

# Finding BERT's Idiomatic Key

Vasudevan Nedumpozhimana and John D. Kelleher

ADAPT Research Center  
Technological University Dublin  
Dublin, Ireland

{vasudevan.nedumpozhimana, john.d.kelleher}@tudublin.ie

06 Aug 2021

# Introduction

- Idiom token identification - identify whether a particular usage of an idiomatic expression is idiomatic or not
- Possible to train a generic idiom token identification model using distributed representations
- BERT - State of the art language model for distributed representations

# Research Questions

- 1 Does BERT encode idiomatic usage information?
- 2 Assuming BERT does encode idiomatic usage then which part of the sentence provides signal to BERT - Where is BERT's idiomatic key?
  - Inside the idiomatic expression?
  - Outside the idiomatic expression - Surrounding context?

# Baseline Experiments

## Data Preparation

- VNIC dataset - 2984 sentences, 56 idiomatic expressions
- Down sampled 20 different versions of balanced data
  - 550 idiomatic and 550 literal samples
- 80% samples for training & 20% samples for testing

## Embedding

- BERT pretrained model
- Sentence embedding - Average final layer embeddings of each token in the sentence

# Probing

## Probing

- Probing using MLP classifier
- Train MLP to predict whether a sentence is idiomatic or literal
- Idiomaticity score - Probability score by MLP
  - Literal sentences - Probability will be close to 0
  - Idiomatic sentences - Probability will be close to 1

# Baseline Results

## Idiomatic sentences

- Mean idiomaticity 0.85
- Ideal idiomaticity 1.00

## Literal sentences

- Mean idiomaticity 0.17
- Ideal idiomaticity 0.00

## Observations

- MLP effectively predicts idiomaticity
- BERT encodes idiomatic usage information

# Masking Experiments

## Masking

- Target Expression - Idiomatic expression
- Random Words - Two other random words

## Masking Strategy

- Word masking - Replace words with [MASK]
- Embedding masking - Exclude embeddings of words from sentence embedding

## Differential Idiomaticity (DI<sub>d</sub>)

- Idiomaticity without masking - Idiomaticity with masking
- Higher absolute value of DI<sub>d</sub> → More idiomatic information in the masked part

# Masking Results

Masking	Idiomatic			Literal		
	Id	DId	p-value	Id	DId	p-value
Baseline	0.85	-	-	0.17	-	-
Target Expn + Word Mask	0.79	0.06 (0.0559)	1.12E-05	0.24	-0.08 (0.0548)	2.83E-07
Target Expn + Emb Mask	0.83	0.02 (0.0105)	1.91E-11	0.19	-0.02 (0.0081)	4.07E-16
Rand Word + Word Mask	0.83	0.02 (0.0411)	0.026	0.17	0.00 (0.0379)	0.854
Rand Word + Emb Mask	0.85	0.00 (0.0053)	0.313	0.17	0.00 (0.0050)	0.378

**Table:** Mean Idiomaticities (Id) and Mean Differential Idiomaticities (DId) and p-values



# Where is Idiomatic Key?

## Target Expression Masking

- Statistically significant impact by using both word masking & embedding masking on both idiomatic & literal sentences

## Random Word Masking

- Idiomatic sentences - Statistically significant impact using word masking but no significant impact using embedding masking
- Literal sentences - No statistically significant impact by using both word masking & embedding masking

## Observations

- Idiomatic key is primarily in idiomatic expression
- Some information is in surrounding context

## In what form is the Idiomatic Key?

### Effect of Incongruity due to Idiomatic usage vs Effect of Disruption due to Masking

- Embedding masking - Less disruption due to masking
- Statistically significant impact with embedding masking of Target Expression
- Embedding masking has less impact than word masking
  - BERT encodes information from masked words in the embeddings of other words

### Observation

- BERT can distinguish incongruity caused by idiomatic usage from disruption caused by masking

# Conclusions & Future Work

## Conclusions

- BERT encodes idiomatic usage information
- Idiomatic key primarily found within the idiomatic expression
- Some information found in surrounding context
- BERT can distinguish incongruity caused by idiomatic usage from disruption caused by masking

## Future Work

- Investigate the presence of idiomatic key in topical content words in the context