Figurative text inference leveraging metaphor interpretation and summarization

Bhagirath Tallapragada
Department of Computer Science
Georgia State University
Atlanta, USA
btallapragada1@student.gsu.edu

Varchaleswari Ganugapati
Department of Computer Science
Georgia State University
Atlanta, USA
vganugapati1@student.gsu.edu

Nikhil Koditala
Department of Computer Science
Georgia State University
Atlanta, USA
nkoditala1@student.gsu.edu

***Abstract*—Information retrieval, though a classical research problem in NLP, continues to be highly relevant due to the variety of domain based challenges associated with it. One of the many forms of information retrieval from text includes summarization. Widely used in news applications such as inShorts to encapsulate news content. In spite of advances made in this field, challenges in key areas such as semantic informativeness, continuity of language remain. Recent advances in the field of figurative text, particularly in identification and interpretation of text involving metaphorical usage of language provides an interesting standpoint to approach analysis and inference from text. In this project we propose to study the interplay of figurative text in day to day language and develop a framework to gauge the unsupervised methods of metaphor identification and interpretation to improve the text summarization quality with focus on semantic informativeness. We focus on text formats involving small sized passage text like news articles, poems, excerpts from books etc and aim to extract intended meaning based on the text. We position the success of our model as a domain independent approach which can find utility in a variety of contexts such as opinion mining, threat identification in surveillance, a means to measure intensity of emotion, text translation etc.**

***Index Terms*—Metaphor identification, text summarization, information retrieval**

I. Introduction

Metaphorical language is prevalent in natural language and it is a significant challenge to infer meaning out of metaphorical text. In literature, researchers have identified metaphors to belong to three broad categories, type 1 metaphors: In this type of metaphors, nouns are identified via a form of verb 'to be'. For example: "She is a fox". Type 2 metaphors: In this form the verb itself is a focus of metaphoric use and acts upon a noun. For example: "The tragedy sucked the life out of him". Type 3 metaphors: In this kind of metaphors, adjective noun phrases convey the metaphoric context. For example: "... sweet child'. Even though these classes do not exhaustively encapsulate every possible metaphoric use case, they definitely identify the three most popular forms of metaphor found in day to day use.

Paraphrasing metaphorical text to literal domains plays a very important role in the model's ability to extract relevant information from target text. As mentioned by Mao et al in [1], one third of typical corpora consists of metaphors. Metaphor interpretation is a difficult task considering that human intuition does not require prior knowledge to generate or interpret a metaphor [8]. Consider an excerpt from a prominent poem by Robert Frost as an example: ***"Two roads diverged in a wood, and I— I took the one less traveled by, And that has made all the difference."*** The excerpt is from the poem ***Road Not taken*** and a human summarization will likely include the authors expression of making a choice between two paths in life. This is semantically far from the literal interpretation of text which conveys the same meaning yet in the literal context of walking. The complex usage of figurative speech is hard to distinguish from the literal meaning intended from the sentence, therefore a classical summarizer is bound to perform poorly in extracting inferred meaning. Similarly, in this case, a sentence level metaphor identification approach may also find itself wanting as very important information about the context is present in a disguised form through the course of the poem such as: ***"I shall be telling this with a sigh Somewhere ages and ages hence: Two roads diverged in a wood and I— I took the one less traveled by,"*** We hypothesize that the summarization processes and metaphor interpretation can be modelled as interdependent components in a framework to better extract the context out of figurative text and in turn be utilized to extract accurate inference. The benefits of such interpretation are relevant across domains. One such future application which we feel can be pipelined involves mining information from conversations in social media that hint towards radicalization. The use of figurative speech in religious texts can be gauged to better understand the picture painted in victim's mind to motivate them towards extremism. Another advanced application could be to identify coded conversations in chat servers that involve terrorism planning etc.

II. Literature Survey

Metaphor identification and interpretation became popular in recent times with growing methodologies using different semantic and statistical techniques. Neuman et al's [16] work is among the most comprehensive studies of metaphor identification in terms of scope of metaphorical phrases and annotated corpora size.

Paul et al [9] demonstrated the role of metaphor usage in human reasoning using a series of psychological experiments.

One of the early works on Metaphor identification is by Fass et al [10] where they have developed a system called met*. met* is a three stage process which uses selectional preference violation and knowledge base to distinguish between metonymy and metaphors. [11] proposed a novel semi-supervised approach where they used noun and verb clustering to learn metaphorical associations. These associations are further used to identify metaphors in text corpus. Shutova et al [12] proposed that metaphor identification not only depends upon the textual features, but also on visual features. Based on this assumption, they used visual embedding along with word-level embedding to build a classifier which outperformed its counterparts which used only word-level embedding.

Mao et al in [1] proposed an unsupervised architecture which identifies and deciphers metaphors at word-level. They have further conducted experiments to evaluate how metaphor processing could be used to support English-Chinese machine translation(MT) and improve the same. They saw a dramatic translation improvement for the metaphorical class for two MT systems used by them for evaluation, which was a 26% improvement for Google Translate and 24% for Bing Translate. [15] describes the challenges involved with automatic text summarization one such problem being dangling anaphors. Ekaterina Shutova reviews computational models of metaphor and describes its usage is ubiquitous in natural language text and it is a serious bottleneck in automatic text understanding [13]. Turney et al proposed a method to utilize abstractness and concreteness of a word context to determine its metaphoricity [17]. Our paper is inspired by and draws primarily from the methodologies presented by [1] and [17] and attempts to relate the ideas presented in both methodologies to leverage identify and interpret metaphors using not only sentence level context but also passage level context.

III. Dataset

The following are the datasets used for various tasks in this project:

- We used the Wikipedia dumps [3] for metaphor identification and interpretation task. We elected this dataset to train the word embeddings and defining the feature vector due to the nature of the data. Wikipedia data provides a collective set of information for a wide variety of concepts. The information is very encyclopedic in nature tends to have formal and polished English. This in a way implies very literal usage of English and is perfect to train the word embeddings as the vector distances between literal and non literal context can be distinguished well (when applied on the test data compiled from a variety of sources).
- To exclusively assess the metaphor identification pipeline, we used the dataset curated by Mohommad et al in [2]. This dataset is a compilation of a collection of sentences with each sentence labelled as metaphor or literal. This dataset further has the term that is being considered as the target term in the given sentence, sense of the term, class as in metaphorical or literal and confidence level. But we consider only the sentence column of the given dataset to evaluate our own model built to identify target words and the meaning. The metrics provided in this project for metaphor interpretation utilize a subset of this data for evaluation of models.
- For the third iteration in metaphor interpretation, we needed to label each word in our corpus (wiki) with a concreteness score. To do this, we relied on the MRC Psycholinguistic Database [19]. The MRC database consists of about 4200 words each labelled with a concreteness score between 100-700. where 700 is the maximum score of concreteness and 100 is the maximum score of abstractness of a term. As mentioned in the methodology section we use the assumption that synonyms of an abstract term are also abstract and vice versa, thereby using the cosine similarities between each word in the corpus and the MRC database words to assign the concreteness score of the word in MRC data that is most similar to the word in our corpus.
- For Text Summarization we use human judgement to evaluate the model. To do this we created a test data comprising of 25 articles/ passages curated based on high occurrence of metaphor usage to evaluate the effect of metaphor identification on text summarization task. We gathered the articles from varied sources across the internet such as the Gutenberg corpus, presidential speeches etc.

IV. Methodology

This section describes the methodologies applied to infer metaphorical context and subsequently assess the impact of inferred meanings in the summarization task. Our implementation contains two pipelines, each explained below for this purpose. We use an unsupervised learning approach for metaphor identification as the unsupervised models are less constrained by localized data and can prevent the need for annotated training data to develop the model. It is also important to note that the innate nature and usage of metaphoric language in english is so diverse that unsupervised methods leveraging the semantic structure of the input text are much more suitable for this purpose. The gains comfortably outweigh the flaws unsupervised models.

For the purpose of summarization, however we use two different techniques that employ both supervised and unsupervised methods. The details are explained in part B.

A. *Pipeline 1: Metaphor Interpretation*

The metaphor interpretation task involves two steps: 1) Metaphor Identification 2) Paraphrasing or replacing the non-literal form of word with its original meaning. In this task we used unsupervised approach to identify and interpret linguistic metaphors at the word-level [1]. We followed an incremental approach while developing the model. In the first iteration, we developed a sentence level metaphor identification model that leverages context-target relationship in a sentence. In the second iteration, we refined this model by enabling a

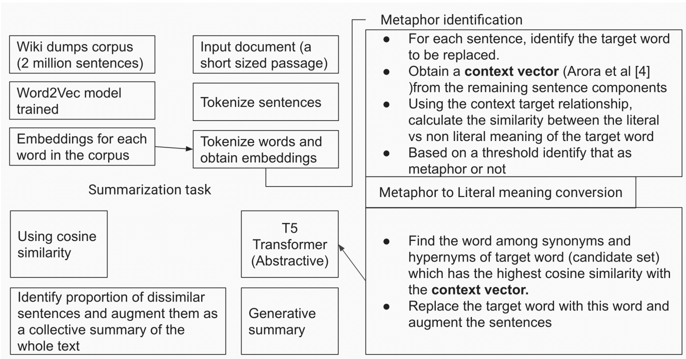

Fig. 1. Sentence level context model flow

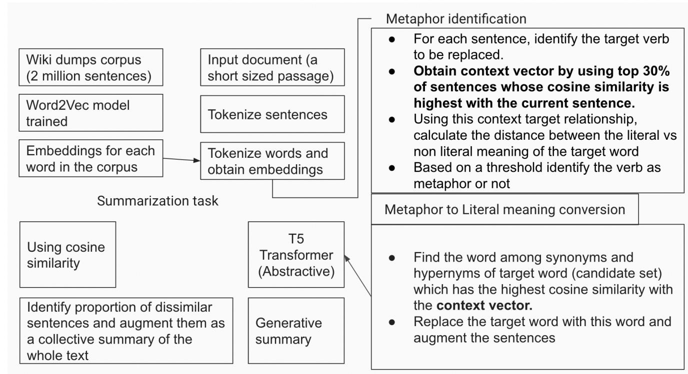

Fig. 2. Passage level context model flow

mechanism to leverage context from relevant parts of the passage instead of focusing solely on the sentence. In the last iteration we attempted to prioritize selection of target words by identifying metaphoric phrases using the abstract and concreteness ratings of the constituent words and determining whether there is a transition between abstract and concrete domains or vice versa [17]. We look at each of these iterations in detail in the following sub sections.

1) Iteration 1: Metaphor Interpretation through sentence level context: In this iteration we developed a model that considers the context-target relationship in a sentence rather than just the metaphorical word. Another way to put it is that, every word in a sentence has some relationship to its context. This Context plays an important role in identifying metaphors in a sentence, for example, in the sentence *"He is a bright star"*, 'bright star' in literal usage refers to a celestial object and the light it emits, however, in referring to a human as a bright star the meaning implied is different. The sense obtained in such a usage is that the referred person has high merit in a specific action or that the person has intelligent attributes that make him stand out. The underlying logic is to use this difference in domains to establish whether a verb is being used metaphorically or not. We calculated the distance between literal and non literal meaning of the target word in the given feature space and then classified it as a metaphor based on a pre-determined threshold theta as mentioned in [1]. The choice of theta in this case is purely experimental and we tried out multiple thresholds to assess our models performance. Fig 1 shows a high level overview of how the model functionality flows and eventually merges into the summarization pipeline explained in part B.

2) Iteration 2: Metaphor Interpretation through passage level context: The first iteration was successful in identifying verbs using sentence level context, however the quality remains susceptible to the choice of threshold theta. In addition, even though we saw encouraging signs in the model's performance on the dataset provided by Mohd et al [2]. it is not entirely equipped to deal with extended metaphors. Extended metaphors are those metaphors whose usage exceeds a single sentence, they can span across a passage in terms of context or can occur at specific points in a passage only to be utilized elsewhere. Extended metaphors are frequent in passage level text and are thereby important for us to decode to test our hypothesis. In order to do this we must take into account significant words occurring across the passage rather than just the sentence that contains the target word. We initially began by looking at all the words of the input passage to construct the context vector. however we realized that this approach is highly susceptible to the noise caused by words that have little or no impact on the context. This resulted in the replaced words being haphazard and unrelatable in many cases. To rectify this we chose to selectively include sentences from the input passage in order to construct the context vector. To do this, we first compute the context vectors of all the sentences in the paragraph based on the following equation proposed by [18].

$$\tilde{c}_s = \sum_{w \in s} \frac{a}{p(w) + a} v_w$$

Once we have list of context vectors corresponding to all the sentences in the paragraph, we calculate the cosine similarity between the target sentence with rest of the sentences present in the paragraph. We pick sentences which have high cosine similarity with the target sentence and build a final context vector from it. The intended assumption being that similar sentences will contain words which are closer to each other in the feature space therefore highly likely to be related to each other. This final context vector, is a weighted average of all selected sentences. This addition now enabled our model to leverage context from words beyond a single sentence in identifying the target word. To validate this model we used a manually curated dataset of 25 sample passage texts from various domains such as news articles, stories, poetry and pop culture. These texts were specifically picked based on high likelihood of extended metaphor usage. the results of the performance are collated in a spreadsheet that is a part of submission. This approach saw a significant change in the quality of the words replaced by the model as the example in fig 4 indicates. The high level overview of models functionality is shown in fig 2 after the passage level context vector creation mechanism is added to iteration 1.

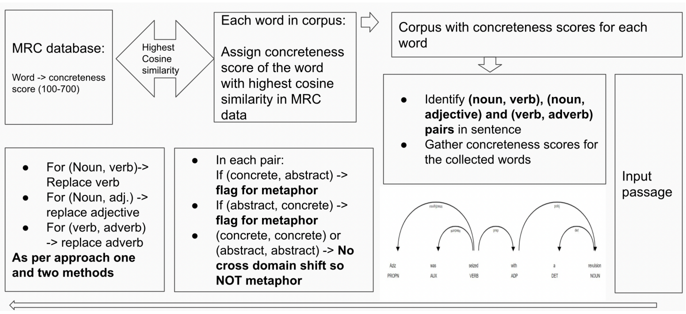

Fig. 3. Flow to utilize concreteness scores for each word in corpus

3) Iteration 3: Utilizing Concreteness/ Abstractness of word pairs in a sentence to identify verbs for replacement: We saw notable improvement in metaphor identification and interpretation using the methods in iteration 2. However this approach too was not a hundred percent noise proof. Particularly in cases where there are only minor dependencies between the target sentence and overall passage context, the model fails to prevent noise from seeping in. Further, we were now handed another dilemma in terms of prioritizing which words should be considered as target words in case there are multiple candidate words within a single sentence. The previous two iterations suffered from this limitation in the sense they are equipped to pick the first word eligible to be considered as a target word.

To solution to this problem is inspired by a methodology provided by Turney et al in [17]. Turney et al argue that metaphors transfer associations from a source domain to a target domain. In specific their work models an algorithm that identifies transfer from a concrete domain to and abstract domain. In English literature, abstract words are described as those words which describe intangible entities such as *"truth, debate, honor etc"*. Notably, these are such words which are existent only in human intellect and do not have a physically identifiable form. Concrete words on the other hand are such words whose existence can be physically perceived as tangible entities such as *"war, book, table etc."*. To explain this, consider the following two sentences: *"He shot down my argument."* vs *"He refuted my argument."*: In both cases, we can infer similar meanings, however in the former, the phrase *"shot down"* transfers knowledge from the concrete domain of war to the abstract domain of a debate. We use this transfer of knowledge from one domain to another to identify and prioritize which words in a sentence should be considered as target words.

The algorithm for this iteration involves the following steps: 1. for each word in the text corpus, identify the word in the MRC psycholinguistic database [18] that has highest cosine similarity. 2. Assign the concreteness score of this word in the MRC database to the word in corpus. 3. Once all the words in our corpus are successfully assigned concreteness scores, store the collected data as E. Now to utilize this data we use the following algorithm: 1. Input the article/ passage to the algorithm. 2. Using Spacy's dependency parsing, identify the (head noun-adjective), (noun- head verb) and (head verb-adverb) pairs for each sentence. 3. For each of the pairs identified in a sentence, fetch the concreteness scores in E via a lookup. 4. once all the pairs have been assigned scores, identify the pairs which convey a transition from one domain to another (abstract to concrete or concrete to abstract). 5. If there are multiple pairs that display such a shift, identify the pairs which are most widely spaced apart in their scores by calculating the difference between their scores. For example in *"But when the time came, Aziz was seized with a revulsion"* the pair (Aziz, seized) has a score difference of 261 where as (time, came) and (came, when) have score difference of 78 and 46. Therefore the phrase selected for replacement is (Aziz, seized). 6. Once the pair is identified for replacement, if the pair is a (noun, adjective) pair then the adjective is replaced, if (noun, verb) pair, then the verb is replaced and lastly if it is a (verb, adverb) pair, then the adverb is replaced. This replacement is done using the context vector computed in iteration 2. The high level overview of this functionality is encapsulated in fig 3.

Even though we recognize that these pairs may not be the exhaustive list pairings in which metaphoric context is implied, based on observation and the referred scientific literature, we infer that these three pairings capture a majority chunk of metaphoric usage in text. In terms of performance, we see a notable change with respect to the previous two iterations. To illustrate the improvement, consider the following sentence from a passage whose context is that of a person mourning the death of his wife: *"And he meditated suicide"*. the identified word to replace is *meditated*. The resulting sentence is *"And he contemplate suicide."*. Clearly the word *contemplated* captures the context properly. Such improvement is felt more consistently across the 25 curated articles when compared with previous iterations. The collective results for all the passages in the dataset are collated in a spreadsheet and color coded (green for reasonable metaphor identification and red where the model fails). A brief summary of the same is presented in table 1 below for reference. Please note that higher priority was assigned to metaphor identification in comparison to the quality of words replaced while assessing. This is due to the fact that metaphor identification is vital for future development where as quality of the words replaced can be refined incrementally as a future work. Also, the replacement of words is treated as an independent pipeline therefore all the words considered for replacement may not feature in the paraphrased text which is fed to summarizer. Even with the notable change, the model is not equipped to deal with some edge cases that include 1. Dead metaphors (metaphors that qualify the theoretic definition of metaphor but are so well trodden and used over ages that they blend into regular use). For example, the models in iteration 2 and iteration 3 still replace the phrase *"The change began after its birth."* to *"The change start after its birth."*. In this case we can see that begin and start are frequently used interchangeably and their difference does not hold huge significance in day to day use.

In other edge cases, the model still suffers with the problem of over classification of metaphors that inevitably results in some wrong words being identified as metaphors that do not warrant a replacement, example: *"Gradually he lost the feeling that his relatives had chosen wrongly for him."* is changed to *"Gradually he lost the feeling that his relatives had vote wrongly for him."*. Keeping aside the quality of word replaced, we can note that the word *chosen* was accurate enough for the given context. Barring such errors the model in third iteration performs reasonably to identify the metaphors across the test data.

B. Pipeline 2: Text Summarization

Text summarization involves summarizing a document in fewer sentences without losing significant details. Text summarization can be further divided into extraction-based summarization and abstraction-based summarization. In extraction-based summarization we extract relevant keywords from a given document and combine them to formulate the document summary. In abstraction-based summarization deep learning models are utilized to paraphrase the sentences and generate a summary. Currently there are many state of the art text summarization methodologies presented in [4], [5] and [6]. Most of the text summarization methods use the semantic meaning of words to generate summary of the document. It is important to stress here that our project is not an attempt to improve the summarization mechanism itself, instead we only assess the impact of metaphor interpretation on various vanilla summarizers both extractive and abstractive. To assess the impact on extractive summarization, we created a cosine similarity based summarizer that takes in an input passage, measures the cosine similarity of each sentence with the rest of the sentences and based on this identifies the most unique sentences of the passage. We consider 30% of the input passage as the size for the generated summary. For the purpose of abstractive summarization we used a pre-trained T5 transformer from Huggingface to generate the summaries. The transformer is also set to consider 30% for the size of the generated summary.

Passage number (from the 25 curated passages mentioned in the proposal)	Total sentences	Expected metaphors, based on human judgement	**Total considered words to replace by model**(excluding words not existing in corpus)	**Accurate metaphors identified and replaced**	**Wrong replacement (mistaken for metaphors)**	Replaced with Same words	Noise words replaced or Blanks (words not existing in corpus or unintended words)
1. (Aziz ..)	28	6	21	3	5	8	5
2. (statutes…)	7	3	7	1	6	1	0
3. (Nietzsche..)	14	3	14	1	5	4	3

Table 1: manual evaluation of metaphor identification post iter-3

The table shows the quality of summary with and without replacement of metaphoric words. The model identified the word *write* as the target. Both the summarizers feature this change and the replaced word is *profile*. verifying the dictionary meanings of the two words it can be observed that *profile* is a word used when writing an article about a person describing a person. The important aspect here is that the model inferred this meaning using the passage level context described in iteration 2 of methodology section. Our submission report comprises of a collated spread sheet that includes all details about the quality and quantity of replaced metaphors replaced in each of the 25 passage texts chosen for assessment. It also features the summary generated by the two models for each text.

V. Experiments and Results

The results of various experiments conducted are shown in the below subsections. The subset that was chosen from Mohommad et al's [2] dataset had a class imbalance with more number of literal sentences compared to the metaphorical ones. However, as our model was unsupervised it does not really effect the training process. During evaluation, we made sure to consider measures like **specificity** and **sensitivity** where sensitivity refers to the true positive rate that summarizes how well the positive class is predicted whereas specificity is the true negative rate that summarises how well the negative class is predicted.

Sensitivity is given by:
TruePositive / (TruePositive + FalseNegative)
Specificity is given by:
TrueNegative / (FalsePositive + TrueNegative)

A. Metaphor Identification without Concreteness Scores

1) The model was evaluated with a **theta** value of 0.65, 0.7, 0.75 and 0.8 and the results in fig 5 show the performance of the model with the changing of theta value. Based on sensitivity and specificity scores, we found theta value with a 0.7 gave a good balance of prediction scores relatively as our prime focus was to label the positive classes quite well. Also, positive class (in our case a metaphor) holds importance as our purpose is to not miss metaphors while parsing the passage text rich in metaphorical content.

B. Metaphor Identification with Concreteness scores assigned

1) This model was evaluated again with the same theta values that were chosen for part A in this section.The results in fig 6 show the performance of the model with the changing of theta value. Here what we noticed is there is a definite change in the classification of text as metaphor or literal. Based on sensitivity and specificity scores, same as in part A of this section, we found theta value with a 0.7 still accounts for a good balance of prediction scores. Also, if we compare the performance of this combination with previous parts A, we can see an improvement in distinguishing between metaphorical and literal texts. Considering an overall view of the metaphor identification, we see that our model grapples with over classification of non metaphorical texts as metaphors i.e. our model performs poorly to identify non metaphoric texts as literal, well.

One of the most profitable uses of history is the narrative of such lives. Having this in mind, it is safe to say that there is no species of writing of more value than biography. It inspires the young to nobler purposes. It develops higher resolves. And it proves an incentive to the laudable imitation of men. Men, who in prominent positions have proved true to principle and duty. It is in this spirit and with this thought in mind, that I undertake to write the story of the life of Dr. Walker.

	On Raw Text	On Metaphor Replaced Text
Cosine Similarity	It is in this spirit and with this thought in mind, that I undertake to **write** the story of the life of Dr. Walker. And it proves an incentive to the laudable imitation of men.	It is in this spirit and with this thought in mind, that I undertake to **profile** the story of the life of Dr. Walker. It inspires the young to noble purposes. And it proves an incentive to the laudable imitation of men.
T5 Abstractive	it is safe to say that there is no species of writing of more value than biography. it inspires the young to nobler purposes, it develops higher resolves, and it proves an incentive to the laudable imitation of men. men, who in prominent positions have proved true to principle and duty, he writes the story of the life of Dr. Walker. i undertake to **write** a biography of dr.walker, the author of his life	the story of the life of Dr. Walker is profiled in this spirit and with this thought in mind, I undertake to **profile** a biography of dr. walker. the biography is the most profitable uses of history. it is fascinating to look at the history of such lives. this is not the same as the edd s..

Fig. 4. Summary Results on Raw text vs Metaphor Replaced Text

C. Summarization results

The impact of metaphor interpreted text showed a mixed set of results. Among the positives, we noted that where the replaced words were accurate enough and featured in the summary of the passage, the summarizers had a reasonable improvement in quality in terms of human inference and readability. This was much more pronounced in T5 abstractive summarizer. The drawback however, is the fact that despite this notable improvement, the limitation remains that there is no guarantee that vanilla summarizers will consider the sentence housing the replaced word for summarization. In other words the change in the vectors comprising the sentence still offered no assurance that the sentence features in the summary.

An example of positive results is mentioned in the below figure 4.

1) Does metaphor interpretation definitively improve summarization?: Based on the results we argue that in most cases the impact is more profound in abstractive summarizer than the extractive summarizer. This can be attributed to the fact that vanilla extractive summarizers are relatively blind to contextual changes (especially unsupervised models like the cosine similarity summarizer). However, we can not definitively claim that the metaphor identification and replacement always leads to improved results in summarization. This can be attributed to multiple factors such as: 1. The quality of words replaced by the metaphor pipeline is crucial to protecting the context of the passage. 2. Limitations in the size of corpus and the dimensionality of vectors created cause deviations from expected behavior in the metaphor replacement task therefore resulting in problems like over classification, non existent words in corpus and replacement with various forms of the same word for example, *mourned* to *mourn*. 3. The limitations mentioned in point 2 affect the summarization pipeline as is it is dependent on the replaced words. 4. Even with success in points 1 to 3, the sentences with replaced text may not feature in the summary for vanilla summarization techniques (not customized). Due to this, we consider our hypothesis regarding summarization does not hold true exhaustively. The encouraging sign however is that the impact of metaphoric interpretation does have significant impact and offers scope for future research to improve upon the limitations.

	theta=0.65	theta=0.7	theta=0.75	theta=0.8
Senstivity	0.5546	0.6806	0.7815	0.8907
Specificity	0.4314	0.3091	0.1962	0.1209

Fig. 5. Metric scores for metaphor identification without concreteness scores

	theta=0.65	theta=0.7	theta=0.75	theta=0.8
Senstivity	0.4759	0.6009	0.7259	0.8413
Specificity	0.5397	0.4227	0.2995	0.2121

Fig. 6. Metric scores for metaphor identification without concreteness scores

VI. Conclusion

Through this paper we explored three different strategies to identify and interpret metaphors, with each strategy improving on the previous one. In addition, once the metaphors were detected and translated we tested our hypothesis of whether interpreting metaphors to their original sense lead to improvement in text summarization in terms of human inference and readability. We concluded that even though the translation impacts and slightly improves the summarization task, particularly when abstractive techniques are used, vanilla extractive summarizers do not necessarily gain from it. This is because they focus exclusively on unique elements of the text rather than leveraging the translated literal meanings of the sentences. Lastly we provided metrics to analyse the performance of our unsupervised model for various theta values to find the sweet spot to accurately identify maximum metaphors in a text.

VII. Future Scope

The initial foray into metaphor interpretation yielded encouraging results and warrants for further study to refine the existing models, our model does not deliver ideal results for many use cases and the given time span could only allow for so much experimentation. As part of future plans, we intend to further experiment and improvise the word pairing pipeline to identify potential metaphors. Our model, indeed, did not fare well with identifying literal classes in text. This is the most important drawback that we want to focus, understand and improvise. Further more special focus needs to be given for the construction of word vectors for this kind of tasks. Our current model uses word2vec to train on wiki dumps for obtaining the word embedding. A custom word vector creation paradigm, something like skip-grams or CBOW, can be introduced to better map the context and usage of a word in a given corpus. This will help us better identify the regular patterns in which the word appears thereby allowing for a much pronounced differentiation between literal and non literal usage of words. Lastly the next step will also involve introducing capabilities that can identify and convert multiple metaphoric uses in a single text and preprocessing abilities to cater to various domains of data.

REFERENCES

[1] Mao, R., Lin, C., amp; Guerin, F. Word embedding and wordnet based Metaphor identification and Interpretation. https://www.aclweb.org/anthology/P18-1113/.

[2] Mohammad, S., Shutova, E., amp; Turney, P. Metaphor as a medium for emotion: An empirical study. https://www.aclweb.org/anthology/S16-2003/.

[3] https://dumps.wikimedia.org/enwiki/

[4] Armen A., Akshat S., Anchit, Naman G., Luke, Sonal G. Better Fine-Tuning by Reducing Representational Collapse. https://arxiv.org/pdf/2008.03156v1.pdf

[5] Weizhen Q., Yu Y., Yeyun, D., Nan D., Jiusheng C., Ruofei Z., Ming Z. ProphetNet: Predicting Future N-gram for Sequence-to-Sequence Pre-training. https://arxiv.org/pdf/2001.04063v3.pdf

[6] Dongling X., Han Z., Yukun, Yu S., Hao T.,Hua W. and Haifeng W. ERNIE-GEN: An Enhanced Multi-Flow Pre-traning and Fine-tuning Framework for Natural Language Generation. https://arxiv.org/pdf/2001.11314v3.pdf

[7] Graff, David and Kong, Junbo and Chen, Ke and Maeda, Kazuaki. English Gigaword. Linguistic Data Consortium, Philadelphia, 2003.

[8] Kesarwani, Vaibhav, Diana Inkpen, Stan Szpakowicz, and Chris Tanasescu. ”Metaphor Detection in a Poetry Corpus.” ACL Anthology. Web.

[9] Paul H. Thibodeau and Lera Boroditsky. 2011. Metaphors we think with: The role of metaphor in reasoning. PLoS ONE, 6(2):e16782, 02.

[10] D. Fass. 1991. met*: A method for discriminating metony my and metaphor by computer. Computational Linguistics, 17(1):49–90.

[11] Ekaterina S., Lin S. and Anna K. Identification Using Verb and Noun Clustering. 23rd International Conference on Computational Linguistics:1002–1010

[12] Ekaterina S.,Douwe K., Jean M. Black Holes and White Rabbits: Metaphor Identification with Visual Features. Proceedings of NAACL-HLT 2016:160–170

[13] Shutova E Models of Metaphor in NLP. In: ACL Anthology.

[14] Shutova and S. Teufel. 2010. Metaphor corpus annotated for source - target domain mappings. In Proceedings of LREC 2010, Malta.

[15] U. Hahn and I. Mani, ”The challenges of automatic summarization,” in Computer, vol. 33, no. 11, pp. 29-36, Nov. 2000, doi: 10.1109/2.881692.

[16] Yair Neuman, Dan Assaf, Yohai Cohen, Mark Last, Shlomo Argamon, Newton Howard, and Ophir Frieder , ”Metaphor Identification in Large Texts Corpora” in PLoS ONE (2013).

[17] Peter Turney, Yair Neuman, Dan Assaf, Yohai Cohen, ”Literal and Metaphorical Sense Identification through Concrete and Abstract Context” -Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, (2011).

[18] Arora, S., Liang, Y., & Ma, T. (2019). A simple but tough-to-beat baseline for sentence embeddings. Paper presented at 5th International Conference on Learning Representations, ICLR 2017, Toulon, France.

[19] https://websites.psychology.uwa.edu.au/school/MRCDatabase/uwa_mrc.htm