How to Cite:
Gupta, A., Sharma, K., & Goyal, K. K. (2022). Measuring of similarity between verb pairs in the biomedical domain: An ontology-based information content perspective. *International Journal of Health Sciences*, *6*(S4), 5099–5112. https://doi.org/10.53730/ijhs.v6nS4.9281

Measuring of similarity between verb pairs in the biomedical domain: An ontology-based information content perspective

Atul Gupta
Bhagwant University, Ajmer, India
Corresponding author email: atul.cdac25@gmail.com

Kalpana Sharma
Bhagwant University, Ajmer, India
Email: kalpanasharma56@gmail.com

Krishan Kumar Goyal
Raja Balwant Singh, Agra, India
Email: kkgoyal@gmail.com

***Abstract*---**Finding the similarity Ontology-based between two verbs in bio-medical ontology is a difficult task as there is no standard dataset available. This paper is focused on verb-based similarity. So, the similarity between two nouns is taken as the benchmark for working on the verb similarity. There is no exact idea that the verb hierarchy of wordnet is capable to calculate the verb similarity between two verbs. The finding of similarity considers three parameters such as path, link, and depth. But in this paper, in addition to the path, link, and depth parameters, we also considered parameters such as stem-similarity weighting, derivation nouns weighting, and gloss similarity weighting. Moreover, we implemented two algorithms namely Rich Hierarchy Exploration and Shallow Hierarchy Exploration on a dataset and found that there is no significant difference in its performance.

***Keywords*---**semantic, similarity, verb similarity, wordnet.

Introduction

Semantic Similarity computation is very important in the area of natural language processing, artificial intelligence, bio-medical, psychology, and cognitive science. Many researchers compute the similarity between nouns using different types of methodologies using WordNet only a few researchers have been focused

Manuscript submitted: 27 March 2022, Manuscript revised: 18 May 2022, Accepted for publication: 9 June 2022

to compute verb similarity (Resnik, Philip, 200 C.E.). The reasons behind this are based on two things: firstly, no standard data sets are available for verb similarity calculation, and secondly, there is no exact idea that the verb hierarchy of WordNet is capable to calculate verb similarity.

The expansion of the work done in (Gupta & Goyal, 2019) on verb similarity where the performance of noun similarity is taken as a benchmark. In this paper, the calculation of the verb is done by two methodologies, the tuning method, and the evaluation partitioning method. For finding the similarity between verbs in WordNet, a well-to-do method based on noun similarity is adapted. Various semantic similarity algorithms that have been developed so far are used to find the semantic similarity between two words in a noun IS-A relationship by using Word-Net as an underlying ontology. Word-Net is a lexical database where words and their semantic relationship are present in more than 200 languages. The words are linked in the Word-Net have various types of relationships like synonyms, hyponyms, and meronyms. Semantic similarity algorithms are classified into two ways path length based as discussed in (Leacock, 2018)(Resnik, 1995) and information content based on(Jiang & Conrath, 1997; Lin, 1997).

The next section of this paper explains two methods for measuring the similarity between two words in detail. Section 3 explains a model-based Multiplicative Method for verb similarity. Different parameters that are used for the computational analysis are discussed briefly in Section 4. The generated results are presented in tabular form and discussed briefly in section 5. Finally, section 6 presents the relevant findings in the form of a conclusion and tell the use of the model for future research in the area of verb similarity.

Methods for Similarity

The similarity between the two words is measured with the help of two knowledge-based methods. They are poor- knowledge method and rich-knowledge method. In both methods, the word knowledge means acquired information that is lexically oriented from a thesaurus, that is pre-handcrafted or corpus learning. This paper tried to explain both the methods and then the similarity between two verbs is computed with the help of rich-knowledge methods.

Poor-Knowledge methods

The poor-knowledge method depends on information, the probability of information that is calculated by corpus, and the internet instead of the knowledge base. This was handled with the cooccurrence occurring of data.

Vector space approach

In this approach, the words which are semantically related are co-occurred in the corpus. This was constructed in a word or document order matrix. i.e: TF*IDF, Hear TF represent for Term frequency and IDF represents for Inverse Document Frequency but the fact of information is word co-occurrence in documents. The similarity is measured by comparing the distance measures with the help of the cosine coefficient as depicted in (Schütze, 1992).

Syntactic dependency approach

In this approach, the relatedness with respect to semantic between two words may lead to their use in grammatical structure. This approach also judges the words, which is done on tagging pos (part-of-speech) in the corpus, sentences shallow parsing, identifying the chunk relationship, and measuring the syntactic with the relationship dependency.

Rich Knowledge methods

The rich knowledge methods are used as a semantic network or a corpus that is tagged semantically, so that similarity between concepts of a word is measured. Various methods that compute similarity uses the semantic distance approach uses thesaurus knowledge or ontology such as WordNet. The famous methodologies which are thesaurus based used to compute semantic similarity can be defined in two different categories. They are the use of semantic links and statistics of the corpus with the distance in the taxonomy.

Edge Counting

The Edge counting approach which is based on the shortest path is derived from a geometrical model which is in generally used in Cognitive Psychology. In this shortest distance means a strong connection between stimuli and response. This method extends by Quillian's semantic similarity model depicted in (Quillian, 1967), where concept means nodes are arranged in the network hierarchy (Hierarchical network) and the similarity between nodes is measured by the number of hops in between nodes. The similarity between words in the thesaurus is defined as:

$$\mathrm{Sim}(c_1, c_2) = 2D - \mathrm{Distt}(c_1, c_2) \qquad (1)$$

In equation 1, D is constant (for example in the WordNet taxonomy maximum depth is 16 if we assume all the hierarchy shared a common node), $\mathrm{Distt}(c_1, c_2)$ count the link between two concepts means the node c_1 and c_2. In edge, counting approach distance is calculated by counting the number of edge traverses from c_1 to c_2. In other words, they are known as the nearest common node (NCN), which is represented as $\mathrm{Distt}(c_1, c_2)$.

In equation 2, the edge counting-based approach which was proposed in (Leacock, 2018) , that computes the verb concept-based similarity when the translation is done from English(verbs) to Chinese in (Leacock, 2018). In this, the similarity between words depends on the weighted sum and all the senses are used in the comparison.

$$\mathrm{Sim}\left(\mathrm{verb}_i, \mathrm{verb}_j\right) = \sum_k w_k \times \frac{2 \times \mathrm{depth}\left(\mathrm{ncn}\left(c_{i,k}, c_{j,k}\right)\right)}{\mathrm{depth}\left(c_{i,k}\right) + \mathrm{depth}\left(c_{j,k}\right)} \qquad (2)$$

- ncn ($c_{i,k}$, $c_{j,k}$) means the nearest common node for the concept $c_{I,k}$, $c_{j,k}$ and the verb of these concepts are $verb_I$ and $verb_j$.
- depth means of the verb concept from the root in the hierarchy,
- w_k denotes the weight of concepts(pairs).

This is one of the best approaches for computing the similarity between two nouns as well as two verbs in the IS-A hierarchical semantic network.

In equation 3, edge counting based approach was proposed in (Resnik, 1995) is presented and adapt the concept of information content approach implemented by Resnik in (Jiang & Conrath, 1997) to measure the semantic similarity between two words and is defined as:

$$\mathrm{Sim}\left(\mathrm{word_i}, \mathrm{word_j}\right) = \mathrm{MAX}\left[\frac{-\log \mathrm{Distt}\,(c_i, c_j)}{2 \times D}\right] = \mathrm{MAX}\left[\log 2D - \log \mathrm{Distt}\,\left(c_i, c_j\right)\right] \quad (3)$$

where, Distt (c_I, c_j) compute the shortest distance between c_i and c_j.

$$\mathrm{Distt}\,\left(c_i, c_j\right) = \mathrm{dep}(c_i) + \mathrm{dep}\left(c_j\right) - 2 \times \mathrm{dep}\left(\mathrm{ncn}\left(c_i, c_j\right)\right) \quad (4)$$

$$\mathrm{Sim}\left(\mathrm{word_I}, \mathrm{word_j}\right) = \mathrm{MAX}\left[\log \frac{2D}{\mathrm{Distt}(c_I, c_j)}\right] \quad (5)$$

Equation 4 and 5 present the concept model of similarity is similar to the approach presented in (Leacock et al. 1998) apart from normalization which is based on log.

Information Content

The author of (Resnik, Philip, 200 C.E.) discussed that the link in the WordNet hierarchy is an equable distance in the edge counting based approach so it doesn't consider semantic (similarity) alteration of a single link. The computation of IC value of nearest node which is common was done with the help of frequency statistics. These statistics are fetched directly from the corpus, not with the help of the edge counting approach. In this approach, the frequency statistics of the nearest common node subsumes the frequency statistics data of nodes subordinate to it. The information was calculated as negative of logarithmic (i.e – log p©). The drawbacks observed in (Resnik, Philip, 200 C.E.) are the information content-based approach is the ncn for every concept pair and the same parent node.

Resnik's work presented in (Resnik, Philip, 200 C.E.) is extended in (Lin, 1997) by combining the work of the edge counting approach and the information content approach. The author tries to enhance the coefficient of correlation on comparison to the judgment done by Human.

Human based correlation coefficient was tried to improve on consider the information content of the concept, conceptual density, depth, and link type. The same is expressed in equations 6 and 7.

$$\mathrm{Distt}\ (c_i, c_j) = IC(c_i) + IC(c_j) - 2 \times IC\ (ncn(c_i, c_j)) \quad (6)$$

where,

- $IC(c_i)$ and $IC(c_j)$ represent for information content value of concept 'i' and 'j'.
- $IC(ncn(c_i, c_j)$ represent for information content value of the nearest common node of concept 'i'

and 'j'.

where,

$$\mathrm{Sim}\ (c_i, c_j) = -\mathrm{Distt}(c_i, c_j) \quad (7)$$

- Sim (c_i, c_j) represent for the similarity between concept 'i' and 'j'.

The author of (Hirst, 2012) tries to find a new concept of computing semantic similarity to determine word sense and that is normalizing form of semantic similarity as discussed in (Lin, 1997) and presented in equation 8.

$$\mathrm{Sim}\left(c_i, c_j\right) = \frac{2 \times IC\left(ncn\left(c_i, c_j\right)\right)}{IC(c_i) + IC\left(c_j\right)} \quad (8)$$

With the help of the abovementioned concepts the proposed model which is used in this paper is described in next section.

A Model-Based On Multiplicative Method

In general, various semantic similarity-based approaches uses WordNet as a corpus and proposed in (Hirst, 2012; Jiang & Conrath, 1997; Leacock, 2018; Lin, 1997; Resnik, 1995). These approaches are presented in the following forms:

where,

$$\mathrm{Sim}(c_1, c_2) = 2X/(Y + Z) \quad (9)$$

$$\mathrm{Sim}(c_1, c_2) = 2X - (Y + Z) \quad (10)$$

X, Y, and Z denotes attributes of the concept c_1, c_2. The similarity between two concepts is defined with the help of concept multiplicatively in (Yang & Powers, 2005a)(Thanh et al.):

where,

Y_t represent for path factor

Z^X represent for link factor

$$\mathrm{Sim}(c_1, c_2) = Y_t \times Z^X \quad (11)$$

Yang and Power in (Thanh et al.) were partially inspired by the algorithm depicted in (Rubenstein & Goodenough, 1965) to detect and correct malapropism, for identical words assign different weights, weight for synonyms/antonyms and

hypernyms/hyponyms. They define the equivalent case so that the concept c_1 and c_2 are same as Y_{td} = 1, X = 0, the antonyms and synonyms are assigned with average weights, Y_{sa} = 0.90, X = 0,
hypernyms/hyponyms and holonyms/meronym are assigned with minimum weight (Y = Y_{hh} = Y_{hm} =
0.85, Z = ZY_{hh} = Z_{hm} = 0.7) if the searching depth(X) is more than 1 then weight being the result of tuned noun based similarity.

The model is evaluated where Human-based similarity judgment is a benchmark and the results evaluated by this model are improved if we compared the result with the previous methods. The calculated value of correlation with average human-based judgment on taking 28 standard noun pair datasets (Jiang & Conrath, 1997) is 0.9210, the correlation is better than any other semantic similarity approaches and it is superior than individual human-judgments (Resnik & Diab, 2000).

This algorithm also validates an independent noun pair dataset of 37 words (Fellabaum et al. 1998) and calculates correlation is 0.874 and also shows the cross-validated result on noun pair datasets of 65 word-pairs discussed in (Fellbaum et al. 1998) so that correlation is 0.895.

A Multi-strategic Verb Method

This method is used to investigate the suitability for the semantic similarity between two verbs, and the hierarchical structure of verbs is there in the WordNet corpus. The WordNet noun hierarchy is very rich in link and complexity where WordNet verb hierarchy is comparatively shallow in hypernymy/troponymy relations, holonymy/meronomy relations are not present in the WordNet Hierarchy. There are 4 nodes maximim when we are computing the distance between two verbs, that's make very difficult for computation of similarity between two verbs (Lesk, 1986).

This paper tries to present an approach, which computes the similarity between two pair of verbs on knowingly the limitation of verb hierarchy of WordNet. The discussed approach is based on the noun-based similarity (Gupta & Goyal, 2019) and almost similar to the approach discussed in (Thanh et al.). The novelty in this paper is by considering the derivational mapping into noun hierarchy while supplementing it into verb hierarchy, use of glosses definition, and the stemming effect. For constructing a verb similarity model using WordNet all the factors like path length, density, depth, etc. are also considered. The suffix removal function was provided with a corpus (WordNet) used for stemming. The computation of similarity between two verbs are done in the following steps:

The similarity between two verbs in the taxonomy of WordNet is computed in the similar way that are used in Noun hierarchy of WordNet (equations 11 and 12) but there is an exception, no correlation is there for holonymy/meronomy relationship. To do this there is a requirement to set up and tuning of parameters in the same way that is done in the noun model.

There are some verbs, that have a noun form as the stem, some noun has a verb form as stem and they are related derivationally. So, the projection is done to the noun hierarchy from verb hierarchy, to enhance the relationship between verbs by introducing Yder i.e discount factor /fusion weight. The verb definition and its glosses give an idea about the relationship with other verbs when no apparent relationships are there in noun and verb (hierarchy). The method proposed in (Pedersen & Banerjee, 2005), calculates the overlapping of target words and the words which select appropriate sense and definition. In the method proposed in (Yang & Powers, 2005a)where the co-occurrence approach is used for similarity calculation. This paper tries to measure the verb similarity by using WordNet corpus, which is not having frequent wordlist like "do", "make' etc. that bring the closed semantic relationship with target words(concepts). The concept of Y_{gls} as discussed in (Thanh et al.) is introduced in this paper for finding the similarity between two verbs.

To find the similarity between two verbs in the verb hierarchy of WordNet, the stemming effect is seen there without consideration of individual sense. This enables us to represents a broad class of relationships between verbs, but the relationship is not as strong as the relationship represents directly by-links. Y_{stm} is also introduced in this paper for finding the similarity between two verbs.

After consideration of the three new factors (Y_{der}, Y_{gls} and Y_{stm}) and already existing depth and link factor, there is a need to tune for the verb hierarchy of WordNet. No adjustment is encouraged because Yang and Power (Thanh et al.) already tuned this in similarity computation in the noun hierarchy of WordNet. So, the new approach is defined as:

$$\mathrm{Sim}(c_1, c_2) = Y_{stm} Y_t \prod_{i=1}^{\mathrm{Distt}(c_i, c_j)} Z_{t_i} \mathrm{Distt}(c_i, c_j) < X \quad (12)$$

$$c = 0, \mathrm{Distt}(c_i, c_j) \geq X \quad (13)$$

$$\mathrm{Sim}_{max}(\mathrm{verb}_1, \mathrm{verb}_2) = \underset{(i, j)}{\mathrm{Max}} \left[\mathrm{Sim}(c_{1,i,}, c_{2,j})\right] \quad (14)$$

Where,
c_1, c_2 denotes concept node where $0 \leq \mathrm{Sim}(c_1, c_2) \leq 1$) so that similarity between two concepts lies between 0 and 1.
der means derived nouns, "gls" means definition.
t = ht (hpernym / troponym) and sa means synonym/antonym.
The stemming factor is denoted by Y_{stm} .If concept (c_1) is a link to concept (c_2) with no stemming then Y_{stm} = I. The depth factor depends on link type and is denoted by Z_t.
The threshold on distance denoted by Z and the maximum value of Z will be five in WordNet verb hierarchy. The shortest distance between c_1 and c_2 is denoted by Distt(c_1, c_2).

The word is strongly related if c_1 is identical to c_2 , Y_{id} = 1 and Distt(c_1, c_2) = 0.If the link is of synonym/ antonym we assigned average weight (e.g. Y_{sa} = 0.91, Distt(c_1, c_2) = 0). In the link type is of hypernym/troponym again we tuned into lowest weight (e.g. Y_{sa} = 0.82). One point to understand if the link is identity type or synonym/antonym type, then it constitutes the complete path and it cannot be part of a multilinked path.

The available facts, that have most of the verbs with polysemous features and assigned the maximum similarity value in all the nouns (n_i) senses concept ($c_{I,j}$) of any word verb (v_i) having polysemous features. The computation of verb similarity in the WordNet corpus has used the Algorithm in (Thanh et al.), where the bidirectional search method is used. First, it finds if there is an identity path or synonym path, if not then discounting it to check if there is a hypernym/troponym path and then for connecting them extra distance need to calculate if it is failed there is a need to redone by further discounting by allowing the connection with derivationally related stemming not only specific senses. The proposed algorithm for the computation of similarity between verbs is presented in the next section

Computation Methods

On Computation of similarity between verbs the dataset, we consider the same as the data set of Yang and Power (Thanh et al.). In our similarity calculation three basic parameters i.e., path, length, and depth are used but we also consider stemming factor, derived noun factor, and gloss factor.

Task performed for verb similarity calculation

Unluckily, there is no standard data set for verbs similarity calculation. A standard data set is available for the calculation of similarity between two verbs. This dataset consists of 20 verbs(words) from selected 80 verbs(words) of TOEFL which was used by Launder and Dumais (1997) in their Multiple Choice Based questions set and 16 verbs (words) from 50 verbs(words) of English as Second language (ESL) used by Tatsuki in 1998[15]. These words are widely used by universities for the enrollment purpose of those who are non- native English speakers or employment in English-speaking countries. Now 36 numbers words are collected from two data sets (i.e. TOEFL and ESL) and all are the answers to the Multiple Choice Based questions (MCQ). There are four numbers of multiple-choice answers are available in each question. These answers are the synonyms and antonyms of those words. So, the total number of verb pairs are available in the data set is 144(i.e.36 x 4). The value of the similarity between the two verbs lies between 0 to 4.

0: words are not related at all 1: words are vaguely related 2: words are indirectly related 3: words are strongly related
4: words are inseparably related

To find the similarity calculation between verbs, first, the verb pairs are sorted in descending order by similarity value in terms of their average score. Then two categories are formed with the help of 26 words by eliminating some words having an average value below 2 for removing imbalance. In the process of making two datasets, 13 pairs of words are assigned to each dataset by maintaining the correlation. These datasets are named as dataset1 and dataset2. The main aim is the optimization of the verb model of the data set. For this, the correlation with the average human score is calculated by optimizing the various parameter of the model. The calculation is purely based on the Greedy approach. In this process, if the difference between parameter values is less than the mid-value then the result is considerable.

The proposed approach is to differentiate the effect of the various parameter of the verb model. In the process of computation of verb similarity, the gloss-level similarity (Y_{gls}) and derived noun-based similarity (Y_{der}) are taken as an independent. The optimal parameterization of the verb hierarchy is first explored and then finds how Y_{der} and Y_{gls} help for finding the similarity between two words.

Tuning process

On the calculation of verb similarity, there are three parameters, factor based on the path (Y), factor based on the link (Z), and the factor based on depth (X) (this factor is basically used to reducing CPU time and it may also serve act as a threshold to stop finding relationship). To find the alternative information sources three parameters are used. They are:
stem similarity weighting(Y_{stm}) derivation noun weighting(Y_{der}) gloss similarity weighting (Y_{gls})

Path Factor in similarity (Y)

The path factor taken here is in the range between 0.5 to 0.95 and is done by incrementing by 0.05. The optimal value of path factor (Y) is 0.85 and that is partially sensitive to its precise value.

Stemming factor in similarity (Ystm)

On seeing the effect of stemming factor in verb similarity on the hierarchy of WordNet. We analyze if the value of Ystm is greater than 0.4 then correlation decreases rapidly, and when the value of Ystm is less than 0.4 then little changes in the value.

Derived noun factor in similarity (Yder)

If the value of derived noun Yder is an increase from 0 to 0.5 then there is little difference in the correlation, but when the value of Yder is greater than 0.5 then the correlation declines slowly. The derived noun value 0.4 is taken as a compromising value. Hence a smaller value in the shallower verb hierarchy is obtained. To maximized the utilization of information in the semantic network larger value will be selected.

Gloss factor in similarity (Ygls)

The rise in the value of Ygls starts at 0.4 and it goes max to 0.9 then it falls.

Link Factor in similarity (Z)

The linking factor (Z) taken here is in the range between 0.3 to 0.7 and tuning is done on incrementing by 0.1. The effect of this in correlation with human-based judgment is observed. The link is of uniform length in the taxonomy of WordNet if Z=1.The performance of this system goes down when Z is greater than 0.6 and performance is maximum at Z=0.5.

Distance factor in similarity (X)

The distance factor plays an important role while computing the similarity between verbs. Initially the value of Y = 0.85 (path factor), Y_{stm} = 0.55 (stemming factor) and Z = 0.55 (link factor). Now for combined path-level, there is a variation in distance limit Z. The maximum distance from a node to another node in the WordNet is at most 5, enlarging the distance factor from 1 to 5 (viz. the distance in BDLS varied from 1 to 10) the model gives the high correlation value (i.e. similarity is high). During the calculation of correlation when the distance is increased from 1 to 2 only then the value of correlation drops, otherwise the correlation value increased. The main goal in this paper to analyze the functions of WordNet verb hierarchy, so for RHE (rich -hierarchy – exploration) use X=6 and SHE (shallow-hierarchy-exploration) use X=2. This paper demonstrates how the variant of RHE calibrates with the proposed model. The algorithmic version of the above discussed computational methods is presented in the next section for easy understanding.

Algorithm used for analysis

The stepwise algorithmic approach of the discussed computational method in section 4 is presented below for easy understanding.

//c_1, c_2 denotes concept node and v_1, v_2 denotes the verbs Repeat all $c_1 \in v_1$ and $c_2 \in v_2$
if c_1 & c_2 are antonyms or synonyms
assign path factor value of antonym or synonym to similarity function
else

end if

assign (path factor value * link factor value) of hypernym, troponym, antonym to the similarity function if the depth value is less than the depth factor of gamma
assign (stem path value factor* path value factor *link value factor of hyper/tropo to
$Sim_{stm}(c_1, c_2)$ & also to $Sim(c_1, c_2)$
end repeat
//procedure for calculation of $Sim_{max}(c_1, c_2)$, where $c_1 \in v_1$ and $c_2 \in v_2$
If def(v_1) is a subset of def(v_2) or def(v_2) is a subset of def(v_1) sim(v_1, v_2) =
$sim_{gls}(v_1,v_2)$=path factor of gloss similarity
else if $der_{noun}(v_1,v_2)$ is not equal to 0

$Sim_{der}(c_1, c_2)$ and also, to $Sim(v_1, v_2)$ assigned by path factor value of (derived noun*$Sim_{noun}(c_1, c_2)$) else
assign $Sim(v_1, v_2)$=maximum similarity of c_1 and c_2
end ifend if

The obtained result of the above algorithm is presented in next section with discussion.

Results and Discussions

The discussed algorithm in section 5 is executed with the help of Rich and Shallow methods independently. In each case, three numbers of data sets are taken. These are 'Data1' contains 65 numbers of the verb as discussed in section 4.1, 'Data2' contains 65 numbers of the verb as discussed in section 4.1. The third dataset named 'Total' is the combination of Data1 and Data2, which contain 130 numbers of verbs as shown in below figure 1.

After executing the algorithm on the dataset (Figure 1) as discusses in section 5 the result is presented in Table 1. The correlation is calculated on tuning the data set is denoted 'r_t' and the correlation calculated on evaluation-set is denoted by 'r_e', where data1 is used as evaluation-set for data2 and data2 is used as evaluation-set for data1 (Thanh et al.).

brag	boast	hail	acclaim	refer	explain	request	levy	anger	approve
concoct	devise	dissipate	disperse	finance	build	arrange	study	approve	boast
divide	split	approve	support	expect	deserve	relieve	hinder	research	distribute
build	construct	impose	levy	terminate	postpone	move	swell	request	concoct
end	terminate	hasten	accelerate	yell	boast	weave	print	boast	yield
accentuate	highlight	rap	tap	swell	curl	swear	think	furnish	impress
demonstrate	show	lean	rest	rotate	situate	forget	resolve	refine	sustain
solve	figure out	make	earn	seize	request	supervise	concoct	acknowledge	distribute
consume	eat	show	publish	approve	scorn	situate	isolate	clean	concoct
position	situate	sell	market	supply	consume	explain	boast	lean	grate
swear	vow	weave	intertwine	clip	twist	ache	spin	postpone	show
furnish	supply	refer	direct	divide	figure out	evaluate	terminate	hail	judge
merit	deserve	distribute	commercialize	advise	furnish	recognize	succeed	remember	hail
submit	yield	twist	intertwine	complain	boast	dilute	market	scrape	lean
seize	take	drain	tap	want	deserve	hasten	permit	sweat	spin
spin	twirl	depict	recognize	twist	fasten	scorn	yield	highlight	restore
enlarge	swell	build	organize	swing	crash	swear	describe	seize	refer
swing	sway	hail	address	make	trade	arrange	explain	levy	believe
circulate	distribute	call	refer	hinder	yield	discard	arrange	alter	highlight
recognize	acknowledge	swing	bounce	build	propose	list	figure out	refer	carry
resolve	settle	yield	seize	express	figure out	stamp	weave	empty	situate
prolong	sustain	split	crush	resolve	examine	market	sweeten	flush	spin
tap	knock	challenge	yield	bruise	split	boil	tap	shake	swell
block	hinder	hinder	assist	swing	break	sustain	lower	imitate	highlight
arrange	plan	welcome	recognize	catch	consume	resolve	publicize	correlate	levy
twist	curl	need	deserve	swear	explain	dissipate	isolate	refer	lean

Figure 1. Data Set (Yang & Powers, 2005b)

It was observed that when the verb model is tuned for the first time on the available dataset, few values are not able to correspond well with each other. Hence there is a need for two folds cross-validation which increases the similarity and reduces the score. It was also observed that tuning is a time taking process, so it is very difficult to perform high ordering cross-validation. As correlation is used to measure the sensitivity of every data sets, so to do tuning on others, we take on a compromise-based tuning on every subset and for later comparison, a human-based judgment method is also used. In the process of finding similarity between two verbs, it was observed that there is not so much difference between RHE (Rich Hierarchy Exploration) and SHE (Shallow Hierarchy Exploration).

Table 1 Correlation table

Rich/ Shallow	Dataset	X	Z	Y	Y_{stm}	Y_{der}	Y_{gls}	r_t	r_e
RHE	Data1	2.0	.49	.79	.39	.11	.91	.846	0.774
	Data2	2.0	.21	.84	.71	.79	.49	.862	.821
	Total	2.0	.49	.81	.49	.74	.59	.807	
SHE	Data1	0.0	.51	.74	.41	.69	.91	.835	.823
	Data2	0.0	.39	.81	.59	.71	.49	.842	.833
	Total	0.0	.51	.79	.51	.74	.61	.823	

The outcome says that only one is notably better than the other in one subject (for native-speaker). The RHE is failed in three subjects as compared with SHE but the result of RHE is better than Human-based judgments.

Conclusion

In the verb model the maximum depth in the verb hierarchy of wordnet is less than X that was computed in the noun model. Also, the link factor Z reduce more quickly on similarity of node, with the path factor discounted relationship of multiple links more severely. With the help of these fact, we are very much sure that verb hierarchy of WordNet is very shallower than noun hierarchy so to find similarity between verbs is not easy. The verb similarity model in the WordNet is not as good as noun similarity model of Yang and Power, reason for this is taxonomy of verb is shallower than noun but another factor is that the taxonomy of verb doesn't into contain part-whole relationship analog to holonym/meronym links of noun-based hierarchy.

The gloss, stem and noun similarity doesn't improve the condition only increased the parameter list to 9, 3 parameters for noun similarity, 3 parameters for verb similarity and 3 parameters for three fallback options. The model doesn't work as imagine so there is a future scope to investigate some other model for verb similarity.

References

Fellbaum, Christiane. "A semantic network of English verbs." WordNet: An electronic lexical database 3 (1998): 153-178.

Gupta, A., & Goyal, D. K. K. (2019). Classification of Semantic Similarity Technique between Word Pairs using Word Net. *International Journal of Engineering and Advanced Technology*, *9*(2), 4397–4402. https://doi.org/10.35940/ijeat.b2961.129219

Hirst, G. (2012). *Lexical chains as representations of context for the detection and correction of malapropisms.*

December.

Jiang, J. J., & Conrath, D. W. (1997). Semantic similarity based on corpus statistics and lexical taxonomy.

Proceedings of the 10th Research on Computational Linguistics International Conference, ROCLING 1997, Rocling X, 19–33.

Leacock, C. (2018). Combining Local Context and WordNet Similarity for Word Sense Identification. *WordNet, January 1998*. https://doi.org/10.7551/mitpress/7287.003.0018

Lesk, M. (1986). Automatic sense disambiguation using machine readable dictionaries: How to tell a pine cone from an ice cream cone. *Proceedings of the 5th Annual International Conference on Systems Documentation, SIGDOC 1986*, 24–26. https://doi.org/10.1145/318723.318728

Lin, D. (1997). *Using syntactic dependency as local context to resolve word sense ambiguity. i*, 64–71. https://doi.org/10.3115/979617.979626

Pedersen, T., & Banerjee, S. (2005). Maximizing Semantic Relatedness to Perform Word Sense Disambiguation.

University of Minnesota Supercomputing Institute Research Report UMSI, *25*(June).

Quillian, M. R. (1967). Word concepts: a theory and simulation of some basic semantic capabilities. *Behavioral Science*, *12*(5), 410–430. https://doi.org/10.1002/bs.3830120511

Resnik, Philip, and M. D. (200 C.E.). Measuring verb similarity. *22nd Annual Meeting of the Cognitive Science Society*, *546*, 2–9.

Resnik, P. (1995). *Using Information Content to Evaluate Semantic Similarity in a Taxonomy. 1.* http://arxiv.org/abs/cmp-lg/9511007

Resnik, P., & Diab, M. (2000). Measuring verb similarity. *Proceedings of the 22nd Annual Meeting of the Cognitive Science Society*, 399–404.

Rubenstein, H., & Goodenough, J. B. (1965). Contextual correlates of synonymy. *Communications of the ACM*, *8*(10), 627–633. https://doi.org/10.1145/365628.365657

Schütze, H. (1992). Dimensions of meaning. *Proceedings of the International Conference on Supercomputing, Part F1297*, 787–796. https://doi.org/10.1109/superc.1992.236684

Yang, D., & Powers, D. M. W. (2005a). Verb similarity on the taxonomy of WordNet. *GWC 2006: 3rd International Global WordNet Conference, Proceedings*, 121–128.

Yang, D., & Powers, D. M. W. (2005b). Verb similarity on the taxonomy of WordNet. *GWC 2006: 3rd International Global WordNet Conference, Proceedings, May 2014*, 121–128.

Thanh Ngoc Dao, Troy Simpson, 2005: Measuring similarity between sentences.

Widana, I.K., Dewi, G.A.O.C., Suryasa, W. (2020). Ergonomics approach to improve student concentration on learning process of professional ethics. *Journal of Advanced Research in Dynamical and Control Systems*, *12*(7), 429-445.

Widana, I.K., Sumetri, N.W., Sutapa, I.K., Suryasa, W. (2021). Anthropometric measures for better cardiovascular and musculoskeletal health. *Computer Applications in Engineering Education*, *29*(3), 550–561. https://doi.org/10.1002/cae.22202

Mukhtar, A. U. S., Budu, B., Sanusi B, Y., Mappawere, N. A., & Azniah, A. (2022). The effect of reproductive health education with multimedia video learning on the improvement of fluor albus prevention behavior young woman pathologist. *International Journal of Health & Medical Sciences*, *5*(1), 75-79. https://doi.org/10.21744/ijhms.v5n1.1841