Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 12, Issue, 03, pp.10480-10483, March, 2020

DOI: https://doi.org/10.24941/ijcr.38217.03.2020

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

RESEARCH ARTICLE

COMPARITIVE STUDY OF S POS TAGGING FOR INTERIOR DESIGNING

*Yashaswini, S. and Shylaja, S.S.

Department of CSE CITECH, Bangalore PES University, Bangalore

ARTICLE INFO

Article History:
Received 24th December, 2019
Received in revised form
20th January, 2020
Accepted 28th February, 2020
Published online 28th March, 2020

Key Words:
Part of Speech Tagging, Scene Generation, POS Accuracy, Interior Designing.

ABSTRACT

Part of speech (POS) tagger assigns a lexical tag to each word in given sentence. Part of Speech Tagging is a subtask Natural Language Processing Systems, identifies the correct tag according to syntax for each word in corpus (6). Part of speech can be used in many scenarios like information extraction, parsing, word semantic disambiguation, question answering, named entities recognition, virtual chats, scene generation etc. While POS tagging works great for commonly used simple English sentences, but it has not yet achieved good accuracy in the field of interior designing due to its morphological characteristics. This paper compares various POS techniques.

Copyright © 2020, Yashaswini and Shylaja. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Citation: ***Yashaswini, S. and Shylaja, S.S.* 2020.** "Comparitive study of s pos tagging for interior designing", *International Journal of Current Research*, 12, (03), 10480-10483.

INTRODUCTION

The input text is a well-trained corpus prepared by extracting interior design arrangement sentences by web scrapping using beautiful soup bs4,these sentences are pre-processed, tokenized and tagged using Parts Of Speech(POS). The process of dividing input sentences into words is tokenizing and assigning each word with the descriptor or tag is POS Tagging (Jurafsky, 2002). The tag may indicate one of the parts-of-speech like a noun, pronoun, verb, adjective, adverb, preposition, conjunction, and interjection. Applications of POS tagging include Machine translation, Natural language text processing, summarization, Multilingual and cross-language information retrieval, Speech recognition, Artificial intelligence and so on (http://www.scs.leeds.ac.uk/ amalgam/ tagsets/brown.html; Mitchell, 1994). Previously the tagging of unannotated text was done by hand which was a tedious, time consuming and error-prone process hence there were efforts to atomize the process of POS Tagging. Efficient tagging before syntactic parsing can reduce ambiguity and complexity of the computational problem. POS tagging extracts useful information about words and their neighbours to derive semantic relationships useful in scene generation and information extraction tasks (Cutting, Doug, 1992). POS tagging has been carried out extensively in English and current POS tagging performance for Brown corpus is at 97% (http://www.scs.leeds.ac.uk/amalgam/tagsets/brown.html).

***Corresponding author:* Yashaswini, S.,**
Department of CSE CITECH, Bangalore PES University, Bangalore.

Challenges of interior designing sentences: Text to 3D scene systems faces several technical challenges. Sentences seem to be incomplete due to the assumption of lots of facts. Eg: There are two desks. Most desks are upright and are placed on the floor but few people would mention this explicitly. This implicit spatial knowledge is critical for scene generation but hard to extract. Emphasis is always on the semantics of objects and their approximate arrangements rather than its geometry. Inferring implicit constraints, spatial relations are the hardest task hence the interface has to leverage spatial knowledge priors learned from an existing 3D scene for instance, given the sentence "there is a table with a cake", should be inferred as a cake is on plate and plate is on table. High-level words such as "in front", "to the left" and "on top of" makes the system complex to understand semantics. 3D scene layout can be automatically generated by applying machine learning techniques to training data. The Scene layout has to follow some interior design guidelines learned from input scene data (Ratnaparkhi, 1996). Object positions, supporting contact points and orientations can be easily modelled by the prior spatial and object co-occurrence relation.

Availability of Corpora: Multiple tag sets have been used for POS tagging of English Corpora. The Penn Treebank uses 45 tags while the Brown Corpus (Mitchell, 1994) made use of 87 tags. Other common tag sets include British National Corpus basic tag set C5 with 61 tags, the enriched C6 tag set with 160 tags and the CLAWS tag set (Petrov *et al*., 2011). Text8 corpus has 100,000,000 characters long and contains 17,005,207 words including 253,854 unique words and 71,290 unique frequent words. For comparison we tend to use brown corpus with our dataset consisting of interior designing sentences.

Pos tagging techniques: POS tagging can be classified into supervised and unsupervised POS tagging model. The supervised POS tagging models requires a tagged corpora which is used for training to extract information about the tags (10),whereas Unsupervised POS Tagging models do not require a pre-tagged corpus, insteadit uses advanced computational methods like the Baum-Welch (David Elworthy, 1994) algorithmto automatically induce tag sets, transformation rules etc. Based on the information collected, probabilistic information needed by the stochastic taggers or contextual rules needed by rule-based systems or transformation based systems (Juan Antonio P´erez-Ortiz,?; Himanshu Agrawal, 2006) are calculated. The rule-based POS tagger uses rules, contextual information and morphological information (each word are built from smallest unit morpheme (http://www.scs.leeds.ac.uk/amalgam/tagsets/brown.html)) to assign POS tags to words. On the other hand, the transformation based tagger uses a pre-defined set of rules, words with high frequencies to generate final POS tags.

A stochastic approach uses frequency, probability or statistics, the simplest stochastic model used most frequently used tag for a specific word in the dataset. In word frequency/n-gram the best tag is calculated based on probability of word from previous tags. The n-gram approach uses a Viterbi Algorithm (Peilu Wang, 2015), which is based on Maximum Likelihood Estimates (MLE). A Conditional Random Field (CRF) framework uses the probabilistic model to segment and labels a sequence of data. The transition between labels can be observed by not only the current but also past and future states (Himanshu Agrawal and Anirudh Mani, 2006). The CRF model can calculate the probability based on the suffix, of the current word, the tags of previous and next words (Pranjal Awasthi, 2006).

Hidden Markov Model is used to build a model based on the joint probability distribution P (word, tag) using probabilistic Finite State Machine (FSM). The tagsare represented as states and the transition between states can be either transition probabilities P(Tagi|Tagi-1) or emission probabilities P(Wordi|Tagi). The states in HMM are hidden due to undetermined nature of tag sequence. In HMM the current tag always depends on the previous n tags (Pranjal Awasthi, 2006). The Maximum Entropy Model (MEM) is defined over (H, X, T), where H is the set of possible word and tag contexts or "histories", and T is the set of allowable tags (http://www.scs.leeds.ac.uk/amalgam/tagsets/brown.html). A Maximum Entropy Part-of-Speech Tagger can run either a Maximum Entropy Markov Model (MEMM) tagger or a cyclic dependency network tagger. This model is convenient for NLP since it combines contextual features and allows experimenters to reuse it by eliminating the need to develop highly customized problem-specific estimation methods. The Memory Based Learning (MBL) Model takes tagged data as input and produces a lexicon and memory-based POS tags as output. MBL consists of two components, memory-based learning component to memorize examples while training and similarity-based performance component to identify similarities between words using any distance metrics (Himanshu Agrawal and Anirudh Mani, 2006). The hybrid approach combines the advantages of the rule-based and stochastic approach.

Words are first tagged probabilistically and then linguistic rules are applied to tagged tokens. There is a significant increase in the accuracy of taggers compared to other taggers. Neural taggers are neural networks that learn the parameters of POS tagger from the training data set using error back-propagation learning algorithm. It outperforms compared to a stochastic method (Raju, S. B., 2002) and other tagging techniques like the Maximum Entropy model (Ratnaparkhi, 1996), Support vector machines, Decision trees, Conditional Random Fields. However, many of these techniques aren't available for tagging of interior design related sentences. Existing literature points to the availability of rule-based tagger, Brill's tagger (Brill, 1992), Stanford's POS Tagger, maximum entropy model and a hybrid model for POS tagging. For interior designing sentences, the best approach is to uses machine learning techniques. A recurrent neural network (RNN) is an artificial neural network where nodes are connected to form a directed graph along a temporal sequence. RNNs can use their internal state (memory) to process sequences of inputs (Juan Antonio P´erez-Ortiz). RNN are capable of holding previous and next state information which helps to preserve the contextual information.

Bidirectional Long Short-Term Memory Recurrent Neural Network (BLSTMRNN) works well for sequential data like speech utterances, text generation or handwritten documents. In this study, we propose to use BLSTM-RNN with word embedding for part-of-speech (POS) tagging task. When tested on the Penn Treebank WSJ test set, a state-of-the-art performance of 97.40 tagging accuracy is achieved. Thus this approach can also achieve a good performance comparable with the Stanford POS tagger for Interior designing application (11).BLSTM RNN mainly usesincorrect/correct tags to make prediction by choosing random words from sentences, the replaced words, tags are predicted as 0 (incorrect) and 1 (correct) for words which are retained without modification. SpaCy's models are the statistical model which predicts part-of-speech tags for words. This prediction is based on the learning happened during training. The text, label, gradient loss and NER are crucial for tagging. The model uses unlabelled text to make a prediction. The feedback on prediction can be provided in the form of an error gradient of the loss function as we know the correct answer. The difference between the training example and the expected output gives a gradient loss (https://nlpforhackers.io/complete-guide-to-spacy/). TextBlob is a Python library used for processing textual data. Its most commonly used API for natural language processing (NLP) tasks such as part-of-speech tagging, noun phrase extraction, language translation, sentiment analysis (https://textblob.readthedocs.io/en/dev/). TextBlob is built on NLTK and Pattern, therefore making it a simple interface for beginners. TextBlob is slower than Spacy but faster than NLTK, but it cannot provide features like dependency parsing and word vectors.

Embedding's from Language Models(ELMo) is a new approach where word vectors are computed on top of a two-layer bidirectional language model (biLM) stacked together. Each layer has forward pass and backward pass. The (biLM) architecture uses a character-level convolutional neural network (CNN) to convert the words into vectors. These word vectors are passed as an input, forward pass to check the context of the given word with a word before it and in backward pass the context is checked with the word after current word.

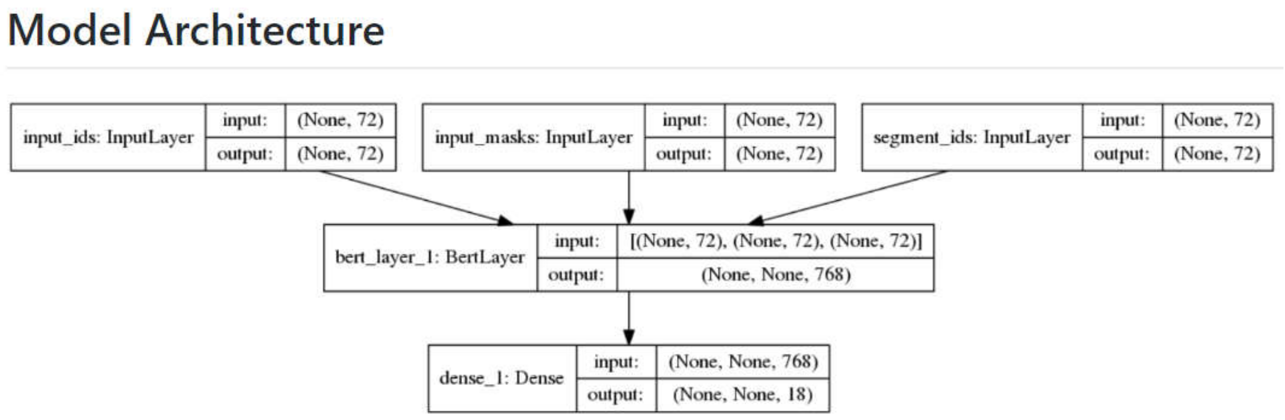

Fig 1.1 architecture of BERT Model

```
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data]   Package punkt is already up-to-date!
Original tokens: ['there', 'is', 'a', 'room', 'with', 'chair', 'and', 'the', 'table', '.']
BERT tokens: ['[CLS]', 'there', 'is', 'a', 'room', 'with', 'chair', 'and', 'the', 'table', '.', '[SEP]']
Converting examples to features  100% 1/1 [00:00<00:00, 53.95it/s]


Word in BERT layer  | Initial word   : Predicted POS-tag
-------------------------------------------------------------
there               | there          : ADV
is                  | is             : VERB
a                   | a              : DET
room                | room           : NOUN
with                | with           : ADP
chair               | chair          : NOUN
and                 | and            : CCONJ
the                 | the            : DET
table               | table          : NOUN
.                   | .              : PUNCT
```

Fig 1.2. POS Tagging Using BERT Pre-trained model

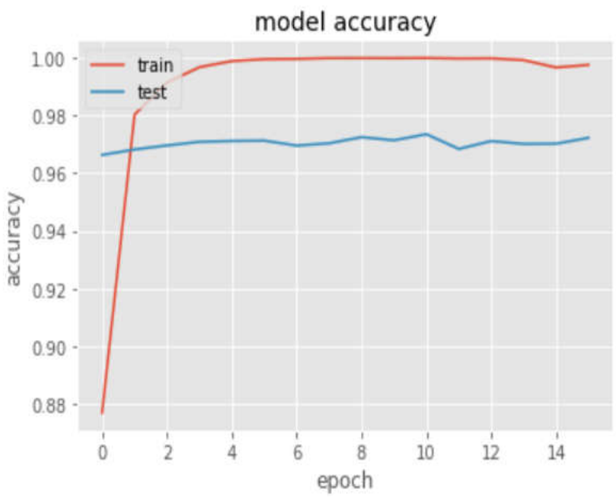

Fig. 1.3(a). model accuracy

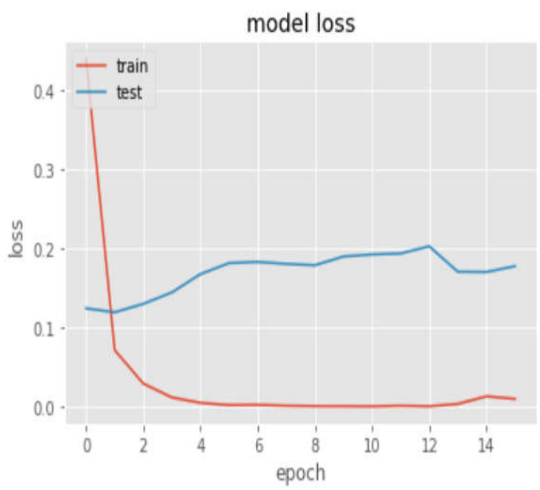

Fig. 1.3(b) model loss

The intermediate word vector is formed has an outcome of a forward and backward pass. The intermediate word vector is passed to the next layer where the weighted sum of the raw word vectors and the 2 intermediate word vectors is calculated considering inner structure of words. For example, the biLM will be able to figure out that terms like beauty and beautiful are related. In traditional word embeddings such as word2vec and GLoVe, word vector was assigned to a particular word itself whereas, in ELMo vector assigned to a token or word is actually a function of the entire sentence containing that word. Thus, the same word can have different word vectors in different contexts (Matthew, 2018). ELMo word vectors also address the issue of Polysemy wherein a word could have multiple meanings or senses. Traditional word embeddings come up with the same vector for the word “read” for both present and future tense. Hence, the system would fail to distinguish between the polysemous words. ELMo word representations take the entire input sentence into the equation for calculating the word embeddings. Hence, the term “read” would have different ELMo vectors under a different context.

BERT stands for Bidirectional Encoder Representations for Transformers. BERT is designed by group at Google AI Language to pre train deep bidirectional representations by jointly conditioning on both left and right context in all layers as shown in the fig 1.1, which makes it suitable for NLP tasks. BERT Transformer uses bidirectional self-attention. BERT's input representation is constructed by summing the corresponding token, segment and position embedding's. Understanding and choosing correct hyper parameters is a challenging task which can either make or break BERT suitable for interior designing sentences. Accuracy of 97.5 is obtained. BERT consists of neural sequence transduction models having an encoder, which maps an input sequence of symbol to a sequence of continuous representations and a decoder structure which generates an output sequence of symbols one element at a time regressively by consuming the previously generated symbols as additional input when generating the next symbol at each step (AshishVaswani). BERT Model works efficiently for POS Tagging Task for interior designing sentences by achieving accuracy of 97.10 % for the sentence "There is a room with chair and the table" as shown in the fig 1.2 The POS Tag are identified. The misclassification rate is around 0.0290 and the metric used for evaluation is f1 score with the value 0.9617. The BERT model is trained on Treebank Dataset, the new sentence is taken from user to determine its POS Tags. The BERT model is tested for 30 epochs consisting of 1950 nodes in the hidden layer for the validation sentence the model outperforms with the accuracy of 97% as shown in fig 1.3(a).The overall loss of the model is around 0.025 by using f1 score as the evaluation metric as shown in fig 1.3(b)

Conclusion

Part of Speech Tagging in interior designing is still at a nascent level due to its complexity and unavailability of resources in terms of dataset. Ample amount of scope exists for creation of corpusand tag sets that capture the challenges involved in scene generation. Existing tagging techniques can be enhanced and hybrid models can be generated to improve the accuracy of the tagging task by using a neural model like BLSTM RNN for word embedding and POS tagging. BLSTM RNN with word embedding is expected as an effective solution for tagging tasks. ELMo seems to solve the context issue related issues for interior design and BERT model gives an accuracy of 97.5%.

Future enhancement

There can be improvement in data pre-processing, by automatically segmenting, stemming, affix-suffix correction and pre recognition of nouns before tagging.

REFERENCES

Cutting, Doug, *et al*., 1992. "A practical part-of-speech tagger." Proceedings of the third conference on Applied natural language processing. Association for Computational Linguistics.

Jurafsky D and Marting J H. 2002. Speech and Language Processing An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition, Pearson Education Series.

The Brown Tagset, available online at: http://www.scs.leeds.ac.uk/amalgam/tagsets/brown.html

Mitchell P. Marcus, Beatrice Santorini and Mary Ann Marcinkiewicz, 1994. "Building a Large Annotated Corpus of English: The Penn Treebank", Computational Linguistics Journal, Volume 19, Number 2, Pages: 313-330,. Available online at: http://www.ldc.upenn.edu/ Catalog/docs/treebank2/cl93.html

Petrov, Slav, Dipanjan Das, and Ryan McDonald. 2011. "A universal part-of-speech tagset." arXiv preprint arXiv:1104.2086.

Dash, Niladri Sekhar. 2008. Corpus linguistics: An introduction. Pearson Education India.

Manning, Christopher D. and Hinrich Schütze. 1999. Foundations of statistical natural language processing.Vol. 999. Cambridge: MIT press.

Ratnaparkhi, Adwait. 1996. "A maximum entropy model for part-of-speech tagging."Proceedings of the conference on empirical methods in natural language processing.Vol. 1.

Juan Antonio P´erez-Ortiz and Mikel L. Forcada, "Part-of-Speech Tagging with Recurrent Neural Networks", Universitatd'Alacant, Spain.

Himanshu Agrawal and Anirudh Mani, 2006. "Part of Speech Tagging and Chunking with Conditional Random Fields", In Proceedings of the NLPAI Machine Learning Competition.

Peilu Wang, Yao Qian, Frank K. Soong, Lei He, Hai Zhao1, 2015. Part-of-Speech Tagging with Bidirectional Long Short-Term Memory Recurrent Neural Network for conference at arXiv on oct.

David Elworthy, 1994. "Does Baum-Welch Re-estimation :Help Taggers?", In Proceedings of the fourth conference on Applied natural language processing, Pages: 53 – 58.

Brill, Eric. 1992. "A simple rule-based part of speech tagger."Proceedings of the workshop on Speech and Natural Language.Association for Computational Linguistics.

Pranjal Awasthi, Delip Rao and Balaraman Ravindran, 2006. "Part Of Speech Tagging and Chunking with HMM and CRF", In Proceedings of the NLPAI Machine Learning Competition.

Complete guide to spaCyhttps://nlpforhackers.io/complete-guide-to-spacy/

Natural Language processing using textblobhttps://textblob.readthedocs.io/en/dev/

Matthew E. Peters, Mark Neumann, MohitIyyer, Matt Gardner, Christopher Clark, Kenton Lee, Luke Zettlemoyer, 2018. "ELMo Deep contextualized word representation"
NAACL.

AshishVaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin 2017. "Attention Is All You Need" 31st Conference on Neural Information Processing Systems (NIPS), Long Beach, CA, USA.

****