Word2Vec Word Embedding-Based Artificial Intelligence Model in the Triage of Patients with Suspected Diagnosis of Major Ischemic Stroke: A Feasibility Study

Aurora Zumbo
Humanitas Research Hospital
Antonio Desai
Humanitas University
Mauro Giordano
University of Campania “L. Vanvitelli”
Pierandrea Morandini
Humanitas Clinical and Research Center - IRCCS
Maria Elena Laino
IRCCS Humanitas Research Hospital
Elena Azzolini
IRCCS Humanitas Research Hospital
Andrea Fabbri
University of Rome "Tor Vergata"
Simona Marcheselli
IRCCS Humanitas Research Hospital
Alice Giotta Lucifero
University of Pavia
Sabino Luzzi
University of Pavia
Antonio Voza (antonio.voza@humanitas.it)
Humanitas University

Research Article

Keywords: Artificial Intelligence, Emergency Department, Major Ischemic Stroke, Word2Vec

Posted Date: May 3rd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1576126/v1

Abstract

Introduction

The possible benefits of using semantic language models in the early diagnosis of major ischemic stroke (MIS) based on artificial intelligence (AI) are still underestimated.

The present study strives to assay the feasibility of the Word2Vec word embedding-based AI model in decreasing the risk of false-negatives during the triage of patients with suspected MIS in the emergency department (ED).

Methods

The main ICD-9 codes related to MIS were used for a 7-years retrospective data collection of patients managed at the ED with a suspected diagnosis of stroke. Data underwent “tokenization” and “lemmatization”. Word2Vec word embedding AI algorithm was used for text data vectorization and a batch strategy was adopted for model training.

The false-negatives were rated from the incorrect attribution of the low-intensity code to patients later diagnosed with MIS.

Results

Out of 649 MIS, the Word2Vec AI algorithm allowed successfully identify 87.69% of them through the recognition of 15 top words, with an area under the curve of 97.2%. The rate of false-negatives related to the implementation of AI model was 0%.

Conclusions

Word2Vec word embedding-based AI model is reliable and effective in decreasing the risk of false negatives of MIS during patients’ triage in the ED. Further studies on larger cohorts are required to validate the model.

Introduction

Major ischemic stroke (MIS) affects over 600.000 patients/year being among the top five causes of death and the first cause of disability in the United States [1]. MIS evolution time accounts for 10 hours on average (range 6–18 hours) and it has been estimated that the patient loses 1.9 million neurons from the diagnosis each minute where MIS is untreated [2]. Misdiagnosis of MIS has been associated with false-positive (stroke mimics) and -negatives (stroke chameleons) up to 26% and 43% of cases, respectively [3].

Randomized trials demonstrated that the best outcome is achievable within 4.5 hours from the onset of stroke [4–8]. Accordingly, an early and accurate diagnosis of possible MIS patients and their aggressive treatment are mandatory [2, 3, 9–12].

While vital, the involvement of human resources as nurses, neurologists, and radiologists has been reported to act itself as a time-limiting step in the stroke triage and imaging pathway, especially because this expertise may be not available at all sites or times [2]. These are the main reasons at the base of a more and more increasing interest toward the automatization of the acute management of MIS. Artificial intelligence (AI) technology has already used in acute ischemic and hemorrhagic stroke imaging [2, 13, 14]. However, the semantic models of representation languages and their potential advantages in the optimization of the MIS management remain still largely underestimated.

The aim of the present study is to test the feasibility of the implementation of the Word2Vec word embedding-based AI model in decreasing the risk of false-negatives during the triage of patients with suspected diagnosis of MIS in the emergency department (ED).

Methods

Data Collection

The study was approved by the Internal Review Board of Humanitas Research Hospital. Patients' data were retrospectively collected from clinical notes at triage of the ED and referred to the timeframe January 2015-August 2021. Admission diagnoses were derived from the assigned International Classification of Diseases 9th revision (ICD-9) code after the first visit. The ICD 9 codes specifically selected because related to an MIS were as follows: 434.01 (cerebral thrombosis with cerebral infarction); 434.90 (cerebral artery occlusion, unspecified without mention of cerebral infarction), 434.91 (cerebral artery occlusion, unspecified with cerebral infarction).

Text Preprocessing

Text data underwent "tokenization" consisting of some preprocessing steps to clean and normalize the variables and separate the paragraphs into words (token). Text words were lowercased and normalized through the removal of punctuation, numbers, and non-ASCII characters. White space character was used as a delimiter for each token, practically transforming the paragraphs into lists of tokens. Stop words, like prepositions and articles, were removed to further clean the texts from undesired tokens.

The last preprocessing step was the "lemmatization", aimed at reducing the number of different tokens. TreeTagger library was used for this step.

Text Data Vectorization

Word2Vec word embedding artificial intelligence algorithm was used for text data vectorization. To produce the embedding, Word2Vec builds a shallow neural network able to predict a word given its context. The values assumed by the intermediate layer during this prediction are then used as embedding for the given word. The embedding dimension N chosen in this setup is 300, meaning that each word is transposed to a numerical vector of 300 dimensions (Fig. 1).

The final vector for each paragraph was obtained averaging the values of the embedding tokens.

Classification and Model Training

Data were split with stratification into the train-test sets with a proportion of 80/20. The model used was a neural network with a hidden layer of 32 neurons. To adjust for the class imbalance, the training process was performed with a batch strategy: in each training round, the two classes were sub-sampled with a replacement for the positive class and without replacement for the negative class. This strategy allowed to present a balanced batch to the model to be trained.

Rating of the False-Negatives

The percentage of the false-negatives was rated in the Word2Vec model and in-use color code triage classification model. The Triages color code used in our hospital is similar to the numeric triage score code, where triage 1 is identified by red code (immediate evaluation), triage 2 and 3 are Yellow code (evaluation by the clinician guaranteed in 15–60 minutes) and triage 4 is green code (evaluation time is 180 minutes).

For this last, the false-negatives were derived from the incorrect attribution of the green code to patients later diagnosed with MIS.

Results

The total number of MIS is 649. The performance on the test set showed that stroke patients were successfully identified with a recall of 87.69% and an area under the curve (AUC) of 97.2% (Fig. 2).

Word2Vec AI algorithm was able to identify top 15 words positively correlated to MIS diagnosis using the cosine similarity as a metric between the average stroke patients text vector and the different word vectors. Dysarthria and aphasia were the text words more strongly correlated with correct diagnosis of MIS (Fig. 3).

The percentage of false-negatives related to Word2Vec AI model and color code classification one was 0% and 5%, respectively (Fig. 4).

Discussion

The present study strived to test the feasibility of the implementation of the Word2Vec AI model in the optimization of the acute management of MIS starting from the suspected diagnosis during patients' triage. The model identified 15 text words highly predictive of MIS and it proved to be highly effective in the prompt and accurate diagnosis with a rate of false negatives of 0%. The model has been confirmed to be even more accurate than the conventional color code triage classification model, which is in use in all the Italian ED nowadays. More than 80% of strokes result from ischemic damage to the brain due to an acute reduction in the blood supply. The goal in the management of acute ischemic stroke is early arterial recanalization to limit the brain damage since the delay in starting the treatment is associated to worst

physical e cognitive outcome, with a high level of disability and comorbidities [2, 15, 16]. Although faster triage, improvements in neuroimaging techniques, thrombolysis, and thrombectomy represent the major advances of MIS management, the overall outcome of patients affected by stroke is still largely dependent on a prompt and accurate diagnosis at admission at the ED [12, 17–22].

Based on our results, the presence of one of the 15 keywords identified by the proposed AI model is associated with a rapid diagnosis of stroke and the performance on the test set shows that stroke patients were successfully identified with a recall of 87.69% and an AUC of 97.2. Dysarthria and aphasia were the text words most importantly correlated with the stroke diagnosis. Noteworthy, the model was able to correctly associate with a suspected diagnosis of stroke also those misspelled text words that were accidentally recorded during the triage. “Disatria” instead of “disartria”, namely dysarthric speech, was an example. Because of a documented lower rate of false negatives, the Word2Vec AI model has proved to be more accurate than the color-code based classification model in-use in our department as many other hospitals in north Europe. The practical implication of this model in daily practice are non-negligible since it may contribute to the optimization of the acute management of patients affected by MIS. In a combined vision where the AI models are integrative rather than substitutive of the human resources, the availability of a computer alert generated by the AI algorithm may be of help to nurses and others to early recognize those patients suspected to be affected by ischemic stroke. Further AI algorithms like that reported in the present study may also be adopted for the hemorrhagic stroke.

One Hot Encoding and Word Embedding are two of the most popular concepts for vector representation in Natural Language Processing. Word2vec is an algorithm created in 2013 that uses a neural network model to identify words associated starting from a big matrix of data set and, once trained, it can select words with similar meaning from the words surrounding it. It represents each word identified by a list of numbers called vectors. The vectors are selected with a simple mathematical function and share a certain level of semantic similarity between the words associated with those vectors [23].

The choice of Word2vec embedding-based AI algorithm lets us work on a big volume of data in a simple way. This algorithm selected words with intrinsic meaning, starting with a numeric vector obtained from a dependent variable. From the numeric vector (whose length is about 300, established by our team) we process data with a statistic model that can interpret artificial neural networks obtained by using Word2vec algorithm.

Another algorithm that could be used because of the easiness of implementation is “One hot encoding”, working in a faster way than Word 2 embedding: every word has its own value in a vector, but in this process, it loses the semantic meaning of the word in a sentence. One hot encoding was one of the first techniques used in artificial intelligence models but with the birth of Word-embedding, it becomes obsolete, especially in scientific fields. Furthermore, by using a one-hot encoding algorithm the size of the embedding vector grows with the vocabulary so it could be difficult to elaborate those data because of the entity of the matrix of embedding obtained, so it doesn't work well in applications that require a large amount of data. Word2vec with its implementation could be a good middle ground even because the

precision of word embedding depends on the volume of the dataset, so it works well on large datasets obtaining the best word embedding with the smallest matrix.

Other models of Artificial Intelligence include GloVe and FastText.

With Word2vec we train a neural network with a single hidden layer to predict a target word based on its context. With FastText each word is composed of character n-gram so it can help to generate better word embeddings for rare words or for out of vocabulary words; a big limit of this algorithm is that it takes longer to do the embedding and as the dataset grows, the memory required grows too, so in this way is similar to One Hot Encoding.

The gloVe is a word embedding technique similar to Word2vec, but it differs from it because it is a count-based model instead of the predictive model. In fact, GloVe focuses on words co-occurrences over the whole corpus, while Word2vec leverages co-occurrence within local context (neighboring words). Glove embeddings relate to the probability that two words appear together.

Word embedding techniques, with respect to count-based methods, on different language tasks such as semantic relatedness, synonym detection, concept categorization, and analogy. With Word2vec we observe large improvements in accuracy at much lower computational cost, i.e. it takes less than a day to learn high quality.

As reported, the need for continuous training of the AI model, by means of the increase of the data collected from other clinical studies, is a key aspect for the further improvement and optimization of the model itself [24, 25].

Limitations of the Study

The first limitation of the present study lies in the exclusion of hemorrhagic stroke or TIA, considering only MIS.

Furthermore, this word embedding-based AI model didn't explore the Vital Signs, which are extremely useful to detect the critical issues of the patient. Using Word2vec we obtained the classification of a word strongly associated with MIS in terms of clinical features, but this algorithm does not work on the definite diagnosis of the disease. With AI models it would be easy to create a warning signal with those “embedded words”, popping up on computers of triage's nurses, but the meaning of that “alert” must be evaluated according to the cases. For example, one of the words most associated with stroke diagnosis according to the model Word2vec is “disorientation” but only in few cases this clinical feature is observed in those patients. Another limitation of the algorithm is that the detection of true positive cases is not well balanced by the identification of true negative rates. It could overestimate the real impact of the disease in triage.

With Word2Vec the word embedding obtained by using the algorithm is “static”, which means that the model has no awareness of the framework in which the word is found. By using recurrent neural

networks, the word embedding could become more dynamic and accurate: this new AI model is able to detect the hidden relationship between inputs as well as to provide a precise sequence prediction of words, giving a high accuracy to results.

Future perspectives could involve dynamic models of word embedding, more refined. In fact, while working on recurrent neural networks the word embedding will help us to obtain more precise results even on false-negative cases, taking all the vectors generated by the algorithm with new technology.

Conclusions

The present feasibility study demonstrated that the Word2Vec word embedding-based AI model was reliable in leading to a suspected diagnosis of MIS during patients' triage in the ED.

It also proved to be more accurate than the color-code-based classification model of triage.

Further studies on larger patients' cohorts are mandatory to validate definitively the proposed AI model.

Declarations

Author Contributions

Conceptualization, A.Z, A.D, A.G.L; methodology, M.G., P.M., L.M.E., software, E.A.; validation, S.L., A.G.L., A.F.; formal analysis, S.M.; data curation, A.Z, S.L., A.G.L., A.S., J.M.S.-A.; writing—original draft preparation, A.Z., A.D, A.V.; writing—review and editing, S.L., A.D., A.V.; visualization, A.V.; supervision, A.V.; project administration, S.L., A.D., A.V. All authors have read and agreed to the published version of the manuscript.

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Compliance with Ethical Standards and Ethical Approval:

All procedures performed in the study were in accordance with the ethical standards of the institution and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Informed Consent:

Not applicable.

References

1. Go, A.S., et al., *Heart disease and stroke statistics–2014 update: a report from the American Heart Association.* Circulation, 2014. **129**(3): p. e28-e292.
2. Saver, J.L., *Time is brain–quantified.* Stroke, 2006. **37**(1): p. 263-6.
3. Liberman, A.L. and S. Prabhakaran, *Stroke Chameleons and Stroke Mimics in the Emergency Department.* Curr Neurol Neurosci Rep, 2017. **17**(2): p. 15.
4. Levine, D.A., et al., *Risk Factors for Poststroke Cognitive Decline: The REGARDS Study (Reasons for Geographic and Racial Differences in Stroke).* Stroke, 2018. **49**(4): p. 987-994.
5. Saver, J.L., et al., *Time to Treatment With Intravenous Tissue Plasminogen Activator and Outcome From Acute Ischemic Stroke.* JAMA, 2013. **309**(23): p. 2480-2488.
6. Emberson, J., et al., *Effect of treatment delay, age, and stroke severity on the effects of intravenous thrombolysis with alteplase for acute ischaemic stroke: a meta-analysis of individual patient data from randomised trials.* Lancet, 2014. **384**(9958): p. 1929-35.
7. Tawil, S.E. and K.W. Muir, *Thrombolysis and thrombectomy for acute ischaemic stroke.* Clin Med (Lond), 2017. **17**(2): p. 161-165.
8. Albers, G.W., et al., *Thrombectomy for Stroke at 6 to 16 Hours with Selection by Perfusion Imaging.* N Engl J Med, 2018. **378**(8): p. 708-718.
9. Fugate, J.E. and A.A. Rabinstein, *Absolute and Relative Contraindications to IV rt-PA for Acute Ischemic Stroke.* Neurohospitalist, 2015. **5**(3): p. 110-21.
10. Ekundayo, O.J., et al., *Patterns of emergency medical services use and its association with timely stroke treatment: findings from Get With the Guidelines-Stroke.* Circ Cardiovasc Qual Outcomes, 2013. **6**(3): p. 262-9.
11. Jia, J., et al., *Accuracy of Emergency Medical Services Dispatcher and Crew Diagnosis of Stroke in Clinical Practice.* Frontiers in Neurology, 2017. **8**(466).
12. Schwamm, L.H., et al., *Intravenous thrombolysis in unwitnessed stroke onset: MR WITNESS trial results.* Ann Neurol, 2018. **83**(5): p. 980-993.
13. Murray, N.M., et al., *Artificial intelligence to diagnose ischemic stroke and identify large vessel occlusions: a systematic review.* J Neurointerv Surg, 2020. **12**(2): p. 156-164.
14. Soun, J.E., et al., *Artificial Intelligence and Acute Stroke Imaging.* AJNR Am J Neuroradiol, 2021. **42**(1): p. 2-11.
15. Prabhakaran, S., I. Ruff, and R.A. Bernstein, *Acute stroke intervention: a systematic review.* Jama, 2015. **313**(14): p. 1451-62.
16. Arenillas, J.F., et al., *Relative cerebral blood volume is associated with collateral status and infarct growth in stroke patients in SWIFT PRIME.* J Cereb Blood Flow Metab, 2018. **38**(10): p. 1839-1847.

17. Oostema, J.A., et al., *Brief Educational Intervention Improves Emergency Medical Services Stroke Recognition.* Stroke, 2019. **50**(5): p. 1193-1200.
18. Gorchs-Molist, M., et al., *An Online Training Intervention on Prehospital Stroke Codes in Catalonia to Improve the Knowledge, Pre-Notification Compliance and Time Performance of Emergency Medical Services Professionals.* Int J Environ Res Public Health, 2020. **17**(17).
19. Oostema, J.A., et al., *The quality of prehospital ischemic stroke care: compliance with guidelines and impact on in-hospital stroke response.* J Stroke Cerebrovasc Dis, 2014. **23**(10): p. 2773-2779.
20. Lin, C.B., et al., *Emergency medical service hospital prenotification is associated with improved evaluation and treatment of acute ischemic stroke.* Circ Cardiovasc Qual Outcomes, 2012. **5**(4): p. 514-22.
21. Yperzeele, L., et al., *Prehospital stroke care: limitations of current interventions and focus on new developments.* Cerebrovasc Dis, 2014. **38**(1): p. 1-9.
22. Brice, J.H., et al., *Stroke: from recognition by the public to management by emergency medical services.* Prehosp Emerg Care, 2002. **6**(1): p. 99-106.
23. Mikolov, T., et al., *Efficient Estimation of Word Representations in Vector Space.* Proceedings of Workshop at ICLR, 2013. **2013**.
24. Powers, W.J., et al., *2018 Guidelines for the Early Management of Patients With Acute Ischemic Stroke: A Guideline for Healthcare Professionals From the American Heart Association/American Stroke Association.* Stroke, 2018. **49**(3): p. e46-e110.
25. Jiang, F., et al., *Artificial intelligence in healthcare: past, present and future.* Stroke Vasc Neurol, 2017. **2**(4): p. 230-243.

Figures

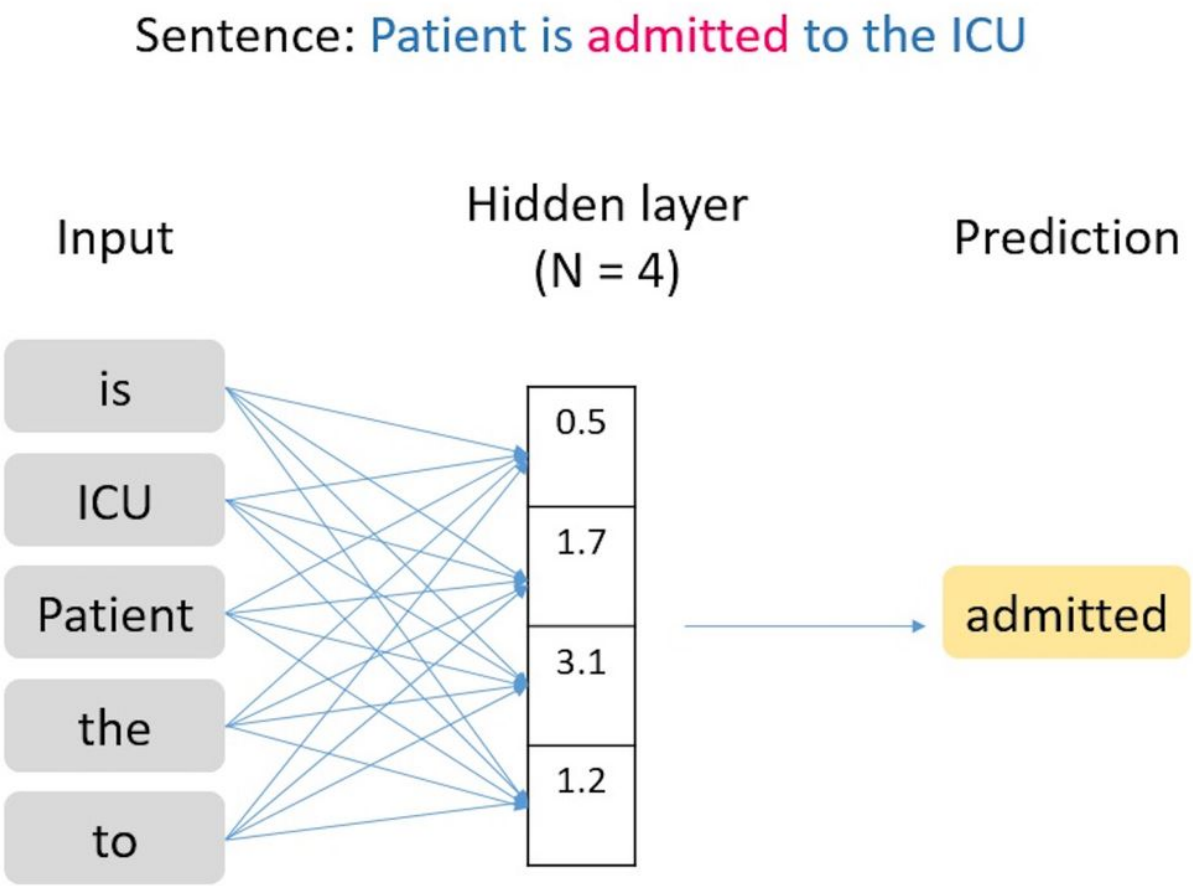

Figure 1

Word2vec Word embedding-based model.

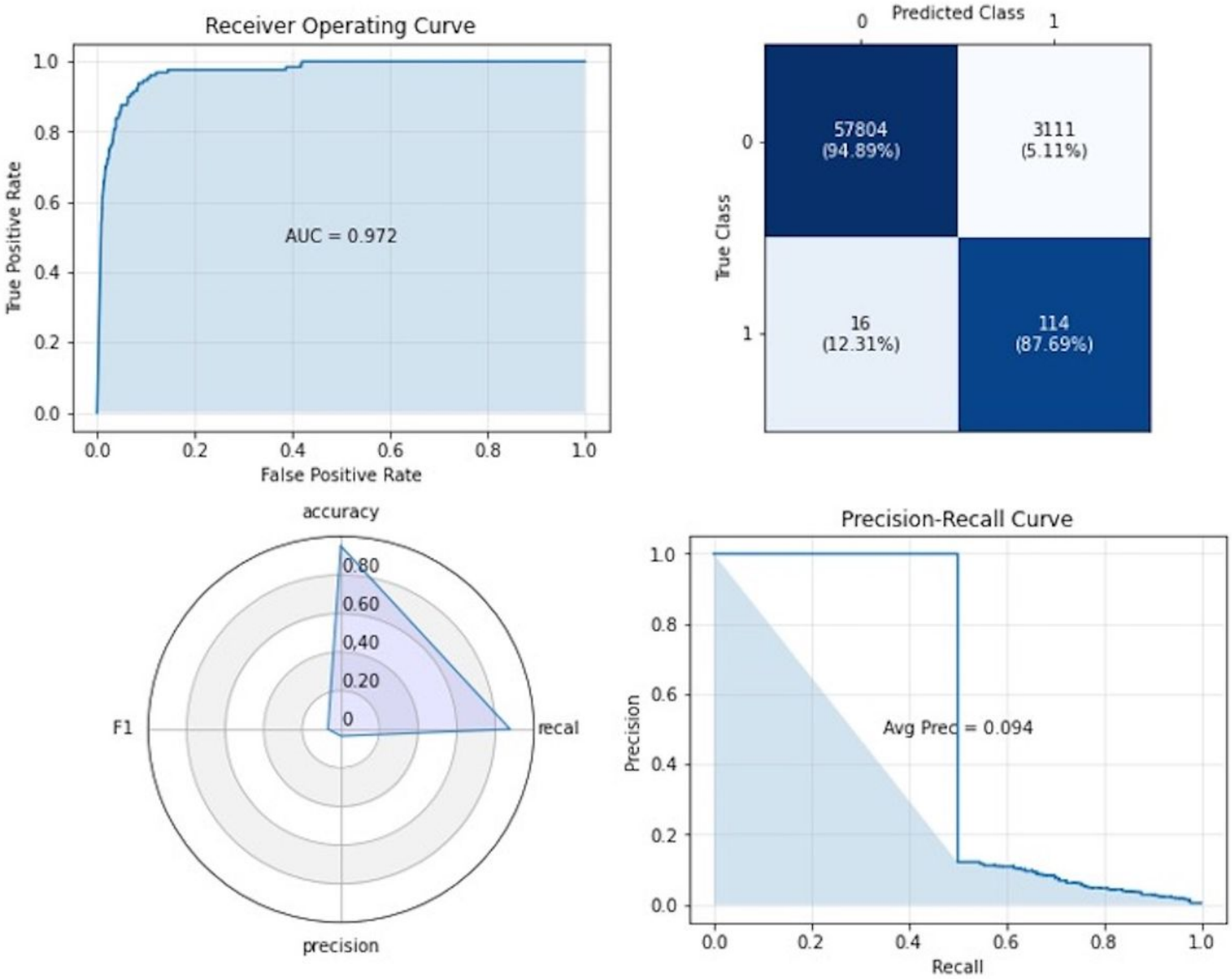

Figure 2

Test set for stroke patients. AUC: Area Under the Curve.

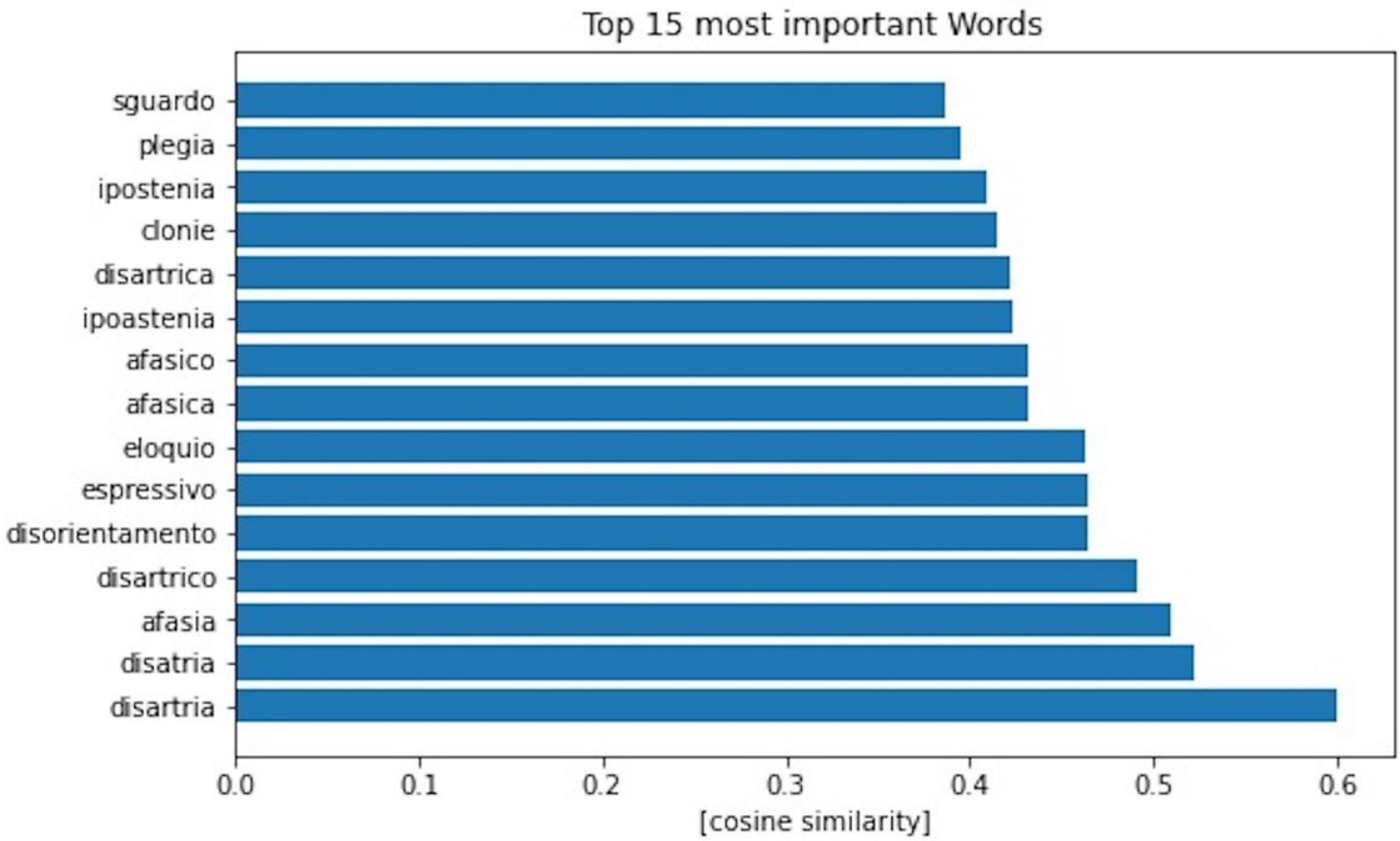

Figure 3

Top words obtained by Word embedding.

Afasia or afasico/a: aphasia/aphasic (masculine and feminine adjective); clonie: clonic movements; disartria/disatria: dysarthria, the second word is misspelled/orthographically wrong; disartrico/a: dysarthric (masculine and feminine adjective); disorientamento: disorientation; eloquio: language; espressivo: expressive, a type of aphasic speech (e.g. expressive aphasia); ipostenia/ipoastenia: weakness, the second word is misspelled/orthographically wrong; plegia: plegy; Sguardo: gaze.

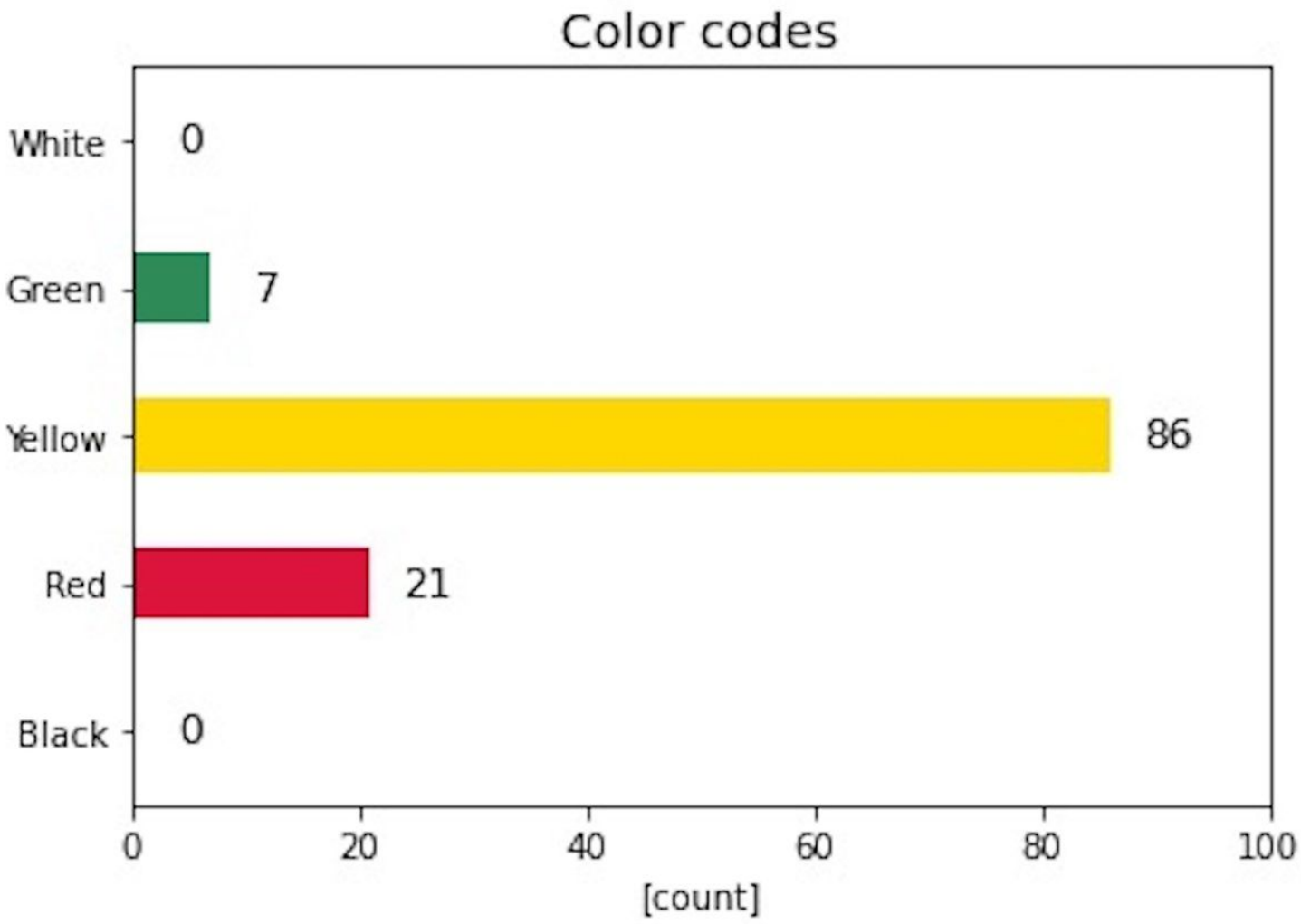

Figure 4

Triages color code.