ISSN 2616-5775 Vol. 6, Issue 8: 146-157, DOI: 10.25236/AJCIS.2023.060818

The application of intelligent algorithms in word discrimination

Zequan Yang[1,*,#], Xinyu Wu[2,#], Xiongfeng Zhang[2,#]

[1]*College of Mechanical Engineering, Chongqing Technology and Business University, Chongqing, 400067, China*
[2]*College of Computer Science and Information Engineering, Chongqing Technology and Business University, Chongqing, 400067, China*
[*]*Corresponding author: yzq_ctbu@163.com*
[#]*These authors contributed equally.*

Abstract: *Words play a huge role in people's communication and transmission of information. The LSTM model is first established in this paper to analyze the changing trend of the number of people in the time series of the data set. According to the model, the linear regression model was used to process the word characteristic values and put them into the least square model for fitting through linear regression, and the MAPE value was obtained, and the comparative test effect was conducted on the value. At the same time, the F statistic was used to test the significance of the regression equation, and the Prob value was obtained. After the comparison of standard values, word attributes did not affect the percentage of the number of people registered in the difficult mode. The characteristic values of 5 words were divided by analysis and input into LR linear regression, XGB, random forest, GA-BP neural network, and Bayesian classifier models for training. It was found that the XGB determination coefficient of the simulated annealing model was 0.506. Finally, BP neural network learning based on a genetic algorithm is used to predict, and the percentage of correct answers to each word is subject to normal distribution results.*

Keywords: *Time Series Prediction, Prediction Model, Bayesian Classifier, Big Data*

1. Introduction

At present, the research status of words in China is mainly divided into the study of word memory, the principle of word memory, memory methods and influencing factors. Research on word learning strategies, learning strategies and effectiveness of word learners[1]. For example, the study of semantics, the study of the meaning of words[2], mainly the hierarchical structure, connection and development of concepts; The study of pragmatics, which studies the way, reason, purpose and effect of words used in the context of language achievement; The study of pronunciation, the study of phonetic symbols, phonemes, syllables, stress, intonation, accent[3].

2. Eigenvalue selection

By analyzing the difficulty of a word from the perspective of how to learn a word, we analyze the origin and structural features of the words [4-5]. Finally, we analyze the pronunciation, structure, usage, length and part of speech of the word, and divide it into the following five characteristics: 'the same word ',' part of speech', 'length', 'the same word of number', 'ordinary word'.

First, we extract the length of words. There is only a small part of words with 6 or 4 letters in the data set. Therefore, the classification of words in this model mainly focuses on the judgment of words with 5 letters. For word parts of speech judgment, the model we choose is the NLTK model in NLP natural language processing.

2.1 Maximum entropy model:

Maxent (maximum entropy model) is a criterion for learning probabilistic models. Figure 1 shows the use of this model. When learning probabilistic models, the model with the highest entropy among all possible models is the best model.

ISSN 2616-5775 Vol. 6, Issue 8: 146-157, DOI: 10.25236/AJCIS.2023.060818

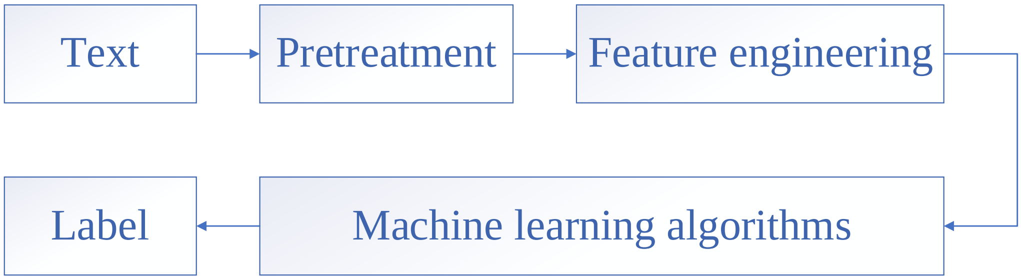

Figure 1: The NLTK model uses steps

At the same time, more concise and convenient categories for word use and labeling are given. The words used by Twitter users in online communication and the words in magazines that people read daily are summarized together and the table about each part of speech is drawn, for example, Table 1.

Table 1: Part of speech ratio

Part of speech	verb	noun	adjective	adverb	article	other
Quantity	1002	2543	839	839	11	237
Proportion	20%	51%	17%	7%	0.2%	5%

We analyze the structure of the words in the data set to find out whether the words are repeated and mark the number of times they are repeated. Often people tend to confuse words with more repeated letters, which increases the difficulty of words to some extent and leads to mistakes.

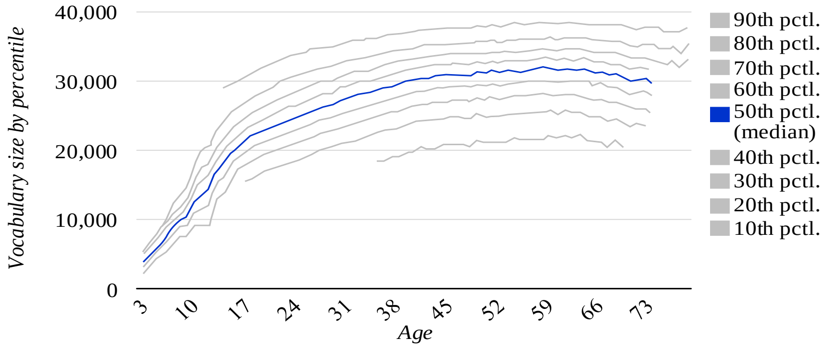

Figure 2: Relationship between age and word vocabulary

According to the famous Ebbinghaus memory curve [6]. The rate of forgetting increases first and then decreases with the passage of time. Therefore, we take whether the words are the words we use in daily life as a measurement index. In addition, we obtain the words people use in daily life from a website report, which contains 20,000 words. Here's a graph of age versus as Figure 2.

3. Basic model

3.1 Time series prediction

By analyzing the questions and interpreting the data, the prediction of the number of reported results on March 2, 2023, is a prediction of the future. We decide to use time series prediction after experimental screening. The statistical words and the number of people answering the questions every day are arranged into a series according to the time sequence, and the change process of the number of people is reflected by analyzing the time series, to predict the number of reported spellers after the next 60 days.

Secondly, data introduction and processing were carried out. The "Daily results from January 7, 2022, to December 31, 2022" file generated by MCM was adopted to process the simple time series in reverse order, and the data set was visualized. In order to accurately divide the prediction range, the average method could not be simply used for processing.

$$y_{\mathrm{i}} = \frac{1}{p}(y_{i-1} + y_{i-2} + y_{i-3} + \ldots + y_{y-p}) \tag{1}$$

ISSN 2616-5775 Vol. 6, Issue 8: 146-157, DOI: 10.25236/AJCIS.2023.060818

The average sliding window data of the last 10 days is used as the data for prediction. The data processing principle of smooth value. Every 10 days, the number of people who report scores in those 10 days is calculated, and the sliding average of the number of screened reports can be obtained.

$$\mathrm{y}_0=\sum_{i=1+day\%10*10}^{i+9}(day_1+day_2+...+\mathrm{day}_i) \tag{2}$$

Time series prediction of deep learning. First, the time series prediction was carried out, and the iteration times of the model were increased by through LSTM[7]. After the iteration was completed, the model verification was carried out. Since the eigenvalue is single, we need to carry out inverse normalization processing for data standardization when we carry out time series prediction:

$$x^{'}=x-\min(\max)-\min x^{''}=x^{'}\times(mx-mi)+mi \tag{3}$$

After processing, we used the performance evaluation Score: mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE), coefficient of determination (R-Squared) and mean absolute percentage error (MAPE) as five indexes to measure the quality of our algorithm.

3.2 Multiple linear regression

Our team obtained different characteristics of words through analysis: ‘the same word’, 'part of speech', 'length', 'the same word of number', and 'ordinary word', using these independent variables and dependent variable 'hard percent' for correlation analysis: A significant analysis was made on the percentage of the attributes of words that affected the number of applicants in the degree of difficulty, and the effect was observed.

$$b=\frac{\sum_{i=1}(x_i-\overline{x})(y_i-\overline{y})}{\sum_{i=1}^{n}(x_i-\overline{x})^2} \tag{4}$$

$$a=\overline{y}-b\overline{x} \tag{5}$$

$$\overline{y}=bx+a \tag{6}$$

$$\mathrm{y}=k_0+k_1x_1+k_2x_2+k_3x_3+... \tag{7}$$

Multiple linear regression model [8] can be expressed as the following formula (4). The linear regression model is established to achieve the fitting of the linear regression model, and the results of fitting and statistical analysis are returned.

Through the establishment of the above model, we can obtain the coefficient of the multiple linear regression model and the constant term of the regression model. The formula (8) of the model is as follows:

$$\begin{gathered}y=0.127\times ordinary\text{ word + }0.001\times the\text{ same word of number}\\ \text{- }0.011\times\text{len-}0.001\times\text{part of speech + }0.0\times\text{the same word}\end{gathered} \tag{8}$$

3.3 Wrapper multiple output regression algorithm

When analyzing the problem, considering that the model to be built uses input variables to predict multiple output variables, we Established Multiple Input-Multiple Output regression model (Multiple input-multiple Output)[9].In the process of multi-output regression, the output depends on the input and is dependent on each other. Multi-output regression support can be added to any regressors using multi-output regressors. This strategy includes fitting each target into one regressor.

3.4 Heuristic algorithm optimization --Simulated Annealing Algorithm

A heuristic algorithm is based on the intuitive or empirical construction of the algorithm, Figure 3 shows the framework diagram of the heuristic algorithmin, the acceptable computing time and space to be solved in each instance of a combinatorial optimization problem to give a feasible solution[10], the feasible solution and the optimal solution cannot be predicted generally. This method can using experience, select methods that have already worked.

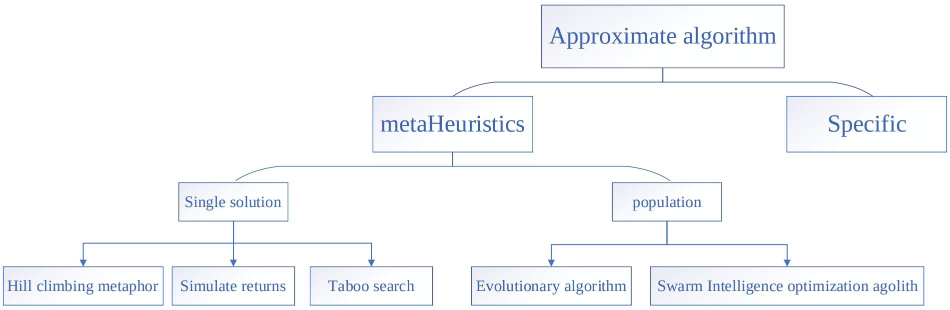

Figure 3: Heuristic algorithm block diagram

Below, our team adopts the simulated annealing algorithm of the heuristic algorithm to search for optimization[11], whose principle is the solid annealing principle. The solid is heated to a full height, and then cooled slowly.

Mathematical principle of simulated annealing acceptance probability:

$$p = \begin{cases} 1 & E_{t+1} > E_t \\ e^{\frac{-(E_{t+1}-E_t)}{kT}} & , E_{t+1} < E_t \end{cases} \tag{9}$$

Initialized first of all, this set of the annealing temperature $T = T_0$, randomly generated an initial solution X_0 and calculate due to the objective function value, setting of initial temperature is 100 ℃ and the termination temperature is 2 ℃; To set up the cooling coefficient, the $T = kT$, k value between 0 and 1, as the temperature drop rate, set the coefficient k is 0.98; Applying random disturbance, x_t again to the current solution in the neighborhood to create a new $x_t + 1$, and calculate the corresponding objective function values $E(x_t + 1)$

$$\Delta E = E(x_t + 1) - E(x_t) \tag{10}$$

If $\Delta E < 0$, accept the new as the current solution, otherwise the probability $e - \Delta E / kT$ determine whether to accept the new; Then set the iteration coefficient for each temperature.

Firstly, the random forest and XGB regression models were established to calculate the importance of features. The model assessment results for random forests are obtained as shown in Table 2.Then we get the model evaluation results in Table 3. Due to the randomness of XGB, the results of each operation are different. If this training model is saved, the subsequent data can be directly uploaded into this training model for calculation and prediction.

Table 2: Random Forest model evaluation results

	MSE	RMSE	MAE	MAPE	R²
Training Set	0.519	0.72	0.35	88.671	0.245
Test Set	0.481	0.694	0.562	58379995169617736	-0.88

Table 3: XGB model evaluation results

	MSE	RMSE	MAE	MAPE	R²
Training Set	8.356	2.891	1.834	31.912	0.506
Test Set	15.433	3.928	3.009	57.426	0.005

3.5 Bayesian Classifier

When we process the data, we find that the probability of answering times of participants in each word game follows a normal distribution, this is shown in Figure 4 so we extract the number of attempts to guess the word in the data set, and carry out visual processing on their number.

ISSN 2616-5775 Vol. 6, Issue 8: 146-157, DOI: 10.25236/AJCIS.2023.060818

Based on the probabilistic analysis of the correct time in these words, Figure 5 shows that the situation follows a normal distribution.

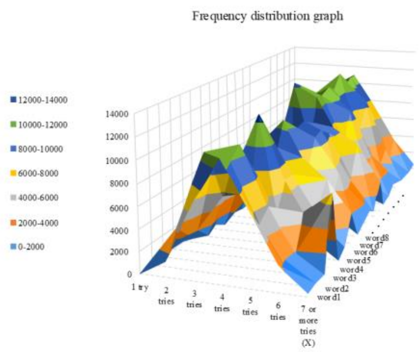

Figure 4: Visual processing of word attempts

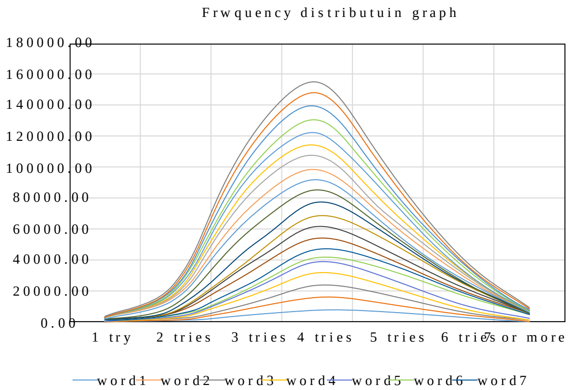

Figure 5: Frequency distribution graph

Therefore, we first need to make an objective and scientific classification of these words and assign labels to them (difficult or not difficult). We decide to use Bayesian classifiers in normal distribution for classification (probability density function)[12] :

$$p(x)=\frac{1}{\sqrt{2\pi\sigma}}e^{\left[-\frac{1}{2}(\frac{x-\mu}{\sigma})^2\right]} \tag{11}$$

For each word, the normal distribution function is used to calculate their corresponding mean and variance. In the case of normal distribution of each data, we integrate and update the original formula(10), Pdf (probability density function of multiple variables) :

$$p(x)=\frac{1}{(2\pi)^{1/2}\left|\sum\right|^{1/2}}\exp\left\{-\frac{1}{2}(x-\mu)^T\sum{}^{-1}(x-\mu)\right\},\quad x\in R^l \tag{12}$$

We assume that the correct number of times for each word follows the correct distribution. And the decision function we use is a function, which is easier to calculate and visually observe the performance of the model.

$$\mu=E\left[X\right]=E\left[x_1,x_2,\ldots,x_l\right] \tag{13}$$

$$\mu=E\left[X\right]=E\left[x_1,x_2,\ldots,x_l\right]=\begin{pmatrix}\sigma_{11}^2 & \cdots & \sigma_{1l}^2\\ \vdots & \ddots & \vdots\\ \sigma_{l1}^2 & \cdots & \sigma_{ll}^2\end{pmatrix} \tag{14}$$

It is assumed that class i is a difficult word and class j is a non-difficult word. According to the decision surface equation (17):

$$\begin{aligned}g_i(x)&=\ln\left[p(x\mid w_i)p(w_i)\right]\\&=\ln p(x\mid w_i)+\ln p(w_i)\end{aligned} \tag{15}$$

$$-\frac{1}{2}\left[(x-\mu_i)^T\sum{}_i^{-1}(x-\mu_i)-(x-\mu_j)^T\sum{}_j^{-1}(x-\mu_j)\right]-\frac{1}{2}\ln\frac{\left|\sum\right|}{\sum_j}+\ln\frac{p(w_i)}{p(w_j)}=0 \tag{16}$$

This function has the form of a closed quadric. In order to simplify the model and eliminate the influence of some noise, we add the following assumptions and introduce Euclidean distance:

ISSN 2616-5775 Vol. 6, Issue 8: 146-157, DOI: 10.25236/AJCIS.2023.060818

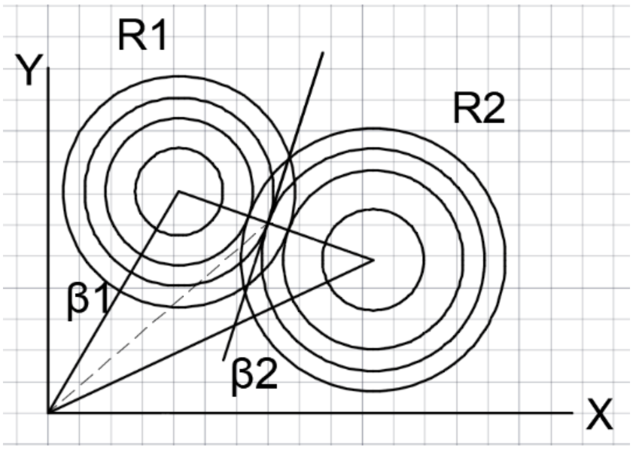

Figure 6: Bayesian classification of Mahala Nobis distance

The covariance matrix is assumed to have equal elements and to be diagonal. As shown in Figure 6, that is, the feature vectors are independent of each other and have equal variance. The Euclidean distances β1 and β2 are introduced and the above formula(15) is replaced by the formula(17). Presents a linear expression that is then substituted into formula (18).In this way, we can classify words to achieve the classification effect that we want.

$$g_i(x) = w_i^T x + w_{io} \tag{17}$$

$$g_{ij}(x) = g_i(x) - g_i(x) = w^T(x - x_0) \tag{18}$$

Finally, enter the label values and use the established model to classify the words in the data set. Constantly adjust the parameters and variance homogeneity test to get the best results.

3.6 GA-BP neural network

BP neural network[13] can learn and store a large number of input-output mode mappings. The learning rule is to use the fastest descent method. The weights and thresholds of the network are constantly adjusted by backpropagation to minimize the sum of squares of error.

We decided to introduce nonlinear function as excitation function. In this way, deep neural network expression ability is more powerful. Sigmoid (logistic), also known as S-shaped growth curves. Functions work better when used in classifiers.

$$f(x) = \frac{1}{1+e^x} \tag{19}$$

Forward propagation is the process of getting information from the input layer into the network. The process of calculating each layer successively to obtain the result of the final output layer.

$$\beta_j = \sum_{h=1}^{q} w_{hj} b_h + \theta_j \tag{20}$$

$$\mathrm{a}_h = \sum_{i=1}^{d} v_{ih} x_i + \theta_h \tag{21}$$

Parameter configuration of BP neural network as Table 4:

ISSN 2616-5775 Vol. 6, Issue 8: 146-157, DOI: 10.25236/AJCIS.2023.060818

Table 4: Activation function

module	build algorithms	involves formulas
genetic code	Floating-point encoding	None
Proportional Fitness Assignment	None	$fit_i = X - loss$
roulette wheel selection	Stochastic universal sampling	$P_i = \frac{fit_i}{\sum_{i-1}^{N} fit_i}$
Cross operator	Uniform crossover	$\begin{cases} X_A^{t'} = aX_B^t + (1-a)X_A^t \\ X_B^{t'} = aX_A^t + (1-a)X_B^t \end{cases}$
$fit_i = X - loss$ Mutation operator	Inversion mutation	$X_t' = U_{\min}' + \gamma \cdot (U_{\max}^t - U_{\min}^t)$
dataset	Proceed sequentially	None
The rest of the parameter settings	Artificial determination	None

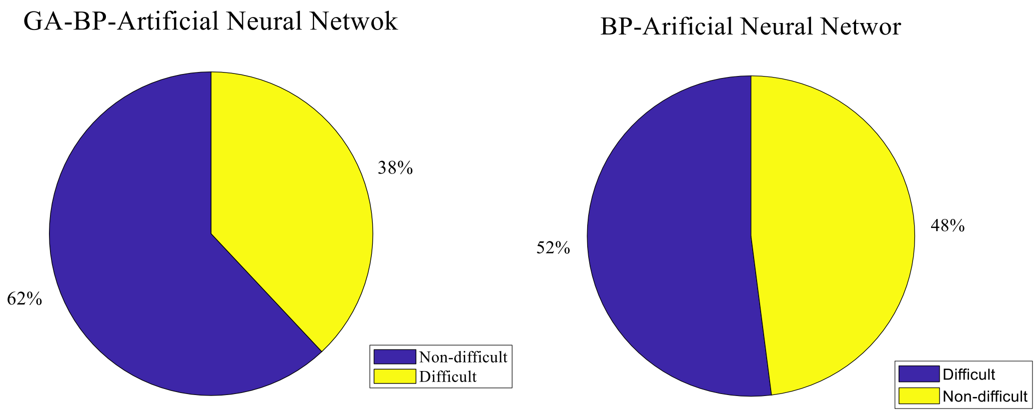

Figure 7: Comparison of neural network effects

Predict and generate classification results as shown in Figure 7. This is obviously compared with the BP neural network with genetic algorithm effect is better.

3.7 F1-score Predicted.

Binary evaluation was used, and accuracy rate, recall rate and accuracy rate were introduced to analyze and verify the reliability of the model. The confusion matrix was established, and TP and FP were defined as positive and negative predictive values respectively, as shown in Table 5:

Table 5: Positive and negative predictive values

True value / Predicted value	correct	negative
correct	TP	FN
negative	FP	TN

These values are substituted into F1-score for calculation. At the same time, the accuracy rate and recall rate of classification model can be taken into account. After several calculations, the score of our model is always close to 0.92.

4. Result analysis

We scientifically selected the following eigenvalues: 'the same 'word', 'partofspeech', 'length','thesamewordofnumber', 'ordinary word’. We took all the words in the dictionary and counted the number of times each letter appeared, as shown in Figure 8 and Figure 9. We found that there were a few letters that appeared most often in words and some letters that appeared almost never in words form

ISSN 2616-5775 Vol. 6, Issue 8: 146-157, DOI: 10.25236/AJCIS.2023.060818

the Table 6. Analyze the probability of occurrence in words and the probability of occurrence in the first letter.

Table 6: Letter difficulty

Easy letter	Difficult letter
a,e,s,t,h,i	z,q,x,j,k

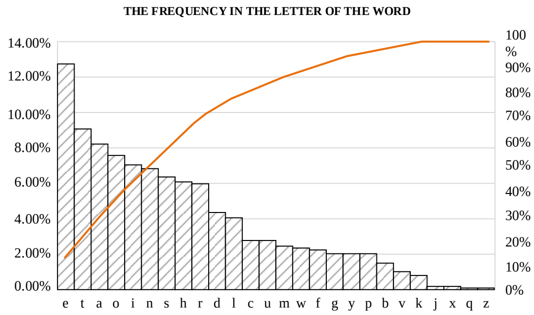

Figure 8: Probability of word occurrence

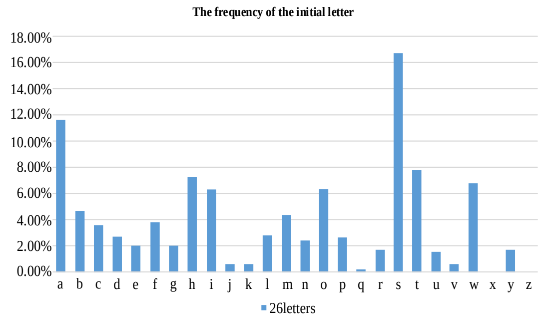

Figure 9: Probability of occurrences in the first letter

4.1 Autoregressive Integrated Moving Average Model

Use the five evaluation indicators in the table to represent the gap between forecast and reality. For root mean square error: As shown in Figure 10, we get a result of 25797.5230. The small value obtained indicates that the gap between the predicted value and the true value is smaller, and the model accuracy is higher. For the average absolute percentage error, a perfect model is indicated at 0%, The MAPE greater than 100% indicates an inferior model. The result we obtained is 8.6838%, which indicates that our model has less error and good fitting effect.

At the same time, as shown in Figure 10, we predict the number of reported results on March 2, 2023, and since we are using 10-day sliding data, we only need to predict 6 data, as shown in Table 7**.**

Table 7: Predict the outcome

Number of Days	January 10th	January 20th	January 30th	February 10th	February 20th	March 2nd
Outcome	22456.6	24805.9	26602.9	27986.2	29056.4	29887.2

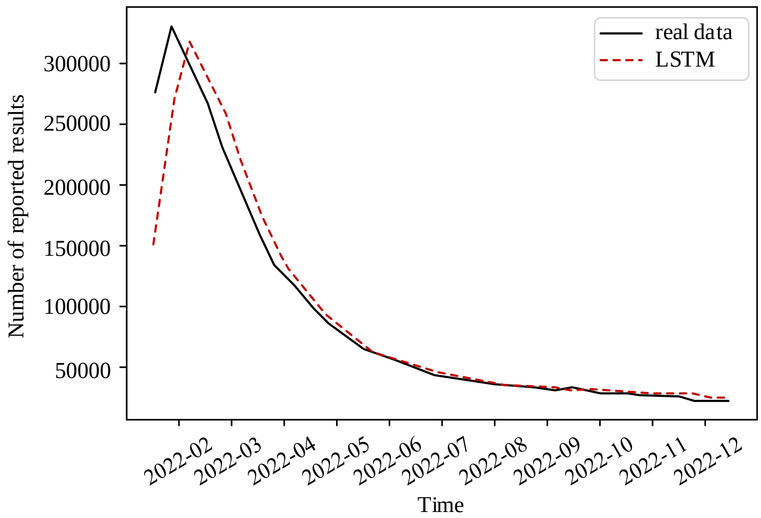

Figure 10: Fitting model training

4.2 Multiple linear regression

The F-statistic is used to test the significance of the overall regression equation. As shown in Table 8, the Prob value of the F test is 0.973, and its significance is greater than 0.05, indicating that the model is invalid and does not show significance. We can conclude that the percentage of difficulty is not

ISSN 2616-5775 Vol. 6, Issue 8: 146-157, DOI: 10.25236/AJCIS.2023.060818

significantly correlated with our dependent variable.

Table 8: OLS regression result

Dep. Variable	hard percent	R-squared	0.002
Model	OLS	Adj. R-squared	-0.012
Method	Least Squares	F-statistic	0.0240
Date	Sat, 18 Feb 2023	Prob (F-statistic)	0.976
Time	12:52:27	Log-Likelihood	562.50
No. Observations	359	AIC	-1117.
Date	355	BIC	-1101.

4.3 Correlation Analysis: Pearson Correlation Analysis

The correlation between the two pairs of data was calculated using Pearson correlation analysis: Firstly, the significant relationship between XY ($P<0.05$) was tested. The correlation coefficient was reanalyzed as positive and negative directions and the degree of correlation; The results of the analysis are summarized as shown in Figure 11.

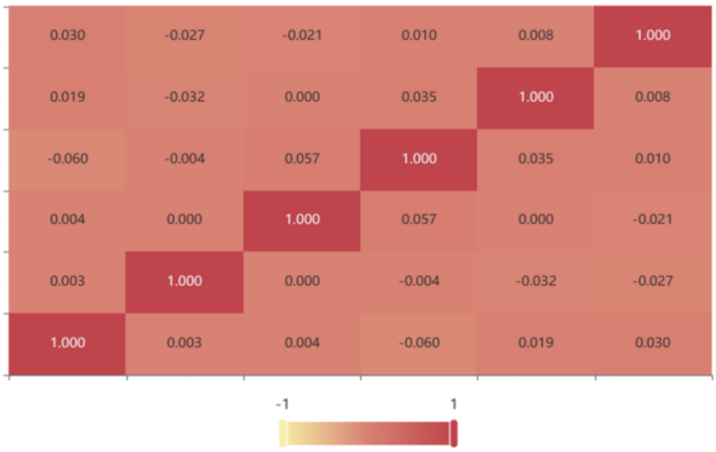

Figure 11: Correlation coefficient heat map

As can be seen from the coefficient table, the degree of correlation, the P value shows less significance.

4.4 Wrapper multi-output regression algorithm

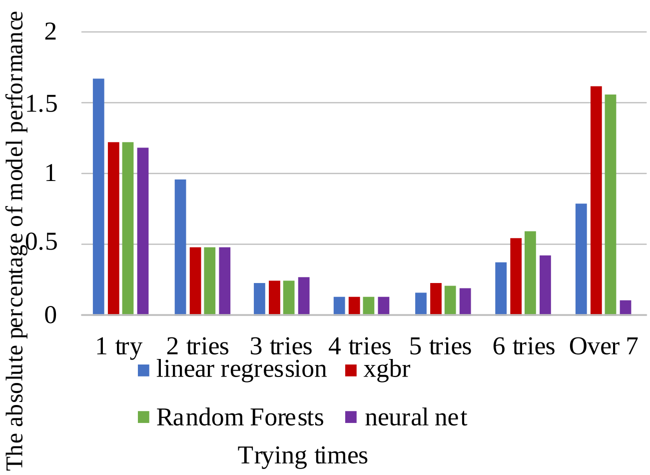

Figure 12: MAPE of model performance

ISSN 2616-5775 Vol. 6, Issue 8: 146-157, DOI: 10.25236/AJCIS.2023.060818

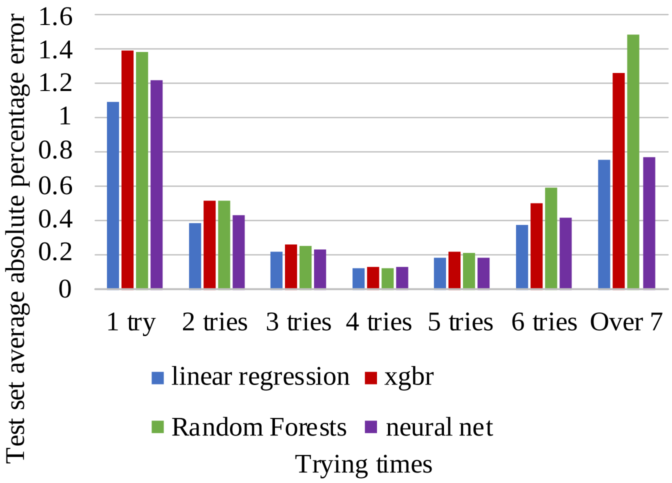

Figure 13: Test set average absolute percentage error

We find that the average absolute percentage error of the training set and the average absolute percentage error of the test set are different for different regressors, and we analyze the performance of each model in the training and test sets respectively, as shown in Figure 12 and Figure 13.

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}-y_i}{y_i}\right| \tag{22}$$

In the regressor we constructed, the Mean Absolute Percentage Error of the model evaluation index normalizes the error of each point, reducing the impact of absolute error caused by individual outliers. We use relative error to compare the accuracy of forecasts by various time series models. It is found that the multi-objective model training effect of LR combination is the best. The correlation percentage of the number of times the word EERIE was in the model on March 1, 2023, is shown in Table 9.

Table 9: Percentage of forecast

1 try	2 tries	3 tries	4 tries	5 tries	6 tries	Over 7
0	4	20	35	27	12	2

4.5 Heuristic algorithm optimization—simulate annealing algorithm.

According to the evaluation results of the above random forest and XGB models, the smaller the values taken by MSE, RMSE, MSE, MAE, and MAPE, the higher the accuracy of the model. R^2The closer the result is to 1 compared to using only the mean, the higher the accuracy of the model, and the better the model performs on the training set.

4.6 Bayesian classifier

Enter the label value into the Bayesian classifier and use the established model formula (17) to classify the words of the dataset.We can get the classification of a word's smallest loss when the sum of the six attempts and the probability of seven or more is between 42 and 40.Through the model results, we label all the words sequentially, with the word "Difficult" as 1 and the "non-difficult" word as 0, and then make a neural network prediction based on the genetic algorithm.

Consider the variance analysis, in order to analyze the fluctuation of the percentage of the score in hard mode, Perform a normality test Perform a Shapiro-Wilk test on the data to see significance: If it does not show significance ($P<0.05$), it is in line with a normal distribution. Secondly, one-sample T-TEST is carried out, and after passing the normality test, the one-sample T-TEST determines whether the P value shows significance ($P<0.05$). If significant, a difference analysis is performed based on the mean and the test value, describing the size of the difference, and the results are shown in Table 8.

4.7 GA-BP neural b network

After normalization and dimension lessness of all eigenvalues, we substitute data into BP neural networks for prediction. At the same time, considering that in the case of only data sets, we make the BP neural network overfit the result and fall into the local optimal solution case. Therefore, the BP algorithm is improved here, and the seed is randomly generated in the neural network when the initial value and threshold of the network can be determined in the BP neural network. This increases the uncertainty of the model. To obtain the optimal value in uncertainty, we introduce a genetic algorithm for the import of initial values and thresholds.

5. Conclusion

Through the results report generated by the Twitter game. We propose a series of models to prove the correlation between word features and the percentage of people reporting scores and the percentage in difficulty mode, Predict the percentage of future word relevance, and how easy it is to classify words.

We analyzed the change in the number of reports and created forecast intervals based on the time series model for the number of results reported on March 1, 2023; In addition, we also use linear regression model analysis to conclude: The Difficult mode property does not affect the percentage of the number of people enrolled in Hard mode. In this regard, we can get the user's enthusiasm for participation, which will decrease over time. Therefore, it is necessary to continuously improve the fun of the game and make more users participate.

Analyzing word difficulty and classifying words, our team uses a normally distributed Bayesian classifier to label each word, and the GA-BP neural network classification model for classification, discusses the difficulty of EERIE words. Try some of the more difficult words, which are more challenging and give users a sense of accomplishment.

We hope that these models can be a satisfactory tool for discovering the mathematical mysteries behind words. Life is full of mathematical wisdom, and we also need mathematics to decorate our lives. In addition, we recommend that while increasing the difficulty of words, you can appropriately increase the interest of memory, and add pictures, music, and other prompts. The selected words can be words that are popular or have popular scientific significance. If there are any further questions or problems with this model, please contact us and we will do our best to explain and improve the model.

References

[1] Hu Guangyao, Zhang Li. Research on English Word Memory Method Based on Roots and Affixes [J]. Curriculum and Teaching Methodology, 2022, 5(4).
[2] Natalia Diachok, Olena Kuvarova, Tetiana Vysotska, Svitlana Korotkova, Iryna Khurtak. A Word and a Word Combination in Language and Speech [J]. Studies in Media and Communication, 2022, 10(3).
[3] Dongbo Zhang, Chin-Hsi Lin, Yining Zhang, Yunjeong Choi. Pinyin or no pinyin: does access to word pronunciation matter in the assessment of Chinese learners' vocabulary knowledge? [J]. The Language Learning Journal, 2019, 47(3).
[4] Lou Zhengwei, Pan Zhigeng, Jin Haiyan, Zhang Beibei. A Method of Combination Word Recognition Based on Word Structure Features [J]. Twelfth International Conference On Graphics And Image Processing (ICGIP 2020), 2021, 11720.
[5] Mangi Kang, Jaelim Ahn, Kichun Lee. Opinion mining using ensemble text hidden Markov models for text classification [J]. Expert Systems with Applications, 2018, 94.
[6] Murre Jaap M J, Chessa Antonio G. Why Ebbinghaus' savings method from 1885 is a very 'pure' measure of memory performance. [J]. Psychonomic bulletin & review, 2022, 30(1).
[7] Xiaoli D. Media Ethics Anomie and Value Construction under Intelligent Algorithms[J].Journal of Huaihai Institute of Technology(Humanities & Social Sciences Edition), 2019.
[8] S. M. Saklain Galib, Sheikh Md. Rabiul Islam, Md. Asadur Rahman. A multiple linear regression model approach for two-class fNIR data classification [J]. Iran Journal of Computer Science, 2020, 4(1).
[9] Sharin Siti Nurhidayah, Radzali Mohamad Khairil, Sani Muhamad Shirwan Abdullah. A network analysis and support vector regression approaches for visualising and predicting the COVID-19 outbreak in Malaysia [J]. Healthcare Analytics, 2022, 2.
[10] Belkourchia Yassin, Es Sadek Mohamed Zeriab, Azrar Lahcen. New Hybrid Perturbed Projected Gradient and Simulated Annealing Algorithms for Global Optimization [J]. Journal of Optimization Theory and Applications, 2023, 197(2).
[11] Zhang Yan, Wang YaJun, Zhang Yong, Yu Tong. Photovoltaic Fuzzy Logical Control MPPT Based

ISSN 2616-5775 Vol. 6, Issue 8: 146-157, DOI: 10.25236/AJCIS.2023.060818

on Adaptive Genetic Simulated Annealing Algorithm-Optimized BP Neural Network [J]. Processes, 2022, 10(7).

[12] Joseph Awotunde, Sanjay Misra, Vikash Katta, Oluwafemi Adebayo. An Ensemble-Based Hotel Reviews System Using Naive Bayes Classifier [J]. Computer Modeling in Engineering & Sciences, 2023, 137(1).

[13] Jin Guangying, Feng Wei, Meng Qingpu. Prediction of Waterway Cargo Transportation Volume to Support Maritime Transportation Systems Based on GA-BP Neural Network Optimization [J]. Sustainability, 2022, 14(21).