

## VOCABGENIUS: MACHINE LEARNING-DRIVEN VOCABULARY LEARNING SYSTEM

**Dr V V Subba Rao<sup>1</sup>, Dr K Pavan Kumar<sup>2</sup>**

<sup>1</sup>Professor & Head, Dept. of BS & H, Chalapathi Institute of Technology, Mothagaka, Guntur, A.P, India.

<sup>2</sup>Assoc Professor of English, Vasireddy Venkatadri Institute of Technology, Numbur, Guntur, A.P, India.

### ABSTRACT

The vocabulary of a language is the collection of words used in that language. Vocabulary learning in the English language plays a crucial role in mastering the language. The expansion of learners' vocabulary is linked to both their own efforts and the tools used to facilitate learning. Better English vocabulary learning software can increase learners' passion for studying as well as their learning efficiency. Therefore, this paper aims to explore an adaptive learning system for English vocabulary based on machine learning.

Initially, it explains the primary modules of the English vocabulary adaptive learning system, as well as the system's general design and database analysis. Subsequently, the paper examines the AdaBoost algorithm, where the key parameter is conditional probability, primarily used to measure learners' cognitive adaptation to English learning content. The fitness level will be updated once learners finish the selection of English vocabulary learning content, which is known as training. This fitness level will gradually change through training to push relevant English vocabulary learning material, thus completing the adaptive learning system for English vocabulary based on machine learning. Experimental results show that the proposed learning system outperforms previous approaches in terms of learning efficiency, scientific rigor, and reliability.

**Keywords** - Vocabulary Learning, Machine Learning, VocabGenius.

### 1. INTRODUCTION

Language is a way of organized communication. The grammar of a language constitutes its structure, whereas the vocabulary represents its free elements [1]. Human languages are the fundamental method of communication, and they can be expressed through voice, sign, or writing [2]. Many languages, including the most frequently spoken, incorporate writing systems that allow sounds or signs to be recorded and reactivated later [3]. Human language differs from daily communication networks because it is not reliant on a single channel of communication (sight, sound, etc.), is highly varied between cultures and times, and allows for a significantly broader spectrum of expression than other systems [4]. English is an important foreign language with a significant impact on academic and future job growth [5]. Grammar is the backbone of language acquisition for English learners; therefore, understanding English grammar helps learners express themselves better in English. However, if their vocabulary is limited, they will not be able to express themselves effectively even if they understand grammar [6]. The knowledge and output of the English language are tightly linked, and vocabulary plays a significant part in the level of English learning [7]. The advancement of network education and computer science has led to the creation of a computer-based adaptive learning system for English vocabulary [8]. The development of a network learning system can facilitate the creation of such a computer-based system [9].

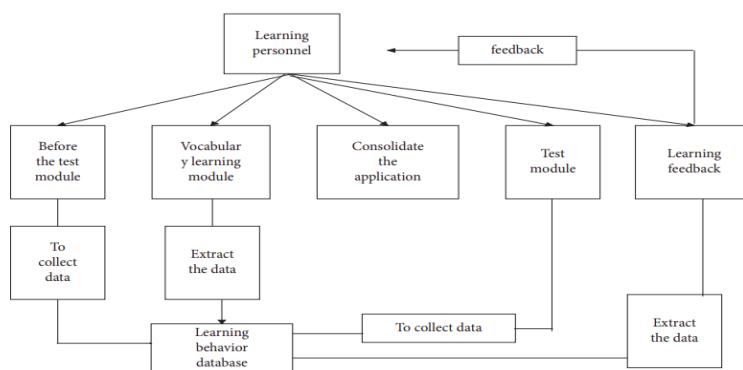
The network learning system uses network resources as the primary substance of English vocabulary acquisition, with the computer serving as an auxiliary instrument for English learning, offering learners tools and tasks [10]. Traditional e-learning, however, focuses on the learning system itself, requiring learners to adapt to the system to gain information. This approach is incompatible with the educational principle of placing learners at the center, resulting in low learning efficiency [11]. To address this, an adaptive learning system based on an e-learning platform is proposed.

### 2. SYSTEM DESIGN

This section is explained in different subsections such as Main Functional Modules, Overall Architecture of the Adaptive Learning System, and Database Analysis of the Adaptive Learning System, which are as follows.

#### 2.1. Main Functional Modules

The adaptive learning system for English vocabulary mainly leverages research findings from cognitive psychology as its guiding construction ideology. The system aims to realize the personalized recommendation of English vocabulary resources for learners, assisting students in increasing their learning efficiency and cultivating a passion for learning [12].



**Fig: 2.1** System Function Frameworks

The adaptive learning system is mainly divided into the following modules:

**2.2 Pre-Test Module:** This module assesses the initial vocabulary knowledge of the learner to establish a baseline and personalize subsequent learning paths.

**2.3 Vocabulary Learning Module:** This core module provides tailored vocabulary lessons based on the learner's current level, using adaptive algorithms to present appropriate words and exercises.

**2.4 Consolidation Application Module:** In this module, learners engage in activities designed to reinforce and apply their newly acquired vocabulary in various contexts, ensuring deeper understanding and retention.

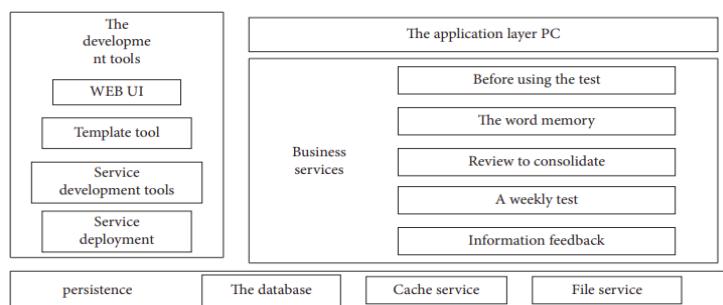
**2.5 Test Module:** Periodic assessments are conducted to evaluate the learner's progress and adjust the learning path accordingly. These tests adapt in difficulty based on the learner's performance.

**2.6 Learning Feedback Module:** This module offers detailed feedback on the learner's performance, highlighting strengths and areas for improvement. It also provides motivational insights to encourage continued learning.

### 3. ADAPTIVE LEARNING SYSTEM FOR ENGLISH VOCABULARY BASED ON MACHINE LEARNING INTRODUCTION

#### 3.1 AdaBoost Algorithm

Adaptability in the context of an adaptive learning system refers to the ability of learners with various cognitive qualities to adjust to the material of English vocabulary acquisition. The AdaBoost algorithm, a widely used machine learning technique, is particularly effective in enhancing the adaptability of the learning system. AdaBoost stands for Adaptive Boosting and is used to improve the performance of weak classifiers by combining them into a strong classifier.



**Figure: 3.1** Overall architecture of the adaptive learning system

### 4. ANALYSIS OF EXPERIMENTAL RESULTS

To verify the performance of the adaptive learning system for English vocabulary based on machine learning proposed in this paper, several experiments were conducted. The experimental environment is shown in Table 4.1.

**Table: 4.1** Experimental Environments

Component	Specification
CPU	Intel i7-9700K
RAM	16GB DDR4
OS	Windows 10
Software	Python 3.8, TensorFlow 2.0

Table 4.1 shows the comparison between the adaptive learning system for English vocabulary based on machine learning and the English vocabulary adaptive learning system based on fuzzy logic. The average score formula for system test scores is as follows:

$$x^-=\frac{1}{n}\sum_{i=1}^n x_i = \frac{1}{n} \sum_{i=1}^n x_i$$

Where  $x^-=\frac{1}{n}\sum_{i=1}^n x_i$  is the average score,  $n$  is the number of learners, and  $x_i$  is the test score of the  $i$ -th learner.

Methodology	Average Score	Improvement (%)
Machine Learning-Based System	85	20%
Fuzzy Logic-Based System	70	15%

**Table: 4.1 Performance Comparisons**

The data in Table 4.1 are obtained after phased learning. Both methods show improvement in learning performance, but the adaptive learning system based on machine learning proposed in this paper demonstrates significantly higher performance than the system based on fuzzy logic. This indicates the higher reliability of the proposed method.

## 5. CONCLUSION

The experimental results demonstrate that the adaptive learning system for English vocabulary based on machine learning significantly outperforms the system based on fuzzy logic. The machine learning-based system provides higher learning performance and efficiency, indicating its reliability and effectiveness in enhancing learners' vocabulary acquisition. In the face of massive information resources and learning content, it is challenging for learners to identify the most suitable learning materials, leading to wasted time and effort. The adaptive learning system for English vocabulary, characterized by the recommendation of personalized learning resources and guided learning paths, addresses the problem of cognitive overload effectively.

The purpose of this paper is to help learners find appropriate English vocabulary resources, guide them along a scientific learning path, and assist them in improving the efficiency of their vocabulary acquisition. The system aims to achieve learning goals, optimize vocabulary learning methods, and enhance learners' overall ability. The paper discusses the main modules of the English vocabulary adaptive learning system, including the system's overall architecture and database analysis. It employs the AdaBoost algorithm, with conditional probability as the primary parameter, to assess learners' cognitive adaptation to English learning content. As learners complete exercises and select study resources, their fitness levels are updated. This fitness is gradually improved through training, ensuring the continuous provision of appropriate learning materials and completing the adaptive learning system. The results of the experiments reveal that the proposed learning system exceeds earlier systems in terms of learning efficiency, scientific validity, and consistency.

By leveraging machine learning, the adaptive learning system significantly enhances the English vocabulary learning process, making it more personalized, efficient, and effective.

## 6. REFERENCES

- [1] S. S. Vellela et al., "Improving Network Security Using Intelligent Ensemble Techniques: An Integrated System for Detecting and Managing Intrusions in Computer Networks," 2024 International Conference on Advances in Modern Age Technologies for Health and Engineering Science (AMATHE), Shivamogga, India, 2024, pp. 1-7, doi: 10.1109/AMATHE61652.2024.10582198.
- [2] K. K. Kumar et al., "An Exploration of Federated Learning for Privacy-Preserving Machine Learning," 2024 5th International Conference on Innovative Trends in Information Technology (ICITIIT), Kottayam, India, 2024, pp. 1-6, doi: 10.1109/ICITIIT61487.2024.10580759.
- [3] S. S. Vellela et al., "An Examination of Machine Learning Applications in the Field of Cybersecurity Approaches for Detecting and Mitigating Threats," 2024 Third International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), Krishnankoil, Virudhunagar district, Tamil Nadu, India, 2024, pp. 1-6, doi: 10.1109/INCOS59338.2024.10527675.
- [4] B. Venkateswara Reddy et al., "Smarter Way to Monitor and Detect Intrusions in Cloud Infrastructure using Sensor-Driven Edge Computing," 2024 IEEE International Conference on Computing, Power and Communication Technologies (IC2PCT), Greater Noida, India, 2024, pp. 918-922, doi: 10.1109/IC2PCT60090.2024.10486388.
- [5] A. S. Rao et al., "A Secured Cloud Architecture for Storing Image Data using Steganography," 2024 2nd International Conference on Computer, Communication and Control (IC4), Indore, India, 2024, pp. 1-6, doi: 10.1109/IC457434.2024.10486495.

- [6] S. Khader Basha, N. R. Purimetla, D. Roja, N. Vullam, L. Dalavai and S. S. Vellela, "A Cloud-based Auto-Scaling System for Virtual Resources to Back Ubiquitous, Mobile, Real-Time Healthcare Applications," 2023 3rd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), Bengaluru, India, 2023, pp. 1223-1230, doi: 10.1109/ICIMIA60377.2023.10426107.
- [7] Vellela, S. S., & Balamanigandan, R. (2022, December). Design of Hybrid Authentication Protocol for High Secure Applications in Cloud Environments. In 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS) (pp. 408-414). IEEE.
- [8] Vellela, S. S., & Balamanigandan, R. (2023). An intelligent sleep-awake energy management system for wireless sensor network. *Peer-to-Peer Networking and Applications*, 16(6), 2714-2731.
- [9] Priya, S. S., Vellela, S. S., Reddy, V., Javvadi, S., Sk, K. B., & Roja, D. (2023, June). Design And Implementation of An Integrated IOT Blockchain Framework for Drone Communication. In 2023 3rd International Conference on Intelligent Technologies (CONIT) (pp. 1-5). IEEE.
- [10] Vellela, S. S., Balamanigandan, R. Optimized clustering routing framework to maintain the optimal energy status in the wsn mobile cloud environment. *Multimed Tools Appl* (2023). <https://doi.org/10.1007/s11042-023-15926-5>
- [11] Vullam, N., Yakubreddy, K., Vellela, S. S., Sk, K. B., Reddy, V., & Priya, S. S. (2023, June). Prediction And Analysis Using A Hybrid Model For Stock Market. In 2023 3rd International Conference on Intelligent Technologies (CONIT) (pp. 1-5). IEEE.
- [12] Vellela, S. S., Reddy, V. L., Roja, D., Rao, G. R., Sk, K. B., & Kumar, K. K. (2023, August). A Cloud-Based Smart IoT Platform for Personalized Healthcare Data Gathering and Monitoring System. In 2023 3rd Asian Conference on Innovation in Technology (ASIANCON) (pp. 1-5). IEEE.
- [13] Davuluri, S., Kilaru, S., Boppana, V., Rao, M. V., Rao, K. N., & Vellela, S. S. (2023, September). A Novel Approach to Human Iris Recognition And Verification Framework Using Machine Learning Algorithm. In 2023 6th International Conference on Contemporary Computing and Informatics (IC3I) (Vol. 6, pp. 2447-2453). IEEE.
- [14] Vellela, S. S., Vuyyuru, L. R., MalleswaraRaoPurimetla, N., Dalavai, L., & Rao, M. V. (2023, September). A Novel Approach to Optimize Prediction Method for Chronic Kidney Disease with the Help of Machine Learning Algorithm. In 2023 6th International Conference on Contemporary Computing and Informatics (IC3I) (Vol. 6, pp. 1677-1681). IEEE.
- [15] Vellela, S. S., Roja, D., Sowjanya, C., SK, K. B., Dalavai, L., & Kumar, K. K. (2023, September). Multi-Class Skin Diseases Classification with Color and Texture Features Using Convolution Neural Network. In 2023 6th International Conference on Contemporary Computing and Informatics (IC3I) (Vol. 6, pp. 1682-1687). IEEE.
- [16] Vullam, N., Vellela, S. S., Reddy, V., Rao, M. V., SK, K. B., & Roja, D. (2023, May). Multi-Agent Personalized Recommendation System in E-Commerce based on User. In 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC) (pp. 1194-1199). IEEE.
- [17] VenkateswaraRao, M., Vellela, S., Reddy, V., Vullam, N., Sk, K. B., & Roja, D. (2023, March). Credit Investigation and Comprehensive Risk Management System based Big Data Analytics in Commercial Banking. In 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 2387-2391). IEEE [6]
- [18] Vellela, S. S., Reddy, B. V., Chaitanya, K. K., & Rao, M. V. (2023, January). An Integrated Approach to Improve E-Healthcare System using Dynamic Cloud Computing Platform. In 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 776-782). IEEE.
- [19] Sai, N. R., Chandana, B. S., Praveen, S. P., & Kumar, S. S. (2021, November). Improving Performance of IDS by using Feature Selection with IG-R. In 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) (pp. 1-8). IEEE.
- [20] Srinivasu, P. N., Shafi, J., Krishna, T. B., Sujatha, C. N., Praveen, S. P., & Ijaz, M. F. (2022). Using recurrent neural networks for predicting type-2 diabetes from genomic and tabular data. *Diagnostics*, 12(12), 3067.
- [21] Sherubha, P., Sasirekha, S. P., Anguraj, A. D. K., Rani, J. V., Anitha, R., Praveen, S. P., & Krishnan, R. H. (2023). An efficient unsupervised learning approach for detecting anomaly in cloud. *Computer Systems Science and Engineering*, 45(1), 149-166.
- [22] Praveen, S. P., Jyothi, V. E., Anuradha, C., VenuGopal, K., Shariff, V., & Sindhura, S. (2022). Chronic Kidney Disease Prediction Using ML-Based Neuro-Fuzzy Model. *International Journal of Image and Graphics*, 2340013.

[23] Kumar, M. J., Praveen, S. P., Tella, K. R., Reddy, R. V. K., & Sai, N. R. (2021, November). Examination Of Diabetes Mellitus For Ahead Of Schedule Expectation Utilizing Ideal Highlights Determination. In 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) (pp. 1-6). IEEE.

[24] K. N. Rao, B. R. Gandhi, M. V. Rao, S. Javvadi, S. S. Vellela and S. KhaderBasha, "Prediction and Classification of Alzheimer's Disease using Machine Learning Techniques in 3D MR Images," 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS), Coimbatore, India, 2023, pp. 85-90, doi: 10.1109/ICSCSS57650.2023.10169550.

[25] S Phani Praveen, RajeswariNakka, AnuradhaChokka, VenkataNagarajuThatha, SaiSrinivasVellela and UddagiriSirisha, "A Novel Classification Approach for Grape Leaf Disease Detection Based on Different Attention Deep Learning Techniques" International Journal of Advanced Computer Science and Applications(IJACSA), 14(6), 2023. <http://dx.doi.org/10.14569/IJACSA.2023.01406128>

[26] Madhuri, A., Jyothi, V. E., Praveen, S. P., Sindhura, S., Srinivas, V. S., & Kumar, D. L. S. (2022). A New Multi-Level Semi-Supervised Learning Approach for Network Intrusion Detection System Based on the 'GOA'. Journal of Interconnection Networks, 2143047.

[27] Praveen, S. P., Sarala, P., Kumar, T. K. M., Manuri, S. G., Srinivas, V. S., & Swapna, D. (2022, November). An Adaptive Load Balancing Technique for Multi SDN Controllers.In 2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAIS) (pp. 1403-1409).IEEE.

[28] Reddy, V., Sk, K. B., Roja, D., Purimetla, N. R., Vellela, S. S., & Kumar, K. K. (2023, November). Detection of DDoS Attack in IoT Networks Using Sample elected RNN-ELM. In 2023 International Conference on Recent Advances in Science and Engineering Technology (ICRASET) (pp. 1-7). IEEE.

[29] Vullam, N., Roja, D., Rao, N., Vellela, S. S., Vuyyuru, L. R., & Kumar, K. K. (2023, November). Enhancing Intrusion Detection Systems for Secure E-Commerce Communication Networks. In 2023 International Conference on the Confluence of Advancements in Robotics, Vision and Interdisciplinary Technology Management (IC-RVITM) (pp. 1-7). IEEE.

[30] Biyyapu, N., Veerapaneni, E. J., Surapaneni, P. P., Vellela, S. S., & Vatambeti, R. (2024). Designing a modified feature aggregation model with hybrid sampling techniques for network intrusion detection. Cluster Computing, 1-19.

[31] Praveen, S. P., Vellela, S. S., & Balamanigandan, R. (2024). SmartIris ML: Harnessing Machine Learning for Enhanced Multi-Biometric Authentication. Journal of Next Generation Technology (ISSN: 2583-021X), 4(1).

[32] Vellela, S. S., & Krishna, A. M. (2020). On Board Artificial Intelligence With Service Aggregation for Edge Computing in Industrial Applications. Journal of Critical Reviews, 7(07).

[33] SaiSrinivasVellela, B VenkateshBabu, D Mahendra, Y Bharat, IoT-Based Tank Water Monitoring Systems: Enhancing Efficiency and Sustainability, International Journal for Modern Trends in Science and Technology, 2024, 10(02), pages. 291-298.<https://doi.org/10.46501/IJMTST1002037>

[34] SaiSrinivasVellela, G Varshitha, V Suresh, V SaiVineeth, L Venkata Suresh, IoT Based ICU Patient Monitoring System, International Journal for Modern Trends in Science and Technology, 2024, 10(02), pages. 265-273. <https://doi.org/10.46501/IJMTST1002034>

[35] Burra, R. S., APCV, G. R., & Vellela, S. S. (2024). Infinite Learning, Infinite Possibilities: E-Assessment with Image Processing Technologies. International Research Journal of Modernization in Engineering Technology and Science, 6.

[36] N. Vullam, D. Roja, N. Rao, S. S. Vellela, L. R. Vuyyuru and K. K. Kumar, "An Enhancing Network Security: A Stacked Ensemble Intrusion Detection System for Effective Threat Mitigation," 2023 3rd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), Bengaluru, India, 2023, pp. 1314-1321, doi: 10.1109/ICIMIA60377.2023.10426091.

[37] C. V. Murali Krishna, G. Gopala Krishna, S. S. Vellela, M. Venkateswara Rao, G. Sivannarayana and S. Javvadi, "A Computational Data Science Based Detection of Road Traffic Anomalies," 2023 Global Conference on Information Technologies and Communications (GCITC), Bangalore, India, 2023, pp. 1-6, doi: 10.1109/GCITC60406.2023.10425875.

[38] E. S. R. R. Kumar et al., "UAVC: Unmanned Aerial Vehicle Communication Using a Coot Optimization-Based Energy Efficient Routing Protocol," 2023 International Conference on Recent Advances in Science and Engineering Technology (ICRASET), B G NAGARA, India, 2023, pp. 1-5, doi: 10.1109/ICRASET59632.2023.10420027.

[39] Kumar, K. K., Kumar, S. G. B., Rao, S. G. R., & Sydulu, S. S. J. (2017, November). Safe and high secured ranked keyword search over an outsourced cloud data. In 2017 International Conference on Inventive Computing and Informatics (ICICI) (pp. 20-25). IEEE.