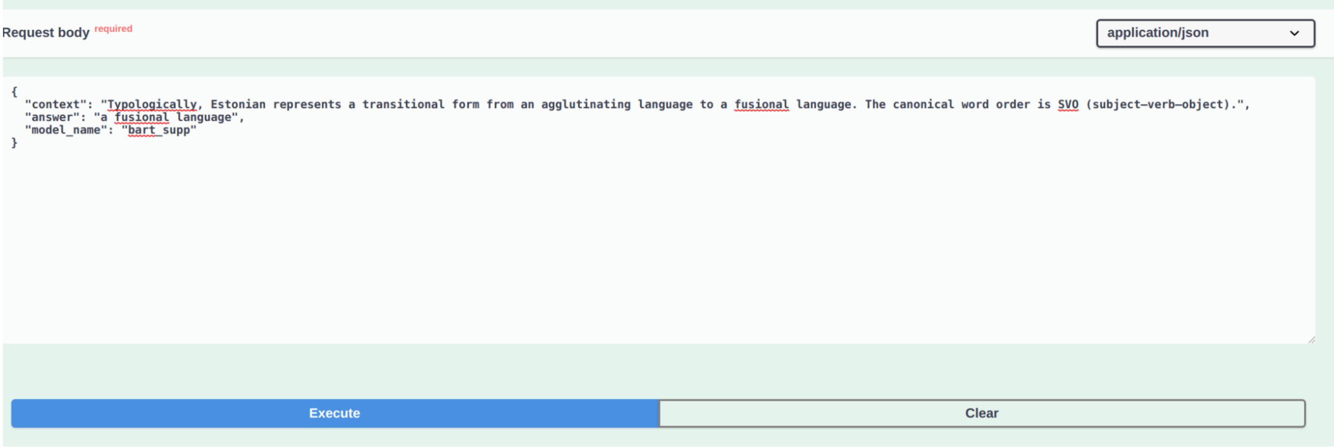
Open browser, go to [http://127.0.0.1:8000/doc](http://127.0.0.1:8000/doc). Go to `Try it out`



Generates a question using specified context and answer. Returns generated question.

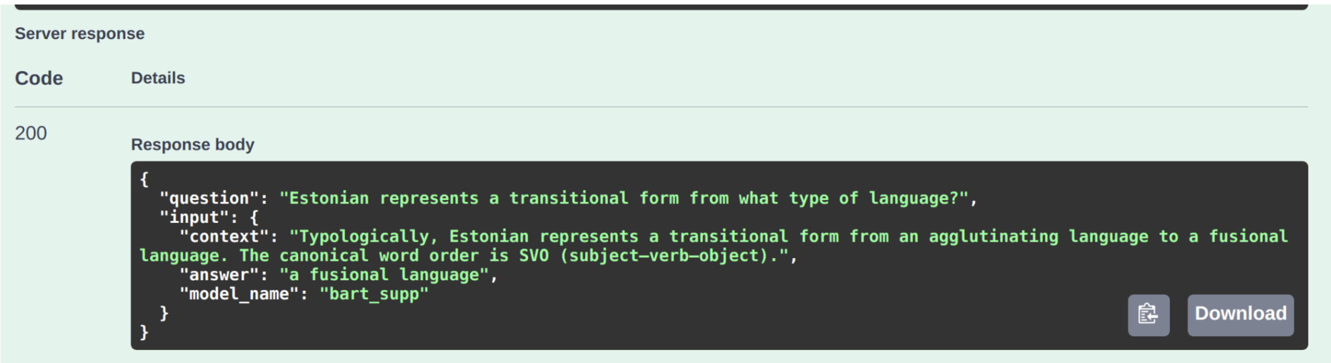
No parameters

Give the `context`, `answer, and model name`. Then execute.



```
{
  "context": "Typologically, Estonian represents a transitional form from an agglutinating language to a fusional language. The canonical word order is SVO (subject–verb–object).",
  "answer": "a fusional language",
  "model_name": "bart_supp"
}
```

You can find the generated question in the response section below that.



```
{
  "question": "Estonian represents a transitional form from what type of language?",
  "input": {
    "context": "Typologically, Estonian represents a transitional form from an agglutinating language to a fusional language. The canonical word order is SVO (subject–verb–object).",
    "answer": "a fusional language",
    "model_name": "bart_supp"
  }
}
```

# References

1. Jurafsky, Dan. Speech & language processing. Pearson Education India, 2000.
2. Sachan, Devendra Singh, et al. "Stronger Transformers for Neural Multi-Hop Question Generation." arXiv preprint arXiv:2010.11374 (2020).
3. Yang, Zhilin, et al. "Hotpotqa: A dataset for diverse, explainable multi-hop question answering." arXiv preprint arXiv:1809.09600 (2018).
4. Su, Dan, et al. "Multi-hop Question Generation with Graph Convolutional Network." arXiv preprint arXiv:2010.09240 (2020).
5. Gupta, Deepak, et al. "Reinforced Multi-task Approach for Multi-hop Question Generation." arXiv preprint arXiv:2004.02143 (2020).