Generating Piano Practice Policy with a Gaussian Process

Alexandra Moringen   ALEXANDRA.MORINGEN@UNI-GREIFSWALD.DE
University of Greifswald

Elad Vromen   ELADVROMEN@GMAIL.COM
Tel Aviv University

Helge Ritter   HELGE@TECHFAK.UNI-BIELEFELD.DE
Bielefeld University

Jason Friedman   JASON@TAU.AC.IL
Tel Aviv University

Abstract

A typical process of learning to play a piece on a piano consists of a progression through a series of practice units that focus on individual dimensions of the skill, the so-called practice modes. Practice modes in learning to play music comprise a particularly large set of possibilities, such as hand coordination, posture, articulation, ability to read a music score, correct timing or pitch, etc. Self-guided practice is known to be suboptimal, and a model that schedules optimal practice to maximize a learner's progress still does not exist. Because we each learn differently and there are many choices for possible piano practice tasks and methods, the set of practice modes should be dynamically adapted to the human learner, a process typically guided by a teacher. However, having a human teacher guide individual practice is not always feasible since it is time-consuming, expensive, and often unavailable. In this work, we present a modeling framework to guide the human learner through the learning process by choosing the practice modes generated by a policy model. To this end, we present a computational architecture building on a Gaussian process that incorporates 1) the learner state, 2) a policy that selects a suitable practice mode, 3) performance evaluation, and 4) expert knowledge. The proposed policy model is trained to approximate the expert-learner interaction during a practice session. In our future work, we will test different Bayesian optimization techniques, e.g., different acquisition functions, and evaluate their effect on the learning progress.

Keywords: computational scaffolding, accelerating motor learning, learning to play piano, practice modes, computational architecture for learning

1. Introduction

It can take many years to master complex motor tasks, such as learning to play the piano (Furuya and Altenmüller, 2013). Such learning is typically guided by an expert (a piano teacher). However, access to experts is limited and expensive, and even learners who do learn with an expert are required to practice alone for many hours a week to achieve mastery.

In order to guide or accelerate learning (i.e., maximize performance improvement in the long run), the learner should be presented with practice targets that take into account their skill level, their learning goal, the characteristics of the piece, and the predicted practice utility. The goal of this work is, therefore, to approach the modeling of the structure that underlies this complex scenario in order to approximate the expert-learner interaction.

One role of a teacher is to assist the learner using scaffolding (Moringen et al., 2021). Scaffolding provides temporary guidance for a learner during the learning process (Zydney, 2012), which can then be gradually removed or faded out. In piano playing, scaffolding can take the form of task simplification, i.e., practicing a particular aspect of performance. For example, this could involve focusing only on pitch (without concern for timing), practicing each hand separately, increasing/decreasing the velocity, or only focusing on timing (by clapping or playing a single note to the rhythm of the piece). The other approach, which we would like to term as *adversarial scaffolding*, is also widely used in training motor tasks and is characterized by artificially increasing the complexity of the task by adding complexity on top (Raviv et al., 2022). One extreme technique professional musicians use is to play the piece in reverse order, from the end to the beginning. Teachers use different training approaches for both beginners (Thompson, 2009) and advanced learners (see Cortot's recommended piano exercises for learning to play Chopin (Cortot, 1986)). The focus on practicing one or several modalities of the complex skill by increasing or reducing the overall task's complexity is termed practice mode (PM) (Moringen et al., 2021).

During practice, we want to optimize the selection of activities the learner performs. To do this, we propose a minimalist framework that can be later extended to a more complex scenario. First, we restrict the actions to two **practice modes**, timing practice and pitch practice. Second, we capture the **state of the learner** by recording their performance before and after learning with a given practice mode. Currently, the state is represented by two error modalities: timing error and pitch error (see a more detailed description in Moringen et al. (2021) and Section 3.1). We assume utility is a function of the learner state, music piece parameters, and action (Tamir-Ostrover et al., 2022). Based on the ground truth recorded during expert-learner interaction, we train a Gaussian Process (GP) (Rasmussen and Williams, 2005) to predict the practice utility value. A GP has multiple advantages over linear regression, such as the ability to employ acquisition functions for decision-making, the possibility of using different kernel functions, and providing a natural way to quantify uncertainties in prediction. Based on the estimated utilities for different practice modes, the long-term goal is to choose the one that corresponds to expert choice.

In this study, we fit a GP into the data acquired from the learner performance and the teacher's decision-making, including the target tempo and the practice mode. The goal is to enable a GP model, based on the learner's state, to generate schedules that mimic the guidance of an expert teacher. By doing that, using the model for guidance should, in the long run, improve the effectiveness of practice, even if the teacher does not guide the learner. We will compare the performance of the GP to a linear regression model to determine whether a GP is necessary for modeling this sort of problem.

2. Experimental Setting

Further details about the experimental setting can be found in Appendix A.

Participants The experimental data were collected from 6 novice piano players aged 21 to 28 (who know how to read sheet music, but are at a beginner level). An experienced piano teacher accompanied them. The same teacher was present for all sessions. The study received ethical approval from the Tel Aviv University Institutional Review Board (IRB)

and participants provided written informed consent before the experiment. The participants received payment for their time. Each session was approximately 1 hour.

Equipment   The learner sat at an electric piano (Nord Piano 4), and the practice tasks that the teacher selected were displayed on a computer monitor placed on the piano notes stand. The MIDI data produced by the learners was transferred to the computer for later analysis. Figure 1 shows a photo of the setup.

Figure 1: The experimental setup is shown in this photo. The learner sits at the piano (Nord piano 4) with a computer monitor showing the notes to play. Using the Python interface, the teacher selects which practice mode to play and when to change the piece.

Selection of piece to play   The piece to play was selected by the teacher, who chose a piece that he thought would be suitable and preferably a bit challenging for the student. Sometimes, when the pieces were too difficult or easy, the piece was changed to a different one after two or three practices. All pieces were for the right hand only.

Practice Modes   In the experiment, only two practice modes were used: timing and pitch. An example of a practice mode can be found in Figure 2.

The **timing practice mode** reduces the pitch of all notes to a single pitch and, by doing so, enables the learner to focus on rhythm and timing only, getting rid of the information about the pitch. The **pitch practice mode** allows the learner to practice only the pitch of the music piece and not pay attention to the timing of the notes. To this end, the participants were instructed to focus on the pitch and not the rhythm. A metronome was used in all trials (with the tempo set by the teacher) except in the pitch practice mode.

Teacher role   The teacher defined each practice unit, defined by the music piece, the practice mode, and the tempo (in beats per minute - BPM). The original piece was played by the learner before and after each practice mode to allow calculation of the utility of practice, defined as the difference in errors before and after practice. After the full piece

Figure 2: The upper picture shows an example of one of the pieces to be played. The lower picture shows the selection of a timing practice mode (PM) - the pitch of all notes is uniform, allowing the learner to focus on the timing.

was played, the teacher determined again which practice mode would then be likely to be most beneficial. The teacher also decided when it was time to move to a different piece. From all the experimental sessions, we acquired 121 tuples of performances, consisting of performance before the teacher-selected practice and performance after the practice mode.

3. Methods: Models and Features

3.1. Calculation of features from user data

This work is inspired by a previous study (Moringen et al., 2021). We provide a brief overview below of the calculated features that we used for modeling based on the recorded performance of the learners: pitch error, timing error, and utility of a practice mode. The raw data of learners playing the piano is recorded in MIDI, a protocol that sends out a message every time a note is pressed or released. To simplify the modeling, in this work we do not use MIDI directly, but extract learner performance features based on the ground truth, i.e., correct performance of the piece.

Pitch error is calculated as follows for a given piece: $error_{pitch}(n) := \frac{\sum_{i=1}^{T} d(n_i) f(n_i)}{\sum_{i=1}^{T} d(n_i)}$, where $n := (n_1, \ldots, n_T)$ is the sequence of notes in the piece; f is a function that returns whether the note pitch was played correctly (0 if correct; 1 if incorrect); d is a function that returns the duration of the given note (e.g., 1.0 for a quarter note and 2.0 for a half note).

Timing error is calculated as: $error_{timing}(n) := \frac{1}{m} \sum_{i=1}^{m} \min\{1, t(n_i)\}$, where t is a function that returns the timing offset for the given note as the offset between the time the note was played and the correct time. $n := (n_1, \ldots, n_m)$ is the sequence of notes played

(excluding missed notes). All error measures were calculated based on performance of the original pieces. This work did not analyze errors during practice with the practice modes.

A central feature of our method is the utility u of practice, which measures how effective a practice mode is. To quantify the **utility** u of a practice unit, we measure the error of the learner's performance, playing the whole music piece before practicing and after practicing in a certain practice mode (*pre* and *post* denote before and after the practice, respectively):

$$u(a, u_\mu) = (1 - a)(p_{pre} - p_{post}) + a(t_{pre} - t_{post}) - u_\mu, \tag{1}$$

where p and t are the pitch and timing errors, respectively, the weight a represents the contribution of each modality's improvement to the overall utility value, and u_μ is the mean utility. Note that both u_μ and a are unknown. In the next section, we describe the Bayesian optimization that we use to find the best values for these parameters. The thinking behind optimizing these parameters is to understand how the teacher combines both of these measures to estimate the utility of a given practice mode.

4. Gaussian process and the scaffolding method

The GP was constructed to align the model's prediction to the teacher's selection. It is formalized as a utility predictor, given the state of the learner before practice, practice mode, and speed of practice:

$$g : (p_{pre}, t_{pre}, PM, BPM) \rightarrow u(a, u_\mu), \tag{2}$$

where u is calculated as in Eq. 1. The left side of the above equation corresponds to the user state and the method of practice, while the right side is trained to reflect the usefulness estimation of the corresponding practice by the teacher. **We optimize the hyperparameters a, u_μ, and the GP kernel so that by comparing the utility estimators for different PMs, the highest estimated utility corresponds to the practice chosen by the teacher.** The objective function we employ in the optimization process uses the utility estimates but only to calculate whether the maximum of the predicted utility measures corresponds to the practice mode selected by the teacher: $PM_{teacher} == arg\,max\{g(PM_{timing}), g(PM_{pitch})\}$. This approach does not consider whether the GP model makes correct predictions of the utility value. Note that the current implementation of the above-mentioned scaffolding algorithm makes a strong simplification by assuming that the teacher maximizes the utility of a practice mode based on both values, without any consideration of the practice history. In future work, we plan to additionally employ the uncertainty of the utility estimator that a GP also provides. It is an open research question, how to integrate the practice context into a computational practice policy generator. At the moment we can answer the following questions based on the current learner state only: "For my current level of errors, what tempo should I select for timing practice of the piece?" or "For my current level of errors, what is the best PM to choose, and at what speed to practice?"

The kernels examined here were the rational quadratic kernel (RatQuad), radial basis function (RBF), and Matérn with $\nu = 5/2$ (Rasmussen and Williams, 2005). The hyperparameters we optimized explicitly with Bayesian optimization were a and u_μ. The other

parameters, multidimensional lengthscale, variance and α (in the case of RatQuad) were optimized with the automatic relevance determination (ARD) using the GPyOpt libary[1]

5. Methods - Linear regression

We compared the GP with linear regression to test whether linear regression is sufficient for predicting which PM to use. We fit a linear regression model to the training data to predict the utility u (Equation 1) as a function of the pitch and timing error before the practice, the practice mode used, and their interactions:

$$u = b_1 + b_2 \cdot PM + b_3 \cdot t_{pre} + b_4 \cdot p_{pre} + b_5 \cdot PM \cdot t_{pre} + b_6 \cdot PM \cdot p_{pre} \tag{3}$$

where $b = \begin{bmatrix} b_1 & b_2 & b_3 & b_4 & b_5 & b_6 \end{bmatrix}$ are the regression parameters fit to the data.

The linear regression was trained to predict the utility on data for all subjects combined. As the utility function has one free parameter (a in Equation 1), we ran multiple linear regressions (using a grid search to test exhaustively values from 0 to 1) to find the one that most accurately predicts the teacher-selected activity assuming that the teacher selects the activity with the greatest utility. We also tested whether including tempo (BPM) improved the prediction accuracy. In addition, we tested whether directly predicting the teacher's selection (i.e., which practice mode - PM) would be more successful using logistic regression:

$$PM = b_1 + b_2 \cdot t_{pre} + b_3 \cdot p_{pre} \tag{4}$$

6. Results

6.1. Gaussian process

Figure 3 shows an example of the practice policy provided by the teacher and policies generated by maximizing the estimated practice utilities predicted by a trained GP for BPM=50, 80, and 100, using a RatQuad kernel. After 50 iterations of Bayesian optimization of a and u_μ, the scaffolding algorithm on average correctly predicts 70% of the practice modes on test data. This computation takes about a minute on a regular PC. Further Bayesian optimization iterations result in a slight increase of accuracy, yielding 75% accuracy.

6.2. Linear Regression

The linear regression was able to predict reasonably well the utility based on the error measures before the practice ($R^2 = 0.357, p < 0.001$). We found that selecting the activity based on the maximum utility was closest to the ground truth (from the piano teacher) when $a = [0, 0.13]$ (i.e., any a value chosen in this range was equivalent). We note that this includes the value of $a = 0$, i.e., only using the pitch data. However, the accuracy of the linear regression was poor (only 38.0%), i.e., less than chance.

When using logistic regression to predict the teacher's selection directly, the accuracy was somewhat higher (69.4%); this increases to 70.2% if BPM is included. In this case, the equation was:

$$PM = -0.283 + 8.124 \cdot t_{pre} - 6.734 \cdot p_{pre} \tag{5}$$

where the timing practice model was selected if $PM > 0.5$.
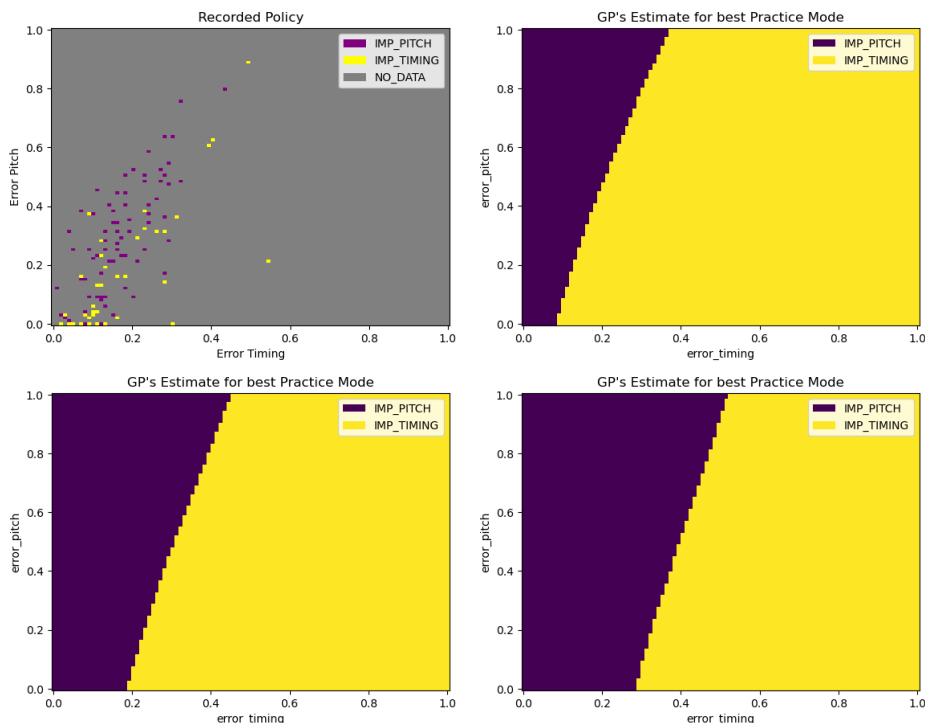
---

1. https://sheffieldml.github.io/GPyOpt/

Figure 3: (Upper left) An illustration of data points recorded during expert-guided piano practice session for all learners and pieces. x-axis denotes timing error; y-axis denotes pitch error before a PM (pitch practice in violet, timing practice in yellow). (Upper right, lower panels) Policies generated by a GP RatQuad kernel, BMP=50, 80 and 100. For slower tempos, the model's policy predicts a higher utility of timing practice; for higher tempos - a higher utility of pitch practice.

7. Discussion

The GP produced a mapping between errors and desired practice modes that makes sense (see Figure 3), while the accuracy of both techniques was similar (approximately 70% for the baseline model and 75% as the best result achieved by the GP model). Using more error measures may show a larger difference between the techniques because it is likely that the relationship between predictor variables and the utility is non-linear (Schulz et al., 2018).

The model has a greater tolerance for small timing errors than pitch errors - when the timing error is low, the pitch PM is always selected. This may be because variability in timing is a typical feature of playing (Vugt et al., 2012), whereas any variability in pitch (which key is pressed) is considered an error.

Future studies will compare using the GP to provide learners with the next PM to a control group who practice with randomly scheduled PMs. The goal is to test whether this use of GP accelerates motor learning of the piano. Another research direction is to use different types of acquisition functions and investigate how similar the schedule generated

by the teacher and by the acquisition functions are, and whether it makes sense for a learner to practice within a Bayesian optimization loop. A further challenge is to find an approach to modeling of the practice context within one session.

8. Data availability

The code for data collection and analysis, and the collected data are available.

Acknowledgments

References

Alfred Cortot, editor. *Chopin: 12 Studies for Piano, Op. 10.* BMG Ricordi, 1986. ISBN 978-1-4803-0456-7. Editor: Cortot, Alfred.

Shinichi Furuya and Eckart Altenmüller. Flexibility of movement organization in piano performance. *Frontiers in Human Neuroscience*, 7:173, 2013. doi: 10.3389/fnhum. 2013.00173. URL http://journal.frontiersin.org/article/10.3389/fnhum.2013. 00173/abstract.

Alexandra Moringen, Sören Rüttgers, Luisa Zintgraf, Jason Friedman, and Helge Ritter. Optimizing piano practice with a utility-based scaffold. *arXiv:2106.12937 [cs]*, June 2021. URL http://arxiv.org/abs/2106.12937. arXiv: 2106.12937.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning.* The MIT Press, Cambridge, MA, November 2005. doi: 10.7551/mitpress/3206.001.0001. URL https://direct.mit.edu/books/book/2320/ Gaussian-Processes-for-Machine-Learning.

Limor Raviv, Gary Lupyan, and Shawn C. Green. How variability shapes learning and generalization. *Trends in Cognitive Sciences*, 26(6):462–483, June 2022. ISSN 1364-6613. doi: 10.1016/j.tics.2022.03.007. URL https://www.sciencedirect.com/science/ article/pii/S1364661322000651.

Eric Schulz, Maarten Speekenbrink, and Andreas Krause. A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions. *Journal of Mathematical Psychology*, 85:1–16, August 2018. ISSN 0022-2496. doi: 10.1016/j.jmp.2018.03.001. URL https://www.sciencedirect.com/science/article/pii/S0022249617302158.

Hila Tamir-Ostrover, Gilad Baruch, Or Peleg, Yonatan Yellin, Maor Rosenberg, Alexandra Moringen, Kathrin Krieger, Helge Ritter, and Jason Friedman. Automatic Evaluation of Aspects of Performance and Scheduling in Playing the Piano. In *Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization*, UMAP '22, pages 276–285, New York, NY, USA, July 2022. Association for Computing Machinery. ISBN 978-1-4503-9207-5. doi: 10.1145/3503252.3531297. URL https://doi.org/10. 1145/3503252.3531297.

John Thompson. *John Thompson's Modern Course for the Piano: the First Grade Book : Something New Every Lession.* Willis Music, Florence, Kentucky, 2009. ISBN 978-0-87718-060-9. OCLC: 838681794.

Floris Tijmen van Vugt, Hans-Christian Jabusch, and Eckart Altenmüller. Fingers phrase music differently: trial-to-trial variability in piano scale playing and auditory perception reveal motor chunking. *Frontiers in Auditory Cognitive Neuroscience*, 3:495, 2012. doi: 10.3389/fpsyg.2012.00495. URL http://www.frontiersin.org/auditory_cognitive_ neuroscience/10.3389/fpsyg.2012.00495/abstract.

Janet Mannheimer Zydney. Scaffolding. In Norbert M. Seel, editor, *Encyclopedia of the Sciences of Learning*. Springer US, Boston, MA, 2012. ISBN 978-1-4419-1427-9 978-1-4419-1428-6. doi: 10.1007/978-1-4419-1428-6. URL http://link.springer.com/10.1007/978-1-4419-1428-6.

Appendix A. Experimental setting

We will conduct an experiment to develop an artificial intelligence algorithm that will assist in learning piano playing. The algorithm will provide accurate recommendations about what the ideal exercise is for the learner based on their mistakes while playing at the current moment. The current experiment is intended to provide a basic ground truth for training the model. In this experiment, the expert will perform the algorithm's decision-making process in its place and will guide the choice of exercises for the learner based on her playing at the current moment. As usual in preliminary computational models, it's important to clarify that this is currently a limited model. The model measures only two errors out of the learning process - pitch and timing errors. We aim to create a successful model based on these two, and later, this can be expanded to include other types of errors.

The current experiment is performed iteratively. First, the expert must choose a piece that suits the learner and is at an appropriate pace. Then, the learner will play the piece in its entirety. The expert will assess her playing and recommend one of two training exercises, according to his professional discretion: Pitch: playing only the correct notes, without regard to rhythm. Timing: playing according to rhythm only, without regard to which note (key) to press. According to the choice, input will be entered into the system, and the software will show the learner the chosen exercise. In this experiment, we will perform each practice once. Afterward, the learner plays the full piece again to assess the improvement in her error. The process will be performed iteratively until the errors in the piece are sufficiently small, in the expert's opinion. Then, we will move on to the next piece.