Predicting Gaps in Usage
in a Phone-Based Literacy
Intervention System

Rishabh Chatterjee, Michael Madaio$^{(\boxtimes)}$, and Amy Ogan

Carnegie Mellon University, Pittsburgh, PA 15213, USA
rishabhc@andrew.cmu.edu, {mmadaio,aeo}@cs.cmu.edu

Abstract. Educational technologies may help support out-of-school learning in contexts where formal schooling fails to reach every child, but children may not persist in using such systems to learn at home. Prior research has developed methods for predicting learner dropout but primarily for adults in formal courses and Massive Open Online Courses (MOOCs), not for children's voluntary ed tech usage. To support early literacy in rural contexts, our research group developed and deployed a phone-based literacy technology with rural families in Côte d'Ivoire in two longitudinal studies. In this paper, we investigate the feasibility of using time-series classification models trained on system log data to predict gaps in children's voluntary usage of our system in both studies. We contribute insights around important features associated with sustained system usage, such as children's patterns of use, performance on the platform, and involvement from other adults in their family. Finally, we contribute design implications for predicting and supporting learners' voluntary, out-of-school usage of mobile learning applications in rural contexts.

Keywords: Machine learning · Dropout · Out-of-school learning

1 Introduction

Access to literacy is critical for children's future educational attainment and economic outcomes [13], but despite an overall rise in global literacy rates, these gains have not been evenly distributed [40]. Educational technologies may help supplement gaps in schooling in low-resource contexts [6,30,32]. However, given that many educational technologies are used in schools [51], children in agricultural communities who are chronically absent from school (e.g., [30]), may be

further denied access to technologies to supplement their learning unless learning technologies are available for use at home (as in [50]).

Côte d'Ivoire is one such context. While enrollment has risen drastically and many more children have access to schooling, nearly a fifth of rural fifth graders are not yet able to read a single word of French (the official national language) [14] and adult literacy rates stand below 50% [25]. Through multiple studies in a years-long research program, we investigated families' beliefs and methods for supporting literacy at home and their design needs for literacy support technology [28], and used these findings as design guidelines to develop an interactive voice response (IVR) literacy system for fostering French phonological awareness [27]. Then, to investigate how and why children and their families adopt and use such a system over several months at their homes, we deployed our IVR system, Allô Alphabet, in a series of studies of increasing size and duration, in 8 rural villages in Côte d'Ivoire [27,29]. We found that there was high variance in the consistency of children's use of the system, with some children who did not access the lessons for several weeks or months at a time [29].

In this paper, in order to understand whether we can predict (and perhaps, ultimately prevent) such gaps before they occur, we explore the efficacy of using system log data to predict gaps in children's system usage. We evaluate the efficacy of multiple models to predict usage gaps for two separate longitudinal deployments of Allô Alphabet and identify features that were highly predictive of gaps in usage. We contribute insights into features that contribute to gaps in usage as well as design implications for personalized reminders to prompt usage for educational interventions in out-of-school contexts. This work has contributions for educational technology usage prediction, as well as for mobile literacy systems more broadly.

2 Related Work

2.1 Educational Technology Used for Out-of-school Learning

While there is prior literature on the use of educational technologies in low-resource contexts [19,34,50], existing solutions are often deployed in schools, where children's use of devices may be controlled by the teacher [39,51]. Given that children in agricultural contexts may have limitations in their ability to consistently access and attend formal schooling [30], there is a need for out-of-school educational technologies for children. Some designers of mobile learning applications suggest children will use their applications to develop literacy skills [17,18,22]. However, as Lange and Costley point out in their review of out-of-school learning, children learning outside of school often have a choice of whether to engage in learning or not—given all of the other options for how to spend their time—a choice which may lead to gaps in their learning [24].

2.2 Predicting Usage Gaps in Voluntary Educational Applications

There is an abundance of prior work on predicting dropout to increase student retention in formal educational contexts like colleges [23,47,48]. Some work

has leveraged Machine Learning (ML) to identify predictors of adult learners' dropout from courses, as in work with English for Speakers of Other Languages (ESOL) courses in Turkey [7]. In addition to this work on predicting dropout from in-person courses, prior work has leveraged ML to identify predictors of dropout from Massive Open Online Courses (MOOCs) [4,33,36,45,53] and distance learning for adult learners [2,12,19,49]. Across these studies, a combination of social factors, like age, finances, family and institutional involvement, etc., and system usage data, like correctness, frequency, response log, etc. were found to be predictive of dropout.

While this is informative, much of this prior work is targeted towards distance learning or use of e-learning portals as part of formal instruction, not informal, out-of-school learning at home. Additionally, the type of learners is different— the majority of MOOC learners are between 18 and 35 years old [9], while we are focusing on children as users, who may have less-developed metacognitive abilities for planning and sustaining out-of-school learning. It thus remains to be seen what factors are useful for predicting gaps in children's literacy education with out-of-school use of learning technology. In particular, we are interested in system usage features as those are more easily and automatically acquired than socio-economic data.

Although there is a dearth of research on predicting gaps in children's usage of educational mobile applications, there is a rich legacy of research on mobile app usage prediction more broadly, primarily for adults (e.g., [20,31,43,44]). In both educational and non-educational application use, the engagement is voluntary, initiated by the user, and designers of such systems want to increase usage and retention. Prior research on *churn* prediction in casual and social gaming applications used machine learning models like Support Vector Machines (SVM) and Random Forests (RF) to model system usage. RF is an ensemble learning method, a category of model that has shown good performance for these predictions [37,42]. *Churn* is defined as using an application and then not continuing to use it after a given period of time [20]. Churn prediction allows systems to develop interventions, like reminders or nudges, which are positively related to increasing user retention [31,43,52]. However, there remain differences between casual and social mobile games and educational mobile applications, including the motivation to use the system and the nature of the data. This leads us to investigate the following research questions:

RQ1: Can we use system interaction data to predict gaps in children's usage of a mobile-based educational technology used outside of school in rural contexts?

RQ2: Which features of the users' interaction log data are most predictive of gaps in system usage of a mobile educational technology?

RQ3: How well does this usage gap prediction approach continue to perform for a replication of the same study in similar contexts?

3   Methodology

3.1   Study Design

This study is part of an ongoing research program [14, 27–29] to support literacy in cocoa farming communities, conducted by an interdisciplinary team of American and Ivorian linguists, economists, sociologists, and computer scientists, in partnership with the Ivorian Ministry of Education since 2016, and approved by our institutional review boards, the Ministry, and community leaders. Based on design guidelines identified through co-design research with children, teachers, and families [28], we developed Allô Alphabet, a system to teach early literacy concepts via interactive voice response (IVR) accessible on low-cost mobile devices ubiquitous in the context (described in more detail in [27, 29]). When a user placed a call to the IVR system, they heard a welcome message in French, an explanation of the phonology concept to be taught in that lesson, and were given a question. For each question, the system played a pre-recorded audio message with the question and response options. Students then pressed a touchtone button to select an answer and received feedback on their responses. If incorrect, they received the same question again with a hint, otherwise a selection of the next question was made based on their level of mastery of the concepts.

In this paper, we use data from two deployments of Allô Alphabet. In the first deployment (Study 1), we deployed Allô Alphabet with nearly 300 families with a child in grade CM1 (mean age = 11 years, SD = 1.5) in 8 villages in Côte d'Ivoire for 16 weeks, beginning in February 2019 [29]. Then we deployed it again in a larger randomized controlled trial with 750 children of similar ages (Study 2), beginning in December, 2019 and ongoing at the time of publication. In the beginning of each study we provided a mobile device and SIM card to freely access the system and a one-hour training session for children and a caregiver, in which we explained the purpose of the study and taught the child and caregiver how to access and use the IVR system (described in more detail in [27, 29]. We obtained 16 weeks of system and call data for Study 1 (February - May, 2019), and equivalent data from the first 8 weeks of the ongoing Study 2 (December, 2019 - February, 2020). For our analysis, we use data from the participants who called the system at least once ($N_1 = 165, N_2 = 408$).

3.2   Data Collection and Processing

The data used in training our models was the same for both Study 1 and 2. Each time a user called the system, the call metadata and the interactions during the call were logged on our database. The metadata included call start and end times (in local time), and the interaction data corresponded to a log of events that occurred during the call, such as attempting quiz questions, correctly completing those questions, parents or other caregivers accessing information (e.g., support messages and progress updates) about their child's usage, and more.

Each record in the data was identified by a unique *user-week*. Because we wanted to use all the data up to (and including) a given week to predict a gap in

usage in the subsequent week, we excluded the final week of system usage from our dataset. For Study 1, we generated a time series with 15 timestamps (one for each week prior to the final week) and data from 165 children for each timestamp ($N_1 = 2475$). For Study 2, we generated a time series with 7 timestamps and data from 408 children for each timestamp ($N_2 = 2856$). Each timestamp corresponded to data up to, and including, the given week. We trained a new model on the data for each timestamp to avoid future bias, i.e., training on future data while predicting the same. Based on prior research on dropout prediction in MOOCs (e.g. [4,33,53]) and churn prediction in mobile applications and social games (e.g. [20,37]) with a focus on features that could be gleaned solely from interaction logs, we used a total of 11 features including *call_duration* (average call duration during the week), *num_calls* (total number of calls), *num_days* (number of days the user called in a given week), *mastery* (percentage of questions attempted correctly), and *main_parent* (number of times a user accessed the main menu for the parent-facing version of the system). A list of all features used in the model and their pre-normalized, post-aggregation means and standard deviations can be found in Table 1. We aggregated the features at the week level, averaging *call_duration* and *mastery*, and summing the others. We decided to average *mastery* and *call_duration* to better represent the non-uniform distribution of lesson performance and call duration across the calls in a given week.

Table 1. Full set of features used in the predictive model

Feature	Explanation	Mean (SD)
sum_correct	Number of questions correct	8.78 (20.90)
sum_incorrect	Number of questions incorrect	10.38 (24.76)
sum_completed	Total number of questions completed	19.16 (44.48)
mastery	Percentage of questions correct	0.19 (0.26)
nunique_unit_id	Number of distinct units attempted	0.46 (0.53)
nunique_lesson_id	Number of distinct lessons attempted	4.34 (9.34)
num_calls	Number of calls	6.39 (12.01)
num_days	Number of days user called system	1.59 (1.83)
start_child_call_flow	Number of times a child began a lesson	3.78 (7.70)
main_parent	Number of times user accessed	2.02 (5.62)
	Parent-facing version of the system	
call_duration	Average call duration in seconds	137.95 (238.91)

3.3   Problem Formulation

We wanted to predict gaps in usage for our users. Given the distribution of usage data in our study which related to the school week, we define a gap as a given

user not calling the system for one week. We thus use this gap as the *positive* class in our model (*base rate* = 65%, i.e., 65% of *user_weeks* have a gap). Because we want to be able to predict for out-of-sample users who might begin calling later in the study (i.e., without prior call log data), we use a population-informed week-forward chaining approach to cross-validation [3]. That is, we held out a subset of users and trained the data for all weeks using a *k*-fold time-series cross-validation [46].

We wanted to use model types that were likely to perform well on smaller, imbalanced datasets as well as models that would allow us to identify feature importance and compare model performance. Prior literature on churn prediction [15,37] identified several model types that might meet these criteria: Random Forests (RF), Support Vector Machines (SVM), and eXtreme Gradient Boosting (XGBoost). Ensemble learning methods (like RF and XGBoost) had been shown to perform well for churn prediction, and SVM's kernel trick had been shown to successfully identify decision boundaries in higher dimensional data. Furthermore, boosted tree algorithms have been shown to perform as well as deep, neural approaches in certain scenarios [11,41], while requiring smaller datasets and compute power, which is of particular interest for predictive models in low-resource, developing contexts [21]. We used Scikit-Learn modules [35] for implementation, and Grid Search for hyper-parameter tuning of the optimisation criterion, tree depth, type of kernel, learning rate, and number of estimators [16].

4   Findings

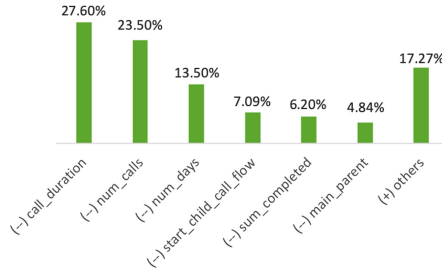### 4.1   Usage Gap Prediction Models for Study 1 (RQ1)

We evaluated the three models (SVM, RF, and XGBoost) models using four metrics—recall, precision, accuracy, and Area Under the Curve (AUC). Of these, we optimised for recall because we wanted to minimize false negatives. That is, we do not want to incorrectly predict that someone will call the next week, and thus miss an opportunity to remind or nudge them to use the system. We report on the mean and standard deviation for the performance metrics for all three models, averaged across all 15 model iterations in Table 2. In Fig. 1, we show the AUC results for each weekly model iteration for all 15 weeks. We found that XGBoost was the best performing model for Study 1, using a tree booster, a learning rate of 0.1, and a maximum depth of 5. We hypothesize that XGBoost performed the best because it was an ensemble learning method (unlike SVM), and used a more regularized model formalization (as opposed to RF), which may be more effective for the nature of our data because it avoids overfitting [5].

Table 2. Performance of Different Models in Study 1: Mean and Standard Deviation

Model	Recall	Precision	Accuracy	AUC
XGBoost	0.93 ($\sigma=0.06$)	0.78 ($\sigma=0.13$)	0.75 ($\sigma=0.12$)	0.68 ($\sigma=0.12$)
SVM	0.92 ($\sigma=0.06$)	0.78 ($\sigma=0.13$)	0.75 ($\sigma=0.12$)	0.65 ($\sigma=0.12$)
RF	0.90 ($\sigma=0.06$)	0.78 ($\sigma=0.13$)	0.74 ($\sigma=0.12$)	0.60 ($\sigma=0.10$)
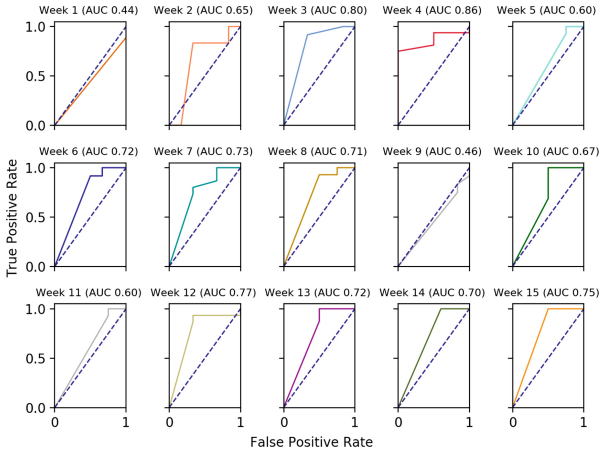
Fig. 1. AUC for each of the 15 iterations of the XGBoost model for Study 1

Fig. 2. Feature importance for Study 1, with direction of the feature in parentheses
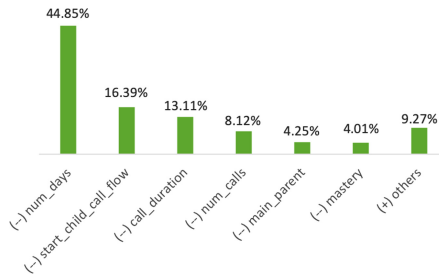
4.2   Feature Importance in Usage Gap Prediction for Study 1 (RQ2)

We next wanted to estimate feature importance in order to identify the features most associated with gaps in usage, to suggest potential design implications for personalized interventions or system designs to promote user retention. The feature importance and the directionality of the top ranked features in the XGBoost model can be seen in Fig. 2. We obtained the direction of the influence of the feature (i.e., either positively or negatively associated) using SHAP (SHapley Additive exPlanation), which allows for post-hoc explanations of various ML models [26]. We find that the most predictive features associated with gaps in

usage are the call duration, number of calls to the system, number of days with a call to the system, and total number of completed questions in a given week—all negatively predictive of gaps (i.e., positively associated with usage).

Table 3. Performance of Different Models in Study 2: Mean and Standard Deviation

Model	Recall	Precision	Accuracy	AUC
XGBoost	0.69 ($\sigma=0.18$)	0.58 ($\sigma=0.19$)	0.66 ($\sigma=0.07$)	0.73 ($\sigma=0.09$)
SVM	0.68 ($\sigma=0.13$)	0.58 ($\sigma=0.19$)	0.66 ($\sigma=0.07$)	0.72 ($\sigma=0.09$)
RF	0.69 ($\sigma=0.24$)	0.56 ($\sigma=0.18$)	0.65 ($\sigma=0.12$)	0.66 ($\sigma=0.11$)

Fig. 3. Feature importance for Study 2, with direction of the feature in parentheses

4.3   Replication of Usage Prediction for Study 2 (RQ3)

In order to evaluate the robustness of our approach, we evaluated the same models on data from the first 8 weeks of Study 2. We used the same features described in Table 1, for the 408 learners with data for the 7 weeks (again leaving out the 8th and final week for testing), as described in Sect. 3.2. We find that our model performance was consistent with the model performance from Study 1. Mean and standard deviation of model performance across the 7 models is reported in Table 3 We find that the AUC values are higher in Study 2 than in Study 1, although recall, precision, and accuracy are lower overall in Study 2. Given that Study 2 (8 weeks) was half the duration of Study 1 (16 weeks), we hypothesize that these prediction performance differences may be due to effects from differences in usage in the beginning of the study. That is, system usage in the first 1–2 weeks of the study was higher than the rest of the duration (for both Study 1 and 2). Thus, the model may fail to minimize the false negatives, as it is inclined to predict that a user will call back, when in reality there may be a gap in usage. The set of important features (seen in Fig. 3) were nearly the same

as in Study 1, but their rank order was different in Study 2, with consistency of calling (operationalized by the number of days called in a given week) being the most predictive feature as opposed to average call duration.

5   Discussion and Design Implications

Contextually-appropriate technologies may support learning outside of school for children in rural contexts with limited access to schooling. However, as prior work has demonstrated, in spite of motivation to learn, a variety of exogenous factors may inhibit children and their caregivers from consistently using learning technologies outside of school, limiting their efficacy [27, 29]. While prior research has developed predictive models of the likelihood of dropout, these approaches have historically dealt with adults dropping out from formal in-person or online courses, each of which may have some financial or social cost for dropping out. These factors may not be relevant for children's voluntary usage of a mobile learning application. In rural, low-resource contexts, mobile educational applications may be more accessible than online learning materials, though there may be additional obstacles to consider (e.g., children's agricultural participation [30]).

We have identified a set of system interaction features that are predictive of gaps in calling. Prior work in predicting dropout of adult learners in online courses found that factors like organizational support, time constraints, financial problems, etc. play an important role in predicting dropout [33]. We extend prior literature by finding that the important features associated with system usage were related to patterns of use, such as the duration of the interactions, their consistency of use (e.g., number of calls and number of days called in a week), as well as features related to their performance on the platform, including the number of questions completed and their overall mastery percent across all questions. In addition, we find that involvement of other family members (operationalized as the number of times the informational menu designed for adult supporters was accessed) is a predictive feature associated with system usage, which had not been accounted for in prior literature on app usage prediction.

Designers of voluntary educational systems can leverage these insights on the impact of learners' consistency of use and patterns of performance on future system usage. First, personalized, preemptive usage reminders may support ongoing engagement with the system. While usage reminders, like SMS texts and call reminders, have been associated with increased usage of mobile learning applications, they are often post-hoc (i.e., sent after a usage gap has already been observed) [38], which may be too late if users have already stopped engaging. Alternatively, sending too many reminders has been associated with a decrease in system usage, perhaps due to perceptions of being spammed [38]. Thus, there is a need for personalized, preemptive interventions based on users' likelihood to not persist in using the system. Researchers can use models trained on the aforementioned features to identify those users who are expected to have a gap in usage in the upcoming week. Furthermore, as we found that family involvement was associated with increased student engagement (following other prior

work that did not use predictive modeling [10,54]), we suggest that parents or guardians also receive personalized messages to prompt children's use of the system.

Second, analysis from both Study 1 and 2 showed that students' mastery (i.e., percentage of questions attempted correctly) was negatively associated with gaps in system usage. We thus hypothesize that users may feel a sense of achievement, or a positive sense of self-efficacy when they answer questions correctly, thus motivating them to continue learning (as in [1,55]). Voluntary educational applications may leverage mechanisms like dynamic question difficulty depending on correctness of responses, or system elements designed to give users this sense of achievement and mastery (e.g., virtual rewards to promote student engagement [8]). Introducing such features may better motivate students to continue using the system.

Finally, we analyzed these features across two studies with similar results. We did find that consistency (measured by number of days called) plays a more important role in shorter studies, as seen in Study 2, while call duration plays a more important role in longer studies, as seen in Study 1. We confirmed this by running post-hoc analyses on 8 weeks of data from Study 1 and found the same result. We see that in the first few weeks of usage, a user's calling pattern, as opposed to the interactions within each call, is more predictive of gaps, while the opposite is true for longer studies. We hypothesize that this may be due in part to the novelty effect, and suggest that over time, students receive more personalized content support in deployments.

5.1   Limitations and Future Work

This study uses system interaction data to predict gaps in children's use of a mobile literacy learning application. However, there may be other relevant information that may be useful for informing usage prediction—including data on children's prior content or domain knowledge (here, French phonological awareness and literacy more broadly), prior experience with similar types of applications (here, interactive voice response used on feature phones), and, more broadly, data on children's motivations for learning and self-efficacy. Future work may explore how to most effectively integrate such data collected infrequently in a survey or assessment with time-series data such as we have used here. In addition, the studies we trained our models on were in rural communities in low-resource contexts, and future work may investigate how predictive models of voluntary educational technology usage may differ across rural and urban contexts, and across international and inter-cultural contexts. Finally, future work may investigate the efficacy of personalized reminders or nudges to motivate increased use of the system and their impact on consistent system usage and learning.

6 Conclusion

Educational technologies have been proposed as an approach for supporting education in low-resource contexts, but such technologies are often used in schools, which may compound inequities in education for children who may not be able to attend schools regularly. However, when ed tech use is voluntary for children to use outside of school, there may be gaps in their usage which may negatively impact their learning, or lead to them abandoning the system altogether—gaps which may be prevented or mitigated using personalized interventions such as reminder messages. In this paper, we explore the efficacy of using machine learning models to predict gaps in children's usage of a mobile-based educational technology deployed in rural communities in Côte d'Ivoire, to ultimately inform such personalized motivational support. We evaluate the predictive performance of multiple models trained on users' system interaction data, identify the most important features, and suggest design implications and directions for predicting gaps in usage of mobile-based learning technologies. We intend for this work to contribute to designing personalized interventions for promoting voluntary usage of out-of-school learning technologies, particularly in rural, low-resource contexts.

References

1. Bandura, A.: Self-efficacy. In: The Corsini Encyclopedia of Psychology, pp. 1–3 (2010)
2. Berge, Z.L., Huang, Y.P.: 13: 5 a model for sustainable student retention: a holistic perspective on the student dropout problem with special attention to e-learning. DEOSNEWS. www.researchgate.net/profile/Zane_Berge/publication/237429805 (2004)
3. Bergmeir, C., Benítez, J.M.: On the use of cross-validation for time series predictor evaluation. Inf. Sci. **191**, 192–213 (2012)
4. Chaplot, D.S., Rhim, E., Kim, J.: Predicting student attrition in MOOCs using sentiment analysis and neural networks. In: AIED Workshops, vol. 53, pp. 54–57 (2015)
5. Chen, T., Guestrin, C.: XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794 (2016)
6. Conn, K.M.: Identifying effective education interventions in sub-Saharan Africa: a meta-analysis of impact evaluations. Rev. Educ. Res. **87**(5), 863–898 (2017)
7. Dahman, M.R., Dağ, H.: Machine learning model to predict an adult learner's decision to continue ESOL course. Educ. Inf. Technol. **24**(4), 1–24 (2019)
8. Denny, P.: The effect of virtual achievements on student engagement. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 763–772 (2013)
9. Glass, C.R., Shiokawa-Baklan, M.S., Saltarelli, A.J.: Who takes MOOCs? New Dir. Inst. Res. **2015**(167), 41–55 (2016)
10. Gonzalez-DeHass, A.R., Willems, P.P., Holbein, M.F.D.: Examining the relationship between parental involvement and student motivation. Educ. Psychol. Rev. **17**(2), 99–123 (2005)

11. Hashim, M., Kalsom, U., Asmala, A.: The effects of training set size on the accuracy of maximum likelihood, neural network and support vector machine classification. Sci. Int. Lahore **26**(4), 1477–1481 (2014)
12. Herbert, M.: Staying the course: a study in online student satisfaction and retention. Online J. Distance Learn. Adm. **9**(4), 300–317 (2006)
13. Ishikawa, M., Ryan, D.: Schooling, basic skills and economic outcomes. Econ. Educ. Rev. **21**(3), 231–243 (2002)
14. Jasińska, K.K., Petitto, L.A.: Age of bilingual exposure is related to the contribution of phonological and semantic knowledge to successful reading development. Child Dev. **89**(1), 310–331 (2018)
15. Jose, J.: Predicting customer retention of an app-based business using supervised machine learning (2019)
16. Joseph, R.: Grid search for model tuning, December 2018. https://towardsdatascience.com/grid-search-for-model-tuning-3319b259367e
17. Kam, M., Kumar, A., Jain, S., Mathur, A., Canny, J.: Improving literacy in rural India: cellphone games in an after-school program. In: 2009 International Conference on Information and Communication Technologies and Development (ICTD), pp. 139–149. IEEE (2009)
18. Kam, M., Rudraraju, V., Tewari, A., Canny, J.F.: Mobile gaming with children in rural India: contextual factors in the use of game design patterns. In: DiGRA Conference (2007)
19. Kemp, W.C.: Persistence of adult learners in distance education. Am. J. Distance Educ. **16**(2), 65–81 (2002)
20. Kim, S., Choi, D., Lee, E., Rhee, W.: Churn prediction of mobile and online casual games using play log data. PLoS ONE **12**(7), e0180735 (2017)
21. Kshirsagar, V., Wieczorek, J., Ramanathan, S., Wells, R.: Household poverty classification in data-scarce environments: a machine learning approach. In: Neural Information Processing Systems, Machine Learning for Development Workshop, vol. 1050, p. 18 (2017)
22. Kumar, A., Reddy, P., Tewari, A., Agrawal, R., Kam, M.: Improving literacy in developing countries using speech recognition-supported games on mobile devices. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1149–1158. ACM (2012)
23. Lam, Y.J.: Predicting dropouts of university freshmen: a logit regression analysis. J. Educ. Adm. **22**, 74–82 (1984)
24. Lange, C., Costley, J.: Opportunities and lessons from informal and non-formal learning: applications to online environments. Am. J. Educ. Res. **3**(10), 1330–1336 (2015)
25. Lucini, B.A., Bahia, K.: Country overview: Côte d'ivoire driving mobile-enabled digital transformation (2017)
26. Lundberg, S.M., Lee, S.I.: A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems, pp. 4765–4774 (2017)
27. Madaio, M.A., et al.: "you give a little of yourself": family support for children's use of an IVR literacy system. In: Proceedings of the 2nd ACM SIGCAS Conference on Computing and Sustainable Societies, pp. 86–98. ACM (2019)
28. Madaio, M.A., Tanoh, F., Seri, A.B., Jasinska, K., Ogan, A.: "Everyone brings their grain of salt": designing for low-literate parental engagement with a mobile literacy technology in côte d'ivoire. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, p. 465. ACM (2019)

29. Madaio, M.A., et al.: Collective support and independent learning with a voice-based literacy technology in rural communities. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1–14 (2020)

30. Malpel, J.: Pasec 2014: education system performance in francophone sub-Saharan Africa. Programme d'Analyse des Systèmes Educatifs de la CONFEMEN. Dakar, Sénégal (2016)

31. Maritzen, L., Ludtke, H., Tsukamura-San, Y., Tadafusa, T.: Automated usage-independent and location-independent agent-based incentive method and system for customer retention, US Patent App. 09/737,274, 28 February 2002

32. McEwan, P.J.: Improving learning in primary schools of developing countries: a meta-analysis of randomized experiments. Rev. Educ. Res. **85**(3), 353–394 (2015)

33. Park, J.H., Choi, H.J.: Factors influencing adult learners' decision to drop out or persist in online learning. J. Educ. Technol. Soc. **12**(4), 207–217 (2009)

34. Patel, N., Chittamuru, D., Jain, A., Dave, P., Parikh, T.S.: Avaaj Otalo: a field study of an interactive voice forum for small farmers in rural india. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 733–742. ACM (2010)

35. Pedregosa, F., et al.: Scikit-learn: machine learning in Python. J. Mach. Learn. Res. **12**(Oct), 2825–2830 (2011)

36. Pereira, F., et al.: Early Dropout prediction for programming courses supported by online judges. In: Isotani, S., Millán, E., Ogan, A., Hastings, P., McLaren, B., Luckin, R. (eds.) AIED 2019. LNCS (LNAI), vol. 11626, pp. 67–72. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-23207-8_13

37. Periáñez, Á., Saas, A., Guitart, A., Magne, C.: Churn prediction in mobile social games: towards a complete assessment using survival ensembles. In: 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA), pp. 564–573. IEEE (2016)

38. Pham, X.L., Nguyen, T.H., Hwang, W.Y., Chen, G.D.: Effects of push notifications on learner engagement in a mobile learning app. In: 2016 IEEE 16th International Conference on Advanced Learning Technologies (ICALT), pp. 90–94. IEEE (2016)

39. Phiri, A., Mahwai, N., et al.: Evaluation of a pilot project on information and communication technology for rural education development: a cofimvaba case study on the educational use of tablets. Int. J. Educ. Dev. ICT **10**(4), 60–79 (2014)

40. Richmond, M., Robinson, C., Sachs-Israel, M., Sector, E.: The global literacy challenge. UNESCO, Paris (2008). Accessed 23 August 2011

41. Roe, B.P., Yang, H.J., Zhu, J., Liu, Y., Stancu, I., McGregor, G.: Boosted decision trees as an alternative to artificial neural networks for particle identification. Nucl. Instrum. Methods Phys. Res., Sect. A **543**(2–3), 577–584 (2005)

42. Schölkopf, B.: The kernel trick for distances. In: Advances in Neural Information Processing Systems, pp. 301–307 (2001)

43. Shankar, V., Venkatesh, A., Hofacker, C., Naik, P.: Mobile marketing in the retailing environment: current insights and future research avenues. J. Interact. Mark. **24**(2), 111–120 (2010)

44. Shin, C., Hong, J.H., Dey, A.K.: Understanding and prediction of mobile application usage for smart phones. In: Proceedings of the 2012 ACM Conference on Ubiquitous Computing, pp. 173–182 (2012)

45. Tang, C., Ouyang, Y., Rong, W., Zhang, J., Xiong, Z.: Time Series Model for Predicting Dropout in Massive Open Online Courses. In: Penstein Rosé, C., Penstein Rosé, P., et al. (eds.) AIED 2018. LNCS (LNAI), vol. 10948, pp. 353–357. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-93846-2_66

46. Tashman, L.J.: Out-of-sample tests of forecasting accuracy: an analysis and review. Int. J. Forecast. **16**(4), 437–450 (2000)
47. Terenzini, P.T., Lorang, W.G., Pascarella, E.T.: Predicting freshman persistence and voluntary dropout decisions: a replication. Res. High. Educ. **15**(2), 109–127 (1981)
48. Tinto, V.: Research and practice of student retention: what next? J. Coll. Stud. Retent.: Res. Theory Pract. **8**(1), 1–19 (2006)
49. Tyler-Smith, K.: Early attrition among first time elearners: a review of factors that contribute to drop-out, withdrawal and non-completion rates of adult learners undertaking elearning programmes. J. Online Learn. Teach. **2**(2), 73–85 (2006)
50. Uchidiuno, J., Yarzebinski, E., Madaio, M., Maheshwari, N., Koedinger, K., Ogan, A.: Designing appropriate learning technologies for school vs home settings in Tanzanian rural villages. In: Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies, pp. 9–20. ACM (2018)
51. Warschauer, M., Ames, M.: Can one laptop per child save the world's poor? J. Int. Aff. **64**(1), 33–51 (2010)
52. Xie, Y., Li, X., Ngai, E., Ying, W.: Customer churn prediction using improved balanced random forests. Expert Syst. Appl. **36**(3), 5445–5449 (2009)
53. Yang, D., Sinha, T., Adamson, D., Rosé, C.P.: Turn on, tune in, drop out: anticipating student dropouts in massive open online courses. In: Proceedings of the 2013 NIPS Data-Driven Education Workshop, vol. 11, p. 14 (2013)
54. Zellman, G.L., Waterman, J.M.: Understanding the impact of parent school involvement on children's educational outcomes. J. Educ. Res. **91**(6), 370–380 (1998)
55. Zimmerman, B.J.: Self-efficacy: an essential motive to learn. Contemp. Educ. Psychol. **25**(1), 82–91 (2000)