Machine learning (ML) modelling techniques for mobile technology-integrated vocabulary learning on Chinese universities EFL students’ adoption

Shiqin Huang\textsuperscript{1,2}, Abdul Rahim Bin Hamdan\textsuperscript{2,*}, Abdul Talib Bin Mohamed Hashim\textsuperscript{2}, Maoxing Zheng\textsuperscript{3,*}

\textsuperscript{1} Faculty of Humanities and Teacher Education, Wuyi University, Wuyishan, 354300, China
\textsuperscript{2} Faculty of Human Development, University Pendidikan Sultan Idris, Tanjung Malim, Perak 35900, Malaysia
\textsuperscript{3} School of Computer Sciences, Baoji University of Arts and Sciences, Baoji 721007, China

\* Corresponding authors: Abdul Rahim Bin Hamdan, abdulrahim@fpm.upsi.edu.my; Maoxing Zheng, ZhengMX1552@163.com

ABSTRACT

Studying new vocabulary is an important goal for language students, as is making the learning of vocabulary a student-centered process. Of course, the goal of investigating what variables affect people’s willingness to embrace new technologies is to better adopt such technologies. The present study will thus look at the elements that have an impact on Chinese EFL college students’ adoption and usage of mobile technology-integrated vocabulary acquisition as a means to promote more learner-centric education. To better forecast human behavior, we successfully included the following elements from the technology adoption literature into the model: attitude toward change brought on by technology usage; attitude toward technology; desire; financial ramifications; aims; past behavior; perceived consistency; positive expected feelings; visibility; and so on. Furthermore, the characteristics, all of which are external factors that describe the traits of optimal technology design, are included in the development of Unified Theory of Acceptance and Use of Technology (UTAUT) using a machine-learning. All these elements help UTAUT progress. Among its many uses, ML-based modeling may be put to good use in enhancing pre-existing explanatory statistical models (UTAUT) by identifying and analyzing hidden patterns and correlations between their many aspects. This is possible because ML-based modeling may enhance traditional statistical explanations (UTAUT).

Keywords: Machine learning, vocabulary learning, mobile technology, EFL students

ARTICLE INFO

Received: 20 July 2023
Accepted: 26 July 2023
Available online: 22 January 2024

COPYRIGHT

Copyright © 2024 by author(s).
Journal of Autonomous Intelligence is published by Frontier Scientific Publishing.
This work is licensed under the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).
https://creativecommons.org/licenses/by-nc/4.0/

1. Introduction

According to data from China’s education statistics there will be about 32 million EFL learners at China’s public universities in end of 2021. When it comes to helping students who are learning English as a second language, one of the biggest concerns is making sure that they are able to communicate effectively in English. Numerous research studies and surveys have shown that the vast majority of language learners see mobile technology for language learning in a positive light\textsuperscript{[1]}. There has been a change in the traditional roles of instructor and student in the language classroom as a result of the popularity of student-centered learning (SCL). I was wondering what it’s like for teachers to go through this change. This topic was explored in Brassinne et al. roles of instructor and student in the language classroom as a result of the popularity of student-centered learning (SCL)\textsuperscript{[2]}. I was wondering what it’s like for teachers to go through this
change. This topic was explored in the research of Andrews and Higson[2], which looked at the transition from a teacher-centered to a student-centered model of education. The research indicates that moving to SCL requires a radical shift in one’s conception of what it means to be a teacher. One of the educators in the survey said that with time, they go from being “content dispensers” to “content resources”. As with any method of instruction, some study instructors found the shift in role more natural than others. Both teachers who took to SCL right away and those who struggled initially claimed long-term benefits from their participation in the program. Teachers were able to build greater connections with their pupils, pay closer attention to the dynamics of the classroom, and tailor their lessons to the needs of their students. However, there is a lack of studies that take a student-centered approach to teaching English vocabulary to EFL students. This would better represent the students’ primary role in the learning process and better develop their independent learning skills.

Since English is rapidly becoming the global language of choice, most Chinese universities now provide instruction in English. As was said before, even if Chinese EFL students have studied English for ten to fifteen years, most of them attend higher education with just rudimentary English language skills. This point has been raised in the preceding debate. The topic of showing that teaching vocabulary to EFL students is a great technique to boost language acquisition in these students and that teaching vocabulary is one of the most crucial talents for human communication[3]. The topic of argued that the ability to write effectively in English as a lingua franca for communication in a variety of fields across cultures had become an absolute requirement in second and foreign language education programs in light of the current trend of globalization and internationalization around the globe[4]. It may correctly indicate students’ level of language proficiency not just in terms of vocabulary and grammar but also in terms of sentence structures, text organization, and logical reasoning. Vocabulary proficiency is a sought-after soft skill among modern employers, who also often express dissatisfaction with recent graduates who lack this skill[5]. One method of determining a person’s level of English proficiency is by studying their vocabulary[6].

The students at certain universities have complained that they are tired of being taught using outdated methods. University students are often left unhappy because they are unable to advance their English skills due to a mismatch between what they need and what is taught. Learning English is hindered when the instructor relies on outdated methods like lecturing and directing the class[7]. Further, one of the numerous ways that technology has changed our lives is the ease with which we may now acquire information. Education is one field that has greatly benefited from technological development. Each nation has its own unique pace of technology adoption. Teaching tools, such as computers, mobile devices, and software, may have a significant impact on student achievement[8]. The use of different technology aids by both students and instructors of foreign languages is crucial. The learning improvement tools made possible by technological progress are much better than those utilized in the more traditional educational approach. Including examples from native speakers in various nations may be an easier way to show how languages are really utilized in light of modern technological advancements. There are many situations in which the use of visual aids like films, images, animation, social media, and mobile applications might be beneficial, but language study is one of the most obvious ones. Even after the virus is cured, the topic of claims that mobile devices will continue to be used as important teaching tools[9]. Technology’s tools have matured to the point where they’re essential for giving clearer and more thorough explanations of complex topics. By having the lecture captured on tape, students have a better chance of fully absorbing the subject. Because of these problems, an idea known as “mobile learning,” or studying English using mobile devices, has evolved. The mobile technology-integrated vocabulary has many advantages, but it also has certain drawbacks that should be considered. The problem of student disengagement has been brought up by many. According to[10], there are a variety of degrees of learner participation in automated language learning technologies, such as the use of intelligent or mobile technology to include vocabulary. Many variables, including the interaction of behavioral, emotional, and cognitive elements, contribute to this level of immersion. Additional research has looked at how students’ beliefs, prior
experiences, and gender play a role in their decision to use consumer information technology. Researchers concluded that gender played a significant role. A user’s propensity to embrace technology is influenced by a number of factors, including how beneficial they find the technology to be and how easy it is to use. Another study looked at how consumers’ prior interactions with a product influenced their decision of what to use for future online monetary exchanges. This debate also illustrates the significance of looking into the components that operate as mediators, since these variables affect students’ intentions to learn the jargon related to mobile technology. It provides more evidence that gender, as well as a variety of other mediating variables, influences students’ decisions. This research aims to fill this second void by investigating the role of students’ attitudes and prior experiences as mediating factors in their intentions to utilize mobile technology-integrated vocabulary in their future conduct. This research aims to fill a second need by collecting data on the factors that influence students’ decisions to use mobile technology-integrated vocabulary. To that end, the present research intends to investigate the impact of moderating factors on students’ lexical adoption of mobile technology.

The purpose of this research was to investigate the behavioral intention of EFL students to acquire a mobile technology-integrated vocabulary, therefore addressing the third research gap. Due to its underutilization in EFL vocabulary-learning research, the UTAUT 2 model is the third area of inquiry that this study intends to investigate. In conclusion, it is worth noting that despite the widespread usage of mobile technology-integrated vocabularies, there is a dearth of empirical information identifying the elements that impact behavioral intention and adoption. It is not apparent whether factors— intrinsic or extrinsic—have a greater impact on students’ ability to successfully use the mobile technology-integrated vocabulary. Three gaps in the current corpus of research highlight the need for conducting a study to better understand the variables that impact EFL students’ use of mobile technology-integrated vocabulary. More specifically, this will make it possible to determine which variables are associated with rapid vocabulary development. The students at a Chinese public university who are taking English as a foreign language (EFL) will be the study’s subjects.

2. Background

It is important to employ many theoretical lenses and conceptual frameworks when studying how students interact with mobile technology-integrated vocabulary. This is a critical factor to add to the list. Research on students’ attitudes toward the terminology and essential features of mobile technology integration has traditionally made use of the technology acceptance model (TAM). Conventional approach models, such as the extended unified theory of acceptance and use of technology, have been used in very few studies to date (UTAUT 2, hereafter). Expected performance, expected effort, social impact, hedonic motivation, enabling conditions, pricing value, and routine are only a few of the elements considered by the UTAUT 2. It has been suggested that UTAUT 2, developed by, is a valid framework for describing how and why different technologies are used and adopted. This is because the model has a higher explanatory power than behavioral use technology, which can be used in a broad range of contexts and has more capabilities, and hence is more universally applicable. In addition, UTAUT 2 is a robust model for clarifying how and why different technologies are used and adopted.

Several factors influence students’ motivation to learn the mobile technology-integrated vocabulary. Moreover, several researchers have noted that the use and adoption of technology is a complex phenomenon that involves a complex interplay of factors, whether internal or external, and that models for examining which factors determine the successful use of technology should be developed to gain a better understanding of this phenomenon. Thus far, there is a lack of evidence from studies that pinpoint which factors significantly affect students’ desire to use mobile technology-integrated vocabulary in their everyday lives. This highlights the fact that more work needs to be done in this area, particularly the necessity for studies that look into what factors motivate EFL students to utilize mobile technology-integrated vocabulary while engaging in
vocabulary acquisition exercises. This study aims to fill this knowledge gap by looking at what motivates EFL Chinese students to adopt mobile technology-integrated vocabulary and what factors they utilize to make this choice. In the field of second language learning, the importance of developing one’s vocabulary has been extensively discussed and studied\(^2\). To become proficient in a second language (L2), one must work to expand his or her vocabulary, which in turn enhances these four language skills: listening, speaking, reading, and writing\(^2\). High-quality word knowledge, which includes knowledge of forms (pronunciation, spelling, morphological, and grammatical word properties) and the knowledge of multiple word meanings across different contexts, is associated with the understanding of the rich and interrelated information that is communicated by that word and plays an essential role in the process of acquiring vocabulary. A student’s vocabulary is one of the most difficult things to master in a language due to the intricacy of understanding a term\(^2\). Because of the low amount of time spent in class on L2 teaching and the even lower amount of exposure to the L2 studied outside of class, this is particularly true. Another difficulty with learning a new language is that many teachers put too much emphasis on rote memorization. They use tedious and unappealing approaches to teaching vocabulary and give their students monotonous and unengaging homework that does not promote student engagement\(^2\). A further unfortunate reality is that most vocabulary-related activities in modern English classrooms do not effectively equip pupils to learn vocabulary on their own initiative and without direct teaching. When it comes to vocabulary, it is more customary for teachers to pick which words their pupils should learn and when, rather than giving them the chance to study on their own. Because this challenge is focused on the educator, it raises related difficulties, such as the fact that educators aren’t always sure which vocabulary phrases to use, how many words should be learned, which examinations to apply, or how often they should grade students. Vocabulary and the ability to learn new words through reading are skills that are put to the test in the process of reading. Poor readers will have considerably greater problems in school if they are not taught decoding and word-learning abilities, even if they have some vocabulary to begin with\(^2\).

3. Method

In order to determine what factors affect students’ decisions to make use of mobile-assisted technologies in higher education, this research draws on the unified theory of acceptance and usage of technology (UTAUT 2). The UTAUT model was created after researchers examined eight well-known theories and models, including the technology readiness assessment (TRA), the technology acceptance model (TAM), the traditional motivational model (MM), the technology planned behavior (TPB), the combined TAM and TPB (CTAM-TPB), the model of PC utilization (MPCU), the innovation diffusion theory (IDT), and the social cognitive theory (SCT). The UTAUT model is now considered the most cutting-edge and robust approach for analyzing many technology implementations and uptakes. We chose the UTAUT model because of its versatility, strong explanatory power of tech use behavior (more than 70%), and capacity to forecast tech use behavior in the future\(^2\). Additionally, the UTAUT 2 model, which was developed by adding to the original UTAUT model with extra variables, is also accessible. The original unified theory of acceptance and use of technology (UTAUT) was created by\(^2\). UTAUT 2 is an expansion of UTAUT that focuses on how students accept and utilize technology. It was modeled after the first iteration of UTAUT\(^2\). Research on UTAUT 2 found a far higher prediction efficiency (74% and 52%) when compared to the diversity in behavioral intention and use behavior seen in UTAUT (40% and 30%, respectively)\(^2\). Students who make heavy use of sophisticated features within a learner-centric framework in this research of Chinese university students who utilize mobile-assisted technologies (mostly offered by app stores) are expected to foot the bill themselves. Students may also be unfamiliar with the benefits of mobile-assisted technology in EFL classroom settings. Paradoxical actions on the responder’s part may result from the UTAUT 2 model’s integration of new variables. Each of the seven factors that make up the UTAUT model—performance expectation (PE), effort expectancy (EE), social influence (SI), facilitating condition (FC), hedonic motivation (HM), price value (PV), and habit—is taken into account here (HT). A number of experts in the field feel that all four of these factors have impacted people’s
BI and UB towards the technology. Also, four additional variables—privacy, trust, personal inventiveness, and information quality—have been added to the original four in the UTAUT model to account for the specifics of mobile-assisted technology. In addition, 11 factors are taken into account in this study, from which the research model is developed. The overall structure of the ideas is shown in Figure 1.

![Figure 1. The conceptual framework.](image)

The currently available models of technology adoption are lacking in human behavior and technology design considerations that are of significant importance (i.e. TAM and UTAUT). Examples of explanatory linear algorithms that are used by both models are multiple linear regression (MLR), structural equation modeling (SEM), and partial least squares structural equation modeling (PLS-SEM). Any added qualities must have a linear relationship to the TAM and UTAUT traits that already exist in order for these approaches to work. When these models are expanded, if a new feature can’t be described by a linear combination of characteristics that already exist, then the model to explain it is expanded without it \[27\]. In spite of the fact that both models make use of explanatory approaches to predict future technology use and adoption, there is one area in which they differ from one another. In addition, the development processes of both TAM and UTAUT contained a number of significant flaws. These flaws included the following: the models relied on a single explanatory linear algorithm to evaluate their modeling power; their strengths were measured utilizing the same metric (i.e. R2); and the models were tested on groups of identical subjects in involuntary settings using the same technology \[28\]. Both of the construction procedures were subjective, therefore it will be difficult to readily modify either model to accommodate new facts. Neither model is structurally predictive, nor is it general, nor does it use a standard set of inputs. Neither model uses a standard set of inputs. Both TAM and UTAUT use the assumption that the previous relationships between aspects in the model are linear and non-monotonic in nature. This is assumed to be the case on account of the fact that both models were developed using linear methodologies like as MLR, SEM, and PLS-SEM \[29\]. We suggest, however, that a combination of data mining (DM) \[30\] and machine learning (ML) \[31\] techniques may be able to extract information from both models and identify new link patterns. We advocate introducing a more objective modeling technique, such as ML, into studies on technology adoption as a means of addressing the inadequacies of both TAM and UTAUT. Because of this procedure, it would be possible to investigate both linear and nonlinear implications on technology acceptance characteristics. In addition, it would be possible to include new characteristics without having to worry about how they relate to the preexisting characteristics of either model. Both of these things would be possible. In addition, a number of other prediction evaluation metrics in addition to linear and nonlinear algorithms should be used in this process in order to determine how effectively different technological acceptance frameworks can foresee results. These frameworks should be tested with a wide variety of subjects,
in an opt-in setting, using a wide range of technologies, and put through their paces. Multiple linear regression (MLR) is the statistical modeling technique that makes predicting a dependent variable as simple as possible (here, technology use). This approach, which assumes that linear interactions exist between the dependent and independent variables, has become very popular as a means of elucidating the connection between inputs and outcomes\cite{31}.

4. Result

4.1. Multiple linear regression

Figure 2, shows the performance of underfitting and overfitting of the machine learning-based UTAUT.

![Learning curve of MLR.](image)

4.2. K-nearest neighbour regression (KNNR)

The results of the KNNR model for the two weight functions are shown in Figures 3 and 4 for the range of one to one hundred neighbors, respectively. It’s important to keep in mind that Mean Absolute Percentage Error MAPE values are a weighted average of the aforementioned neighbors and the whole dataset.

![Learning curve of KNNR uniform.](image)
4.3. Decision tree regression (DTR)

The best Mean absolute error (MAE) and Mean squared error (MSE) results for DTR are shown in Figures 5 and 6, respectively. Both graphs demonstrate that the model’s underfitting and overfitting were managed.
4.4. Multilayer perceptron regression (MLPR)

Machine learning-based UTAUT was modeled with the use of a logistic function and limited Broyden, Fletcher, Goldfarb, and Shanno (LBFGS) optimizer, the results of which are shown in Figure 7.

![Figure 7. Learning curve of MLPR logistic.](image)

4.5. Support vector regression (SVR)

The effectiveness of SVR in modeling machine learning-based UTAUT was examined using a variety of parameters, including an error penalty parameter C ranging from 1 to 10 and an epsilon value of 0.10. Given that the maximum R2 value was achieved by SVR when the POLY function was utilized, just the POLY function’s performance is shown in Figure 8.

![Figure 8. Learning curve of SVR polynomial.](image)

4.6. Ranking of personal technology acceptance model’s features using sensitivity analysis

After tracking down the coefficients for every conceivable variable, we ranked them from smallest to largest. Past behavior was the most significant indicator of future technology use, whereas free will played a far less significant role. Each variable’s impact on the predicted outcome is shown in the coefficient column (i.e., use behavior). Each variable’s coefficient is divided by the sum of all coefficients for the 37 variables to get the normalized coefficient column. That’s why we can pinpoint how each factor affected the model.

5. Discussions

There are a few notable outliers, but SEM is used extensively throughout TAM and UTAUT literature to make predictions about the spread of various technologies. However, as SEM is a tool for explanatory modeling, it should be used to explore the connections between the various determinants of user behavior. In
this research, we used a data-driven, ML-based modeling approach to create a predictive UTAUT (as opposed to an explanatory one). Given that current technology adoption explanatory models are theory-driven, it is expected that ML-based models would provide a more accurate reflection of reality. In fact, it may be impossible to determine the optimal performance of machine learning-based UTAUT without first investigating a variety of linear and nonlinear approaches. Therefore, we used a variety of techniques to evaluate the efficacy of machine learning-based UTAUT and choose the most effective model.

The proposed machine learning-based UTAUT achieves the first, second, third, and fifth aims since no characteristic has a weight of 0 since SVR-POLY provided the most reliable predictions, we will use its variable ranking in the following analysis. Out of the 37 criteria used to predict technology adoption, previous behavior is the most important. To accomplish objectives 3 and 6, this result indicates the impact of a data-driven approach on the current body of literature. This result runs counter to what has been found in the literature on TAM and UTAUT, which argues that behavioral intention is the strongest predictor of technology use behavior. Furthermore, the topic of claim that psychological and TAM research supports the conclusion that behavioral intention is the strongest predictor of technology use is challenged by these findings[17]. Ajzen agrees with the findings that prior actions are the most telling of future outcomes. TAM was derived from TPB, which did not include the strongest predictor of future behavior—past behavior—because it did not have a strong linear connection with behavioral intention. Unlike other studies that only looked at TAM and UTAUT, our work uses both linear and non-linear algorithms, such as SVR-POLY, to demonstrate that past behavior is crucial for forecasting future technology usage.

In TAM and UTAUT, the behavioral intention component was predicted to be essential in comprehending why end users were eager to use technology, yet it was ranked only eighteenth in terms of importance. Thirdly, we show that the recommended data-driven approach has an impact on the literature on technology acceptance by comparing the rankings of a few additional criteria with those in the literature. The importance of factors such as desire, perceived ease of use, habit, technology self-efficacy, visibility, perceived enjoyment, result demonstrability, technology playfulness, privacy, attitude toward technology use, financial consequences, security, service quality, safety, attitude toward change brought about by technology use, perceived consistency, and image were ranked higher than behavioral intention, indicating that these factors are more important in predicting technology use. The seventeen characteristics listed after users’ prior actions provide credence to the idea that users’ stated intentions to make use of technology are less important than is often assumed. Another important outcome of our sensitivity analysis is a rating of five characteristics: how people feel about and use technology; how visible and consistent these technologies are; how consistent they feel about using them; and how confident they feel about their own technical abilities. Attitude toward technology was proven to influence the forecast of technology use at UTAUT based on machine learning, despite just being recently presented in the literature. Regardless of how they feel about using technology, the subjective norm was eliminated from TAM due to the inability to isolate the effects of attitude on behavior from those of behavior on attitude. Using SVR-POLY, we found that there is a substantial and distinguishable influence of one’s attitude toward technology usage and one’s subjective norms on forecasting technology use. Because they did not have significant linear relationships with behavioral intention, attitudes toward technology use, visibility, perceived consistency, and technical self-efficacy were all left out of the UTAUT model. In contrast to other models, SVR-POLY was able to accurately capture the influence of these four factors on technology adoption, ranking them higher than the intention characteristic. Two objectives have been met with this result.

6. Conclusion

This paper modeled and evaluated 37 features of UTAUT based on machine learning that have an impact on consumers’ decisions about their own in-home technology. Existing work on TAM and UTAUT is dominated by qualities related to human behavior; however, there is a paucity of clear human behavior and
technology design considerations. We successfully incorporated the following characteristics from the literature on technology acceptance into both models to improve their ability to predict human behavior: openness to change, interest in technology, willingness to make sacrifices for the greater good, perceived consistency, positive anticipated emotions, and the ability to see the results of one’s actions. In addition to improving TAM and UTAUT, machine learning-based UTAUT does so by including the exterior features of desirable technical design qualities like compatibility, flexibility, functionality, mobility, navigability, safety, service quality, and technology quality. By capturing complex underlying patterns and interactions across TAM and UTAUT’s current features, ML-based modeling may, among other capabilities, improve existing explanatory statistical models. New features might be added more easily if this approach is used. Adopting DM and ML strategies not only improved the technical landscape but also uncovered hitherto unseen phenomena. Unique (non-linear, monotonic, and non-monotonic) relationships between UTAUT features were discovered in the offered study. The present literature does not provide an explanation for the formation of these novel linkages or a discussion of the implications of their creation with regard to use behavior. We see this as a chance for researchers to learn more about the ways in which users’ perceptions change and develop over time.

Author contributions

Conceptualization, HS and ARBH; methodology, HS; software, MZ; validation, ARBH and ATBMH; formal analysis, HS and MZ. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

The authors declare no conflict of interest.

References

25. Brassinne, K., Reynders, M., Coninx, K., & Guedens, W. (2020). Developing and implementing GAPc, a gamification project in chemistry, toward a remote active student-centered chemistry course bridging the gap between precollege and undergraduate education. Journal of Chemical Education, 97(8), 2147-2152.
31. Sultana, J. (2020). Determining the factors that affect the uses of Mobile Cloud Learning (MCL) platform Blackboard-a modification of the UTAUT model. Education and Information Technologies, 25(1), 223-238.