The Lifecycle Change Model based on Wordle analysis and study

Tianchang Zou*, Muzi Li and Xinyue Li

Department of food science and engineering, Beijing Technology and Business University, Beijing, China

*Corresponding author: jerryzou07@sina.com

Abstract. Wordle is a popular game, and I did a detailed analysis of the Wordle dataset provided by the New York Times to better understand the game. First, we proposed a lifetime change model to describe the number of reported results over time, and through analysis we found that the property of "whether a word contains repeated letters" affects the fraction percentage of the difficult pattern. We then constructed a GloVe-LBP neural network to predict the reported percentage of a given solution and performed a detailed analysis of the model confidence, loss function, training adequacy, and generalization ability. Next, we constructed a K-Means and z-score classification model to classify the difficulty of words, through which the "EERIE" difficulty level was level 5. The relationship between the number of reported results and the number of difficult patterns in the data set was concluded and analyzed.

Keywords: Lifecycle Change Model; GLoVe-LBP Neural Network; K-means.

1. Introduction

Wordle Is a popular puzzle currently provided daily by the New York Times. The player tries to solve the puzzle by guessing a five-letter word six times or less, receiving feedback on each guess. For this version, every guess must be a real English word. Speculation that is not identified as text by the competition is not allowed. There are two total modes in the game: Hard mode and Normal mode. The Wordle note on the New York Times website states that the color of the tiles will change after you submit the text.

2. Lifecycle Change Model

2.1. Lifecycle Model Building

Lifecycle Model is a software development model used to describe the life cycle of a software project, that is, the entire process from development to release, maintenance and elimination. As a fast and short-lived game, Wordle can be described using a life cycle model, which mainly includes the following phases: Introduction Stage, Growth Stage, Maturity Stage, Decline Stage.

According to the above analysis, the growth of the number of reported results over time can be divided into 4 phases. Rising Phase, Falling Phase 1, Falling Phase 2, Falling Phase 3.

For the data of each stage, the exponential function is used to fit the data separately. The formula of the exponential function is as follows:

$$y = a_0 e^{-\frac{x}{a_1}} + a_2 \tag{1}$$

Where a_0, a_1, a_2 are the parameters to be optimized respectively. The parameter results of each stage are shown in the table below:

Table 1. Index parameter table

	a_0	a_1	a_2
Rising Phase	-416331.89	24.43	489396.96
Falling Phase1	165962059.87	47591.50	-165535549.46
Falling Phase2	636472.31	48.53	28631.22
Falling Phase3	70602288.52	636716.35	-70541791.71

The established lifecycle change model is shown in the figure below:

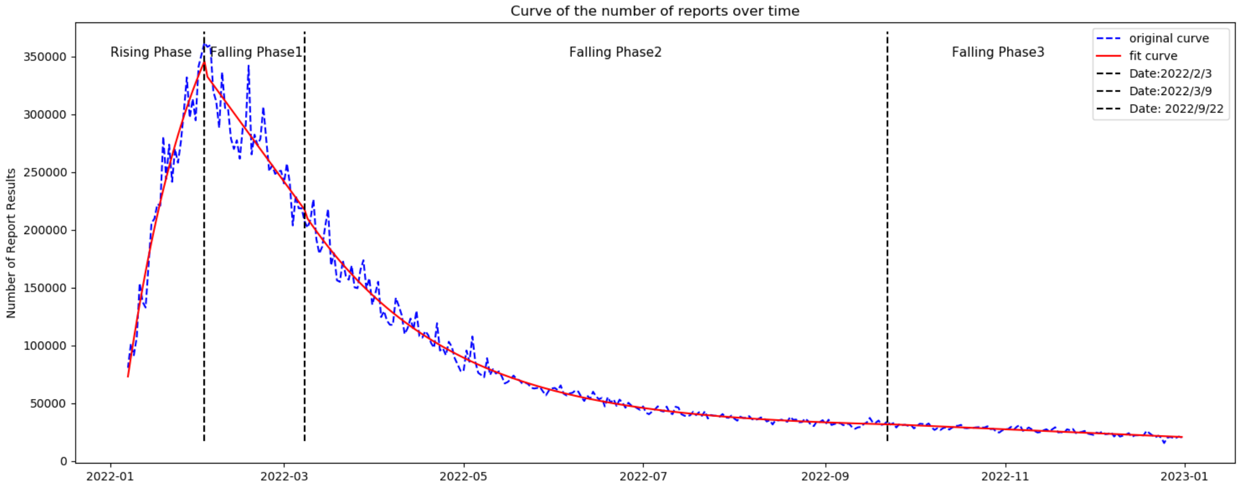

Fig. 1 Lifecycle change model

2.2. Effect of Attributes on Scores Reported in Hardmode

When we tried the Wordle game, we found that if there are repeated letters in the word, it may affect the number of our attempts, so we divide the words in the data set into two attributes according to whether there are repeated letters in the word: Repeated letter words and No repeated letter words. The classification is shown in the table below.

Table 2. Attribute Table

Attribute	Number
Repeated letters	100
No repeated letters	256

We count the percentage of words that report scores under this attribute, as shown in the table below.

Table 3. Attribute Table

	Try 1	Try 2	Try 3	Try 4	Try 5	Try 6	More
Repeat letters	0.003439	0.036143	0.175189	0.31618	0.275006	0.154102	0.039861
No repeat letters	0.007063	0.069822	0.244761	0.323144	0.22051	0.1089	0.02717

In order to analyze the above table more intuitively, we have drawn the following figure.

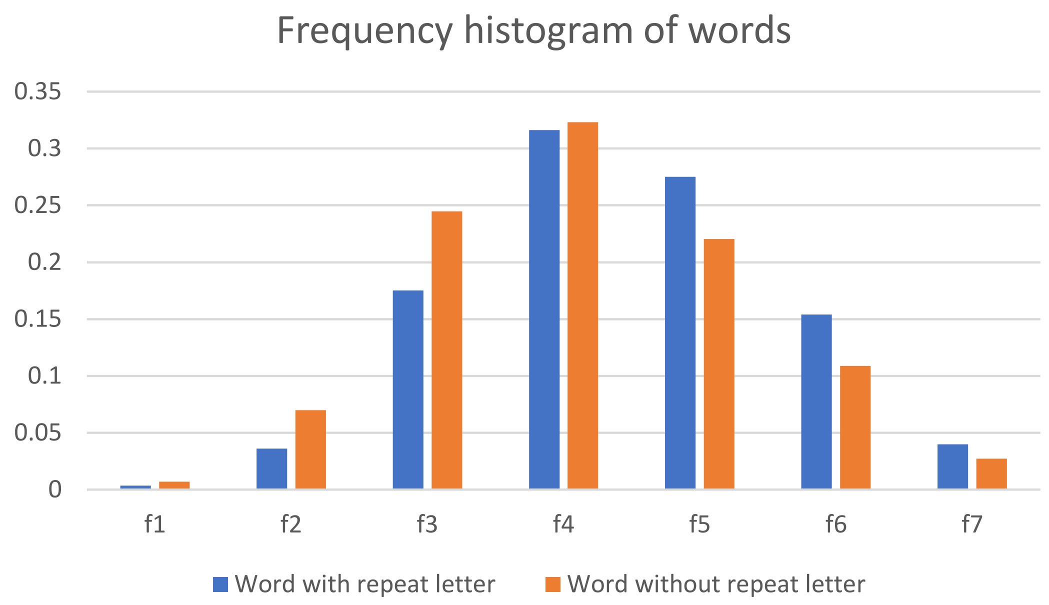

Fig. 2 Predict the number of participants image

It can be seen from the figure below that when the number of attempts is 1-4, the proportion of words without repeated letters is higher, indicating that words without repeated letters are easier to guess; when the number of attempts is 5-7, words with repeated letters A higher proportion of words means that it is more difficult to guess words that contain repeated letters.

To sum up, we believe that "whether a word contains repeated letters" has an impact on the reported score percentage in hard mode. When the word contains repeated letters, the difficulty factor of the game is higher for the player, the percentage of trying 5-7 times will be higher than when there are repeated letters in the word, and the percentage of trying 1-4 times will be higher than when there are repeated letters in the word Low.

3. GloVe-LBP Neural Network Prediction Model

3.1. GloVe Model

To better utilize the textual information, this paper uses the word embedding model GloVe to represent each word as a low-dimensional dense word vector. GloVe (Global Vectors) is a pre-trained word vector model based on matrix decomposition, which was developed by a research team at Stanford University.

The following is the algorithm flow of the GloVe Model:

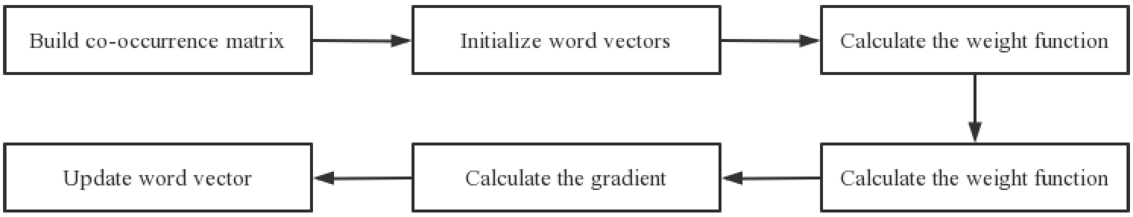

Fig. 3 The algorithm flow of the GloVe Model

3.2. Leaky BP Neural Network

Leaky BP (LBP) neural network is an improvement of the traditional BP neural network, mainly to solve the problem that the training process of traditional BP neural network tends to fall into local minima. The algorithm was originally proposed in 1994 and was invented by Jaakko Hollmen and Eduardo R. Caianiello. The LBP neural network used in this paper is shown below.

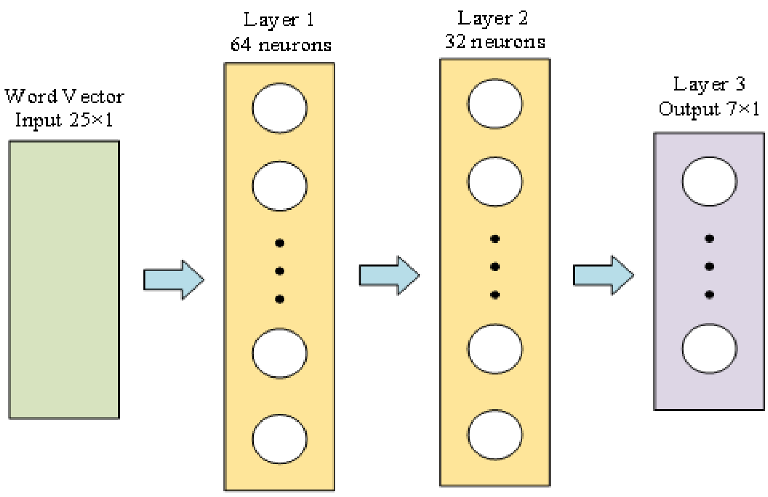

Fig. 4 LBP neural network

The term "Leaky" in Leaky BP neural network refers to the use of Leaky ReLU as the activation function which is shown below. Leaky ReLU is similar to the traditional ReLU activation function, but it does not completely output zero when the input is negative. Instead, it retains a small negative slope, which can prevent the "neuron death" phenomenon and thus improve the training efficiency and generalization performance of the network.

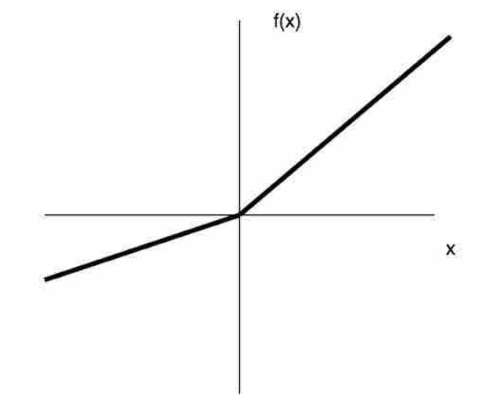

Fig. 5 Leaky Relu

LBP neural network usually consists of an input layer, one or more hidden layers,and an output layer. Each layer contains multiple neurons, and the neurons in adjacent layers are connected with weighted connections. The input layer receives input features, the output layer produces prediction results, and the hidden layers transform and extract the input features to generate more discriminative features. Each neuron receives the weighted input signals from the neurons in the previous layer and applies a nonlinear transformation through an activation function to produce the output of the neuron.

Before training, the 356 words in the wordle were effectively processed, and the GloVe model was used to convert the words into 25-dimensional word vectors, which serve as the input to the LBP neural network model. The output of the model is a 7-dimensional vector, and the loss function is calculated using the mean squared error and percentage labels (1, 2, 3, 4, 5, 6, X).

In order to address the issue of the label percentages not summing up to 100% in certain cases and improve the prediction performance of the LBP neural network model, this paper normalizes label into [0,1]. A random seed of 42 was used, and the train_test_split() function was used to split 80% of the data into the training set and the remaining 20% into the test set.

3.3. Prediction Analysis

The prediction results of "EERIE" on March 1, 2023 using the trained LBP neural network model are shown in the figure below, and the correlation percentages of (1, 2, 3, 4, 5, 6, X) are (0%, 4%, 11%, 28%, 27%, 18%, 12%).

Table 4. Prediction result

	1 try	2 tries	3 tries	4 tries	5 tries	6 tries	7 or more tries (X)
EERIE	0	4	11	28	27	18	12

This result can be considered scientifically reasonable, mainly due to the following points:

3.3.1. Powerful fitting ability

BP neural network has strong fitting ability, and LBP neural network as an improved model of BP neural network further enhances the fitting ability of the model. The model selected in this paper has 3 layers, the first layer has 64 nodes, the second layer has 32 nodes, and the third layer has 7 nodes, a total of 3975 nodes, which can meet the percentage prediction of only 356 samples.

3.3.2. Reasonable loss function

Introduce the MSE mean square error and put it into the formula.

3.3.3. Adequate training

In order to improve the training effect of the LBP neural network and avoid overfitting, this paper sets a smaller learning rate of 0.001 and a larger number of training rounds of 10,000 rounds. After sufficient training, the loss of the model in the training set and test set is shown in the figure below. The LBP neural network model can predict the relative percentage of the word "EERIE" well.

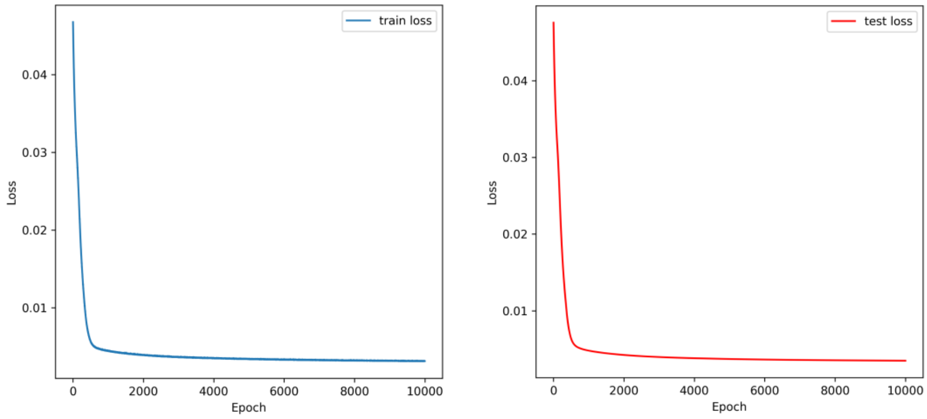

Fig. 6 Loss diagram

3.3.4. Excellent generalization ability

Due to the use of Leaky Relu function and Dropout technology, the model has excellent generalization ability. After 10,000 rounds of training, the gap between Train loss and Test loss is only 0.0003, as shown in the figure below. Therefore, it is certain that the LBP model can more accurately predict the percentage of the word "EERIE" in the non-dataset.

```
Epoch [9980/10000], Train Loss: 0.0032, Test Loss: 0.0035
Epoch [9981/10000], Train Loss: 0.0032, Test Loss: 0.0035
Epoch [9982/10000], Train Loss: 0.0032, Test Loss: 0.0035
Epoch [9983/10000], Train Loss: 0.0032, Test Loss: 0.0035
Epoch [9984/10000], Train Loss: 0.0032, Test Loss: 0.0035
Epoch [9985/10000], Train Loss: 0.0032, Test Loss: 0.0035
Epoch [9986/10000], Train Loss: 0.0032, Test Loss: 0.0035
Epoch [9987/10000], Train Loss: 0.0031, Test Loss: 0.0035
Epoch [9988/10000], Train Loss: 0.0032, Test Loss: 0.0035
Epoch [9989/10000], Train Loss: 0.0031, Test Loss: 0.0035
Epoch [9990/10000], Train Loss: 0.0032, Test Loss: 0.0035
Epoch [9991/10000], Train Loss: 0.0032, Test Loss: 0.0035
Epoch [9992/10000], Train Loss: 0.0032, Test Loss: 0.0035
Epoch [9993/10000], Train Loss: 0.0031, Test Loss: 0.0035
Epoch [9994/10000], Train Loss: 0.0032, Test Loss: 0.0035
Epoch [9995/10000], Train Loss: 0.0031, Test Loss: 0.0035
Epoch [9996/10000], Train Loss: 0.0031, Test Loss: 0.0035
Epoch [9997/10000], Train Loss: 0.0032, Test Loss: 0.0035
Epoch [9998/10000], Train Loss: 0.0032, Test Loss: 0.0035
Epoch [9999/10000], Train Loss: 0.0032, Test Loss: 0.0035
Epoch [10000/10000], Train Loss: 0.0032, Test Loss: 0.0035
```

Fig. 7 Gap between Train loss and Test loss

4. K-Means and Z-score Classification Model

4.1. Calculate Evaluate Grade

For each word, calculate the ratio of the number of guesses to the total number of attempts to get the percentage of attempts. This can reflect the difficulty of guessing the word.

$$Grade = \frac{P(t1)+P(t2)+P(t3)+P(t4)+P(t5)+P(t6)}{P(t1)+2P(t2)+3P(t3)+4P(t4)+5P(t5)+6P(t6)+P(t7)} \tag{2}$$

The smaller the calculated score, the more difficult the word is.

4.2. Z-score

After calculating the data, it is found that the difference between the data scores is too small to accurately reflect the difficulty of the word. Therefore, we consider normalizing the data.

The z-score normalization formula is as follows:

$$z = \frac{x-\mu}{\sigma}, \mu = \frac{\sum_{i=1}^{n} X_i}{n}, \sigma = \sqrt{\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{n-1}} \tag{3}$$

Where x is the original data, μ is the mean of the original data, σ is the standard deviation of the overall sample space and z is the normalizing data.

After z-score, each data point in this dataset was transformed to a value in a standard normal distribution, allowing for comparable analysis. The smaller the median score, the harder it is to prove the word.

4.3. K-Means clustering

4.3.1. K-Means Algorithm

K-Means is a clustering algorithm. The basic idea is to iteratively update the center point of the cluster until it reaches a convergent state. The basic steps of the K-Means algorithm are as follows:

Fig. 8 K-Means Algorithm

4.3.2. Classification Result

Words are divided into several difficulty levels according to the standardized score of the percentage of attempts.

The classification results obtained by the K-means algorithm are shown in the figure below.The classification situation is reflected in scatter plot and pie chart respectively.

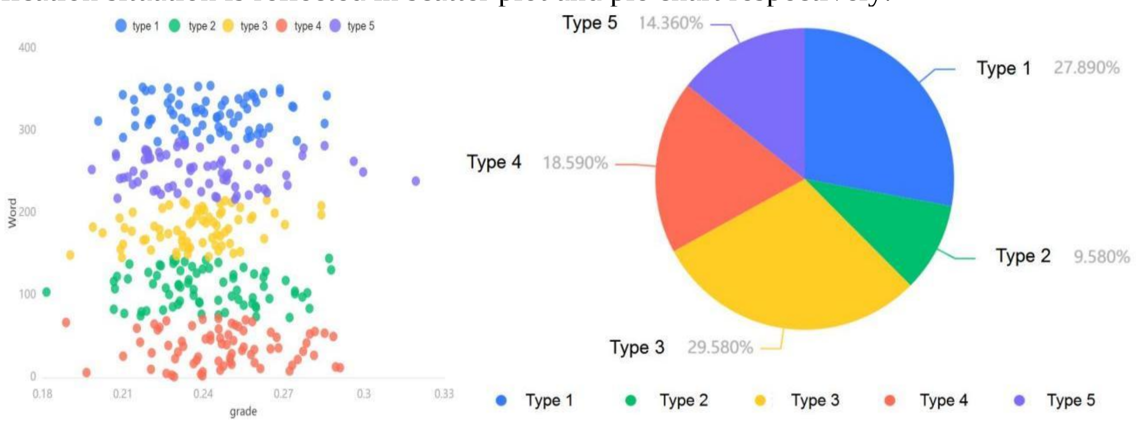

Fig. 9 Classification result map

The words are divided into five difficulty levels. The percentages for the five classes are: 27.890%, 9.580%, 29.580%, 18.590%, 14.360%.

According to the results, the difficulty is analyzed. see table below:

Table 5. Central Score Table

Type	central score
1	-0.6177378273838383
2	1.9111710187941175
3	0.1470813933428572
4	0.8597771091666666
5	-1.490443239941177

The lower the score, the more difficult the word. According to the above table, we divide the difficulty into five categories, and the order of installation from easy to difficult is: Level 1 (Type 2), Level 2 (Type 4), Level 3 (Type 3), Level 4(Type 1) and Level 5 (Type 5).

4.3.3. Judging ‘EERIE’ Difficulty

The title asks us to predict the difficulty of ‘EERIE’. According to the established classification model, the difficulty range table shown in the following table is obtained.

Table 6. Difficulty range table

Mean± Standard Deviation				
Type 3	Type 1	Type 4	Type 5	Type 2
0.147±0.209	-0.618±0.237	0.86±0.231	-1.49±0.384	1.911±0.455

According to the prediction results about ‘EERIE’ obtained in 6.3, substitute into the classification model. The calculated grade of the word is 0.1880341, and the standardized result is -2.5314881. According to Table 7.2, the word belongs to Tpye 5, and the difficulty is Level 5.

5. Interesting Features of Data Set

Analyze and compare the Number in hard mode and Number of reported results in the wordle data set, draw Hard Mode Users Proportions (Number in hard mode / Number of reported results) and Number of reported results line chart, as shown in the figure below, you can find interesting feature. It can be found that in the game life cycle of wordle, Hard Mode Users Proportions, that is, Number in hard mode / Number of reported results, grows with the growth of Number of reported results, and can still be maintained at 8-10% during the recession period. Analyzing the figure below, the following characteristics can be found:

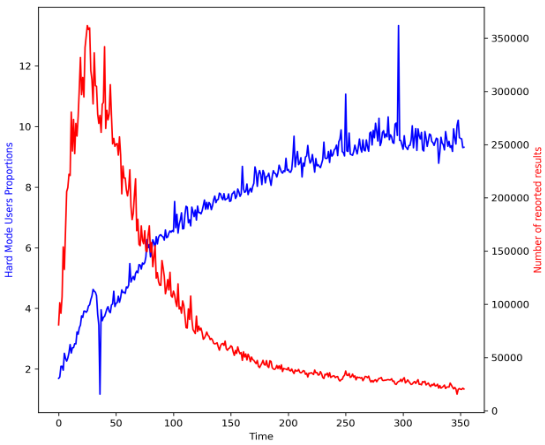

Fig. 10 Percentage Change Chart of Hard Mode Users to Reported Users

As the Number of reported results increases, Hard Mode Users Proportions also increase. The reasonable explanation is that with the increase of the total number of players in the game, some users are willing to spend more time and energy to challenge the hard mode after getting familiar with the wordle rules, so as to obtain better self-feedback and rewards. In addition, it is also possible that the hard mode is more attractive and challenging, causing some users to switch to the hard mode.

Hard Mode Users Proportions stabilized between 8-10%, with no significant growth. The possible explanation is that when the proportion of hard mode users reaches about 10%, more users may choose to give up this mode, or no longer play the game. It may also be because the difficulty of hard mode is higher, which is beyond the ability of most users.

During the recession period of the wordle game, the Number of reported results dropped significantly, but the Hard Mode Users Proportions remained between 8-10%. The possible reason is that hard mode users are more loyal to wordle, and they are more willing to participate in wordle to improve their abilities and challenge their limits. Therefore, when the total number of users drops, the percentage of hard mode users remains the same, probably because they are more inclined to engage in the game for a long time, and it is more difficult to switch to try other games.

6. Model Evaluation

6.1. Strengths

The Lifecycle Model divides the number of report results into 4 segments and fit them respectively. The model can deeply explain the mechanism of the number of report results varying with time, and achieve high-precision forecasting at the same time.

The Glove-LBP neural network is used to establish the prediction model, and the performance of the model is fully evaluated. The model has excellent performance in prediction accuracy and robustness.

“Attempt Percentage” is proposed, which is defined as the ratio of the number of guesses to the total number of attempts for every word. This definition is combined with the K-means model, which has a good effect in dividing the difficulty level of words.

6.2. Weaknesses

In the Glove-LBP classification model, only the uncertainty factors affecting the results are qualitatively analyzed, and the uncertainty is not quantitatively measured. This may be because the neural network is a black box model, and it is difficult to quantitatively analyze its uncertainty.

References

[1] LIU H, LI S, HUANG X. Adaptive surrogate model coupled with stochastic configuration network strategies for time-dependent reliability assessment [J]. Probabilistic Engineering Mechanics, 2023, 71.

[2] SOLAK S, ALTINISIK U. A new method for classifying nuts using image processing and k-means plus plus clustering [J]. Journal of Food Process Engineering, 2018, 41(7).

[3] WEI X, SHANG J. Mobile value chain collaboration for product diffusion: Role of the lifecycle [J]. Expert Systems with Applications, 2023, 215.

[4] YU S-S, CHU S-W, WANG C-M, et al. Two improved k-means algorithms [J]. Applied Soft Computing, 2018, 68: 747-55.

[5] Kelly, B. F., & Laws, K. R. (2022). Wordle: A game to improve word retrieval in aphasia. Aphasiology, 36(1), 25-41. doi: 10.1080/02687038.2021.1926095

[6] Linsley, D., & Waring, R. (2021). Investigating the effectiveness of Wordle as a vocabulary teaching tool for EFL learners. Innovation in Language Learning and Teaching, 15(3), 288-300. doi: 10.1080/17501229.2019.1614439

[7] Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A K-Means Clustering Algorithm. Journal of the Royal Statistical Society. Series C (Applied Statistics), 28(1), 100–108. https://doi.org/10.2307/2346830

[8] Jain, A. K. (2010). Data clustering: 50 years beyond K-means. Pattern Recognition Letters, 31(8), 651–666. https://doi.org/10.1016/j.patrec.2009.09.011