DYSLEXIC READING ASSISTANCE WITH LANGUAGEPROCESSING ALGORITHMS

Sharada Lakshmanan[1]

[1]Equifax, Atlanta, US

ABSTRACT

The Dyslexic reading challenge is not completely resolved till date with advance learning algorithms. The need for breaking complex words and aiding the children to remember with hints is the need of the hour. The development of reading assistant with breakdown of complex words, using hints to remember backword and forward iterations, creative use of word cloud and using deep learning techniques to effectively tokenize and assist the struggling readers.

KEYWORDS

Language processing algorithms, deep learning , tokenization.

1. INTRODUCTION

The usage of learning algorithms like Bag Of Words , BERT, Reinforcement Learning Algorithms effectively to break down complex words and segment text to aid children with reading difficulties to guide them in amore simplistic and illustrative manner. We have to focus on breaking difficult words along with a contextual understanding of the sentences in paragraphs. The algorithm uses word embeddings for finding hidden connections between words in sentences.It can create words and sentences vectors from hidden states. It also ensures the vectors are contextually dependent without misleading the challenged dyslexic readers. The main focus of the algorithm is to bring utmost simplicity to the challenged readers.

Throughout the learning years people with learning disorders , ADHD Spectrum ,dyslexic children and adults suffer learning tough subjects like literature, history and so on.

The need of the modern-day challenges is to effectively use AI related technologies to facilitate the children and adults to read with enhanced simplicity. As of this year 2022 the reinforcement learning algorithms are able to achieve 27% that enhances students' learning rate with dyslexia to the face-to-face approach. So, making the algorithms with gamification and reward-based model will encourage the disabled user communities.

The witness and comparison of school text books and their simplification of learning with modern day aides are still not sufficient to help the students to get into the expected learning rate.

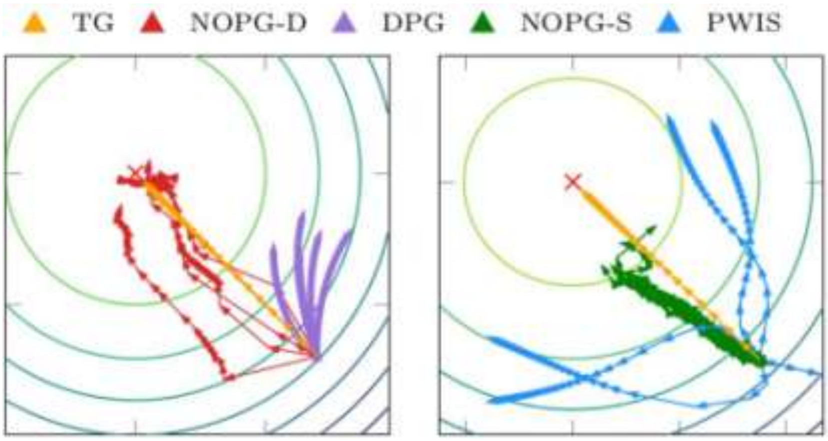
The study has been made in both Indian C.B.S.E text books, Cambridge publication books, American public schools.

For example, I had to simplify the complexity of Cambridge English to simpler language with an aid of a special teacher along with help of pictures and audio-based books to bring down the simplicity. Thru further usage of interactive quiz to recall the learning and also the use of gamification was required for intervention.

The US public effectively use audio books, interactive learning with quiz, animation and further gamification techniques to break the complexity of the language.

The author of this article has specifically witnessed the problems from toddler to high school grade students reading problems, the intervention techniques used by special teachers, therapist, use of audio books, gamification techniques, using art and pictorial representation to fasten the learning, motivational tools used by gamification methodology and also experimenting and comparing the results used in Natural Language processing algorithms with deep learning algorithms.

The results of the students between using the gamification, technology aides were compared and contrasted with studies without the aides.



The developments in artificial intelligence and autonomous learning have shown exemplary results in tasks like board games and computer games. However, the applicability of learning techniques remains mainly limited to simulated environments.

The major cause of this inapplicability to real-world scenarios is the general sample-inefficiency problems and inability to guarantee the safe operation of state-of-the-art reinforcement learning.

Methods: We administered two versions of a probabilistic selection task to 40 schizophrenia patients and 31 control subjects, using difficult to verbalize stimuli (experiment 1) and nameable objects (experiment 2). In an acquisition phase, participants learned to choose between three different stimulus pairs (AB, CD, EF) presented in random order, based on probabilistic feedback (80%, 70%, 60%). We used analyses of variance (ANOVAs) to assess the effects of group and reinforcement probability on two measures of contingency learning. To characterize the preference of subjects for choosing the most rewarded stimulus and avoiding the most punished stimulus, we subsequently tested participants with novel pairs of stimuli involving either A or B, providing no feedback.

Results: Control subjects demonstrated superior performance during the first 40 acquisition trials in each of the 80% and 70% conditions versus the 60% condition; patients showed similarly impaired (<60%) performance in all three conditions. In novel test pairs, patients showed decreased preference for the most rewarded stimulus (A; t = 2.674; p = .01). Patients were unimpaired at avoiding the most negative stimulus (B; t = .737).

2. LANGUAGE PROCESSING ALGORITHMS

BERT stands for "Bidirectional Encoder Representations from Transformers," .This gives importance to context, tokenization and understand the eye movement of the imbalanced readers.

In English 'Content' means either information or state of satisfaction. The word 'Bold' means the letter style or the character of person. BERT analyses and parses the words the same way as human does. BERT is capable of analysing chaotic data.

A subset of machine learning, deep-learning requires the use of neural networks similar to the neurons in the human brain. These networks sift through data in a hierarchical fashion, which allows the algorithm to process information nonlinearly. This allows BERT to analyze data more quickly while also learning from its previous experiences with other sets of data. In theory, this means that BERT will continually improve its ability to understand the context of data

2.1. BERT use AI and Natural Language Programming to Understand Complex data

Older versions of NLP would omit certain words from a long sentence, and the results that do not match the intention of the reader. BERT takes ALL the words into account of the context of the query. It then uses an artificial intelligence (AI) algorithm and Natural Language Programming (NLP) to 'understand' the intent of the reader and deliver the true results the AI 'thinks' the User is interested in. Only one in ten results will typically change on the average. The more complex the information the more results may change due to BERT.

Long information or complex sentences are a nightmare for AI to understand. Language and the meaning of words change as they are used in a sentence. An early word in the search has a meaning as you read it, but it changes when you are done reading it in the full context in the full sentence.

Older versions of the Algorithm would ignore or devalue certain words in a long-tail query. Because of this, many times the results were incorrect as language is very fluid and every word HAS meaning in a query. BERT seeks to provide more accurate results.

Example 1: If someone searches for "2022 India traveler to usa need a visa".

With BERT Update, algorithm is able to grasp that it's related to Indian traveling to USA. . So, BERT is now focussing more on Long Tail Keywords.

2.2. Reinforcement Learning Algorithms

There are three approaches to implement a Reinforcement Learning algorithm.Value-Based:

In a value-based Reinforcement Learning method, you should try to maximize a value function V(s). In this method, the agent is expecting a long-term return of the current states under policy π.

Policy-based:

In a policy-based RL method, you try to come up with such a policy that the action performed in every state helps you to gain maximum reward in the future.

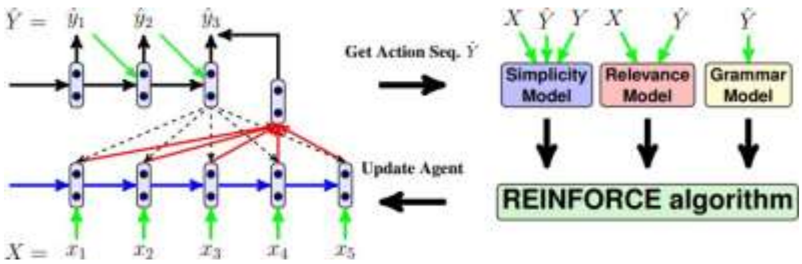
2.3. Reinforcement AI Terminology

Here are some important terms used in Reinforcement AI:

- Agent: It is an assumed entity which performs actions in an environment to gain some reward.
- Environment (e): A scenario that an agent has to face.
- Reward (R): An immediate return given to an agent when he or she performs specific action or task.
- State (s): State refers to the current situation returned by the environment.
- Policy (π): It is a strategy which applies by the agent to decide the next action based onthe current state.
- Value (V): It is expected long-term return with discount, as compared to the short-term reward.
- Value Function: It specifies the value of a state that is the total amount of reward. It is an agent which should be expected beginning from that state.
- Model of the environment: This mimics the behavior of the environment. It helps you to make inferences to be made and also determine how the environment will behave.
- Model based methods: It is a method for solving reinforcement learning problems whichuse model-based methods.
- Q value or action value (Q): Q value is quite similar to value. The only difference betweenthe two is that it takes an additional parameter as a current action.

Here an agent is dyslexic and uses Reinforcement AI for reward

2.4. Reinforcement Learning to simplify sentences and improve relevance



Deep reinforcement learning simplification model. X is the complex sentence, Y the reference (simple) sentence and Yˆ the action sequence (simplification) produced by the encoder-decoder model.

deep reinforcement learning to simplify questions, like splitting complex questions and substitutes difficult words with common paraphrases

For learning sentence-level transformations, a model requires instances of original sentences and their corresponding simplified versions. In this section, we present the most commonly used resources.

2.5. Reinforcement Learning simplification examples

Main - Simple English Wikipedia

The Simple English Wikipedia (SEW)[1] is a version of the online English Wikipedia (EW)[2] primarily aimed at English learners, but which can also be beneficial for students, children, and adults with learning difficulties (Simple Wikipedia 2017b). With this purpose, articles in SEW use fewer words and simpler grammatical structures. For example, writers are encouraged to use the list of words of Basic English (Ogden 1930), which contains 850 words presumed to be sufficient for everyday life communication. Authors also have guidelines on how to create syntactically simple sentences by, for example, giving preference to the subject-verb-object order for their sentences, and avoiding compound sentences (Simple Wikipedia 2017a).

3. Dataset Analysis with Simplification Metrics

SARI (**S**ystem output **A**gainst **R**eferences and **I**nput sentence) was introduced by Xu et al. (2016) as a means to measure "how good" the words added, deleted, and kept by a simplification model are. This metric compares the output of an SS model against multiple simplification references and the original sentence.

The intuition behind SARI is to reward models for adding *n*-grams that occur in any of the references but not in the input, to reward keeping *n*-grams both in the output and in the references, and to reward not over-deleting *n*-grams. SARI is the arithmetic mean of *n*-gram precisions and recalls for add, keep, and delete; *the higher the final value, the better*. Xu et al. (2016) show that SARI correlates with human judgments of simplicity gain. As such, this metric has become the standard measure for evaluating and comparing SS models' outputs.

Considering a model output O, the input sentence I, references R, and $\#_g(\cdot)$ as a binary indicator of occurrence of n-grams g in a given set, we first calculate n-gram precision $p(n)$ and recall $r(n)$ for the three operations listed (add, keep, and delete):

$$p_{add}(n)=\sum_{g\in O}\min(\#_g(O\cap\bar{I}),\#_g(R))\sum_{g\in O}\#_g(O\cap\bar{I}),\#_g(O\cap\bar{I})=\max(\#_g(O)-\#_g(I),0)$$

$$r_{add}(n)=\sum_{g\in O}\min(\#_g(O\cap\bar{I}),\#_g(R))\sum_{g\in O}\#_g(R\cap\bar{I}),\#_g(R\cap\bar{I})=\max(\#_g(R)-\#_g(I),0)$$

$$padd(n)=\sum g\in O\min(\#g(O\cap\bar{I}),\#g(R))\sum g\in O\#g(O\cap\bar{I}),\#g(O\cap\bar{I})=\max(\#g(O)-\#g(I),0)$$

$$radd(n)=\sum g\in O\min(\#g(O\cap\bar{I}),\#g(R))\sum g\in O\#g(R\cap\bar{I}),\#g(R\cap\bar{I})=\max(\#g(R)-\#g(I),0)$$

$$p_{keep}(n)=\sum_{g\in I}\min(\#_g(I\cap O),\#_g(I\cap R'))\sum_{g\in I}\#_g(I\cap O),\#_g(I\cap O)=\min(\#_g(I),\#_g(O))$$

$$r_{keep}(n)=\sum_{g\in I}\min(\#_g(I\cap O),\#_g(I\cap R'))\sum_{g\in I}\#_g(I\cap R'),\#_g(I\cap R')=\min(\#_g(I),\#_g(R)/r)$$

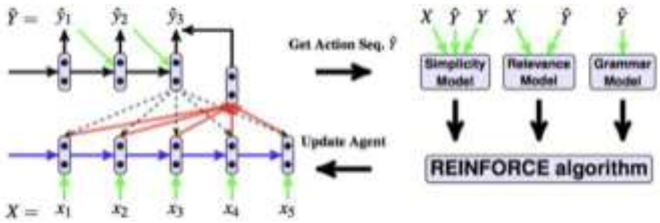
$$p_{del}(n)=\sum_{g\in I}\min(\#_g(I\cap\bar{O}),\#_g(I\cap\overline{R'}))\sum_{g\in I}\#_g(I\cap\bar{O})\ \#_g(I\cap\bar{O})=\max(\#_g(I)-\#_g(O),0)\ \#_g(I\cap\overline{R'})=\max(\#_g(I)-\#_g(R)/r,0)$$

3.1. Neural Sequence-to-Sequence

In this approach, SS is modeled as a sequence-to-sequence problem, and tackled normally with an attention-based encoder-decoder architecture. The encoder projects the source sentence into a set of continuous vector representations from which the decoder generates the target sentence. A major advantage of this approach is that it allows training of end-to-end models without needing to extract features or estimate individual model components, such as the language model. In addition, all simplification transformations can be learned simultaneously, instead of developing individual mechanisms as in previous research.

3.2. RNN for Sentence simplification

Without significantly changing the standard RNN-based architecture described before, some research has experimented with alternative learning algorithms with which the models are trained.



The agent reads the original sentence and takes a series of actions (words in the vocabulary) to generate the simplified output. After that, it receives a reward that scores the output according to its simplicity, relevance (meaning preservation), and fluency (grammaticality). To reward simplicity, they calculate SARI in both the expected direction and in reverse (using the output as reference, and the reference as output) to counteract the effect of having noisy data and a single reference; the reward is then the weighted sum of both values. To reward relevance, they compute the cosine similarity between the vector representations (obtained using a LSTM) of the source sentence and the predicted output. To reward fluency, they calculate the probability of the predicted output using an LSTM language model trained on simple sentences.

For learning, the authors used the REINFORCE algorithm whose goal is to find an agent that maximizes the expected reward. As such, the training loss is given by the negative expected reward:

$$L(\theta)=-E_{(\hat{y}_1,\ldots,\hat{y}_{|\hat{Y}|})\sim P_{RL}(\cdot|X)}[r(\hat{y}_1,\ldots,\hat{y}_{|\hat{Y}|})]L(\theta)=-E(\hat{y}1,\ldots,\hat{y}|\hat{Y}|)\sim PRL(\cdot|X)[r(\hat{y}1,\ldots,\hat{y}|\hat{Y}|)]$$
(21)

where P_{RL} is the policy, given in our case by the distribution produced by the encoder-decoder Equation and $r(\cdot)$ is the reward function.

An encoder-decoder is trained in a parallel original-simplified corpus to obtain probabilistic word alignments (attention scores α_t) that help determine whether a word should or should not be simplified. For these lexical simplifications to take context into consideration, they are integrated into the RL model using linear interpolation following Equation (22), where P_{LS} is theprobability of simplifying a word.

$$P(y_t|y_{1:t-1},X)=(1-\eta)P_{RL}(y_t|y_{1:t-1},X)+\eta P_{LS}(y_t|X,\alpha_t)$$

4. Results - Reinforcement Learning for Training Accuracy

The training accuracy increases with the time. It seems that previous experiences are helping the agent to increase the training accuracy.

5. Conclusion

We developed a reinforcement learning-based text simplification model, which can jointly model simplicity, grammaticality, and semantic fidelity to the input. We also proposed a lexical simplification component that further boosts performance. Overall, we find that reinforcement learning offers a great means to inject prior knowledge to the simplification task achieving good results across three datasets. In the future, we would like to explicitly model sentence splitting and simplify entire documents (rather than individual sentences). The live comparison between reading and school performance with and without using aides, gamification, technology was performed. The results were compared for improvement for learning disabilty, ADHD spectrum and also a case study extracted for schizophrenia (SZ) ,selective impairment was done. There should be lot of drastic changes in this field as the degree of learning disorder, sprctrum disorders are highly varying factors . The need to ensure meeting the learning needs with changing trands of natural language processing field in AI is the need and high priority task of researchers to aid the struggling human population. The motto should be "**Let all learn easier** ".

References

[1] Sentence Simplification with Deep Reinforcement Learning | dress (xingxingzhang.github.io)
[2] https://xingxingzhang.github.io/dress/
[3] Understanding BERT — Word Embeddings | by Dharti Dhami | Medium
[4] https://medium.com/@dhartidhami/understanding-bert-word-embeddings-7dc4d2ea54ca
[5] What is reinforcement learning? The *complete* guide - deepsense.ai
[6] https://deepsense.ai/what-is-reinforcement-learning-the-complete-guide/
[7] https://link.springer.com/article/10.1007/s10458-022-09548-8
[8] https://pubmed.ncbi.nlm.nih.gov/17300757/
[9] Deep Reinforcement Learning Training Accuracy - Stack Overflow
[10] https://stackoverflow.com/questions/56446182/deep-reinforcement-learning-training-accuracy https://www.sciencedirect.com/science/article/abs/pii/S0925231221015265
[11] Data-Driven Sentence Simplification: Survey and Benchmark | ComputationalLinguistics | MIT Press
[12] https://direct.mit.edu/coli/article/46/1/135/93384/Data-Driven-Sentence-Simplification-Survey-andShort Biography