TEXT ANALYSIS IN MONGOLIAN LANGUAGE

Chuluundorj Begz

University of the Humanities, Ulaanbaatar, Mongolia

ABSTRACT

The relevance of textual analysis appears in numerous case studies across fields of social, business and academic communication. A central question in multilingual research is to develop a universal concept representation using a variety of models. Typologically different languages may have differing numbers of values for the same concept. There is an increased interest how typologically different languages encode morphology and syntax features across different layers. Language specific features as subject-verb agreement, flexibility of word order, morphological type require more parameters to represent. Cross-lingual abstractions of morphosyntactic concepts serve as a basis for disentangling latent grammatical concepts across typologically diverse languages.

KEYWORDS

neural network, transformer architecture, three layers, matrixes, vector (tensor) model, dot product, encoding and word embedding.

1. INTRODUCTION

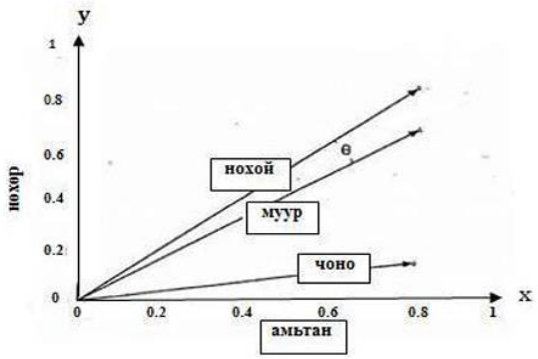
At the present, content analysis of text has practical relevance in all areas of social communication.Researchers have developed different techniques and algorithms for content analysis. The vector models, particularly dot product attention offers new opportunities for optimization of verbal communication and data processing, allowing for the analysis of relationships between words and texts.

Neural network models with multiple layers, are widely used for extracting features from text data.Recurrent neural network with the attention mechanism as a deep learning model facilitate feature extraction in a multiple-task social communication converting sequential data input into a specific output.(8)

Tensor modeling in neural networks is used to represent and manipulate data, where a vectors are derived from the tokens embedding through learned transformations.(10) In high dimensional vector space semantic relationships between words (concepts) is point of analysis in word embedding. Metric space (Euclidean, Minkowski, Hausdorff, Hilbert etc.) is important starting point for interpretation of psycho-cognitive operations as an operations of mental grammar, rules of verbal thinking.

Quantifying distances and vector representations in metric spaces enhances the ability to analyze textual data effectively. A mapping of a metric space to a Euclidean space might be useful to embed sequences into the Euclidean distance in linear semantic relationships in neural word embedding. Embedding as a nearest neighbor classification, cluster analysis, multi-dimensional scanling is be based on similarity measures between objects.

DOI: 10.5121/ijnlc.2025.14401



$$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

нөхөр - friend,
амьтан - animal,
нохой - dog,
муур - cat,
чоно - wolf.

Fig 1. Semantic space between words

Distance between words in the sentence reflects ratios of co-occurrence probabilities to encode meaning components in word vector space. Interpretation of the sentences illustrates mapping of event components in syntax structures:

Giving	Багш хүүд ном өгөв.
Taking	Хүү багшаас ном авав.
Throwing	The girl threw the ball to the boy.
Catching	The girl caught the ball from the boy.

Vector-based analysis of sequence regularities in above named structures proposes that implicit statistical knowiedge in working memory brings to light the relevance of intrinsic and extrinsic features of an object to verbal conition.

Association-based semantic models must be effectively applied to analysis of differences in word embeddings in typologically different languages. (6.9)

In the sentence “Хүү шатар тоглов” each word represented as a vector:

query-тоглов, key-хүү, value-шатар.

The query vector tells us what “тоглов” is seeking to understand: who or what is playing (тоглов< хүү, шатар). Key vector for “хүү” represents the information these tokens hold, value vector represents the actual content. We will focus on “тоглов” seeking information from“хүү” and “шатар”. To determine how relevant “хүү” and “шатар” to “тоглов”, we calculate the keys

$$\left(s\,(x_i) = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}} \right)$$

of the other tokens using the softmax function to convert these similarity scores into attention weights.

These weights determine how much attention “тоглов” should pay to “хүү” and “шатар”. Using these attention weights, we compute the weighted sum of the value vectors from “хүү” and “шатар”.

Difference in distance between the components in syntax constructions SVO and SOV is reflected in vector representation of words in the embedding space. In Mongolian language direct

object animacy and causative verbs are more relevant for contextual embedding. Computing the dot product between two vectors gives a value close to 1 as they are semantically closer.
Input: I bought a book.

Би нэг ном авсан.

	I bought a book				**Би нэг ном авсан.**				
Word Embedding	v_0	v_1	v_2	v_3	Word embedding	v_0	v_1	v_2	v_3
Positional encoding matrix.	p_0	p_1	p_2	p_3	Positional encoding matrix	p_0	p_1	p_2	p_3

R (index of token)

Би	0
нэг	1
ном	2
авсан	3

p_{00}	p_{01}	p_{02}	p_{03}
p_{10}	p_{11}	p_{12}	p_{13}
p_{20}	p_{21}	p_{22}	p_{23}
p_{30}	p_{31}	p_{32}	p_{33}

Above named structures with a differences in positional encoding present an object for modeling with sinusoidal function.

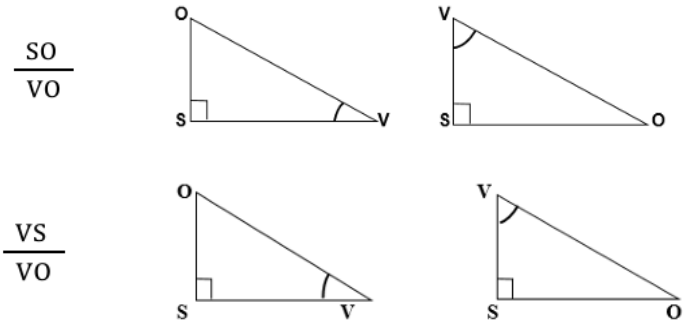
Neural representation of a word is mostly shaped by its syntactic category and unique semantic representation.

Parts of sentence attend to a fixed-size window and perception of a sentence depends on the attention transformer parts of sentence directly interacting across distant positions. In that's way sine and cosine functions play important role in modeling syntax structures as a sentence where words or tokens interact across distant positions.

Using Cosine and Sine functions to analyse of syntax structures SVO and SOV:

Brain opened the letter-SVO

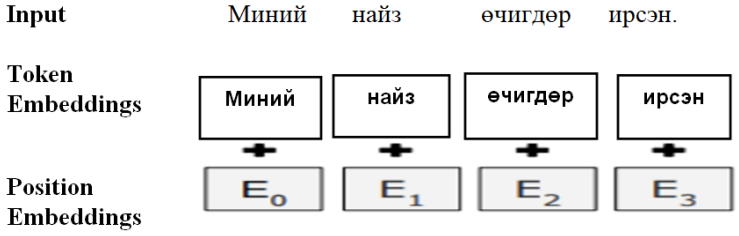
Хүү нэг ном авсан-SOV



The wave or signal coming from the vertical coordinate, or the a signal from horizontal coordinate must be used to modeling syntax structures in typologically different languages in relation to attention structure. The sine and cosine functions can be used to build waveform,

which help to describe a strength of the relationship between the parts of a sentence. Dot product in combination with applying cosine and sine functions to encoding positions presents an effective technique to analyze SVO and SOV structures reflecting the differences in human verbal cognition. This idea must be applied to analysis other complex syntax structures in typologically different languages with following contribution to developing LLMS.

Simple way to associate token and positional embeddings is based on co-occurrence matrix in linear meaning.



The model has a token embedding table to represent all the possible tokens in the input, it also has a positional embedding table with all the possible that the model supports.
Asymmetry in mapping, correlation between attention window and sentence structure also have importance in the light of tensor transformations, Tensor models present effective way to interpret, complex sentence as a object-extracted relative clauses, subject-extracted relative clauses, left-branching and right-branching constructions.

Хүү зам буруу заасан алдаагаа ойлгожээ.
Зам буруу заасан алдаагаа хүү ойлгожээ.

Syntax structures related to prototype-plus-distortion phenomenon (central prototype and radial prototype categories) present special case for embedding in human mental space depending on typological differences of the language.

Эгч талх зүсэв. The sister cuts the bread.

Radial Categories	**Force**=voluntary agent: **Цас хаалга даржээ - Салхинд хаалга хаагджээ.** The wind closed the door.
	Instrument: Чулуу цонх хагалжээ. The hammer broke the window.
	Neutral: Дэмбээ ууланд авирав. Peter climbed the mountain.

Applying the positional embedding to modeling different ways of encoding complex syntax structures, is important to alternate between cosine and sinusoidal functions to distinguish between odd and even indices of embedding dimension of a word. According to researchers, sinusoidal function is more applicable to positional embedding with multiple dimensions.(4) In multidimensional embedding space, is important to determine centers and spreads of topic-related words using methods like clustering algorithm for radial basis function neural network.

Vector multiplication-based modeling presents effective technique to embedding metaphorical, non-linear verbal structures where semantic transformations in combination with pragmatic meaning must be described in terms of torque (cross product). This model serves as a basis to develop the transformer-based cross-lingual embedding models.

Attention mechanism which uses weights the importance of different words stimulates the semantic transformation within neuro-semantic field serves as a basis to develop the model for embedding in non-linear dimensions of human mental space.

The attention mechanism which weights the importance of different words stimulates the semantic transformation within neuro-semantic field stimulating a development of more efficient attention mechanisms.

2. CONCLUSIONS

Embeddings have become a valuable tool in neuro-cognitive research for modeling human language in comparative perspectives.

Multimodal embeddings in shared representational space provide a computational framework for human verbal cognition. By representing words as vectors in multidimensional space, word embeddings in typologically different languages encapsulate the cohesion between words.

Representation of verbal structures on continuous vector space provides advanced tools for generating human language. The transformer-based neural network model offers new possibilities to developing learning methods in shared vector space.

Representation of words in different languages as dense vectors in a high dimensional space offers new ways of mirroring the complexities of human verbal cognition in the digital space.

Comparative analysis of linguistic features as a flexibility of word order, morphological type provides insights into cross-lingual word embeddings. Vector based shared space between different languages supports for developing knowledge transfer models in different applications.
Concepts of superposition, interference can be applied to analysis of metaphorical structures in typologically different languages. Principles that govern complex semantic representation must be integrated into new architectures of high dimensional semantic space based on quantum algorithms.

REFERENCES

[1] Alexander P., Solveiga V., Vivian Griff Griffiths., Jochen Braun. (2015). Transformation priming helps to disambiguate sudden changes of sensory inputs. *Elsevier*, 116

[2] Daniel Ibanez., Michal Albin. (2025). *Encoder-decoder models for language a processing*

[3] Evan D.Anderson., Aron K.Barbey. (2022). Investigating cognitive neuroscience a theories of human intelligence: A connection-based predictive modeling a approach. *Human brain mapping,* pp. 11-12

[4] Hong Liang., Xiao Sun., Yunlei Sun., Yuan Gao. (2018). Text feature extraction based on deep learning. *EURASIP. Journal of Wireless Communication and Networking*

[5] Jon Sprouse., Norbert Hornstein. (2016). Syntax and cognitive neuroscience of syntactic structure building. *Science Direct*, p.168

[6] Maithi Bhide. (2016). Style on multi-document pure institute of computer technology. India summarization technique. *IJCST*. 4*(3)*, p. 376

[7] Manning G., Sebastian Raschka. (2025). *Build a large language model*. pp. 7,171

[8] Momtazi, S., Rahbar, A., Salami, D., Khanijazani, I. (2019). A joint semantic vector representation model for text clustering and classification. *Journal of AI and data mining*, p.446

[9] Olushola A. (2024). Mathematical modeling of neural networks: Bridging the gap between mathematics and neurobiology. *Word Journal of Advanced Engineering Technology and Science*, 518, p. *1*

[10] Peter D. Bruza., Zheng Wang., Jerome R, Busymeyer. (2015). Quantum cognition: a new theoretical approach to psychology. *Trends in Cognitive Sciences*, 18

[11] Shuiwang, J., Yaochen Xie., Hongyang Gao. (2020). *A mathematical view of attention models in deep learning*. A&M University, Texas

[12] Vaibhav Sharma. (2025). *Scaling attention on transformers: sliding window & chunked attention.* p. 85

[13] Zuccon, G., Azzopardi, L.A., Rijsbergen, C.J. (2009). *Semantic space: Measuring a the distance between different subspaces.* University of Glasgow, 2009, p. 230