Indonesian Journal of EFL and Linguistics

eISSN: 2503-4197, pISSN: 2527-5070
Available online at:
www. indonesian-efl-journal.org
doi: http://dx.doi.org/10.21462/ijefl.v9i1.758

Pattern Recognition of EFL University Students' Online Behaviors through Data Science: Any Investment on English Language Components or Skills?

Maryam Arabshahi[1], Zeinab Kafi[2*], Dennis Krummacker[3], Hans D. Schotten[4]

[1]*Intelligent Networks Research Group, German Research Center for Artificial Intelligence (DFKI GmbH), Germany.*
maryam.arabshahi@dfki.de

[2]*Assistant Professor, Department of English, Khorasan Institute of Higher Education, Mashhad, Iran.*
kafizb@gmail.com

[3]*Intelligent Networks Research Group, German Research Center for Artificial Intelligence (DFKI GmbH), Germany.*
dennis.krummacker@dfki.de

[4]*Institute for Wireless Communication and Navigation, RPTU University of Kaiserslautern-Landau, Germany.*
schotten@rptu.de

Abstract

Artificial intelligence (AI) has played a crucial role in many fields of study, and its merits and demerits have been investigated by related scholars. However, it seems a somewhat gray area when implementing AI in teaching and learning. This could be mainly because AI developers know little about learning varied sciences and they also lack pedagogical knowledge for implementing AI in teaching (Luckin & Cukurova, 2019). Therefore, the current research aimed at gathering data regarding the online behaviors of 21 Iranian EFL university students through installing an application on their mobile phones for a period of eight days. Afterwards, the collected data were analyzed through Data Science methods and varied patterns of behavior were recognized. The results exhibited that messaging and social media applications are two great parts of the students' life. Also, there existed some common patterns regarding the applications students owned and the amount of time they spent on them. As a small case in point, the results of the study cannot be generalized; however, the outcomes highlight the proof that in an EFL context, serious steps need to be adopted by course designers and stakeholders in order to cultivate a greater and purposeful use of online learning applications, websites etc. among EFL university students so as to enhance the teaching-learning process. This gains prominence because any AI-mediated learning approach is believed to enhance students' second/foreign language motivation as well as self-regulation in learning. AI is actually promising a revolution in the realm of language education.

Keywords: *Data Science, EFL Students, Online Behaviors, Pattern Recognition, Wireless Network Traffic Data*

* Corresponding Author
Email: kafizb@gmail.com

1. INTRODUCTION

Artificial Intelligence (AI) is a field of science which simulates human intelligence by writing intelligent computer programs to solve different problems. "Intelligence" includes the ability to learn, reason, generalize and infer meaning (McCarthy, 2004). AI is like a comprehensive toolbox which could help us solve a variety of problems. Usually, the AI algorithms are better in performance in the case of accuracy and precision in comparison with many conventional methods. AI has one subset which is called Machine Learning (ML). Machine learning is the ability of a computer program to learn about a subject and after that act regarding what it has learned. It means for example given the past events, the program could be able to guess the future, generalize and adapt itself with new situations and information and eventually get better with practice (Haixiang et al., 2017). Likewise, machine learning has one subset which is called Deep Learning (DL). Deep learning methods are used for more complex problems like image/video processing, voice recognition, text recognition, etc.

Talking about the realm of teaching and learning, teachers' experiences, attitudes, and interests regarding AI can be a great facilitator or barrier for implementing AI in education. That is why teachers are often referred to as "the most crucial stakeholders in AI-based teaching" (Clik, 2022). Therefore, more attention must be invested on AI-based education from teachers' perspective. As one of the researchers in the current study is an expert teacher in the field of TEFL (Teaching English as a Foreign Language), this pattern of thought invoked the researchers to probe the online behaviors of Iranian EFL students in order to accumulate some preliminary information over how university students use their data when they are online.

Additionally, delving into the concept of AI and education from learners' perspective, it has been proved that those who are aware of how to use AI and technology-based techniques in their language learning, can eventually turn into independent learners who can put their learning anxiety and self-regulation under control (Hsu et al., 2023). Therefore, teachers and students form both sides of the continuum and any research regarding EFL students and teachers in the field of AI and education could benefit both groups.

The study by Ling (2023), illustrates an instance of research regarding AI-based education and the EFL context. The researcher investigated the impact of Artificial Intelligence on English Language achievement, L2 motivation and self-regulated learning and the outcomes depicted positive results in the experimental group. Other educators have employed means of technology in order to create virtual environments through which learners can actively engage in the language learning process in a facilitated way (Loncar et al., 2023; Nguyen & Le, 2023).

Regarding English language skills and sub-skills, the influence of AI has also been probed. As a case in point, a study by Hsu et al. (2023) assessed the effects of AI language learning tools on EFL learners' vocabulary knowledge through two groups of control and experimental, and the results exhibited the outperformance of those who utilized AI tools for the sake of enhancing vocabulary knowledge. Another similar study focused on investigating the same effect on EFL learners' speaking skills, and once again, there existed proof that AI users outperformed in speaking proficiency(Junaidi, 2020). Considering the already discussed, to the best knowledge of the researchers, no study in an EFL context has so far shed light on the extent to which EFL university students use their time online and which patterns of online behavior they possess when they go online. This is crucial because it provides a kind of preliminary knowledge about the actual online

habits of EFL learners/university students which helps to make sound educational decisions based on the recognized patterns of behavior.

Chatbots, as an example of AI, have recently been used in education for purposes such as intelligent tutoring and varied automated grading systems. Nonetheless, to the best knowledge of the researchers, employing AI to locate the online behaviors of students in academic settings has not been the center of any research so far. Additionally, quite a lot of the research regarding AI and education merely focuses on systematic reviews/mini-reviews of the available literature, and few empirical studies are traceable.

Moreover, regarding the importance of data science, it can be said that it is an interdisciplinary academic field that uses statistics, scientific methods, and algorithms to explore the data and extract knowledge and insights from it. Data science has different steps including framing the problem, collecting the data, understanding the data, data cleaning, visualization and exploratory data analysis, model building, evaluation, and deployment. There are different tools and methods for each of these steps; AI algorithms are one of them. CRISP-DM (CRoss Industry Standard Process for Data Mining) is one of the first and most famous data science standards for data science (Wirth et al., 2000).

Therefore, to address these research gaps, the current study shed light on probing probable patterns of the online behavior of Iranian EFL university students. Since AI-based education is yet under development and way behind common expectations, the current research can be a contribution that allows teachers and learners to participate in the process of designing, creating, developing and integrating various AI-based educational programs and study patterns.

Accordingly, the scope of the study covered the following research questions and the outcomes are discussed in the 'results' section:

RQ1. How much time do the participants usually spend on their phones?
RQ2. Which applications are frequently used by the EFL participants?
RQ3. What is the data usage of the participants while using their phones every day?
RQ4. How many different applications does each student have on his/her phone?
RQ5. Do the participants use any English-learning applications on their phones?

2. LITERATURE REVIEW

2.1 AI in Education

Clik et al. (2022) probed teachers' use of artificial intelligence applications and machine learning methods for analyzing teachers' data. They concluded that AI offers teachers varied opportunities such as improved planning, implementation and assessment of their teaching. The authors also concluded that teachers could have influential roles in the development of AI technology. Lastly, the study singled out some of the challenges of applying AI methods in teaching practice.

Other research areas regarding the application of AI in education (Luckin et al., 2016) support the idea of personalization of the learning process as well as student collaboration. Profiling students' backgrounds as well as monitoring their progress (Cohen et al., 2017; Gaudioso et al., 2012; Swiecki et al., 2019) have also been the scope for other relevant research in the field.

Another perspective has been centering around the issue that AI will soon replace teachers in classrooms. However, it sounds more reasonable to instead consider the opportunities and additional support that AI can provide in order to strengthen the role of teachers (Hrastinski et al., 2019). For instance, AI-based machines can help teachers perform their orchestrator role in the learning and teaching process (Dillenbourg, 2013).

Chatbots which are one common means of AI, can be employed in the classroom setting for developing learners' language skills such as speaking, listening, reading and writing (Gayed et al., 2022). Accordingly, the use of chatbots in university EFL settings has been the subject of some studies (Klimova & Seraj, 2023). The core contribution of their study revealed the chatbot's potential in applying and integrating the existing concepts in EFL teaching and learning including CEFR, mind mapping, or self-regulatory learning theory.

Investigating how AI image recognition technologies and self-regulated learning can affect vocabulary acquisition, self-regulation and learning anxiety in the language classroom was the focus of another study by Hsu et al. (2023). The results of the study exhibited that the two groups of experimental and control, did not have any significant difference regarding their learning anxiety and self-regulation; however, the experimental group outperformed the control group with regard to their vocabulary acquisition.

2.2 Analysis of Network Traffic Data

Spatiotemporal information extracted from the mobile wireless networks has been studied by Parwez et al., (2017). Call Detail Record (CDR) of many end-users has been analyzed to find anomalous behavior of mobile wireless network.

An analysis on mobile user data has been performed to understand users' behavior based on different parameters. The characteristics of data usage of different user groups during different time intervals have been compared with each other. It has been suggested by the authors that the result of the analyzed mobile data could be used by the Mobile Network Operators (MNOs) in different ways including new pricing mechanisms for data usage, efficient usage of network resources, dynamic advertisement pricing, efficient social content caching scheme, and split billing. It is also mentioned that research on social networks shows that users with social bonds have a corresponding mobility behavior and interest in the content (Bulut et al., 2015).

It has been discussed by Verkasalo (2010) that precise mobile data of uses behavior and experience is needed for today's development of products and promotion measures. A framework has been developed which collects data at the users' devices and analyzes it on their side.

The mobile traffic data of different users has been collected within a week by Yang et al. (2014). The data has been studied from three aspects: data usage, mobility pattern, and application usage. Mobile users were classified into different groups, and their resource consumption was analyzed. It was investigated that both the data usage and the mobility pattern are strongly linked to the user's access to the applications. Also, the users have been clustered through their application usage behavior, and the applications have been categorized by the manner in which they engage the users.

An empirical analysis on end-users' searches for information in the browser of mobile devices has been done (Vojnovic, 2008). Focus of this study is on temporal dynamics, semantics, and topics of queries. The users' data has been collected over a week. It has been discovered that Human search for information via mobile devices exhibits a spread of arrival time between requests that follows a potential rule up to one day and declines exponentially thereafter.

Patterns extracted from the network traffic data of large-scale cellular towers in an urban environment have been explored (Xu et al, 2016). It has been claimed that by detecting and modeling the network traffic patterns, the Internet Service Providers (ISPs) could customize the resource allocation strategies and be more flexible, instead of applying the same old methods for each cell tower. It has also been explained that the traffic of a cell tower is influenced by different parameters, such as time of the day and location of the cell (residential, business, etc.).

A survey on applying deep learning methods for network traffic monitoring has been done (Abbasi et al., 2021). As well, the challenges and open issues in this field were discussed. It was explained that due to various reasons, network administrators are willing to evaluate a network, such as security challenges, Quality of Service (QoS), and improving resource consumption. The authors describe that there are two types of methods for analyzing the network traffic data: active methods for analyzing the performance of a network online, passive methods for monitoring the network for post-event situations like troubleshooting. Network traffic data could be monitored with Shallow Packet Inspection (SPI) methods or with Deep Packet Inspection (DPI).

Predictive analysis of network traffic data has been done by Trinh et al. (2018). For this purpose, Long Short Term Memory (LSTM) has been applied to the traffic data of an LTE base station. The authors mention that in order to satisfy the needs of applications which require Machine to Machine (M2M) communications, the networks should be improved in terms of latency, capacity and context awareness. To do so, it is needed to study and learn the traffic demands of the networks.

It is discussed by Trinh et al. (2017) that the difficulties of getting access to real mobile network data are mainly because the network operators won't release full datasets due to users' privacy concerns. The traffic data captured from an LTE network using a sniffer was analyzed. The analysis shows the temporal and spatial behaviors of different LTE cells. Also, a stochastic Markov Model to characterize the traffic patterns in the network was created.

3. RESEARCH METHODOLOGY

3.1 Participants

Since the data collection procedure required participants to install an application on their phones to keep a record of their online behavior, the consent of the participants was quite needed at the very beginning phase of the research. Thus, convenient sampling was applied and prior to that, the researchers thoroughly explained the purpose, importance, and the data collection procedure of the study to a group of around 70 students. Afterwards, a list of the willing participants was prepared, and data collection started accordingly. Also, the researchers confirmed that participants may seek and catch up on the study procedure in case of any further interest or inquiry.

From among the 70 participants who attended the session for presenting the study design and purpose, 21 Iranian EFL female and male B.A university students volunteered. Majority of them were majoring in English Language Teaching and English Language Translation studying at Khorasan Institute of Higher Education, and Imam Reza University, Mashhad, Iran. Their ages ranged from 18 to 25. Their English language background was 2 to 5 years.

3.2 Instrumentation

Regarding the data collection, the researchers needed a concise record of the online activities of the participants for a period of eight days. For this purpose, one mobile application (PCAPdroid) was

opted. PCAPdroid is a privacy-friendly open-source application that tracks the network traffic data in Android phones[†]. The application and its efficiency for data collection was previously assessed by the researchers.

To help the participants install and use the application properly, a short audio-visual video clip was prepared and sent to the participants via a Telegram group. The participants received a reminder to turn on the app every morning and turn it off at night. This process lasted for 8 days from 08.04.2023 till 15.04.2023. The app recorded their daily network traffic data and provided a .PCAP format dataset. Maintaining normal, routine-based online behavior was highly recommended to the participants.

Once the eight-day data recording through the application was over, the participants were asked to send the output of the application via the Telegram group to the researchers for further data analysis. After that, Wireshark[‡] was used to convert the .pcap data into .csv for further data analysis with Python. Python libraries for data science including pandas, NumPy and scikit-learn and also AI method Clustering were employed for analyzing the collected data.

3.3 Study Limitations

Considering the limitations of the study, two aspects need to be mentioned. Firstly, since the data collection required the direct access of the researchers to the participants, students who took part in the research were mostly those studying at two universities in Mashhad, Iran: Khorasan Institute of Higher Education and Imam Reza International University. There existed few participants from other universities; however, the number was not significant.

Moreover, the serious internet restrictions in Iran, which oblige participants to use VPN every single time they want to connect to some platforms such as U tube, Instagram and some other non-local apps or websites, put the online behaviors of the Iranian participants under its partial influence. As a result, the online behavior of the students could have been a little more different if such limitations and restrictions did not exist.

In addition, PCAPdroid is an *Android* application and does not have an iOS version. Additionally, to the best knowledge of the researchers to this date, there is no application for capturing the network traffic data which works on both Android and iOS. And because the Android phones are more common, it was decided to stay with the Android version. This caused some drop-out in the original number of voluntary participants (around 10 students were removed from the range of participants and the final data was collected from 21 participants).

3.4 Data Analysis

The researchers employed JupyterLab as a platform for programming with Python. Using pandas, NumPy, scikit-learn and Matplotlib, the collected data was analyzed and some overall interesting patterns in the data were detected including:

- Majority of the participants mostly used similar applications, which means there were some apps that were common in every cellphone and also favored by all participants.

[†] https://github.com/emanuele-f/PCAPdroid
[‡] https://www.wireshark.org/

- Students spent most of their time online using social networks: Instagram stood at the top. Messaging apps stood in the second place of the spent time (WhatsApp and then Telegram). And the third place belonged to the streaming apps like YouTube.
- Another observation was that the participants consumed more network traffic data on Saturday (which is the first day of the week in Iran)

However, the results for each of the research questions are discussed respectively hereafter.

4. RESULTS

4.1 Activation Time of the Applications and Applications Categories Used by the EFL Participants

The outcomes for the first two research questions are depicted in Table 1 and Figures 1 and 2. One of the early assumptions was that students in this age range might have many games on their cellphones and spend a lot of time playing; however, to the researchers' surprise, this was not the case and only few students had few games on their cellphones, and they didn't spend much time on them in comparison to other applications.

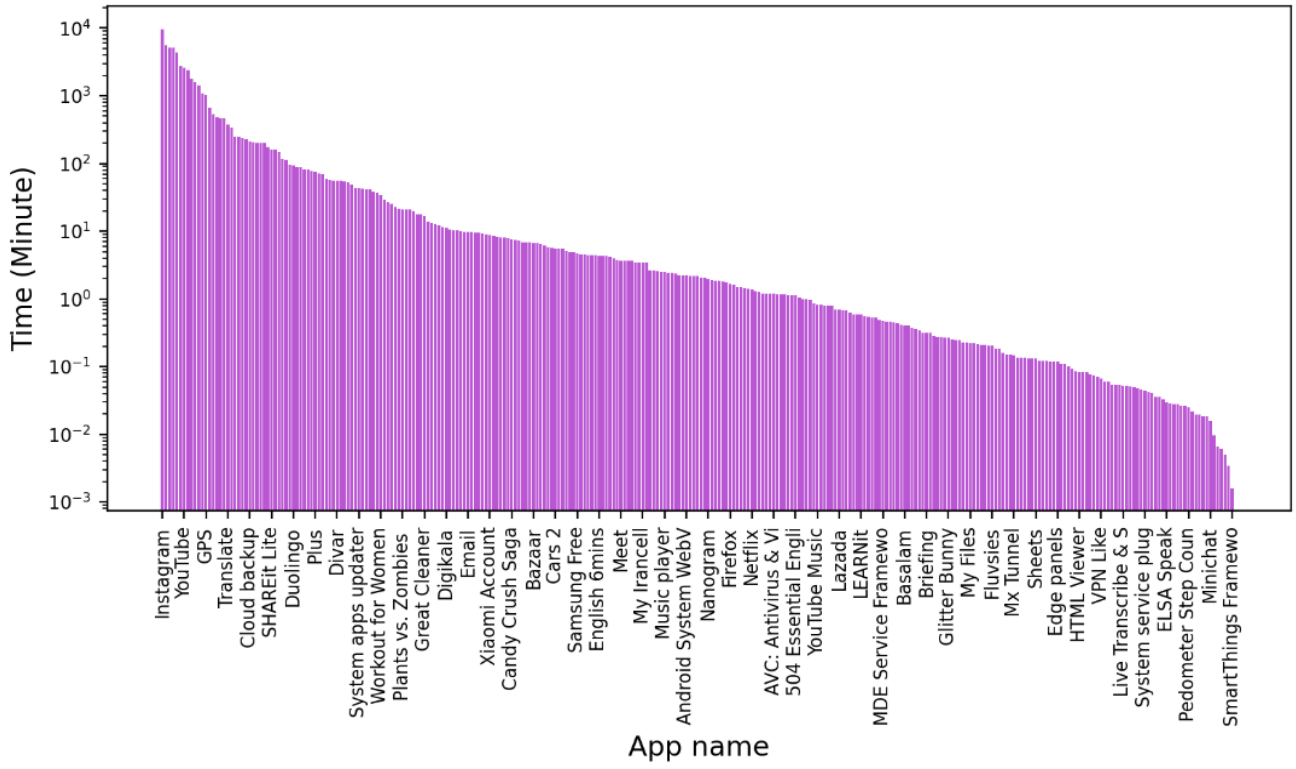

Figure 1. *Activation Time of the Applications*

Table 1: Applications which were activated the most

App Name	Time (Minutes)
Istagram	9676.758707
WhatsApp	5624.028766
Google Play Service	5186.1539
Telegram	5082.778435
Google	4405.837891
Google Play Store	2729.959344
YouTube	2597.952032
netd	2401.794334
Meta Services	1769.699063
Lantern	1578.619766
Snapchat	1410.583053
Android	1082.266138
GPS	1039.130547
Chrome	667.096122
Soroush+	529.48733
msa	484.440071
Xiaomi service fram	469.648904
Root	464.537042
Translate	371.996087
Samsung Internet	332.341615

Fig. 1 shows which apps have been used and how much time has been spent on them in the whole target group. Table 1 displays the top 20 apps which were used widely. Instagram has been used the most frequently and is on top of the list with great distance from the others. WhatsApp, Telegram and YouTube are also among the top 7 used apps. In addition, Google-related apps were activated very often too. Fig. 2 elaborates a clustering of distribution of different categories of applications students used during the data collection period.

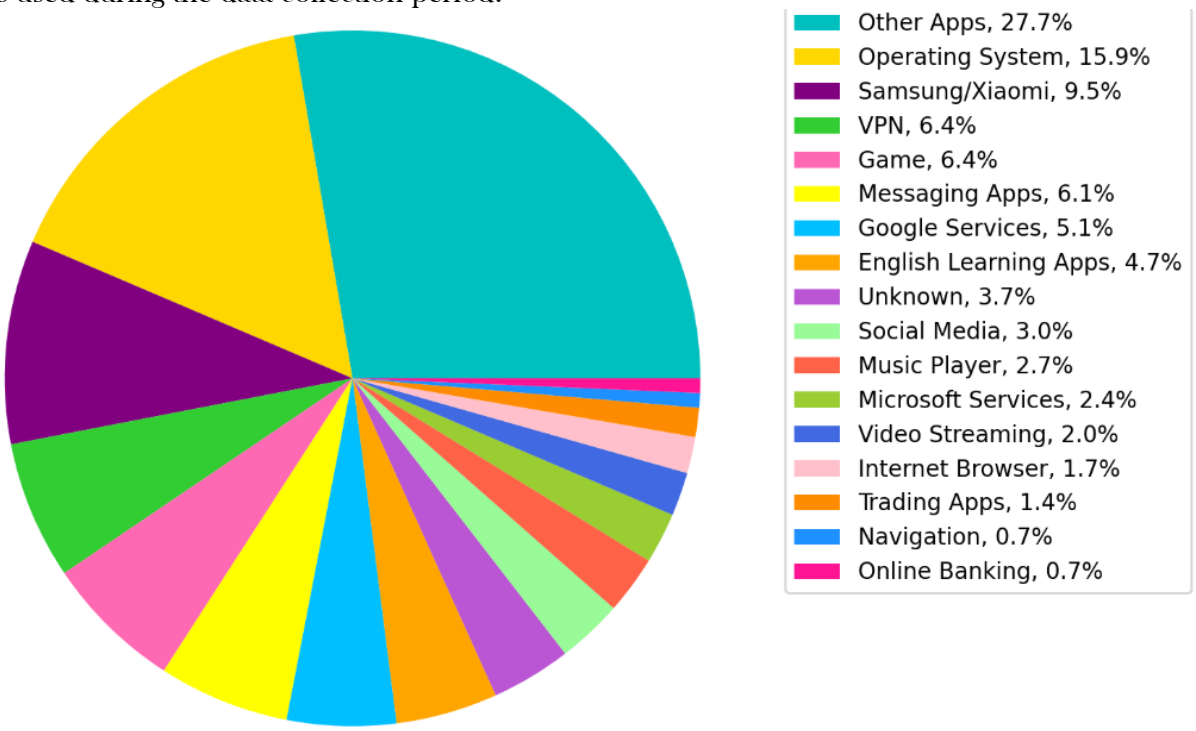

Figure 2. *Applications Categories*

As Fig. 2 illustrates, each category consists of different number of applications. There are 17 different categories in total. Some categories like "Operating Systems", "Samsung/Xiaomi", and "Google Services" are mostly the applications which were not used by the students, but they are the ones running in the background of the phone. 27.7% of the applications are in the category "Other Apps". These were the applications which were not easy to assign to a special category. For example, weather app, pdf creator app, etc.

"Operating System" was the second biggest category and contained 15.9% of the applications. Also "Samsung/Xiaomi" contained 9.5% of all the categories. 3.7% of the applications belonged to the "Unknown" category. This means that the PCAPdroid could not recognize them. It is usually due to the bad internet connection.

"VPN" took 6.4 percent of the categories. This is a proof to the fact that though internet status in Iran is under severe restrictions, the youth are determined enough to stay connected to the online world! Once again, reference can be given to Table 1 which shows the top 20 apps which were activated the most among all the participants. The time has been given by minutes. Here, it is revealed that there are some apps which belong to the operating system and they were activated for a great amount of time. These are the apps that users do not use actively, but they run on the background of the operating system.

4.2 The Data Usage of the Participants while Using their Phones Every Day

The results regarding the third research question are elaborated in Figure 3.

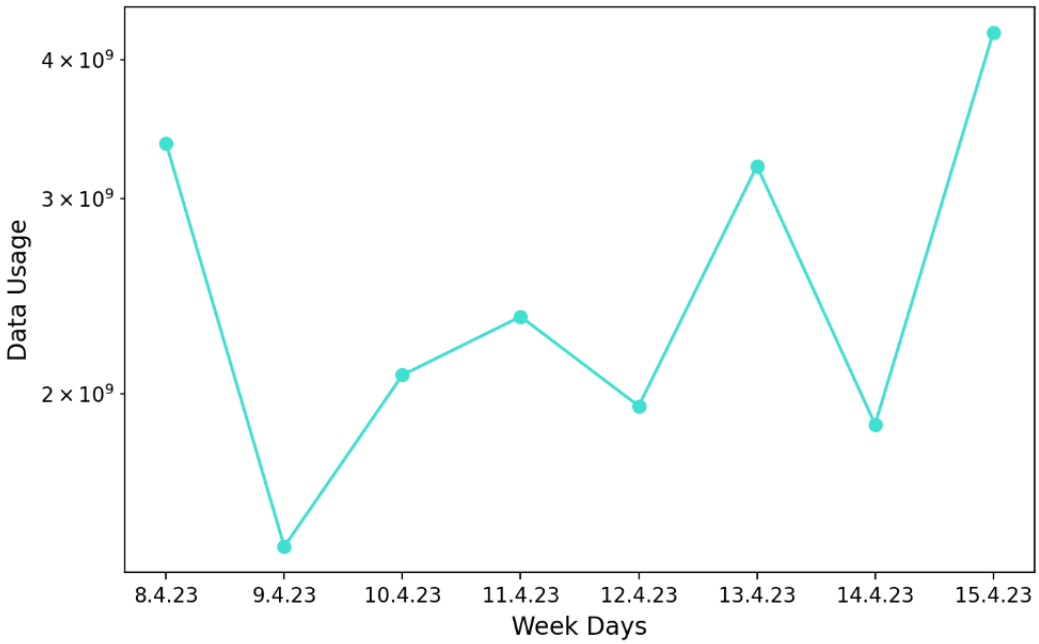

Figure 3. *Daily Network Traffic Data Usage.*

Figure 3 reveals the daily network traffic data of all the students. It is observed that students consumed the most amount of data on both Saturdays which is the first weekday in Iran. Also, the data consumption was the least on Sunday.

4.3 Number of Applications Used

The data analysis regarding the fourth research question is depicted in Fig. 4.

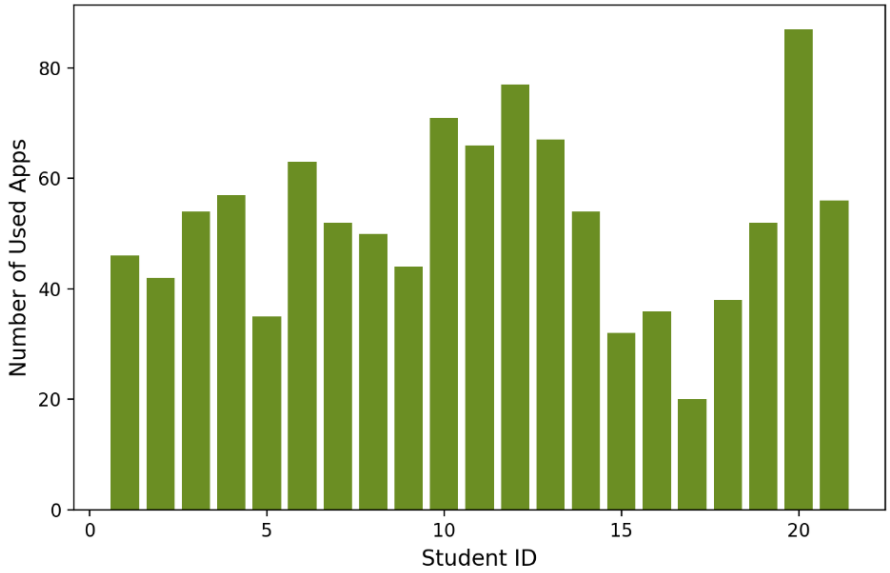

Figure 4. *Number of Applications used by each Student*

Fig. 4 reveals that the number of applications each student used during the time their network traffic data was captured, embraced a wide range, between 20 to 87 applications in total. There were 296 different applications used by all the students in total throughout this study which puts forward the assumption that EFL university students have the habit of managing many of their educational and private affairs of their life online and through remote means.

4.4 English Learning Assistant Applications used by the Students

The results regarding the last research question are illustrated in Fig. 5.

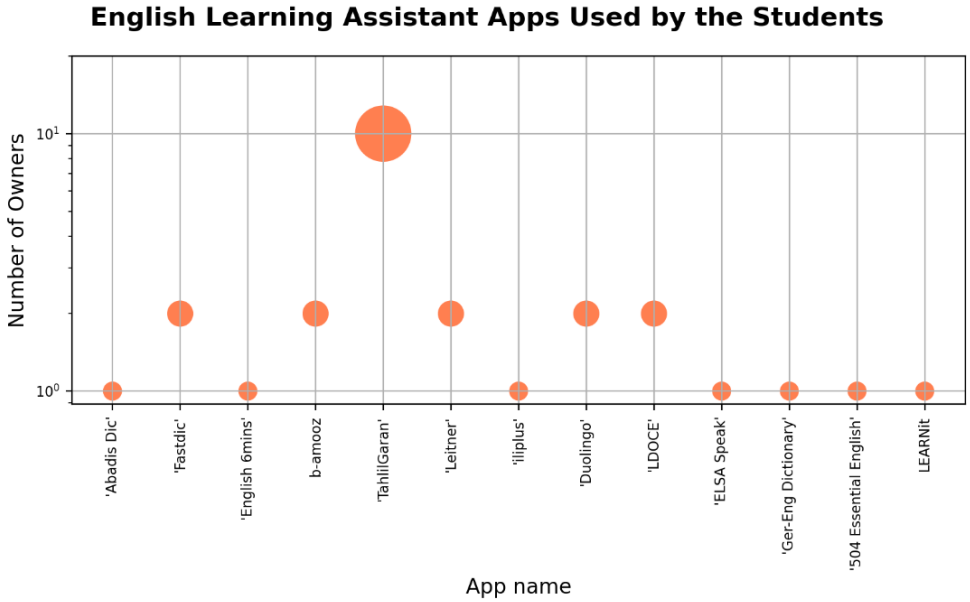

Figure 5. *English Learning Assistant Applications used by the Students*

Fig. 5 illustrates the different English learning assistant applications used by the participants. There existed 13 applications used in total. There did not exist many commonalities among the used applications, except for one which is called "TahlilGaran" which is a comprehensive dictionary available on both offline and online formats. This application was used by 10 students. In general, four of the students did not have any English learning applications. Eight of them owned one English learning application. Eight participants had two of such applications. And one student had three English learning applications. Considering the 21st century needs and evolution in language learning which is taking place both in online and offline modes, the results depict that more investment has to be observed by policy makers in general and EFL teachers in particular, regarding the design of more online-oriented learning habits for EFL students.

5. DISCUSSION

In this research, the wireless network traffic data of 21 EFL university students within 8 days was collected. Different data science and AI methods were used in order to visualize and analyze the data. Different patterns regarding the number and kinds of applications which the students had used and also the amount of time they spent on their phones were detected.

This study revealed that in the target group, most of the students already use at least one learning assistant application. This, along with the amount of time they spend on their phone shows that they are familiar with such applications, and they are motivated to use them. Additionally, this research exhibited that data science approaches and especially AI methods could be extremely helpful to understand students' online behaviors. It could extract useful patterns and reveal insightful information about learners. This is supported by other similar researches in the field. In his study, Utami et al., (2022), probed pattern recognition in online learning and asserted that pattern recognition can be applied in the non-verbal aspect assessment system of teaching skills.

According to the revealed outcomes, in an EFL setting and among EFL university students, there are commonalities regarding the 'day' during which students used the most amount of data. Moreover, some apps such as Instagram, WhatsApp, and Telegram were the most widely used which stands for the fact that the online behavior and data usage of Iranian EFL university students are largely centered around the purpose of daily communication and social exchange. As also supported by the results of the current study, it can be suggested that educational decision makers such as textbook developers, university officials, teachers etc. should design engaging strategies through which students as well as teachers spend as much time for study purposes and educational reasons as they spend for social exchange. Specifically, in case such AI-driven apps or means of learning can be run within the aforementioned social platforms. Such as using Instagram for learning various English language skills or components through AI. Overall, there seems to be the vital need for the use of technology and AI-based teaching and learning techniques in any educational context, Iran is also no exception. This is mainly because teachers' and students' use of emerging technologies can make a major contribution to the development of 21st-century practices in schools (Valtonen et al. 2021).

To the best knowledge of the researchers, the results of all the studies carried out so far support the use of computers and AI in any realm of education. However, the critical standpoint which has to be given considerable attention, is the caution needed to be taken when designing or picking a particular AI tool for learning language skills or language components. The issue lies in the fact that countless number of apps and AI tools have recently flooded the field of English Language Teaching

(ELT) and sticking to such tools and means does not always guarantee better language achievement or better teaching quality. This means, careful needs analysis has to be carried out regarding what learners are aimed to adopt, who they are, their individual differences, the content of the subject matter, the ages etc. in order to witness a better integration of them in the field of English language teaching and learning (Alhalangy & AbdAlgane, 2023; Marandi, 2011). This is mainly because, following the fad, merely because it is available out there, does not necessarily label us as professionals in the field of English language teaching or learning.

From a pedagogical standpoint, our study underscores the critical importance of cultivating awareness, among both learners and educators, regarding the merits and demerits of technology, particularly AI. Armed with such awareness, individuals can transform into independent, self-critical, and self-regulated participants, actively steering and monitoring their own teaching and learning processes. However, some aspects regarding the use of AI and technology within English teaching learning context should consistently be considered, such as the ease, technical aspects, and flexibility of such means (Utami et al., 2022).

6. CONCLUSION

Comparing the use of AI in education to other fields such as business & medicine, there are still few studies available. (Borges et al., 2020; Luckin & Cukurova, 2019). One of such reasons could be because many decision makers such as teachers, educators and textbook publishers resist the use of AI (EdTechXGlobal Report, 2016, as stated in Clik, I. et al., p. 619). Therefore, the results of the conducted research can act as the groundwork for further research in the area. One of the highlight outcomes of the study centers around the issue that though the participants used apps related to learning English skills and components, greater investment on the part of EFL students for learning English language skills and English components through means of AI have to be encouraged by EFL teachers and educators. This has also been claimed by other researchers in the field. For instance, Zhai and Wibowo (2023) reported that the use of AI in learning English as a Foreign Language (EFL) can effectively improve university students' English language skills such as reading, writing and the listening abilities.

Moreover, there is enormous potential for research on various aspects of the students' lifestyle which could be studied. Data science could help analyze and evaluate students' educational performance. Machine learning methods could help predict their study results. Also, different tools could be developed using AI methods which could help students through their study as a kind of peer or supplementary source. In addition, teachers could also benefit from AI-driven tools and assistants to help them with teaching and evaluating students.

The overarching implication of our study is the pressing necessity for the integration of technology and AI-driven educational techniques, an imperative not exclusive to Iran but applicable globally. In embracing emerging technologies, both teachers and students can significantly contribute to the evolution of 21st-century educational practices (Abbasi et al., 2021).

In conclusion, our research not only unravels pertinent insights into the online behaviors of Iranian EFL university students but also advocates for a paradigm shift in educational practices. The judicious use of technology and AI stands as an imperative for fostering a dynamic and adaptive learning environment in the 21st century.

7. FUTURE RESEARCH DIRECTIONS

The outcomes of the research could be advantageous for AI course developers and stakeholders in that they can locate the educational needs and necessities of EFL university students and design needs-based learning applications and tools which can be run through the online platforms that students have the habit of using them daily.

Additionally, gathering data from larger groups of participants majoring in other fields of study and drawing a kind of comparison among them would bring about thought-provoking insights for further research.

Besides, research can focus on teachers as the core scope of educational settings. Their knowledge of AI-based systems and strategies and their impact on students' success can also be investigated.

Also, using machine learning and artificial neural networks could extract and learn different features of students' online behaviors. These features could be for example the amount of time each student spends on his phone, the share of time for each application category, if s/he uses any educational applications, etc. Knowing so, the ML model could be able to predict the educational performance of the students.

Moreover, analyzing the role of learning-assistant applications in the students' educational performance could be studied using data science and AI methods. This could help know if those applications help the students with the improvement or cause them time and stress.

In summary, the future research outlined here not only seeks to refine our understanding of AI in educational contexts but also strives to make tangible contributions. These efforts aim to inform the development of targeted interventions, enhance pedagogical strategies, and ultimately propel the integration of AI into education to new heights.

ACKNOWLEDGMENT

The authors acknowledge the financial support by the German Federal Ministry for Education and Research (BMBF) within the project Open6GHub {16KISK003K}.

8. REFERENCES

Abbasi, M., Shahraki, A., & Taherkordi, A. (2021). Deep learning for network traffic monitoring and analysis (NTMA): A survey. Computer Communications, 170, 19-41.

Borges, A. F., Laurindo, F. J., Spínola, M. M., Gonçalves, R. F., & Mattos, C. A. (2020). The strategic use of artificial intelligence in the digital era: Systematic literature review and future research directions. *International Journal of Information Management,* 102225. https://doi.org/10.1016/j.ijinfomgt.2020.102225

Bulut, E., & Szymanski, B. K. (2015, June). Understanding user behavior via mobile data analysis. In 2015 IEEE International Conference on Communication Workshop (ICCW) (pp. 1563-1568). IEEE.

Cohen, I. L., Liu, X., Hudson, M., Gillis, J., Cavalari, R. N., Romanc-zyk, R. G., ... & Gardner, J. M. (2017). Level 2 Screening with the PDD Behavior Inventory: Subgroup Profiles and Implications for Differential Diagnosis. *Canadian Journal of School Psychology, 32*(3-4), 299-315. https:// doi. org/ 10. 1177/ 08295 73517 721127

Celik1, I., Dindar, M., Muukkonen, H. & Järvelä, S. (2022). The Promises and Challenges of Artificial Intelligence for Teachers: a Systematic Review of Research. *TechTrends* (2022) 66:616–630. https://doi.org/10.1007/s11528-022-00715-y

Dillenbourg, P. (2013). Design for classroom orchestration. *Computers & Education*, 69, 485–492. https:// doi. org/ 10. 1016/j. Compe du.2013.04.013

EdTechXGlobal. (2016). EdTechXGlobal report 2016—Global EdTech industry report: a map for the future of education and work. Retrieved from http:// ecosy stem. edtec hxeur ope. com/ 2016-edtech-report

Gaudioso, E., Montero, M., & Hernandez-Del-Olmo, F. (2012). Supporting teachers in adaptive educational systems through predictive models: A proof of concept. *Expert Systems with Applications, 39*(1), 621–625. https://doi.org/10. 1016/j.eswa.2011.07.052

Gayed, J. M., Carlon, M. K. J., Oriola, A. M., and Cross, J. S. (2022). Exploring an AI-based writing assistant's impact on English language learners. *Comp. Educ. Art. Intel.* 3:100055. doi: 10.1016/j.caeai.2022.100055

Haixiang, G., Yijing, L., Shang, J., Mingyun, G., Yuanyue, H., & Bing, G. (2017). Learning from class-imbalanced data: Review of methods and applications. Expert systems with applications, 73, 220-239.

Hrastinski, S., Olofsson, A. D., Arkenback, C., Ekström, S., Ericsson, E., Fransson, G., ... & Utterberg, M. (2019). Critical imaginaries and reflections on artificial intelligence and robots in post digital K-12 education. *Post digital Science and Education, 1*(2), 427-445. https://doi.org/10.1007/s42438-019-00046-x

Hsu, T.C., Chang, C. & Jen, T. (2023). Artificial Intelligence image recognition using self-regulation learning strategies: effects on vocabulary acquisition, learning anxiety, and learning behaviors of English language learners, Interactive Learning Environments, DOI: 10.1080/10494820.2023.2165508

Junaidi, J. (2020). Artificial intelligence in EFL context: rising students' speaking performance with Lyra virtual assistance. *Int. J. Adv. Sci. Technol. Rehabil., 29*, 6735–6741.

Klímová, B. and Ibna Seraj, PM (2023) The use of chatbots in university EFL settings: Research trends and pedagogical implications. *Front. Psychol.* 14:1131506.doi: 10.3389/fpsyg.2023.1131506

Ling, W. (2023). Artificial intelligence in language instruction: impact on English learning achievement, L2 motivation, and self-regulated learning. Front. Psychol., *14*. https://doi.org/10.3389/fpsyg.2023.1261955

Loncar, M., Schams, W., and Liang, J. S. (2023). Multiple technologies, multiple sources: trends and analyses of the literature on technology-mediated feedback for L2 English writing. *Comput. Assist. Lang. Learn. 36*, 722–784. doi: 10.1080/09588221.2021.1943452

Luckin, R., & Cukurova, M. (2019). Designing educational technologies in the age of AI: A learning sciences-driven approach. *British Journal of Educational Technology, 50*(6), 2824–2838. https:// doi. Org/10.1111/bjet.12861

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson Education.

Marandi, S. (2011). CALL 101: Some basics any CALL practitioner needs to know. *Roshd, 26* (1), 70-77.

McCarthy, J. (2004). *What is artificial intelligence*? Universal Publishers.

Nguyen, L. Q., and Le, H. V. (2023). Enhancing L2 learners' lexical gains via Quizlet learning tool: the role of individual differences. Educ. Inf. Technol. 28, 12143–12167. doi: 10.1007/s10639-023-11673-0

Parwez, M. S., Rawat, D. B., & Garuba, M. (2017). Big data analytics for user-activity analysis and user-anomaly detection in mobile wireless network. IEEE Transactions on Industrial Informatics, 13(4), 2058-2065.

Swiecki, Z., Ruis, A. R., Gautam, D., Rus, V., & Williamson Shaffer, D. (2019). Understanding when students are active-in-thinking through modeling-in-context. *British Journal of Educational Technology, 50*(5), 2346–2364. https:// doi. org/ 10. 1111/ bjet. 12869

Trinh, H. D., Giupponi, L., & Dini, P. (2018, September). Mobile traffic prediction from raw data using LSTM networks. In 2018 IEEE 29th annual international symposium on personal, indoor and mobile radio communications (PIMRC) (pp. 1827-1832). IEEE.

Trinh, H. D., Bui, N., Widmer, J., Giupponi, L., & Dini, P. (2017, October). Analysis and modeling of mobile traffic using real traces. In 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC) (pp. 1-6). IEEE.

Utami, P., Hartanto, R. and Soesanti, I. (2022). A Brief Study of The Use of Pattern Recognition in Online Learning: Recommendation for Assessing Teaching Skills Automatically Online Based. *ELINVO (Electronics, Informatics, and Vocational Education), 7*(1),48-62

Valtonen, T., Hoang, N., Sointu, E., Näykki, P., Virtanen, A., Pöysä-Tarhonen, J., Häkkinen, P., Järvelä, S., Mäkitalo, K., & Kukkonen, J. (2021). How pre-service teachers perceive their 21st-century skills and dispositions: A longitudinal perspective. *Computers in Human Behavior, 116*, 106643. https:// doi. org/ 10. 1016/j. chb. 2020.106643

Verkasalo, H. (2010, June). Analysis of smartphone user behavior. In *2010 Ninth International Conference on Mobile Business and 2010 Ninth Global Mobility Roundtable (ICMB-GMR)* (pp. 258-263). IEEE.

Vojnovic, M. (2008, March). On mobile user behaviour patterns. In 2008 IEEE International Zurich Seminar on Communications (pp. 26-29). IEEE.

Wirth, R., & Hipp, J. (2000, April). CRISP-DM: Towards a standard process model for data mining. In Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining (Vol. 1, pp. 29-39).

Xu, F., Li, Y., Wang, H., Zhang, P., & Jin, D. (2016). Understanding mobile traffic patterns of large scale cellular towers in urban environment. IEEE/ACM transactions on networking, 25(2), 1147-1161.

Yang, J., Qiao, Y., Zhang, X., He, H., Liu, F., & Cheng, G. (2014). Characterizing user behavior in mobile internet. *IEEE transactions on emerging topics in computing*, *3*(1), 95-106.

Zhai, C. , and Wibowo, S. (2023). A systematic review on artificial intelligence dialogue systems for enhancing English as foreign language students' interactional competence in the university. *Computers and Education: Artificial Intelligence*,*4*. https://doi.org/10.1016/j.caeai.2023.100134