Rodrigo, M. M. T. et al. (Eds.) (2021). Proceedings of the 29th International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

From Hello to Bye-Bye: Churn Prediction in English Language Learning App

Daevesh SINGH*, Rumana PATHAN, Gargi BANERJEE & Ramkumar RAJENDRAN
IDP in Educational Technology, Indian Institute of Technology Bombay, India
*daeveshsingh@iitb.ac.in

Abstract: Mobile phones and apps have changed the landscape of e-learning and have revolutionised the way people learn a second language by facilitating anytime-anywhere learning, game-based resources and socially interactive learning activities. Despite these features and affordances, these language learning apps suffer a fate of high churn rates. In this paper, we examined the churning behaviour of learners in the context of a language learning app called Hello English. We applied descriptive analytics to analyse the behavioural differences between churners and non-churners and studied their interaction with the app to early-predict churning behaviour. Our findings indicate that non-churners interact with the mobile app more frequently compared to churners. Also, the trained machine learning classifiers can predict learner churning behaviour with a high recall value (0.824) and F1 (0.778). This churn detection will enable the app developers to provide intervention for learner retention.

Keywords: Language learning, churn rate, interaction behaviour, descriptive analytics

1. Introduction

The popularity of mobile phones has immensely increased in the last two decades due to various factors like portability, unobtrusiveness, ease of use, affordability, and personal adaptation (Chih & Shih, 2011; Sharples, 2000). The installation of applications (apps) on these devices further enhances their usability and capability and have made them a preferred tool for entertainment, business and learning (Godwin, 2011; Papadakis et al., 2020). These apps, which were initially designed for assisting productivity (emails, calendar etc.), are now being explored in other areas such as gaming, online shopping, social media, and education (Papadakis et al., 2020). The apps developed for educational purposes provide enormous educational resources, interactive activities, challenge-based learning through games, puzzles and collaborative activities, which have changed the landscape of e-learning, especially in the field of Mobile-Assisted Language Learning (MALL) (Burston, 2014; Chih & Shih, 2011; Kim & Kwon, 2012; Miangah & Nezarat, 2012).

MALL has revolutionised the way people learn a second language by facilitating anytime-anywhere learning, collaborative learning activities, socially interactive and game-based resources (Kim & Kwon, 2012). Despite these features and affordances, MALL faces a plethora of challenges, such as mismatch between pedagogy and technology resulting in fragmented language practice (Pareja et al., 2013), focus on receptive language skills denying the opportunities for socio-cognitive activities (Kim & Kwon, 2012), and low motivation leading to high attrition rates (Elaish et al., 2019). These issues point out the need to employ behavioural analytics to understand how learners interact with these language learning apps (LLAs). Analysing the learner behaviour can help us throw light upon the various features of the LLAs that learners are exploring. For example, the time they are devoting to use these LLAs, how they interact with other learners in LLA, and their progress. Such information about the learner can help us in addressing the issues stated above. Although these issues are crucial to address, this study aims to only focus on analysing learner behaviour to explore the churn behaviour of the learners. Analysing churn behaviour is important as research shows that educational apps have the highest user churn rates among other mobile apps (Pham & Wang, 2016).

App users are said to be churned when they become inactive for a certain period or altogether stop using the app by uninstalling it. The method of identifying churners is called churn prediction or churn behaviour prediction. In our study we have defined churn as week long period of inactivity. Churn prediction is crucial as it will help the app developers devise a strategy to bring back the churners. This is important since churning has a detrimental impact on learning, as the learner's engagement with the app for a more extended period is crucial for learners (Kim & Kwon, 2012; Burston, 2014).

In this paper, we explored the interaction behaviour of 700 randomly chosen learners of an English language learning app called Hello English (HE). The main goal of this paper is to:

1. Find the differences in the interaction behaviour of the churners and non-churners.
2. Use this interaction behaviour to train Machine Learning (ML) models to predict churn behaviour.
3. Early predict the churn behaviour using the best model.

To find the differences in the interaction behaviour of the churners and non-churners, we applied descriptive analytics and compared the average activities per day for both the groups. We used ML models and trained these models on the interaction data of 7 days (observation period) to predict the churn behaviour. Finally, we performed early prediction using ML algorithm. We found that 1) non-churners, on average, performed more activities as compared to churners, 2) the ML model (LR) used for prediction had high recall (0.824) and F1 (0.778) value which is comparable to churn prediction models used in other domains, and 3) this model could early predict the churning behaviour of learners with a recall (0.704) and F1 (0.675) value from the fourth day onwards.

In the sections that follow, Section 2 examines the existing literature on churn prediction in mobile apps, helping us operationalise churn in Section 3. Section 4 provides the details of the learning environment. Section 5 and 6 describes the details of the methods used in this study and the results obtained. Finally, in Section 7, we present the conclusion and limitations of the study.

2. Background and Related Work

Existing research on user retention for mobile apps has progressed in three primary directions, they are:

1. Why learners uninstall apps
2. Finding solutions to enhance user retention
3. interpret the usage behaviour of churners

In the first direction, research studies have found a host of reasons that contribute to the low user retention for mobile apps. In one such study, Ickin et al. (2017) surveyed users and identified reasons that affected user retention like intrusive advertisements, boredom, lack of updates of the app, high memory allocation, and low popularity in their friend circle.

In the second direction, research focused on finding the solutions to enhance user retention. These studies examined many factors affecting user engagement. They suggested that integrating gamified components in the learning design, experiential learning, incorporating digital leaderboards and badges, built-in social media, and frequent release of updates leads to an increase in engagement (Pechenkina et al., 2017; Pham & Wang, 2016).

In the third direction, research studies are focused on analysing the usage behaviour of churners for making app recommendations (Shang et al., 2017), providing scaffolding to learners (Pishtari et al., 2019), or developing models that predict the churning. Although in academia, behavioural analytics is ubiquitous and is a subject of widespread interest ranging from learning, affect detection, dropout and engagement (Nishane et al., 2021; Rajendran et al., 2013, 2018; Pathan et al., 2019, 2020). We did not find any research study focused on developing models for predicting churn in an LLA to the best of our knowledge. Thus, despite mobile educational apps having the lowest retention rates (Pham & Wang, 2016), there is a paucity of literature that explores behaviour patterns of churners in the context of mobile educational apps. In this paper, we focus on interpreting the usage behaviour of churners in the context of a commercial English language learning app.

Although several LLA's exist, we did not find any research studies related to churn prediction. Thus, we reviewed the research articles on churn prediction in non-educational apps to understand churn and the different features used for churn prediction. The study of Hadiji et al. (2014) predicted churn in mobile games by analysing data of 50,000 randomly selected players from 5 different mobile games up to a specific date (cutoff date). The authors' defined churn in two ways, i.e. "hard churn" and "soft churn". In the case of hard churn, a player without any session after the cutoff date was termed as the churner, and in the case of soft churn, a player with a low number of sessions was considered churned. The feature set included universal features (game-independent features) such as the number of sessions, days, the time elapsed since the last session, average playtime per session and average time between the sessions. They also incorporated four "economy features", such as the number of purchases. The four ML classifiers employed were LR, Neural Networks (NN), Naive Bayes (NB), and Decision Tree (DT). DT predicted churn with high accuracy for some games. However, a relatively low accuracy was reported in games with maximum players who churned after playing a few sessions.

Similarly, the study by Kim et al. (2017) is based on churn prediction of three mobile and online casual games using log data of 193,443 players. The authors have defined churn using observation period and churn prediction period. The user is considered churned if they remain inactive in the churn prediction period after playing at least one session in the observation period. The models included three traditional ML algorithms LR, Gradient Boost (GB), and Random Forest (RF), along with two Deep learning algorithms, i.e., Long Short Term Memory (LSTM) and Convolutional Neural Networks (CNN). A total of ten universal features like the number of sessions played, the time between first and last session, the time between the consecutive sessions etc., along with game-specific features like the number of purchases made, the price paid etc., were extracted. The results show that LSTM outperformed other models in churn prediction for one game with an AUC (i.e. area under the receiver operating characteristic curve) of 0.792. In another game, gradient boosting performed best with an AUC of 0.728. Likewise, in the third game, LR and GB outperformed others with an AUC of 0.842.

In another study, Runge et al. (2014) predicted churn for high-value players of two social games with a data set of 18445 players. In this study, the players are said to churn if they remain inactive for more than 14 days. The authors categorised the data into three types, i.e. in-game data (log data of the player during the play session), revenue data (revenue generated by the player), and player profile data (profile data of the player). The feature set includes features like days in-game, last purchase, days since last purchase etc. The binary classification of churners and non-churners was done using four ML algorithms, viz. NN, SVM, NB, and LR. NN outperformed other models in both games with an AUC value of 0.815 in one game and 0.930 in another.

To summarise, the research papers presented above are based on churn prediction of non-educational app users by analysing the interaction data generated over a period of time. The way churn is operationalised in these studies vary due to the difference in the context. Similarly, the feature sets used in these studies are universal (such as the number of sessions) and app-specific (such as the frequency of using a particular resource). Likewise, the best performing ML models are different in each of these studies. Hence, defining churn and developing a feature set in the context of LLA is a significant research gap that this paper addresses. Since there is no single ML model that outperforms other models in the literature, we are motivated to explore multiple ML models.

3. Defining Churn

There are several ways in which churn is operationalised in the literature, the simplest one being the app's uninstallation. However, this way of defining churn seems inappropriate in the educational context since existing studies have identified the positive impacts of apps on language learning (Burston, 2014; Kim & Kwon, 2012). Therefore, the learning gain of the learners is dependent on the time spent on the app. Hence, being inactive on the app is equivalent to uninstalling the app. As a result, we define churn as a specific period of inactivity.

We did not find any particular method in the literature informing us about choosing an optimal churn period. Also, there is a lack of uniformity in the duration of the churn period reported in the literature. Hence, we define churn as a week-long period of inactivity because our target audience is school-going children (i.e. 14-18 years). Thus, a period of one week will ensure that we monitor them on both working days and weekends. This definition also enabled us to stratify the learners into two groups, i.e. churners and non-churners.

The churn prediction problem requires the interaction data of the learner with the app. As a result, we use seven days of interaction data of the learner, starting from the first day the learner interacts with the app, followed by seven calendar days. This seven day period is called the observation period. For instance, if a learner installs the app on 01 August and interacts with the app on the same day, the period from 01-07 August is considered an "observation period". The seven day period after the observation period (i.e. 08-14 August) is termed the churn period. If a learner does not interact with the app in this period, they will be called churners. A non-churner is a learner who interacts at least once with the app during the churn period.

4. Learning Environment: Hello English

Hello English (HE) is an English learning application specifically designed for second language learners in India to learn English. The application, launched in October 2014, encompasses learning activities that are interactive, personalised, and contextualised to local learning. It encapsulates all four aspects of language acquisition: reading, writing, listening and speaking. The learning activities in HE are equipped with advanced voice recognition technology that allows learners to interact with the app and hold real-life conversations. The interactive learning activities consist of games (individual and multiplayer), activities to practise speaking and contextual learning tools that leverage news, sports, and entertainment. Moreover, most of the app's features can be accessed offline, which saves data expenses for learners and helps make learning a seamless experience.

As mentioned above, learners can perform several activities in HE, which can be broadly grouped under eight categories. That is, learn a lesson (L), practise reading (PR), practice speech (PS), play a game (PG), play reward activity (RA), take a test (TT), respond to a quiz (RQ), and seek help (SH). The app also captures learner data, such as the number of coins for each activity. The detailed description of each category of action is narrated in Table 1.

Table 1. *Categorising Learner's Interaction in Hello English Mobile App and their Description*

	Category of actions	Description
1	Learn (L)	Learner accesses a lesson that involves all the four skills of language learning, i.e. listening, speaking, writing and reading
2	Practice reading (PR)	The learner reads various contextual articles such as current news, articles on entertainment, sports, etc.
3	Practice speech (PS)	The learner practices conversation using inbuilt voice recognition technology and holds real-life conversations.
4	Play a game (PG)	The learner plays various games such as rearranging jumbled words, games that reinforce learning from the preceding lesson etc.
5	Reward activity (RA)	The learner earns coins as rewards for solving different activities.
6	Take a test (TT)	The learner takes a summative assessment after learning a certain number of lessons. These are mainly grammar and vocabulary tests.
7	Respond to a quiz (RQ)	Learner responds to quiz
8	Seek help (SH)	The learner wants to connect with an educator or wants to share his achievement score on social media
9	Number of Coins (C)	It refers to the coins earned after the successful completion of an activity.

5. Research Methodology

The research goal of this paper is to:

1. Identify the difference in the interaction behaviour of churners and non-churners.
2. Use the learner's interaction behaviour to predict whether the learner will churn or not.

3. Reduce the observation period up to 4 days to early-predict churn.

In this section, we first describe the dataset and provide information about the data collection process. The following subsection gives details about data pre-processing procedures used in our analysis. The last subsection informs about the labelling of data.

5.1 Dataset

The dataset analysed in this paper for churn prediction is obtained from randomly selected 700 learners of HE. These learners installed the application between July 2019 and September 2019 and were aged between 14 to 18. At the time of app installation, the learner's consent was sought, and they were informed about all the terms and conditions they need to agree to for using the app. We collected the data comprising all the activities (as described in table 1) completed by the learners during their observation period.

5.2 Data Pre-processing and Analysis

The data consisted of 700 learners who interacted with the language learning application for a minimum of two sessions. The number of days for which they interacted with the app is different for different learners, varying from 1 to 90 days. Out of these 700 learners, more than 90 per cent of learners had more than ten sessions. In the data-preprocessing stage, we discarded the possibly erroneous data (e.g. sessions that lasted for days). We then normalised the data by rescaling the features to the range between 0 and 1 using max-min normalisation. This normalisation is crucial as the data had varying scales. Finally, we used this processed data for extracting features.

To perform descriptive and predictive analytics, labelling learners was crucial. So, we provided labels to these learners as per the definition of the churn described in section 3. The churners were learners with zero sessions in the churn period. Out of the 700 learners, 347 were churners, and the remaining 353 were non-churners.

6. Results and Discussion

This section presents the result of the study by 1) providing a comparison of the interaction behaviour of churners and non-churners, 2) describing features extracted from the interaction data and performance of ML models employed to predict learner churn behaviour, and finally, 3) reporting the best performing model to early to predict learner churn behaviour.

6.1 Comparing Interaction Behaviour of Churners and Non-Churners

In order to compare the interaction behaviour of churners and non-churners, we computed the average number of actions per day per learner in each category of activities. For instance, we calculated the total number of 'practice game' activities per day per learner by dividing the total number of practice games by the product of the number of active days and the number of learners in that group. By ‘active days’, we mean days on which learners perform at least one activity. Figure 1 represents the frequency distribution of the average actions per day per learner for churners and non-churners. The frequencies represent the average number of actions per day per learner in each group. On average, churner has completed only 7.04(SD=3.5) number of activities per day compared to non-churners who have completed 9.80(SD=5.9) number of activities per day. Similarly, churners have accessed only a 3.23(SD=2.68) number of lessons per day on average as compared to 4.27 (SD=4.7) number of lessons by non-churners. We obtained similar results for practice games as well.

Overall, non-churners have accessed more activities per day on average than churners in all the action categories. Our following subsection used these differences in interaction behaviour to predict the learner's app uninstallation behaviour.

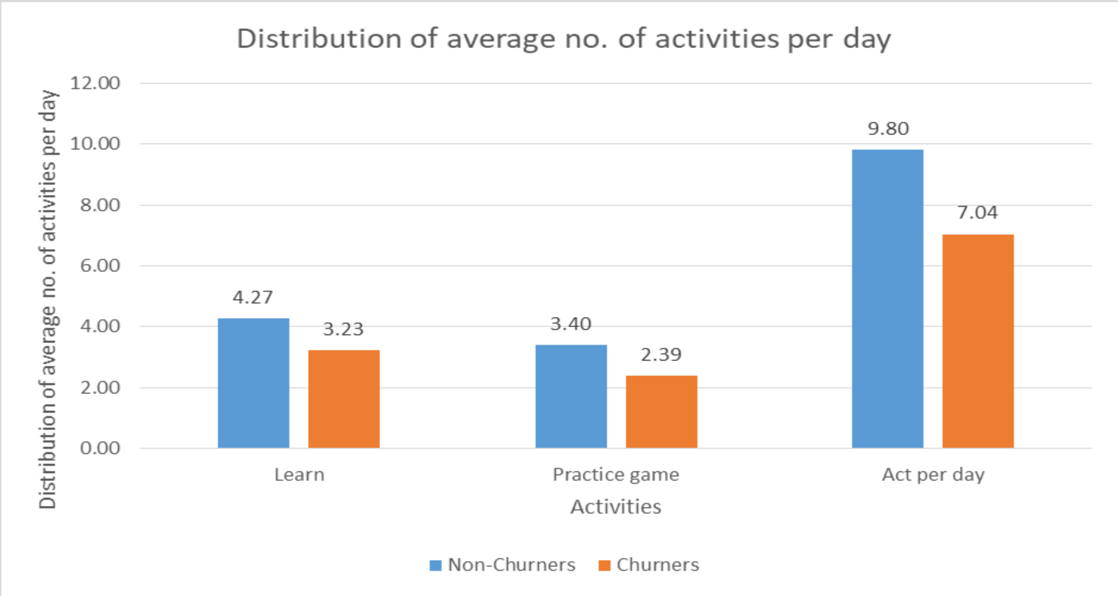

Figure 1. Frequency Distribution of per day per User Activities for "Churners" and "Non-churners"

6.2 Predicting Learner's Churn Behaviour

To predict the churn behaviour, we computed the features using the actions categorised in table 2. Table 2 describes all the features developed from the log interaction data of learners during the observation period.

Table 2. *List of Features Extracted from Interaction Data and their Description*

	Category of actions	Description
1	Active_days	Days in the observation period on which the learner did at least one activity.
2	Inactive_days	Days in the observation period on which the learner did not interact with the app.
3	Lesson/day	The number of lessons learned is divided by the number of active days.
4	PR/day	The number of reading lessons practised divided by the number of active days.
5	PS/day	The number of speech lessons practised is divided by the number of active days.
6	PG/day	The number of games played is divided by the number of active days.
7	RQ/day	The number of responses to the quiz divided by the number of active days.
8	RA/day	Number of reward activities divided by number of active days.
9	C/day	Number of coins earned divided by number of active days
10	A/day	Total number of activities divided by the number of active days.
11	C/activity	Number of coins earned per activity.

We developed these features along two dimensions, namely the learner activity dimension and the reward activity dimension. For example, in the learner activity dimension, we extracted features such as number of lessons learned per day (L/day), number of reading lessons practised per day (PR/day), number of speech lessons practised per day (PS/day), number of games played per day (PG/day), number of responses to quiz per day (RQ/day), and the total number of activities per day.

Similarly, we considered features such as reward activity per day (RA/day), number of coins per day (C/day), and the number of coins per activity (C/activity) as features emerging from the reward dimension. We extracted these features for both groups and developed a classifier using three major classification algorithms proven efficient for small datasets (Sharma & Paliwal, 2015). The classification algorithms used in this research are RF, NB, and LR.

Table 3 shows the result of the model's prediction using 10-fold cross-validation on three different classifiers. We stratified the data at the student level. The results indicate that the LR algorithm performed better than other algorithms in terms of all the evaluative metrics, i.e. F1 score (0.778), precision (0.800), recall (0.824) and accuracy (0.824).

Table 3. *Performance of Different ML Models in Predicting Churners using the Hello English App Users' 7-Days Interaction Data*

Models	F1	Precision	Recall	Accuracy
Random Forest	0.746	0.730	0.777	0.801
Naïve Bayes	0.740	0.772	0.720	0.720
Logistic Regression	0.778	0.800	0.824	0.824

6.3 Early Prediction of Learners Churn Behaviour

The LR model outperformed the other models in the churn prediction task using seven day observation period. We, therefore, used the LR algorithm for early prediction. This model also used 10-fold cross-validation and student-level stratified data. The performance of the LR algorithm using learners' interaction data with different duration of the observation period ranging from day 4 to day 6.

Since we are making an early prediction of churning, it is crucial to identify the set of learners who have churned. Recall as a measure signifies the proportion of all churners that the model accurately predicted. Therefore, we emphasise recall score because it is preferable not to miss any user about churn, even if the model flags some non-churners as churners. Similarly, the F-score is used to measure the model's performance by considering both precision and recall. It is the harmonic mean of precision and recall. Hence, we plotted the recall and F-score values for different days, i.e. days 4-6 (refer to figure 2 and table 4), to understand how early we can make a reasonable prediction of the churning, comparable to other works in different domains. We did not further reduce the observation beyond this as then the learner interaction data would reduce, which might make the model pick up noise.

Table 4. *The Performance of the Logistic Regression Model using Learners' Interaction Data with Different Duration of the Observation Period Ranging from Day 4 to Day 6*

Days	F1	Precision	Recall	Accuracy
4	0.675	0.685	0.704	0.704
5	0.726	0.750	0.809	0.808
6	0.753	0.773	0.813	0.812

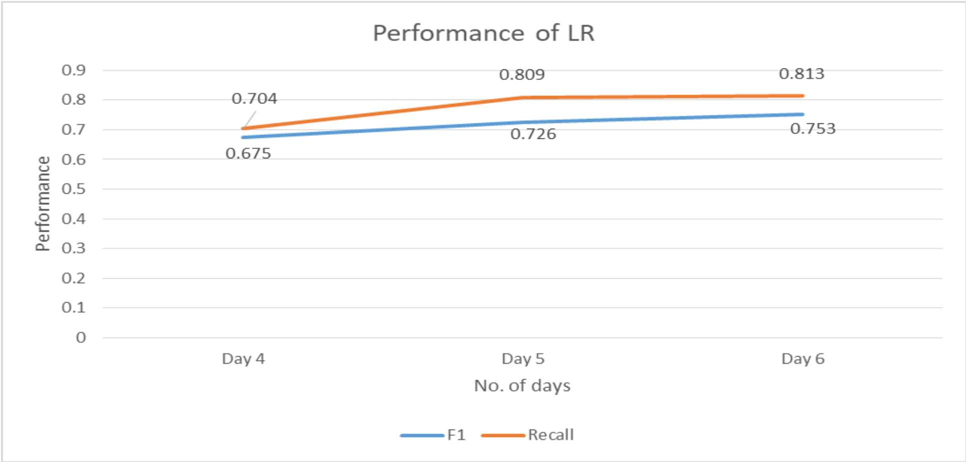

Figure 2. The Performance (F1 Score and Recall) of the Logistic Regression Model using Learners' Interaction Data with Different Duration of the Observation Period (Day 4 To Day 6).

We observed that from day four onwards, we could predict the churning behaviour of learners with a high recall value (0.704) and F1 = 0.675. This result indicates, with data from the first four days, we can predict the learners who will churn and develop focused activities to persist the learners to continue using the app.

7. Conclusion and Limitations

This article is an attempt to make use of learner interaction data for predicting churn behaviour. The significance of our work lies in the fact that it is the first time an educational app is used for churn prediction. Although the study explores the interaction behaviour of learners in a specific LLA called "Hello English", many of the features extracted are generic (features that are not app-specific such as the number of sessions, number of days the app is used etc.). They hence could be easily extended to other LLAs. We analysed the interaction behaviour of churners and non-churners at three levels of granularity:

1. comparing frequencies of various actions performed by them
2. predicting learners churning behaviour using interaction data of 7 days
3. early prediction of churning using different observation periods starting with four days periods.

We found that non-churners accessed a higher number of activities as compared to churners. We could also predict learners churning behaviour using the LR classification algorithm with a reasonably good recall and F-score. Similarly, we have also shown that early prediction is possible with a four-day observation period, making our work significant. Early detection of these learners allows mobile app developers to target a specific learner group and implement engagement measures to retain the learner group.

Although early prediction allows us to predict learner behaviour decently, this approach does not inform us why the learners uninstalled the app. Our future goal is to understand the churners better by identifying why learners uninstall or become inactive for a longer duration. We would like to perform qualitative studies with focussed interviews to identify the reasons for uninstallation and provide informed decisions to the learners. Also, the procedure described in this paper can be further enhanced by using other classification algorithms such as deep neural networks with more data from learners to improve the models' performance further.

Acknowledgements

We would like to thank CSF for funding this project and also Hello English Team for providing data.

References

Burston, J. (2014). MALL: The pedagogical challenges. *Computer Assisted Language Learning*, *27*(4), 344-357.

Chih-Ming, C., & Shih-Hsun, H. (2011). Personalised intelligent mobile learning system for supporting effective english learning [J]. *Educational Technology & Society*, (3), 153-180.

Elaish, M. M., Shuib, L., Ghani, N. A., & Yadegaridehkordi, E. (2019). Mobile English language learning (MELL): A literature review. *Educational Review*, *71*(2), 257-276.

Godwin-Jones, R. (2011). Mobile apps for language learning. *Language Learning & Technology*, *15*(2), 2-11.

Hadiji, F., Sifa, R., Drachen, A., Thurau, C., Kersting, K., & Bauckhage, C. (2014, August). Predicting player churn in the wild. In *2014 IEEE Conference on Computational Intelligence and Games* (pp. 1-8). Ieee.

Kim, H., & Kwon, Y. (2012). Exploring smartphone applications for effective mobile-assisted language learning. *Multimedia-Assisted Language Learning*, *15*(1), 31-57.

Kim, S., Choi, D., Lee, E., & Rhee, W. (2017). Churn prediction of mobile and online casual games using play log data. *PloS one*, *12*(7), e0180735.

Miangah, T. M., & Nezarat, A. (2012). Mobile-assisted language learning. *International Journal of Distributed and Parallel Systems*, *3*(1), 309.

Nishane, I., Sabanwar, V., Lakshmi, T.G., Singh, D., Rajendran, R., Learning about learners: Understanding learner behaviour in software conceptual design TELE., To be appeared in the Proceedings of *International Conference on Advanced Learning Technologies (ICALT), 2021*

Papadakis, S., Vaiopoulou, J., Kalogiannakis, M., & Stamovlasis, D. (2020). Developing and exploring an evaluation tool for educational apps (ETEA) targeting kindergarten children. Sustainability, 12(10), 4201.

Pareja-Lora, A., Arus, J., Mart ın Monje, E., Read, T., Pomposo-Yanes, L., Rodrıguez-Arancon, P., Barcena, E. (2013). Toward mobile assisted language learning apps for professionals that integrate learning into the daily routine. In L. Bradley & S. Thou€esny (Eds.), 20 years of EUROCALL: Learning from the past, looking to the future. Proceedings of EUROCALL 2013 EUROCALL Conference, Evora, Portugal (pp. 206–210). Dublin: Research Publishing

Pathan, R., Rajendran, R., & Murthy, S. (2020). Mechanism to capture learner's interaction in VR-based learning environment: design and application. Smart Learning Environments, 7(1), 1-15.

Pathan, R., Shaikh, U., & Rajendran, R. (2019, December). Capturing learner interaction in computer-based learning environment: design and application. In 2019 IEEE Tenth International Conference on Technology for Education (T4E) (pp. 146-153). IEEE.

Pechenkina, E., Laurence, D., Oates, G., Eldridge, D., & Hunter, D. (2017). Using a gamified mobile app to increase student engagement, retention and academic achievement. *International Journal of Educational Technology in Higher Education*, *14*(1), 1-12.

Pham, P., & Wang, J. (2016, October). Adaptive review for mobile MOOC learning via implicit physiological signal sensing. In *Proceedings of the 18th ACM international conference on multimodal interaction* (pp. 37-44).

Pishtari, G., Rodríguez-Triana, M. J., Sarmiento-Márquez, E. M., Terasmaa, J., Kori, K., Kangur, M., ... & Puusepp, L. (2019, September). An overview of learning design and analytics in mobile and ubiquitous learning. In *International Conference on Web-Based Learning* (pp. 312-319). Springer, Cham.

Rajendran, R., Iyer, S., Murthy, S., Wilson, C., & Sheard, J. (2013). A theory-driven approach to predict frustration in an ITS. *IEEE transactions on learning technologies*, *6*(4), 378-388.

Rajendran, R., Munshi, A., Emara, M., & Biswas, G. (2018). A temporal model of learner behaviors in OELEs using process mining. In *Proceedings of ICCE* (pp. 276-285).

Runge, J., Gao, P., Garcin, F., & Faltings, B. (2014, August). Churn prediction for high-value players in casual social games. In *2014 IEEE Conference on Computational Intelligence and Games* (pp. 1-8). IEEE.

S. Ickin, K. Petersen, and J. Gonzalez-Huerta, "Why do users install and delete apps? a survey study," in International Conference of Software Business. Springer, 2017, pp. 186–191.

Shang, J., Wang, J., Liu, G., Wu, H., Zhou, S., & Feng, Y. (2017, November). App Uninstalls Prediction: A Machine Learning and Time Series Mining Approach. In *International Conference on Neural Information Processing* (pp. 514-522). Springer, Cham.

Sharma, A., & Paliwal, K. K. (2015). Linear discriminant analysis for the small sample size problem: An overview. International Journal of Machine Learning and Cybernetics.Sharma, A. and Paliwal, K. K., "Linear discriminant analysis for the small sample size problem: An overview," International Journal of Machine Learning and Cybernetics, 2015.

Sharples, M. (2000). The design of personal mobile technologies for lifelong learning. *Computers & education*, *34*(3-4), 177-193.