# Teaching Students About Machine Learning Through a Gamified Approach

Prapa Rattadilok
*School of Computer Science*
*University of Nottingham*
Ningbo, China
prapa.rattadilok@nottingham.edu.cn

Chris Roadknight
*School of Computer Science*
*University of Nottingham*
Ningbo, China
chris.roadknight@nottingham.edu.cn

Li Li
*School of Computer Science*
*De Montfort University*
Leicester, UK
p14150783@my365.dmu.ac.uk

***Abstract*—The teaching of machine learning requires a range of tools and techniques to engage students and allow them to relate the processes involved to real world situations that they have previously experienced. One way to facilitate this learning process is to integrate the learning into a game situation, which is by definition fun to engage with and offers immediate rewards. This research shows that by collecting the student's behaviour and actions as they engage with well-known game software, the learning of key machine learning concepts can be enabled. It is also shown that customising of learning can be made possible by the use of gamification.**

***Keywords—machine learning; gamification; personalised learning; universal design for learning; machine teaching***

## I. Introduction

In today's world, artificial intelligence (AI) has significant impacts on the growth and the productivity in many industries. AI related subjects such as machine learning are therefore currently being taught to university students during their second and third year in many courses including those outside of computer science. However, teaching machine learning is not trivial. The two key challenges commonly faced by many teachers who teach machine learning related contents are: 1) the difficulty in teaching technologically driven subjects to students who may not have much interest towards technology [1] and 2) the inflexibility of the existing educational environment which relies on textbooks and worksheets that are rigid and unengaging [2].

Setting up machine learning worksheets require labelled data sets. Although there are many applications where raw data is plentiful but the creation of a labelled dataset or the process of labelling existing datset, can be costly and time consuming. Many researchers attempt to tackle this issue through crowdsourcing [3][4], using semi-supervised learning [5] or semi-automatically generating labelled data using machine learning technique itself [6]. As a result of this time-cost factor, many machine learning teachers end up having to use existing data depositories in their worksheets. For example, UCI [7] or Kaggle [8]. Consequently from the use of irrelevant or uninteresting data sets, students become disengaged or the difficulty in working through the worksheet itself is increased significantly by the difficulty of the data sets that are available.

In today's education, unlike many other industries, there are very few AI-based learning systems in classrooms or homes. Research has shown [9][10] learning performance or student's achievement with a particular subject content depends directly on the student's interest and indirectly on the teacher's responsiveness. The potential for being able to tailor learning to individual needs and interests would greatly impact the student learning experience. For example, providing access to personalised digital contents, or engage students in meaningful way.

Machine teaching [11][12][13], the inverse problem of machine learning, is one of several emerging research fields in AI. While machine learning focuses on creating new algorithms and improving the accuracy of the "learners", machine teaching focuses on the "teacher" to increase the efficacy of the teachers given the learners. This machine teaching could be an answer to further enhance personalised learning and universal design for learning (UDL) [14][15], given that most learning contents are now digitalised.

This paper discusses the use of a modified version of an online mobile game called "Clash of Clans". The short term goal is to to utilise the gaming environment and gamify student's learning process, specifically on the learning of machine learning techniques, in order to: 1) attract students who may not have much interest towards technology through gamification; 2) allow the students to learn, explore and observe instantaneously the effect of their chosen machine learning techniques though active learning. The long term goal is to automatically generate labelled data sets which can be used to improve personalised learning and UDL through the use of machine teaching.

## II. Personalised Learning

Personalised learning is a system that responds to individual students, by creating their structured learning path according to their needs, interests and aspirations [16][17][18]. Students should receive teaching tailored to their needs based on their learning stage not age. The system whilst stretching and challenging the more able students, should also ensure that no child falls behind. This extends the educational concept of differentiation to include learning preferences, learning pace and learning experience of different individuals [19].

The Department for Education and Skills (DfES) articulate five key components of personalised learning [20]. Personalisation needs elements of choice. However, too much choice can lead to anxiety and confusion. Choice for personalised learning can be in self-management and self-provision. By giving students self-support, linking them up in a peer-to-peer way, provide them with better tools, learning can take place beyond the classroom. Data analytics should be used to deliver quality teaching according to the diagnoses e.g. each student's learning needs, their strengths and weaknesses. It was suggested that effective learning analytics should be executed at multiple timepoints, and not only just assessment of learning at the end of the course. [19].

## III. Universal Design for Learning

The early emphasis of universal design was obstacle-free environment [14]. Universal design for learning (UDL) was developed based on the observed disconnectibity between diverse student population, learning strategies and digital media where "one-size-fits-all" curriculum would no longer produce the academic achievement being sought [15]. The three core components of UDL are: multiple means of representation i.e. ways they can learn; multiple means of expression i.e. ways they can demonstrate what they have learnt; and multiple means of engagement i.e. ways to motivate and interest them to learn [21].

## IV. Clash of Clans

Clash of Clans is an online strategy mobile game developed by Supercell [22]. Each player can customise their village and build an army to attack other players' village. Buildings can be categorised into three types: supply, sentry and barrack. Supply buildings collect and store virtual resources which can be used to build buildings and train the army. Sentry buildings watch and defend any incoming attacks. Barracks train the army and attack other players' village in order to steal their resource.

There are two main objectives to the game: 1) to defend the village using the sentry building and 2) to attack the other villages using the army. For this paper, the focus is only on the attack objective. At the end of each attack i.e. a maximum of three minutes, the game will display the result of the attack in 3-star rating and the percentage of the opponent's village being destroyed. Fig. 1 displays an example where 52% of the opponent's village were destroyed.


Fig. 1. CoC's attack result.

For the 3-star rating, one star is given when the attacker destroyed the village's town hall, one star is given when the attacker destoyed over 50% of the village and the final star is given when the attacker annihilate the village. Fig. 1 indicates that the defender's town hall was not sucessfully destroyed by this attack.

There are 19 different possible types of soldier in any army, where each type has different stats e.g. hit points, damage per second, housing space, movement speed, favourite targets. Fig. 2 demonstrates all 19 types of soldier.

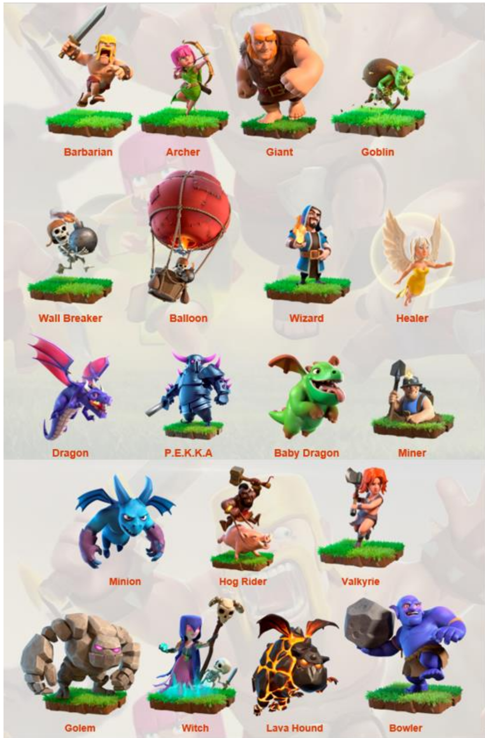


Fig. 2. CoC's types of soldiers.

Fig. 3 illustrates an attack interface. Soldiers can only be deployed outside of the red highlighted area on the attacking interface. By selecting a type of soldier within an attacking army, the player can then tab on any other gaming area to place the soldier and begin the attack.

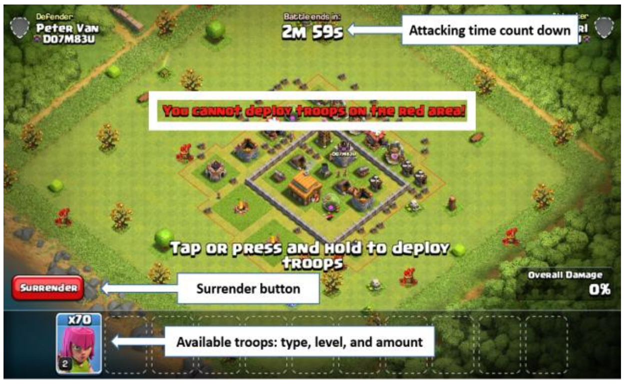


Fig. 3. CoC's attacking interface.

The level of success of any attack depends on a number of factors. In addition to the strength of the defender's village, examples of other factors include: the size of the army, the selected type(s) of soldier, where the soldiers are placed, when the soldiers are placed, how many soldiers are placed in different positions. The combination of these possible options make it suitable to utilise machine learning techniques to optimise the attacking strategies.

## V. Traditional Machine Learning Classrooms

The structure of many machine learning classes in universities are pretty similar to those that are available online. The main difference is the lack of instructors and peer presence. Many online courses now provide certificates from well-established universities. For example, Andrew Ng's is one of the most well-known machine learning online classes on Coursera [23].

Using Andrew Ng's course as a reference for the purpose of discussion, his course consists of a number of lectures, practical sessions and coursework. The syllabus contains a number of mathematical concepts, machine learning descriptions, and application examples. The course itself involves the use of mathematically oriented software such as Octave [24] and MATLAB, where the interface (Fig. 4) can easily put-off many experienced data scientists, let alone second or third year university students.

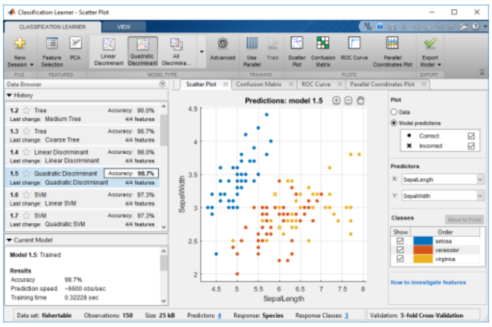


Fig. 4. MATLAB's example interface [25].

During practical sessions, students follow a set of instructions to complete the task, ranging from data cleasing, applying different machine learning techniques, and visualisation of the outcome. Most comments are positive but there are some that indicates the difficulties/simplicities in following complex mathematical concepts and the use of new software. For example, "As a person who has a decent background in math and somewhat good basis in programming, sometimes I felt lost as to why he went and picked some method to work with because sometimes he tries not to get into his reasoning to keep the course easy and simple to understand.", or " I struggled with the most is familiarising myself with the octave language especially when it comes to vectors and vectorised operations. Therefore I suggest that you add an additional helpful material along with the pdf to guide the student in the assignments because sometimes it can get tricky to express a mathematical formula in octave language while also using vectorisation" [23].

## VI. Gamifying Traditional Machine Learning Classrooms

The history of game bots can be traced back to the end of the 20th century when the first video game was first introduced. [26]. In the traditional Role-Playing-Games (RPGs), Non-Player-Characters (NPCs) are an example of game bots that interacts with the players, e.g. providing quests and information. In the fighting games or the team sports games, game bots are used as the competing opponents for players.

As shown in [23], typically the teaching of machine learning involves two main components: ensuring the understanding of a machine learning technique and the usage of such machine learning technique. The understanding part usually involves giving an introduction about a particular machine learning technique, and how to apply that particular machine learning technique on sample data sets using a particular machine learning software tool. The usage part usually depends on the nature of the machine learning technique itself. For example, some of the machine learning techniques such as neural networks [27] will stereotypically involve how the developed model can be used; other machine learning techniques such as clustering [28] will characteristically involve how to interpret the outcome and analyses.

iGaME (In class Gamified Machine learning Environment) is a game bot that has been purposefully developed for the teaching of machine learning algorithms in the class rooms. iGaME can be used to develop student's understanding of a machine learning technique, by allowing students to develop a gaming strategy by applying a selected data set to the CoC. For each game bot execution, iGaME accept one gaming strategy from the student. iGaME repeatedly applies the gaming strategy to CoC according to the number of times specified by the students. At the end of the execution, iGaME produces an output file, which records the outcome of each time the gaming strategy is applied. The student can iteratively improve their gaming strategies from their intuitions or methodologically using the output file from iGaME for data analyses.

### A. *Teaching Difficult Technologically Driven Subject to Non-Technical Students*

Based on our survey to compare the interfaces of traditional machine learning software tools and CoC, CoC is much more visually attractive than traditional machine learning software tools. iGaME can be used to reduce technological barriers for some students. Whilst some may argue that not all students like to gamify their learning process, iGaME is not aimed at replacing mathetically oriented software for the teaching of machine learning, but to support the learning process e.g. can be used along side traditional tools or as an options for less-technical students.

When a gaming strategy is applied to CoC via iGaME, the student can watch the outcome of their strategy as it unfolds in real-time. Timeliness of the feedback is vital in motivating students in the process of competency achievements. Feedback that shortly follows an assessment allows students to reflect on their own performance while it is fresh in their minds, whether this be regarding their strengths and weakness. This further builds on their capabilities and addresses deficient areas [29].

A worksheet is given to a student to develop a strategy to deploy 70 soldiers using only 70 tabs on screen and the tabs must be in the valid attack area. On a 1280 x 720 screen resolution, there are 921,600 pixel. Each of those pixels has a colour associated with it. The student chose to use these pixel colours as the input for iGaME. The student attempted three iterations of pixel selections. In the first iteration, pixels were randomly selected. The second and third iterations use previously successful selections to drive the pixel selections for the current iteration. The outcome of using pixel colours as the strategy for deploying 70 soldiers is shown in Table I.

TABLE I. IGAME OUTCOME ON USING PIXEL COLOURS

| | *1st iteration* | *2nd iteration* | *3rd iteration* |
|---|---|---|---|
| avg | 45.13 | 40.73 | 43.97 |
| std | 6.69 | 12.95 | 9.00 |
| max | 61.42 | 81.42 | 81.42 |

In Table I, the values represent the average, standard deviation and maximum number of valid attempts i.e. deploying a soldier in a valid attacking area. While the average number of valid attempts is slightly worse in the third iteration when comparing to the first iteration, the maximum number of valid attempts in the third iteration is much better.

The student was given a table (Table II) obtained from human players to be used as a benchmark. "Novice" represents the players who have no mobile gaming experience. "Intermediate" represents the players who have some mobile gaming experience but not CoC. "Expert" represents the players who have played CoC for more than a year.

TABLE II. HUMAN PLAYERS' OUTCOME

| | *Novice* | *Intermediate* | *Expert* |
|---|---|---|---|
| avg | 87.14 | 97.14 | 99.05 |
| std | 8.69 | 5.19 | 0.82 |
| max | 92.86 | 100.00 | 100.00 |

Table II illustrates the average, standard deviation and maximum number of valid attempts by novice, intermediate and expert human players i.e. deploying a soldier in a valid attacking area. In general, the performance of using pixel colour as an attacking strategy is incomparable to any human testing groups. The average of successful deployment is 43.97% in the third iteration, which is nearly 50% lower than the average of the successful deployments i.e. 87.14% in novice human players However, when using pixel colour as the strategy, some of the successful attacks have better performances than the performances of some attacks in the novice human testing group.

The student expressed his enjoyment of using iGaME in comparison to traditional machine learning tool. The simplicity of selecting the gaming strategy and how close it could do when comparing to the human players' means that he is likely to explore more strategies himself outside of classroom to observe the outcome further and study machine learning algorithms further.

### B. *Inflexibility of Existing Machine Learning Educational Environments*

iGaME adds challenges to widen the learning scope for more-technical students. Fig.5 illustrates the difference between human's tab and swipe gesture on mobile screen. Supplementary tasks can be provided, for example, to clease data recorded by human players to exclude those that contain swipe.

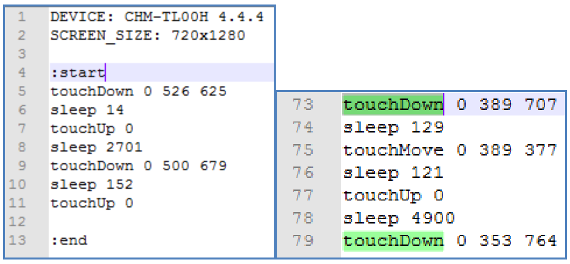


Fig. 5. Human gestures: on-screen tab (left) and on-screen swipe (right).

A number of other different scenarios can be generated from iGaME during the learning sessions to expand what students can learn from other students. As a result, this process will also enhance peer-to-peer learning [30].

### C. *Enhancing Personalised Learning and Universal Design for Learning – The Future of Machine Teaching in Teaching*

The larger the number of digitalised content a course has, the larger the number of data items that can be recording and later used to analyse students' learning behaviour. iGaME can be used to record what students do in-class and outside-of-class, monitor their learning pathways, how many tasks were completed on-time, how many supplementary tasks were attempted. When this data are mapped to students' learning profiles, assessment from the course itself and the other coures, the teacher can analyse this data and tailor the future content for the incoming students.

iGaME can also generate further data sets for students to use and learn from. The students can even use their own records of what they had chosen to learn or which data sets they have used in the past to recommend their future data sets they may enjoy. The teachers can exploit iGaME data to handpick which datasets to provide to different students or groups of students to work on according to their learning profile to optimise personalised learning.

From a UDL perspective, data about learning preferences can help the teacher to focus and enhance preferred learning tools, identify students that are less engaged in order to make sure that no one falls behind.

## VII. EVALUATION PLANS

In order to further evaluate iGaME, the plan is to use iGaME in the actual machine learning lectures and labs in order to: 1) collect more sample data sets i.e. student developed strategies, which can then be used to allow non-technical students to gain understandings of machine learning by interactive examples; 2) generate more supplementary tasks from the collected data sets allowing technical students to work on advanced machine learning processes e.g. data cleansing, missing data handling; 3) collect feedback from both technical and non-technical students based on their experience of traditional machine learning tools and iGaME.

Student satisfaction surveys will be conducted at the beginning and the end of the semester. The activities within the module will alternate weekly between using traditional machine learning software tools and iGaME. Effectiveness of iGaME will be measured according to the quantity as well as how well students complete the main tasks and the supplementary tasks in each week. This will also help in managing supplementary tasks and adaptively adjusting the difficulty and the number of the provided tasks to each student group.

## VIII. CONCLUSIONS

This paper discusses a modified mobile game called Clash of Clans (CoC) as a teaching tool for machine learning courses. iGaME is a game bot developed to accept gaming strategies from the users and repeatedly apply such strategies to the game and record the outcome of the applied strategies.

The simplicity of iGaME and the visual attractiveness of CoC is shown to minimise the burden in learning technologically driven content specfically for non-technical students. Because iGaME records various interactions with the students, data can be used in two ways. Firstly, the students can expand their learning from their own learning records, making traditional machine learning classrooms more flexible. Secondly, when combining iGaME data with student profiles, the teacher can analyse and use the information to enhance personalised learning, UDL and in the future machine teaching for teaching.

## REFERENCES


[1] B. Vlist, R. Westelaken, C. Bartneck, H. Jun, R. Ahn, E. Barakova, F. Delbressine, and L. Fejs. "Teaching Machine Learning to Design Students, " Internaltional Conference on Technologies for E-learning and Digital Entertainment, pp. 206-217, 2008.

[2] B. P. Woolf, H. C. Lane, V. K. Chaudhri, and J. L. Kolodner. "AI Grand Challenges for Education," Special Issue on Intelligent Learning Technologies, AI Magazine, June 2013.

[3] J. C. Chang, S. Amershi, and E. Kamar. "Revolt: Collaborative Crowdsourcing for Labeling Machine Learning Datasets," CHI Conference, pp. 2334-2346, 2017.

[4] Dbrain. "The blockchain platform to lavel data and to build AI apps". [Online]. Available at: https://dbrain.io/DbrainWhitepaper.pdf. (Accessed 24 August 2018).

[5] N. V. Chawla and G. Karakoulas. "Learning from Labeled and Unlabeled Data: An Empirical Study across Techniques and Domains," Journal of Artificial Intelligence Research, vol. 23, pp. 331-366, 2005.

[6] A. Ratner, C. D. Sa, S. Wu, D. Selsam, and C. Re. "Data Programming: Creating Large Training Sets, Quickly," Conference on Neural Information Processing Systems, 2016.

[7] UCI Machine Learning Repository. [Online]. Available at: https://archive.ics.uci.edu/ml/datasets.html. (Accessed 24 August 2018).

[8] Kaggle. [Online]. Available at: https://www.kaggle.com/datasets. (Accessed 24 August 2018).

[9] J. L. Abrantes, C. Seabra, and L. F. Lages. "Pedagogical affect, student interest, and learning performance," Journal of Business Research, vol. 60(9), pp. 960-964, 2007.

[10] K. A. Renninger and S. Hidi. "Student Interest and Achievement: Developmental Issues Raised by a Case Study," The Development of Achievement Motivation, pp. 173-220, 2002.

[11] P. Y. Simard, S. Amershi, D. M. Chickering, A. E. Pelton, S. Ghorashi, C. Meek, G. Ramos, J. Suh, J. Verwey, M. Wang, and J. Wernsing. "Machine Teaching, a New Paradigm for Building Machine Learning Systems". [Online]. Available at: http://www.microsoft.com/en-us/research/wp-content/uploads/2017/07/machine-teaching.pdf. (Accessed 28 August 2018).

[12] W. Liu, B. Dai, A. Humayun, C. Tay, C. Yu, L. B. Smith, J. M. Regh, L. Song. "Iterative Machine Teaching," International Conference on Machine Learning, 2017.

[13] M. Weimer. "Machine Teaching: a Machine Learnng Approach to Technology Enhanced Learning," PhD Thesis, Angenommen vom Fachbereich Informatik der Technischen Universitat Darmstadt, 2009.

[14] D. H. Rose and A. Meyer. "Teaching Every Student in the Digital Age: Universal Design for Learning," Education Tech Research Dev, 55, pp.521-525, 2007.

[15] D. L. Edyburn. "Universal Design for Learning," Special Education Technology Practice, 7(5), pp. 16-22, 2005.

[16] J. Sebba, N. Brown, S. Steward, M. Galton, and M. James. "An Investigation of Personalised Learning Approaches used by Schools," Research report 843, University of Sussex, 2007.

[17] S. McCarthy-Fry. "Personalised Learning a Practical Guide," Department for children, schools and families. [Online]. Available at: https://www.essex.gov.uk/Business-Partners/Partners/Schools/One-to-one-tuition/Documents/Personalised%20Learning%20a%20practical%20 guide.pdf. (Accessed 24 August 2018).

[18] Complex Needs, "Personalised Learning A Policy Overview". [Online]. Available at: http://www.complexneeds.org.uk/modules/Module-4.1-Working-with-other-professionals/All/downloads/m13p070b/personalised_learning_ policy_overview.pdf. (Accessed 24 August 2018).

[19] E. Bartle. "Personalised Learning: an Overview". [Online]. Available at: https://itali.uq.edu.au/filething/get/1865/Personalised_learning_overview_Final_16_Mar_15.pdf, March 2015. (Accessed 24 August 2018).

[20] National College for School Leadership. "Leading Personalised Learning in Schools, Helping individuals grow". [Online]. Available at: http://dera.ioe.ac.uk/5397/1/media-578-c9-leading-personal ised-learning-in-schools.pdf. (Accessed 24 August 2018).

[21] CAST. "Universal Design for Learning". [Online]. Available at: http://www.udlcenter.org/sites/udlcenter.org/files/updateguidelines2_0.pdf .(Accessed 24 August 2018).

[22] Supercell. [Online]. Available at: http://supercell.com/en/. (Accessed 24 August 2018).

[23] A. Ng. "Machine Learning". [Online]. Available at: https://www.coursera.org/learn/machine-learning#syllabus. (Accessed 24 August 2018).

[24] Octave. [Online]. Available at: https://www.gnu.org/software/octave/. (Accessed 24 August 2018).

[25] MATLAB. [Online]. Available at: https://uk.mathworks.com/help/stats/classificationlearner-app.html. (Accessed 24 August 2018).

[26] J. E. Laird and M. Lent. "Human-Level AI's Killer Application – Interactive Computer Games," AI Magazine, 22, 2001.

[27] D. Kriesel. "A Brief Introduction to Neural Networks". [Online]. Available at: http://www.dkriesel.com/_media/science/neuronalenetze-en-zeta2-2col-dkrieselcom.pdf. (Acceessed 23 September 2018).

[28] P. Tan, M. Steinbach, A. Karpatne, and V. Kumar. "Introduction to Data Mining," Second Edition, Pearson, 2005.

[29] T. Zehra, M. Tariq, S. K. Ali, A. Motiwala, J. Boulet, "Challenges of providing timely feedback to residents: Faculty perspectives," Journal of Pakistan Medical Association, 65(10), 1069-1074, 2015.

[30] M. Andrews and N. Manning. "A Guide to Peer-to-Peer Learning – How to make peer-to-peer and learning effective in the public sector?," Effective Institutions Platform. [Online]. Available at: https://www.effectiveinstitutions.org/media/The_EIP_P_to_P_Learning_Guide.pdf. (Accessed 24 August 2018).