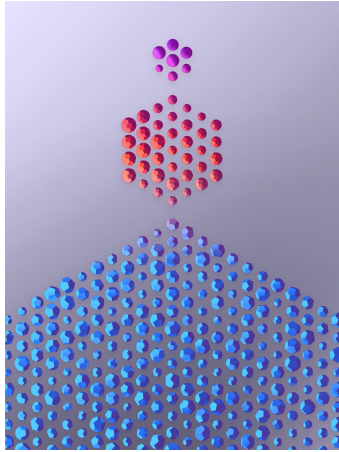


Introduction to BERT

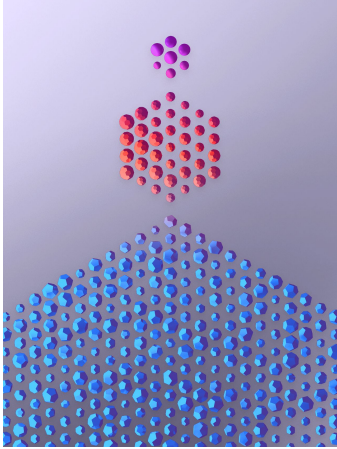
Transformer Based Models



OpenAI's GPT

Analytics
Vidhya

Transformer Based Models

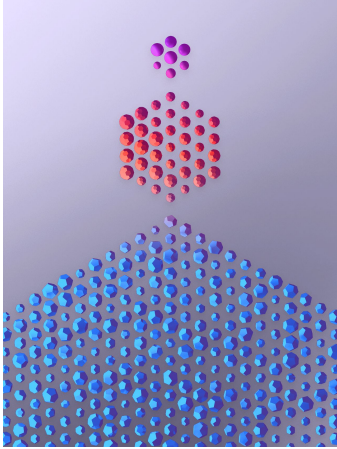


OpenAI's GPT



Google's BERT

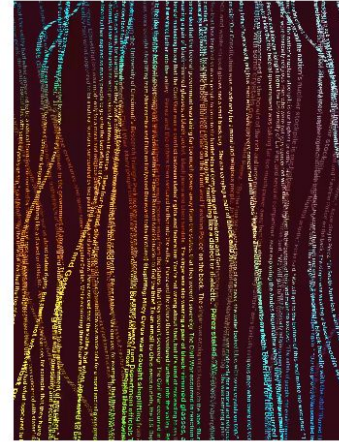
Transformer Based Models



OpenAI's GPT



Google's BERT



OpenAI's GPT-2

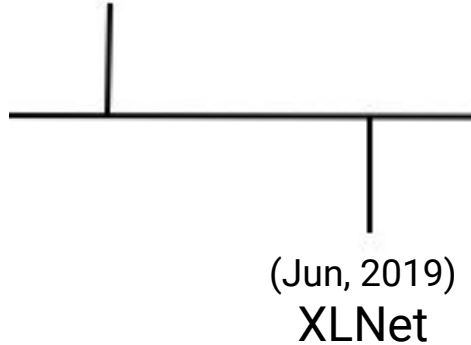
Transformer Based Models

Transformer-XL
(Jan, 2019)



Transformer Based Models

Transformer-XL
(Jan, 2019)



Analytics
Vidhya

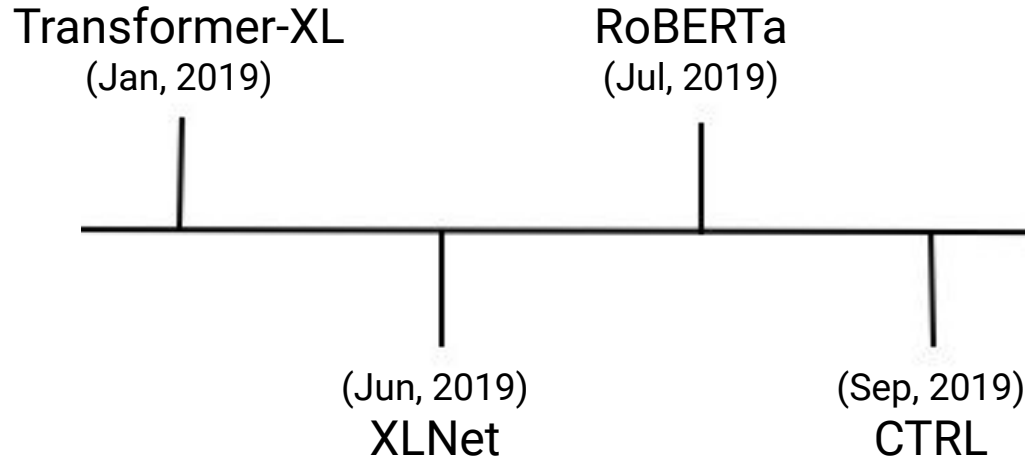
Transformer Based Models

Transformer-XL
(Jan, 2019)

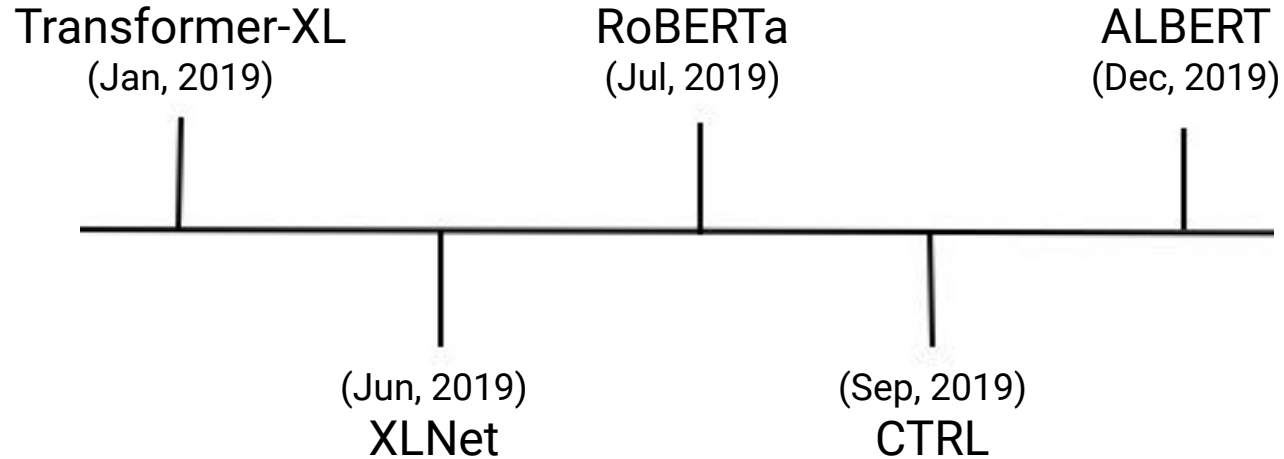
RoBERTa
(Jul, 2019)

(Jun, 2019)
XLNet

Transformer Based Models



Transformer Based Models



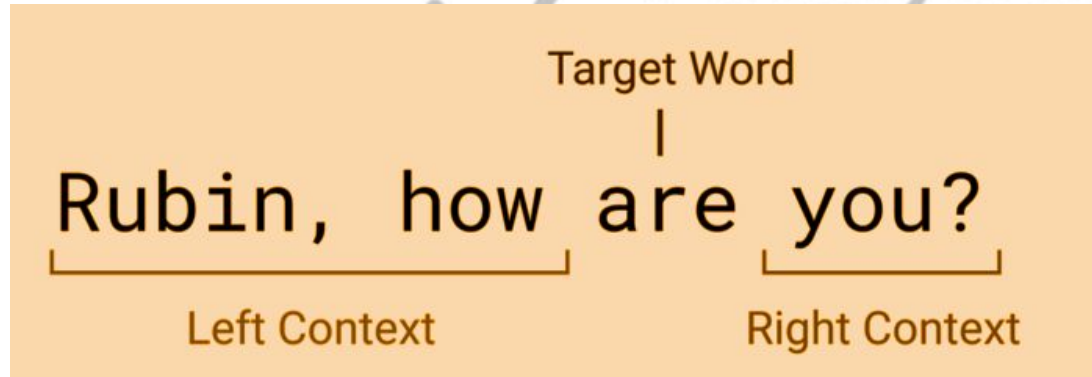
BERT

- **BERT** - Bidirectional Encoder Representations from Transformers



BERT

- **BERT** - **Bidirectional** Encoder Representations from Transformers



BERT

- **BERT** - Bidirectional **Encoder Representations** from Transformers



BERT

- **BERT** - Bidirectional Encoder Representations from Transformers
- BERT is pre-trained on a large corpus of unlabelled text
 - Wikipedia(that's 2,500 million words)
 - Book Corpus (800 million words)



BERT

- **BERT** - Bidirectional Encoder Representations from Transformers
- BERT is pre-trained on a large corpus of unlabelled text
 - Wikipedia(that's 2,500 million words!)
 - Book Corpus (800 million words)
- BERT Models:
 - BERT-base (12 Encoder Stack)
 - BERT-base (24 Encoder Stack)



Performance of BERT

- **GLUE (General Language Understanding Evaluation) Benchmark**

- It is a collection of datasets used for training, evaluating, and analyzing NLP models relative to one another

Rank	Model	Score	CoLA	SST-2	MRPC	STS-B	QQP	MNLI-m	QNLI	RTE
1	BERT: 24-layers, 1024-hidden, 16-heads	80.4	60.5	94.9	85.4/89.3	87.6/86.5	89.3/72.1	86.7	91.1	70.1
2	Singletask Pretrain Transformer	72.8	45.4	91.3	75.7/82.3	82.0/80.0	88.5/70.3	82.1	88.1	56.0
3	BiLSTM+ELMo+Attn	70.5	36.0	90.4	77.9/84.9	75.1/73.3	84.7/64.8	76.4	79.9	56.8

Applications of BERT

- Generate Text Embeddings



Applications of BERT

- Generate Text Embeddings
- Text Classification



Applications of BERT

- Generate Text Embeddings
- Text Classification
- Named Entity Recognition



Applications of BERT

- Generate Text Embeddings
- Text Classification
- Named Entity Recognition
- Question & Answering





Thank You