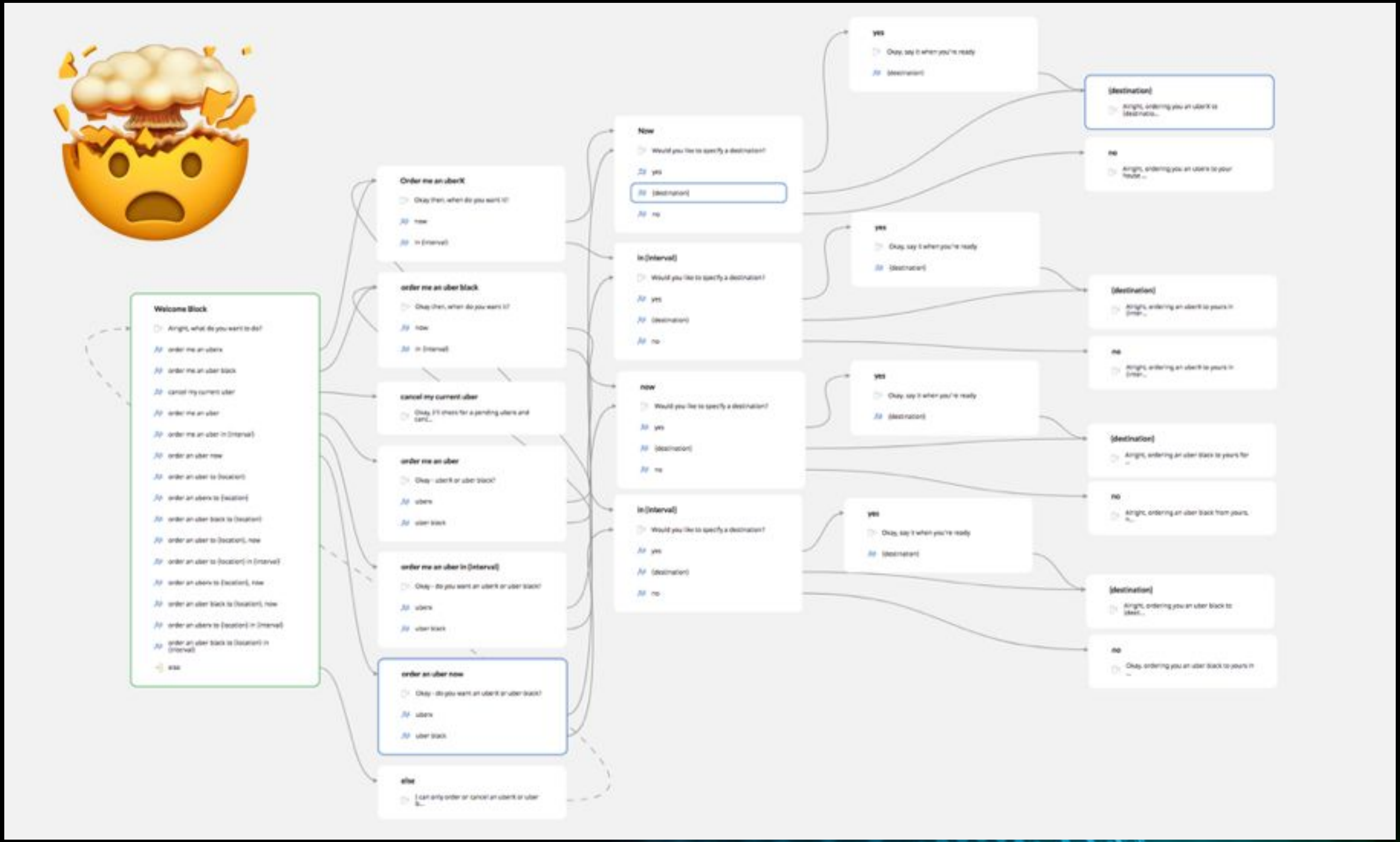
TATIANA SHAVRINA
R&D TEAM, NLP

Sberbank
Artificial Intelligence

Для науки и не только: как мы используем библиотеку DeepPavlov каждый день.

2 года назад...

regex
rules
key-words



DeepPavlov

DeepPavlov

```
## Справка по DeepPavlov
Определились с веткой?
...
### Установка чит-чата
git clone https://github.com/deepmipt/DeepPavlov.git
cd DeepPavlov
git checkout --track origin/feat/chitchat_vs_odqa
virtualenv -p python3.6 dp
source ./dp/bin/activate
pip3 install -r requirements.txt
$ git clone https://github.com/facebookresearch/fastText.git
$ cd fastText
$ pip install .
python setup.py install
python -m deeppavlov install deeppavlov/configs/classifiers/config_chitchat_vs_odqa_ft.json
python -m deeppavlov download deeppavlov/configs/classifiers/config_chitchat_vs_odqa_ft.json

/datadrive/class4/DeepPavlov/utils/settings/server_config #- поменять порт

python -m deeppavlov riseapi deeppavlov/configs/classifiers/config_chitchat_vs_odqa_ft.json
```

Многообразие диалоговых систем

- Chit-chat
- Dialog managers, Dialog State Tracking
- ODQA
- Skills
- Domain Classification
- Emotional State, Sentiment, Personalization
- Slot-filling
- Text generation
- Few-shot classifiers

Deep Pavlov

Не только чат-боты!

- Морфология и синтаксис
- Sentiment
- NER
- SQuAD
- BERT, ELMo
- классификаторы:
  - оскорбления
  - намерения
  - информативное/неинформативное сообщение
- диалоговые агенты
- чит-чат на общие темы

2019

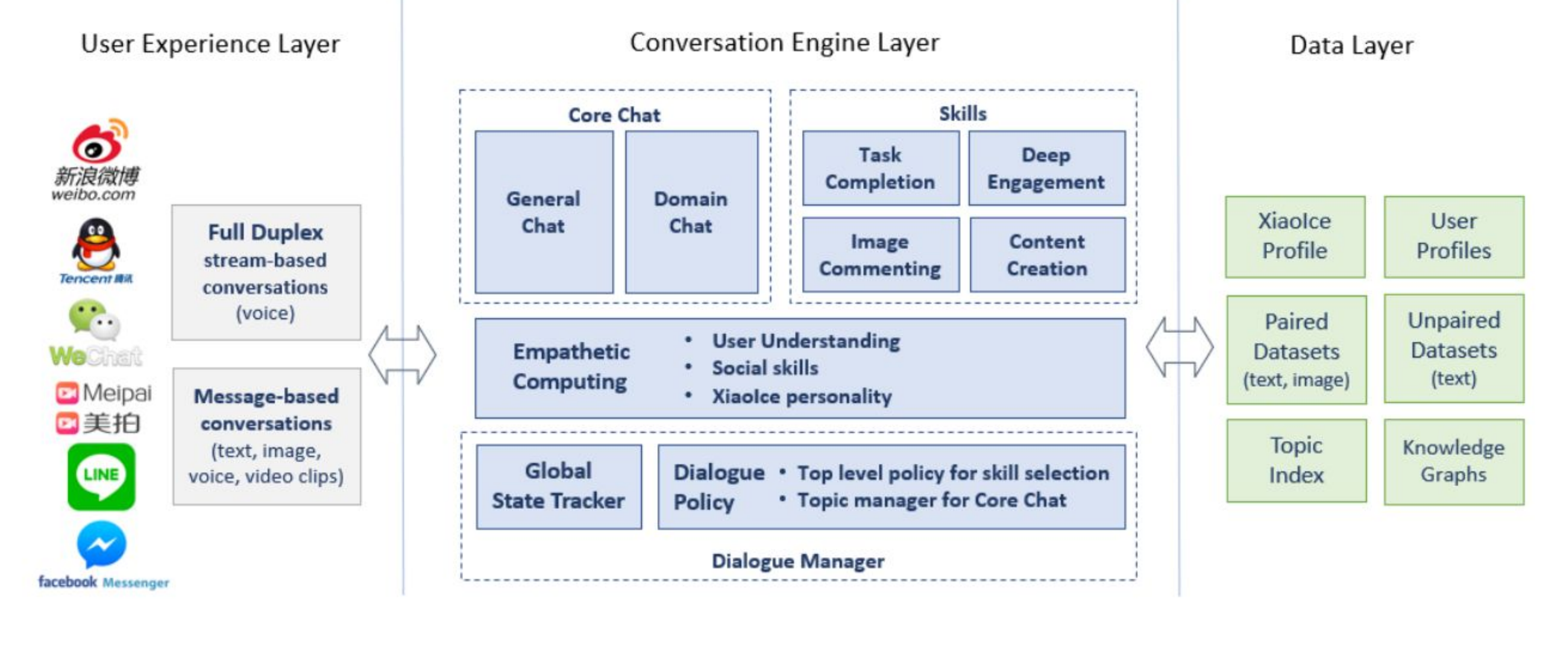
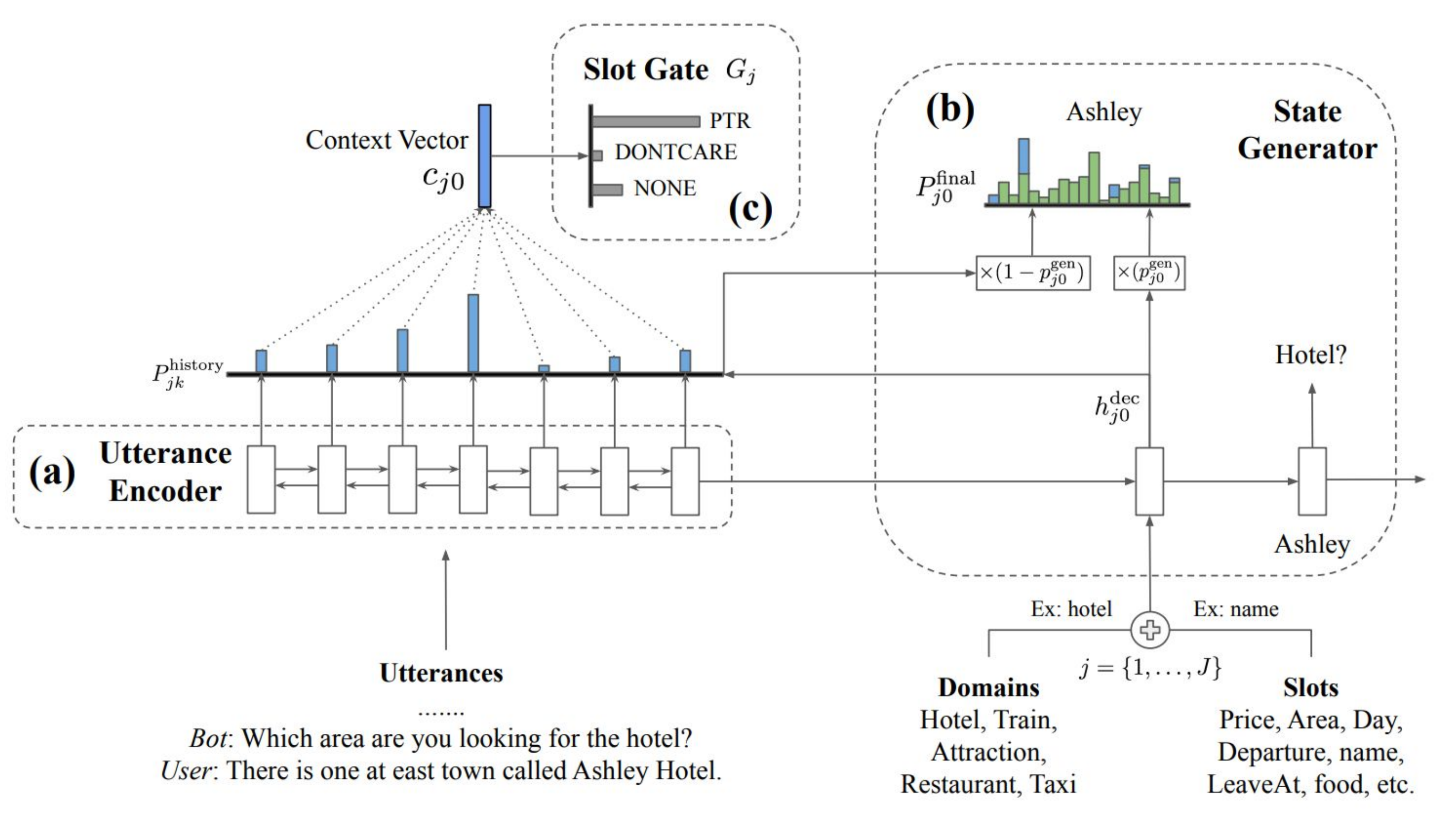


Figure 4: XiaoIce system architecture.

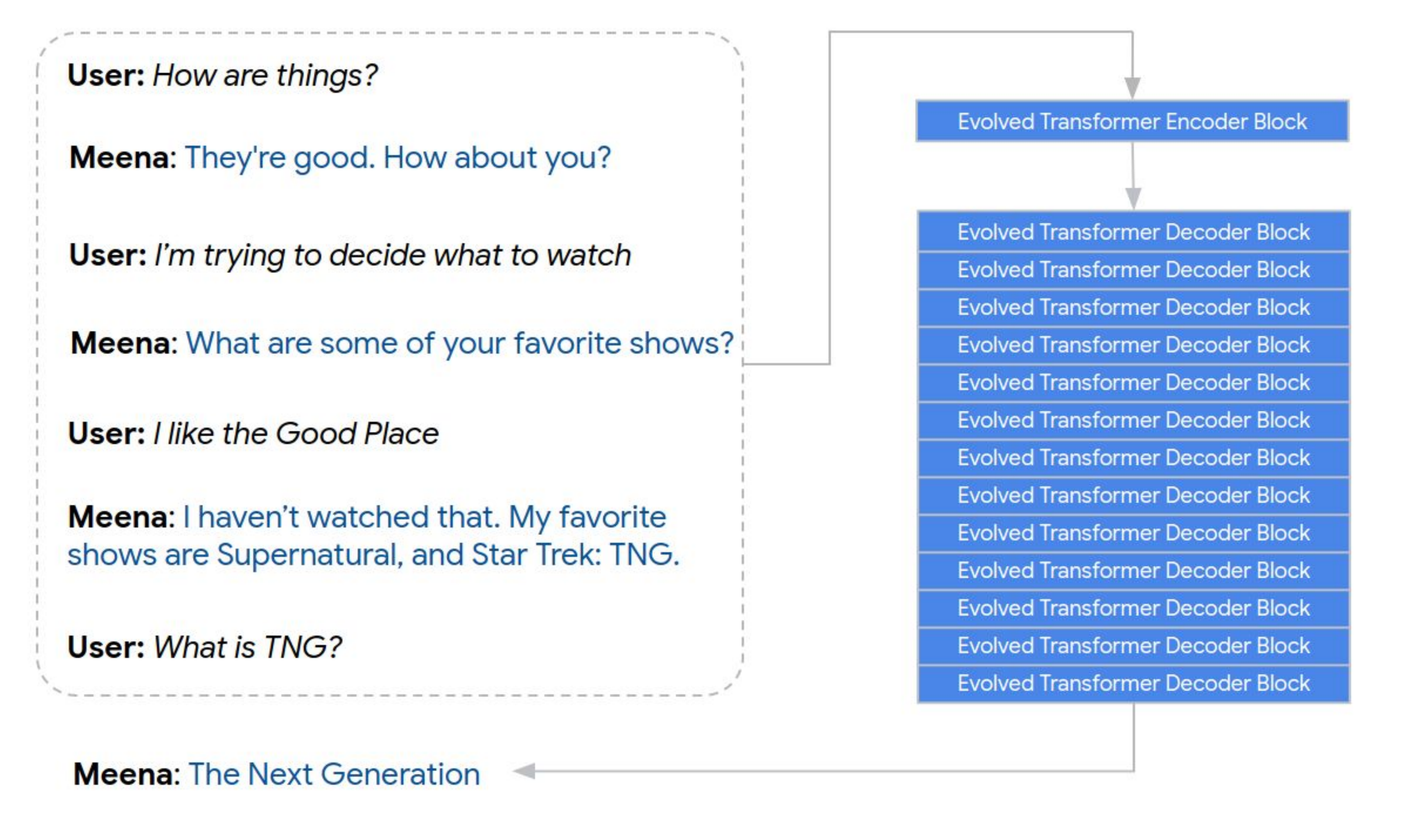
The Design and Implementation of XiaoIce, an Empathetic Social Chatbot

2019



Transferable Multi-Domain State Generator for Task-Oriented Dialogue Systems

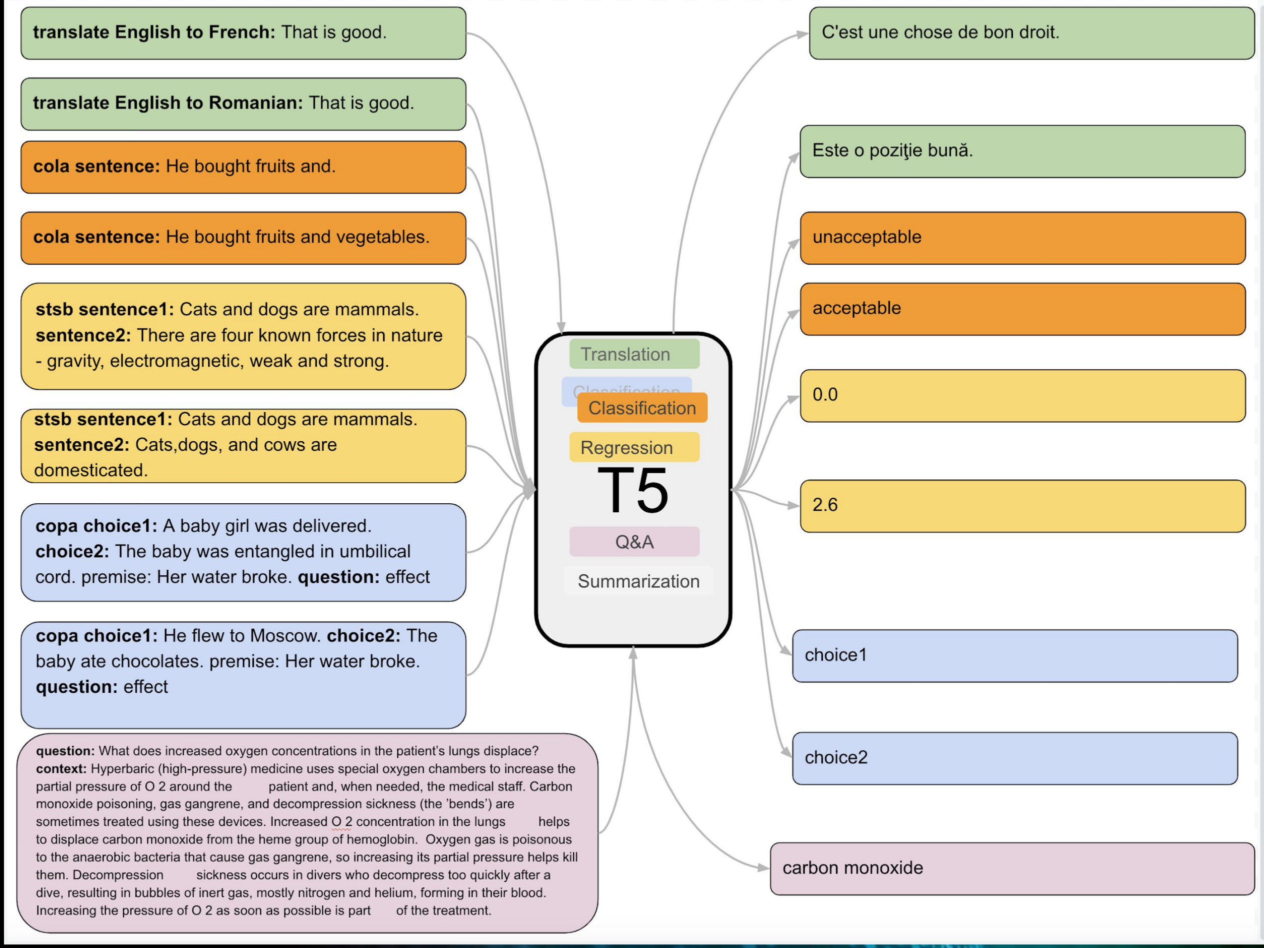
2020



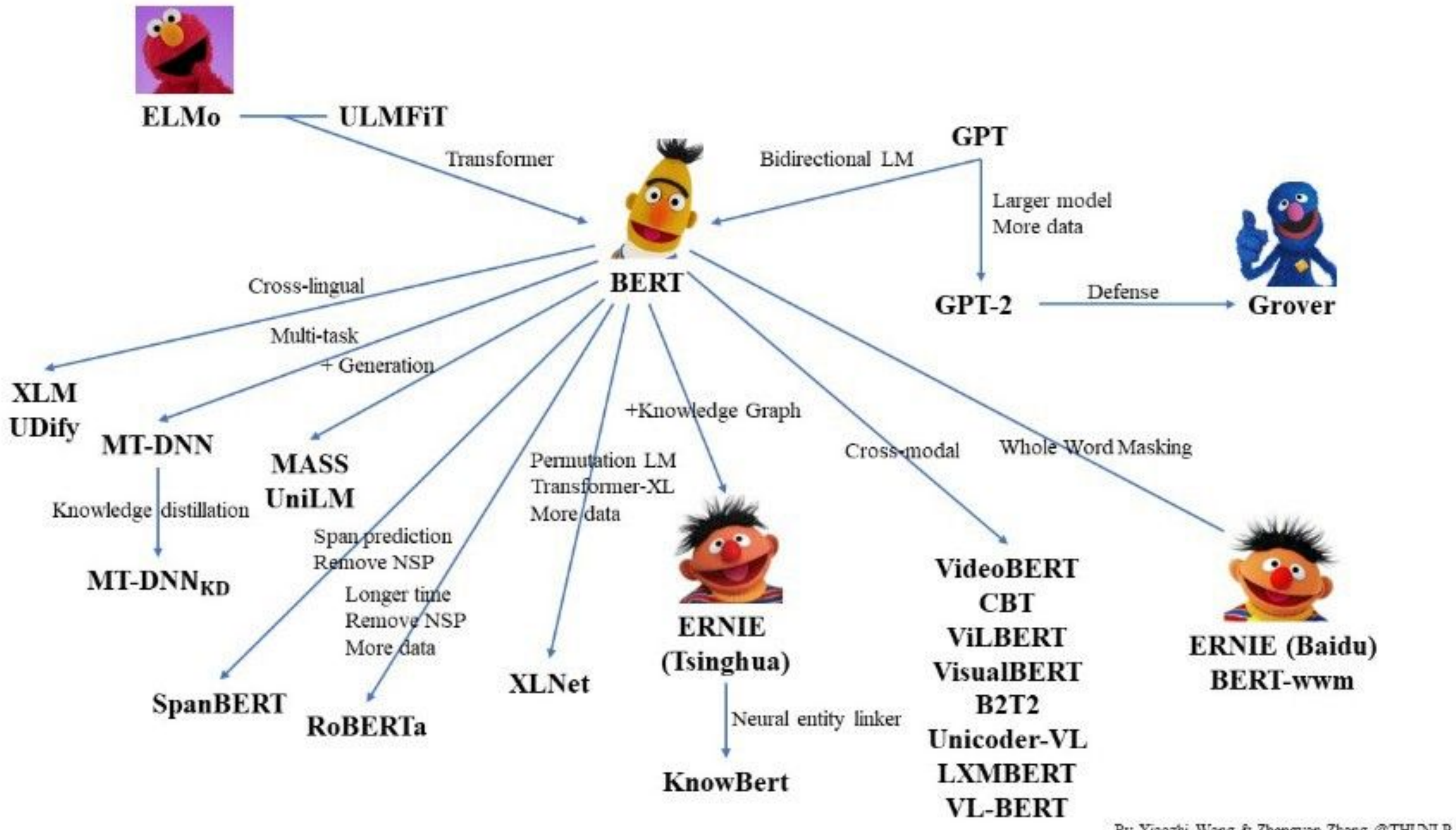
Meena - Towards a Human-like Open-Domain Chatbot

2020 - Text is enough

T5 - Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer



Universal Models
& Transformers


ELMo
ULMFiT
Transformer
BERT
GPT
Bidirectional LM
Larger model
More data
GPT-2
Defense
Grover
Cross-lingual
Multi-task
+ Generation
XLM
UDify
MT-DNN
Knowledge distillation
MT-DNN$_{KD}$
MASS
UniLM
Span prediction
Remove NSP
SpanBERT
Longer time
Remove NSP
More data
RoBERTa
Permutation LM
Transformer-XL
More data
XLNet
+Knowledge Graph
ERNIE (Tsinghua)
Neural entity linker
KnowBert
Cross-modal
VideoBERT
CBT
ViLBERT
VisualBERT
B2T2
Unicoder-VL
LXMBERT
VL-BERT
Whole Word Masking
ERNIE (Baidu)
BERT-wwm
By Xiaozhi Wang & Zhengyan Zhang @THUNLP

DeepPavlov Pretrained Embeddings

Pre-trained embeddings

BERT

We are publishing several pre-trained BERT models:

- RuBERT for Russian language
- Slavic BERT for Bulgarian, Czech, Polish, and Russian
- Conversational BERT for informal English
- and Conversational BERT for informal Russian

Description of these models is available in the BERT section of the docs.

License

The pre-trained models are distributed under the License Apache 2.0.

Downloads

The models can be run with the original BERT repo code. The download links are:

Description	Model parameters	Download link
RuBERT	vocab size = 120K, parameters = 180M, size = 632MB	[rubert_cased_L-12_H-768_A-12]
Slavic BERT	vocab size = 120K, parameters = 180M, size = 632MB	[bg_cs_pl_ru_cased_L-12_H-768_A-12
Conversational BERT	vocab size = 30K, parameters = 110M, size = 385MB	[conversational_cased_L-12_H-768_A-
Conversational RuBERT	vocab size = 120K, parameters = 180M, size = 630MB	[conversational_cased_L-12_H-768_A-

ELMo

We are publishing `Russian language ELMo embeddings model` for tensorflow-hub and `LM model` for training and fine-tuning ELMo

DeepPavlov Pretrained Embeddings

Downloads

The `TensorFlow` models can be run with the original BERT repo code while the `PyTorch` models can be run with the HuggingFace's Transformers library. The download links are:

Description	Model parameters	Download links
RuBERT	vocab size = 120K, parameters = 180M, size = 632MB	[tensorflow], [pytorch]
Slavic BERT	vocab size = 120K, parameters = 180M, size = 632MB	[tensorflow], [pytorch]
Conversational BERT	vocab size = 30K, parameters = 110M, size = 385MB	[tensorflow], [pytorch]
Conversational RuBERT	vocab size = 120K, parameters = 180M, size = 630MB	[tensorflow], [pytorch]
Sentence Multilingual BERT	vocab size = 120K, parameters = 180M, size = 630MB	[tensorflow], [pytorch]
Sentence RuBERT	vocab size = 120K, parameters = 180M, size = 630MB	[tensorflow], [pytorch]

ELMo

We are publishing `Russian language ELMo embeddings model` for tensorflow-hub and `LM model` for training and fine-tuning ELMo as LM model.
ELMo (Embeddings from Language Models) representations are pre-trained contextual representations from large-scale bidirectional language models. See a paper Deep contextualized word representations for more information about the algorithm and a detailed analysis.

License

The pre-trained models are distributed under the License Apache 2.0.

Как мы решали ЕГЭ

AI Journey 2019: решить экзамен по русскому языку

26 заданий - с открытыми и закрытыми вариантами ответов (59% от оценки)
+ сочинение по тексту (41% от оценки)

Типы заданий:

- орфография
- логика
- семантика
- пунктуация
- орфоэпия (ударения)
- морфология и синтаксис
- составление / генерация текста

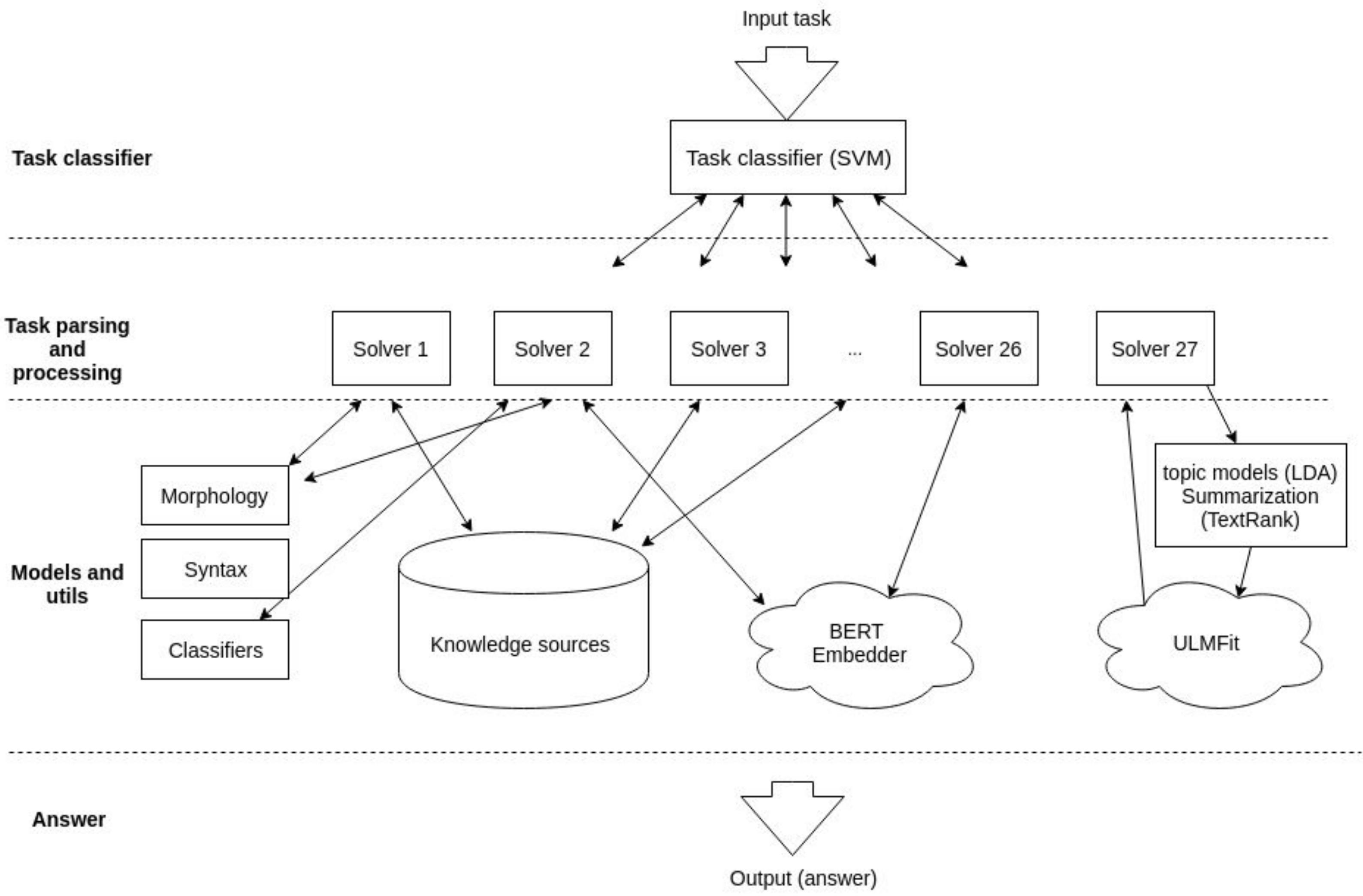
Как мы решали ЕГЭ

№ 9. Укажите варианты ответов, в которых во всех словах одного ряда пропущена безударная чередующаяся гласная корня. Запишите номера ответов.

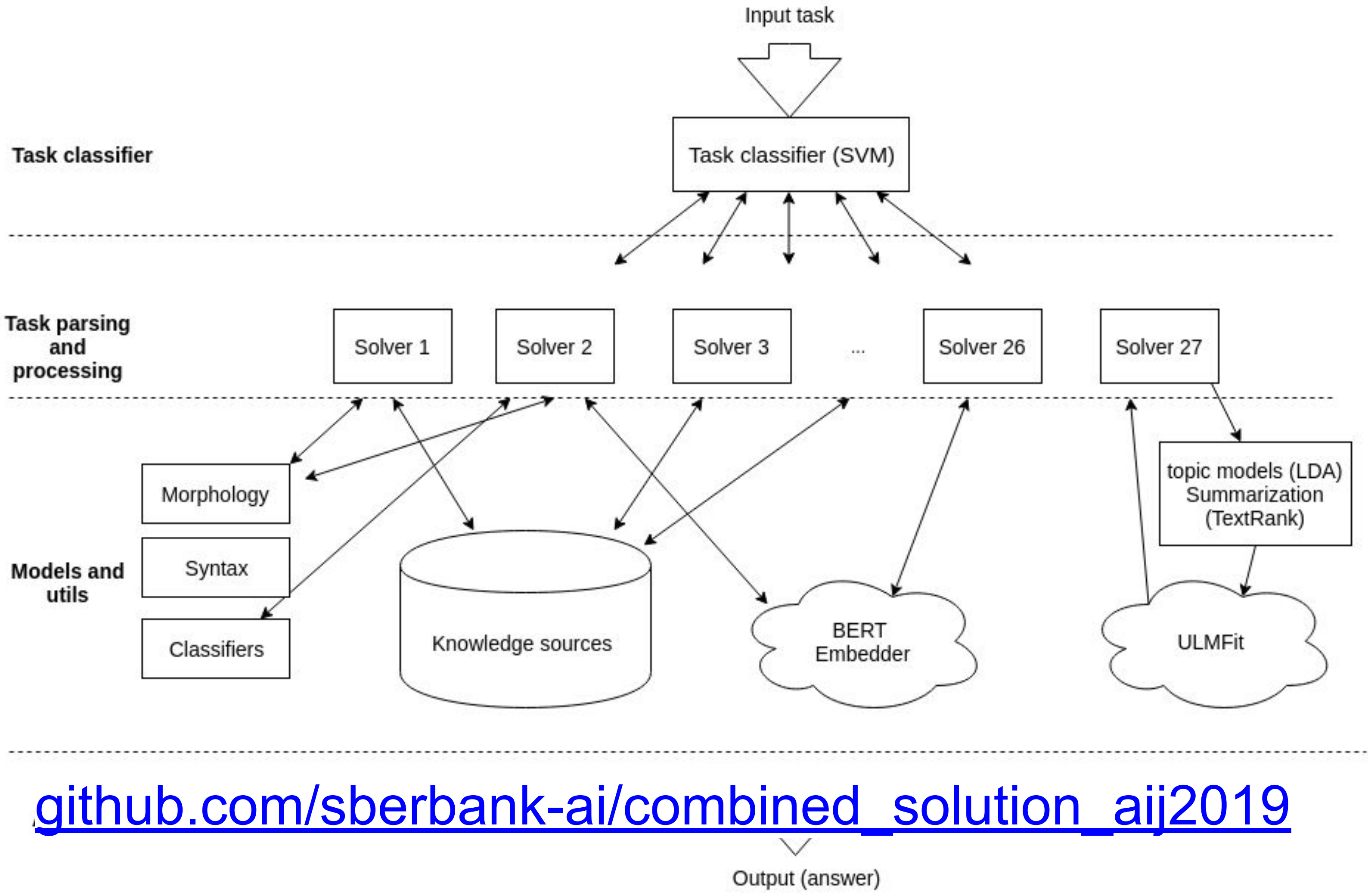
1) зап..рать, р..стение, прил..гательное
2) сп..раль, заст..лить, к..мфорт
3) б..режок, ф..рмат, затв..рдеть
4) предв..рительный, прид..рожный, зам..чать
5) тв..рительный, з..рница, пл..вец

Ответ: 1, 5

BASELINE


Input task
Task classifier
Task classifier (SVM)
Task parsing and processing
Solver 1
Solver 2
Solver 3
...
Solver 26
Solver 27
Models and utils
Morphology
Syntax
Classifiers
Knowledge sources
BERT Embedder
topic models (LDA) Summarization (TextRank)
ULMFit
Answer
Output (answer)

BASELINE



github.com/sberbank-ai/combined_solution_aij2019

Половина заданий ЕГЭ решена BERTом

Semantics

Task 1: one or more sentences containing thegeneral information on the task text with 5 choices provided (Answer type: multiple choice).
Solution: all of the sentences relevant to the text areexpected to be close by cosine similarity of BERT embeddings; returning the two closest options (Est.accuracy:0.7).

Task 3: select the most relevant word meaning inthe given context with 5 choices provided (Answer type: choice).
Solution: the problem is treated as a masked task and solved with BERT model by predicting the missing token (Est. accuracy:0.7)

Половина заданий ЕГЭ решена BERTом

Logic

Task 2: fill in a gap between sentences or textparts with the most relevant logical connector or a conjunction without any options provided (Answer type: text).
Solution: a combination of BERT’s maskedtasks, morphological analysis by pymorphy2 and a custom list of candidates (Est. accuracy:0.53).

Task 22: select one or more statements relevant to a task text content with 5 choices provided (Answer type: multiple choice).
Solution: training a binary classifier over BERT embeddings that outputs 1 if a choice is relevant to the task text (Est. accuracy:0.63).

Половина заданий ЕГЭ решена BERTом

Grammar

Task 5: select and replace an incorrect wordwith a paronym within 5 sentences (Answer type: text).
Solution: dictionary lookup to retrieve potential candidates; thesecandidates are used as input for BERT masked token prediction to score potential replacements (Est.accuracy:0.66)

Task 8: matching the correspondence of 5 grammatical error types with 9 provided sentences (Answer type: matching).
Solution: parsing the question with a rule-based approach; multi-class classification over BERT embeddings. (Estimated accuracy:0.86)

Половина заданий ЕГЭ решена BERTом

Spelling

Task 13: select one out of 5 sentences in which the specified word is written separately with the previous one (Answer type: text)
Solution: using BERT for binary classification (Est. accuracy:1.00).
Task 14: select one out of 5 sentences in which two specific words are written separately in the given context (Answer type: text).
Solution: getting the impossible spellings with morphology analysis and scoring the remaining candidate spellings with BERT after replacing the candidate word with a [MASK] token (Est. accuracy:0.83).
Task 15: select gaps corresponding to the specified spelling, typically “Н” or “НН” letters in a suffix (Answer type: text).
Solution: a BERT classifier trained to predict whether a masked gap stands for a letter combination mentioned in the task defi-nition (Est. accuracy:0.83)

Половина заданий ЕГЭ решена BERTом

Punctuation

Tasks 17-20: restore sentence punctuation and select the gaps with a comma in the given context(Answer type: multiple choice).
Solution: Masked BERT’s output to decide if this placeholder replaces a comma, also some probability thresholds are used to evict corner cases
(Est. accuracy:0.90,0.56,0.86,0.9 respectively)

Discourse and text analysis

Task 26: one-to-one matching of 4 sentences with 9 out of 40 possible versatile literary means (Answer type: matching).
Solution: using BERT for multi-class classification; combining the baseline approach with the rules (Est. accuracy:0.75)

contest.ai-journey.ru/ru/leaderboard

#	Команда	Тест (max 34)	Сочинение (max 24)	Итого (max 100)	Последнее решение	Попыток
1	qbic	19,93	14,25	59,77	01 ноября 2019, 23:56	48
2	Bilbo Bagging	20,90	12,67	58,47	01 ноября 2019, 18:09	170
3	Magic City	14,50	16,33	55,63	01 ноября 2019, 23:47	115
4	borsden	15,77	13,33	53,40	29 октября 2019, 05:54	74
5	Ololosh AI	16,70	10,58	50,93	30 октября 2019, 21:06	150
6	nice	12,63	13,83	49,70	01 ноября 2019, 16:56	40
7	Niw	21,20	4,75	49,20	01 ноября 2019, 19:45	111
OOC	CDS_team	17,17	5,42	45,23	01 ноября 2019, 18:02	21
8	stickman	17,07	5,00	43,77	01 ноября 2019, 14:57	70
8	Orcs	13,00	8,92	43,77	01 ноября 2019, 23:58	19

Eventually
Russian AI ...
got 69 out of 100
for the exam!

exBERT.net

[***exBERT: A Visual Analysis Tool to Explain BERT's Learned Representations***](#)

exBERT — это интерактивный инструмент, представленный компанией IBM, который позволяет пользователям исследовать что и как выучил трансформер в процессе создания языковой модели.

Работает с одним предложением!

Demos. AllenNLP

[AllenNLP Interpret: Explaining Predictions of NLP Models]

1) набор методов интерпретации, применимых к большинству моделей, доступных в библиотеке:
    - маскирование языковых моделей
    - анализ тональности
    - SQuAD
    - Textual Entailment
    - etc.
2) API для разработки новых методов интерпретации
3) многократно используемые веб-компоненты для визуализации результатов

Можно загружать свои модели!

Demo Usage

Enter text or

Choose an example...

Premise

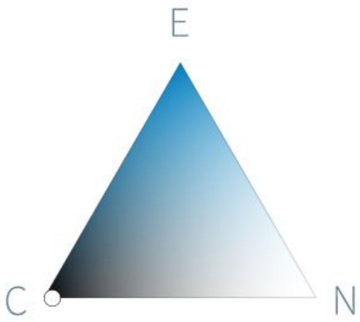
Two women are wandering along the shore drinking iced tea.

Hypothesis

Two women are sitting on a blanket near some rocks talking about politics.

Summary

It is **very likely** that the premise **contradicts** the hypothesis.



Judgment	Probability
Entailment	0%
Contradiction	97%
Neutral	3%

Model Interpretations *What is this?*

SuperGLUE GLUE Paper Code Tasks Leaderboard FAQ Diagnostics Submit

Leaderboard Version: **2.0**

	Rank	Name	Model	URL	Score	BoolQ	CB	COPA	MultiRC	ReCoRD	RTE	WiC	WSC	AX-b	AX-g
	1	SuperGLUE Human Baselines	SuperGLUE Human Baselines		89.8	89.0	95.8/98.9	100.0	81.8/51.9	91.7/91.3	93.6	80.0	100.0	76.6	99.3/99.7
+	2	T5 Team - Google	T5		89.3	91.2	93.9/96.8	94.8	88.1/63.3	94.1/93.4	92.5	76.9	93.8	65.6	92.7/91.9
	3	Zhuiyi Technology	RoBERTa-mtl-adv		85.7	87.1	92.4/95.6	91.2	85.1/54.3	91.7/91.3	88.1	72.1	91.8	58.5	91.0/78.1
	4	Facebook AI	RoBERTa		84.6	87.1	90.5/95.2	90.6	84.4/52.5	90.6/90.0	88.2	69.9	89.0	57.9	91.0/78.1
	5	IBM Research AI	BERT-mtl		73.5	84.8	89.6/94.0	73.8	73.2/30.5	74.6/74.0	84.1	66.2	61.0	29.6	97.8/57.3
	6	SuperGLUE Baselines	BERT++		71.5	79.0	84.8/90.4	73.8	70.0/24.1	72.0/71.3	79.0	69.6	64.4	38.0	38.0 .4
			BERT		69.0	77.4	75.7/83.6	70.6	70.0/24.1	72.0/71.3	71.7	69.6	64.4	23.0	97.8/51.7
			Most Frequent Class		47.1	62.3	21.7/48.4	50.0	61.1/0.3	33.4/32.5	50.3	50.0	65.1	0.0	100.0/50.0
			CBoW		44.5	62.2	49.0/71.2	51.6	0.0/0.5	14.0/13.6	49.7	53.1	65.1	-0.4	100.0/50.0
			Outside Best		-	80.4	-	84.4	70.4/24.5	74.8/73.0	82.7	-	-	-	-
	-	Stanford Hazy Research	Snorkel [SuperGLUE v1.9]		-	-	88.6/93.2	76.2	76.4/36.3	-	78.9	72.1	72.6	47.6	-

Click on a submission to see more information

Спасибо за внимание!

Leaderboard Version: **2.0**

Rank	Name	Model	URL	Score	BoolQ	CB	COPA	MultiRC	ReCoRD	RTE	WiC	WSC	AX-b	AX-g
1	SuperGLUE Human Baselines	SuperGLUE Human Baselines		89.8	89.0	95.8/98.9	100.0	81.8/51.9	91.7/91.3	93.6	80.0	100.0	76.6	99.3/99.7
2	T5 Team - Google	T5		89.3	91.2	93.9/96.8	94.8	88.1/63.3	94.1/93.4	92.5	76.9	93.8	65.6	92.7/91.9
3	Zhu												58.5	91.0/78.1
4	Fac												57.9	91.0/78.1
5	IBM												29.6	97.8/57.3
6	SuperGLUE Baselines	BERT++		71.5	79.0	84.8/90.4	73.8	70.0/24.1	72.0/71.3	79.0	69.6	64.4	38.0	[illegible]
		BERT		69.0	77.4	75.7/83.6	70.6	70.0/24.1	72.0/71.3	71.7	69.6	64.4	23.0	97.8/51.7
		Most Frequent Class		47.1	62.3	21.7/48.4	50.0	61.1/0.3	33.4/32.5	50.3	50.0	65.1	0.0	100.0/50.0
		CBoW		44.5	62.2	49.0/71.2	51.6	0.0/0.5	14.0/13.6	49.7	53.1	65.1	-0.4	100.0/50.0
		Outside Best		-	80.4	-	84.4	70.4/24.5	74.8/73.0	82.7	-	-	-	-
-	Stanford Hazy Research	Snorkel [SuperGLUE v1.9]		-	-	88.6/93.2	76.2	76.4/36.3	-	78.9	72.1	72.6	47.6	-

38.0

Click on a submission to see more information