ITMO UNIVERSITY

Saint Petersburg, Russia

Application of Natural Language Processing in Generating English Grammar Exercises

Shivam Mehta

Supervised by Ivan Smettanikov, Chris Develder, Thomas Demeester, Johannes Deleu, Nicolas Staelens

ITMO UNIVERSITY

Supervisor	Ivan Smettanikov

GHENT UNIVERSITY

Supervisor	Chris Develder
Consultant	Johannes Deleu
Consultant	Thomas Demeester

televic education

Supervisor	Nicolas Staelens

ITMO UNIVERSITY

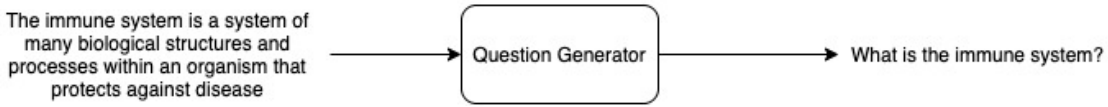
Objective of this study

- Propose and Analyze architectures for generation of Reading Comprehension **(RC)** Questions
- Propose and analyze algorithms and architectures for generation of Fill-in-the-blank **(FITB)** based language learning exercises

ITMO more than a UNIVERSITY

RC Question Generation

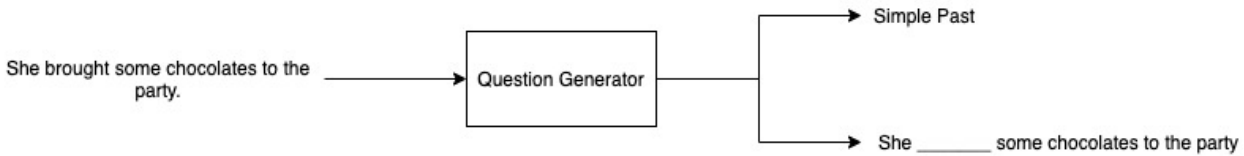
- Generation of Reading Comprehension Questions from the Sentences



ITMO UNIVERSITY

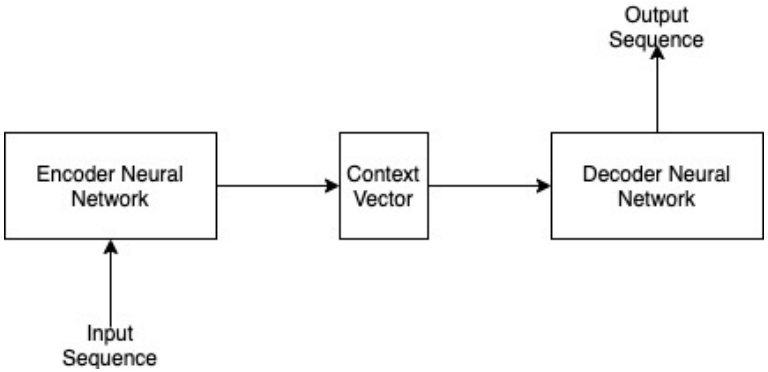
Fill-In-The-Blank Exercise Generation

- Different from RC
- Determine the position of blank in sentence or predict the blank word



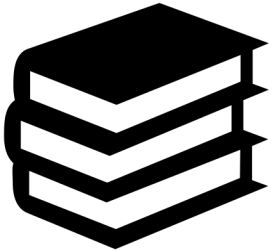
ITMO more than a UNIVERSITY

RC Question Generation Using Sequence to Sequence



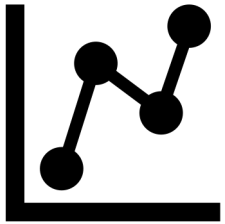
Literature Study

- Learning to Ask: Neural Question Generation for Reading Comprehension [*Du el at*] (2017)
- Neural Question Generation from Text: A Preliminary Study [*Zhou et al.]* (2017)
- Neural Models for Key Phrase Extraction and Question Generation [*Subramanian et al.*] (2017)
- Machine Comprehension by Text-to-Text Neural Question Generation [*Yuan et al.*] (2017)
- Unified Language Model Pre-Training for Natural Language Understanding and Generation [UniLM] [*Dong et al.*] (2019)
- Question Generation for Question Answering [*Dual et al.*] (2017)
- Question Generation by Transformers [*Kriangchaivech et al.*] (2019)



Evaluation Metrics

- **BLEU Score:**
  - Bi-Lingual Evaluation Understudy
  - N-grams overlapping between machine translation and reference translation.
  - Compute precision for n-grams of size 1 to 4
  - Used as Evaluation in previous attempts to generate Question
- **METEOR**
  - Uses F1 Measure to compute the values based on mapping of Unigrams
- **$ROUGE_L$**
  - Used more often in text summarization, takes account for Longest Common Subsequence.



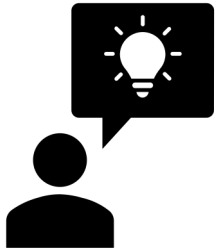
Dataset

SQUaD

- Stanford Question Answering Dataset (SQuAD)
- Reading Comprehension Dataset
- Consisting of Questions posed by crowd workers on a set of Wikipedia articles.
- Dataset Characteristics
  - Training Set: 70,484 Question and Answer Pairs
  - Validation Set: 10,570 Question and Answer Pairs
  - Test Set: 11,877 Question and Answer Pairs

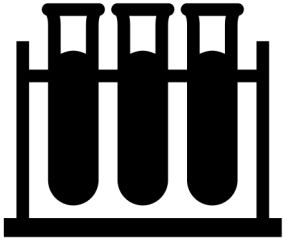
Objectives

- Generate Reading Comprehension Based Questions
- Introduce Convolutional Seq2Seq Approach to generate Reading Comprehension Questions
- Comparative Analysis of performance of different architectures for RC Question Generation Task



ITMO UNIVERSITY

Experimentation

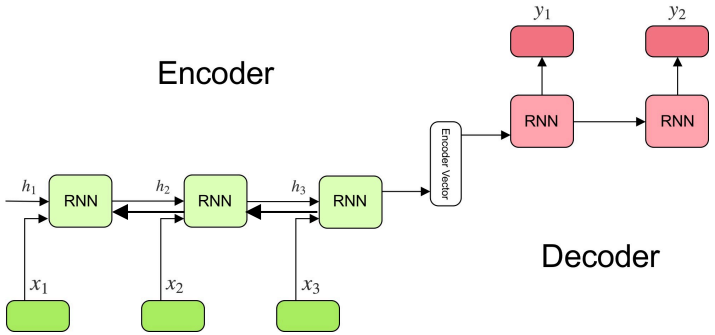


ITMO more than a UNIVERSITY

Experiment

Recurrent Sequence To Sequence

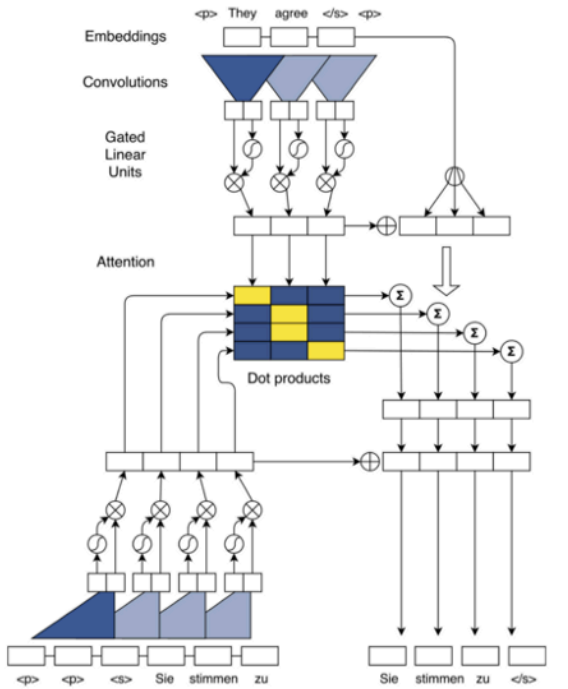
- Sequence to Sequence Learning with Neural Networks (2014) by Sutskever, I
- BiDirectional Encoder based on RNN Neural Network
- Used GloVe Embeddings
- Uses Attention to align the input and output sequences



Experiment

Convolutional Sequence To Sequence

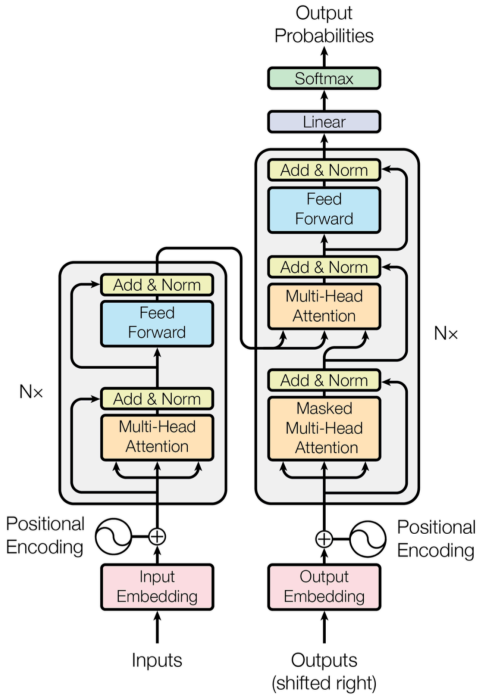
- Introduced in Convolutional Sequence to Sequence Learning (2017) by Gehring, J
- Encoder & Decoder contains Gated Linear Units & Residual Connections
- Used GloVe Embeddings
- Parallel Convolutions leads to faster training



Experiment

Transformer Encoder Decoder Network

- Introduced in Attention Is All You Need (2017) by Vaswani, A
- Encoder contains self attention and feed forward network
- Used GloVe Embeddings
- Decoder contains self attention, encoder-decoder attention layer and a feed forward network



ITMO UNIVERSITY

Result of Question Generation

Model	BLEU1	BLEU2	BLEU3	BLEU4	METEOR	$ROUGE_L$	Training Time (s)
Vanilla RNNSeq2Seq	31.34	13.79	7.36	4.26	9.88	29.73	--
RNN Seq2Seq	**34.17**	**18.51**	**11.78**	**7.92**	**12.97**	**33.72**	**3361**
CNN Seq2Seq	**32.23**	**16.65**	**10.34**	**6.78**	**12.57**	**32.35**	**1044.3**
Transformer Encoder-Decoder	27.74	11.32	5.96	3.48	7.24	27.2	2652.2

ITMO more than a UNIVERSITY

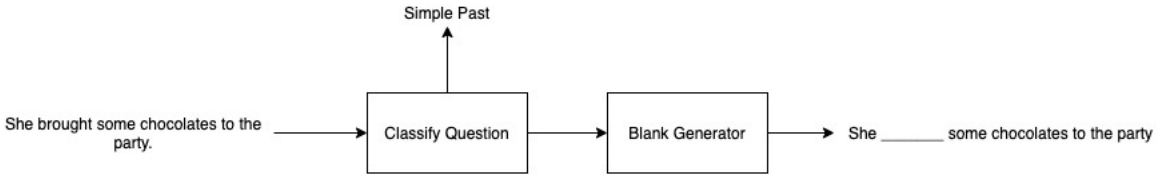
Example of Generated Questions

Sentence	Madonna was a pop singer, she was raised in a methodist household	
Ground Truth	What was Madonna raised in ?	
RNN	**CNN**	**Transformer**
Madonna was raised in what religion?	What was Madonna raised in?	What was the name of the name of the first child that was born ?

Example of Generated Questions

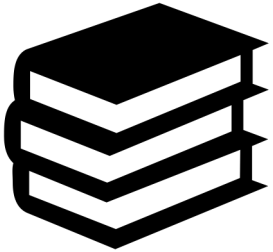
Sentence	Thomas Newman returned as spectre's composer	
Ground Truth	Who wrote the music for spectre?	
RNN	**CNN**	**Transformer**
Who returned as spectre's composer?	What was the name of the spectre's composer ?	What was the name of madonna 's career ?

Generation of Fill in the Blank Questions



Literature Study

- Gap-fill Tests for Language Learners: Corpus-Driven Item Generation [*Perez-Beltrachini et al.*] (2010)
- Measuring Non-native Speakers' Proficiency of English by Using a Test with Automatically-Generated Fill-in-the-Blank Questions [*Sumita et al.*] (2005)
- Learning to Automatically Generate Fill-In-The-Blank Quizzes [*Taylor et al*] (2018)
- Automatic Generation of Fill-in-the-Blank Programming Problems [*Terada et al*] (2019)

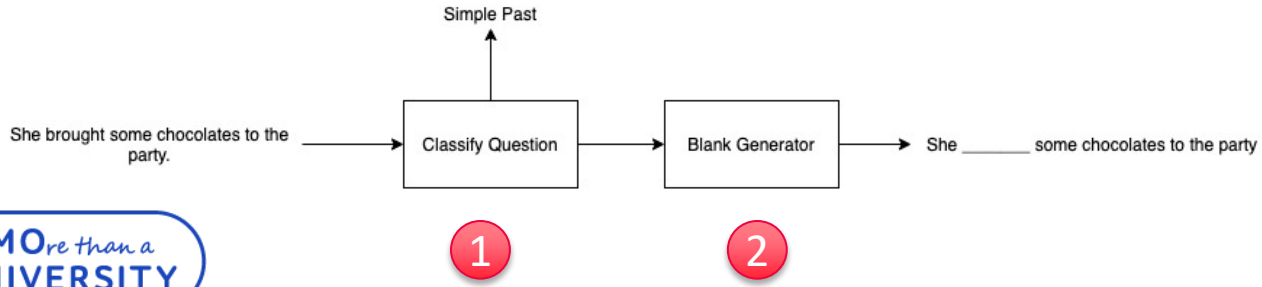


Dataset

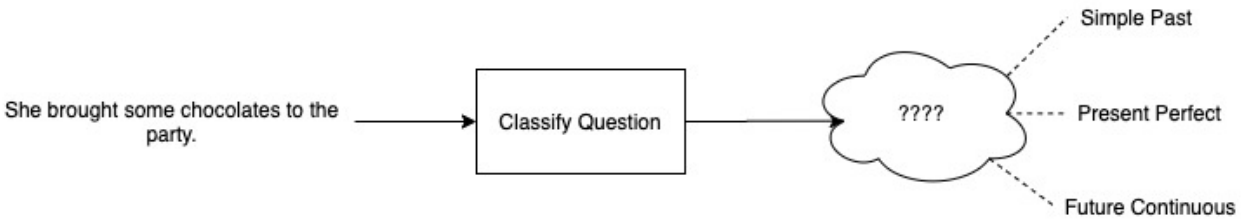
- Scraped dataset from online sources
- The dataset consist of columns: Question, key, answer, Type of Question
  - Trainset: 569 Question, Key, Answer and Type of Question Tuples
  - Test Set: 101 Question, Key, Answer and Type of Question Typles
- Type of Verb Conjugation Exercises Present: 11
- Example:
  - **Question:** Mother _____ the cake for my birthday party
  - **Key:** bring
  - **Answer:** Mother brought the cake for my birthday party
  - **Type of Question:** irregular verb

Methodology

- Two Step Approach
  1. Classify the nature of question
  2. Generate possible location of Blank
     - Sequence Labeling
     - Sequence To Sequence Generation



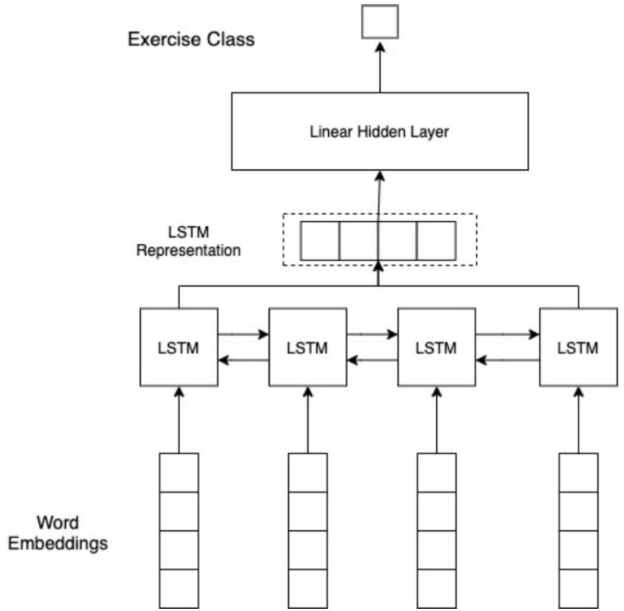
Classification of Nature of Exercise



Experiment

Bi-Directional LSTM Classifier

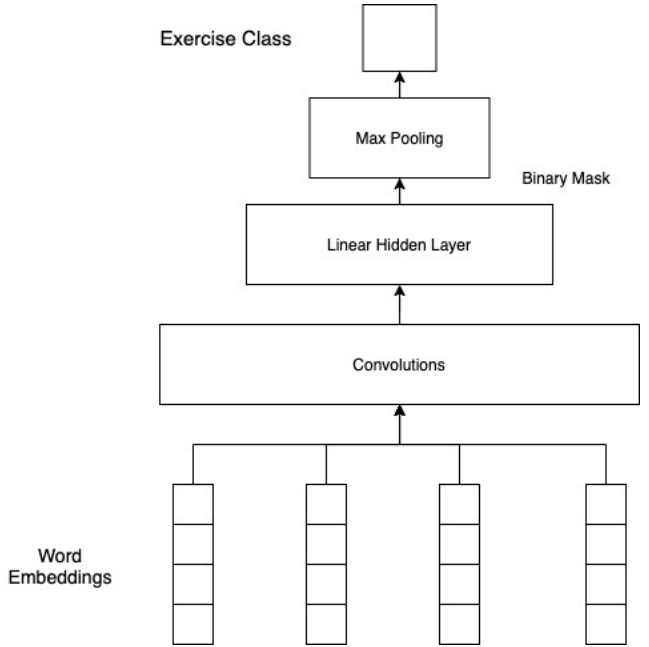
- Concatenating last hidden states of BiDirectional LSTM
- Predict the Nature of Exercise that could be generated from a given sentence
- Number of Trainable Parameters: 443,147



Experiment

Convolutional Classifier with Masking

- 1D - Convolutional Kernels of size 1, 3, 5
- Added a hidden layer after convolutions
- Trained using binary masks
- Number of Trainable Parameters: 199,115



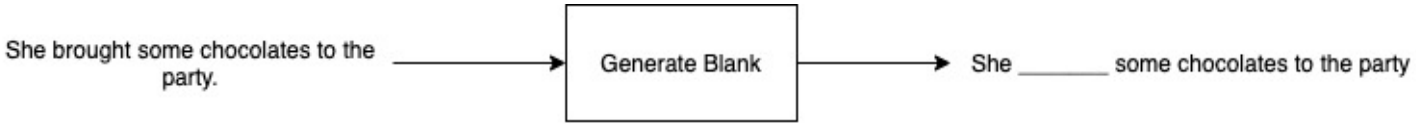
ITMO UNIVERSITY

Result of Question Generation

Model	Classification Accuracy
Naive Bayes	78.22%
Support Vector Machine	81.19%
Logistic Regression	80.20%
Random Forest	76.24%
XGBClassifier	74.26
LSTM Classifier	79.24%
CNNClassifier	**84.29%**

ITMO More than a UNIVERSITY

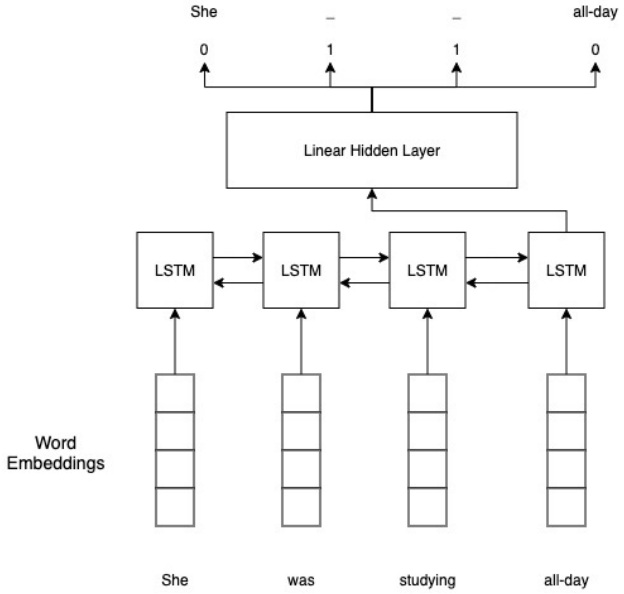
Fill in the Blank Generation



Fill in the Blank Generation

Sequence Labeling

- Classify that the token is a potential blank or not
- 0 is that the token is not a blank and 1 that the token is a blank
- Used RNN Based Classifier to classify each token



Fill in the Blank Generation

Sequence Labeling

- Compared our results with Automatic Generation of Fill-in-the-Blank Programming Problems [*Terada et al*] (2019)

Model	Precision	Recall	F1 Measure	Accuracy
FITBPP	0.66	0.73	0.69	--
Ours	0.61	**0.88**	**0.72**	**76.94%**

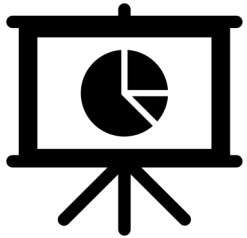
Fill in the Blank Generation

Sequence to Sequence Generation

Model	BLEU1	BLEU2	BLEU3	BLEU4	METEOR	ROUGE$_L$
NQG	43.09	25.96	17.50	12.28	16.62	39.75
RNN Seq2Seq	38.81	25.51	15.63	10.12	8.67	**44.6**
CNN Seq2Seq	37.10	23.12	12.02	6.77	8.13	**41.85**

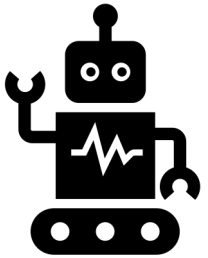
Conclusion

- Generated Reading Comprehension English Learning Exercise Questions as a Sequence to Sequence Problem with RNN based Seq2Seq Models and CNN based Seq2Seq Models
- CNN based Seq2Seq model performs comparably to generate Questions with RNN based models with faster training
- Fill in The Blank Exercises as a Sequence Labeling with an RNN based classifier
- Fill in the Blank Exercises as RNN and CNN based Seq2Seq models



Future Work

- Better evaluation metrics for FITB using Sequence2Sequence.
- Generation of more type of exercises with Sequence labelling
- Integration of Distractor Generation



Thank You for your Attention

GHENT UNIVERSITY

IT'sMOre than a UNIVERSITY

televic education