

Logistic Knowledge Tracing Tutorial: Practical Educational Applications

Philip I. Pavlik Jr., Luke G.
Eglington*, Meng Cao, and Wei
Chu
University of Memphis, *Amplify Edu-
cation Inc.
ppavlik@memphis.edu,
leglington@amplify.com,
mcao@memphis.edu

ABSTRACT

Logistic Knowledge Tracing (LKT) is a framework for combining various predictive features into student models that are adaptive, interpretable, explainable, and accurate. While the name logistic knowledge tracing was coined for our R package that implements this methodology for making student models, logistic knowledge tracing originates with much older models such as Item Response Theory (IRT), the Additive Factors Model (AFM), and Performance Factors Analysis (PFA), which exemplify a type of model where student performance is represented by the sum of multiple components each with some sort of feature computed for the component. Features may range from the simple presence or absence of the component to complex functions of the prior history of the component. The LKT package provides a simple interface to this methodology, allowing old models to be specified or new models to be created by mixing and matching components with features. We will provide concrete examples of how the LKT framework can provide interpretable results on real-world datasets while being highly accurate.

Keywords

Keywords are your own designated keywords. Choose around 3–5 keywords, separated by commas.

1. INTRODUCTION

The task of knowledge tracing has become ubiquitous in educational data mining and intelligent learning system design. While there is still no great agreement on the best way to do knowledge tracing, it has become a popular method, and knowledge tracing models of some sort occur in most adaptive learning systems due to the need for a model to capture the complexity of learning well enough to react to it with pedagogical decisions, i.e., instruction.

Knowledge tracing originates in the cognitive modeling of the early 60s but has since spawned many varieties. Early models were often described in terms of Markov models, and Bayesian Knowledge Tracing dominated for many years [e.g., 7]. Over the years, more options have been revealed [25]. More recently, deep learning approaches have also been shown to be accurate for predicting learning [26].

P. I. P. Jr., L. G. Eglington, M. Cao, and W. Chu. Logistic knowledge tracing tutorial: Practical educational applications. In B. Paaßen and C. D. Epp, editors, *Proceedings of the 17th International Conference on Educational Data Mining*, pages 1009–1012, Atlanta, Georgia, USA, July 2024. International Educational Data Mining Society.

© 2024 Copyright is held by the author(s). This work is distributed under the Creative Commons Attribution NonCommercial NoDerivatives 4.0 International (CC BY-NC-ND 4.0) license.
<https://doi.org/10.5281/zenodo.12730033>

This tutorial will not spend much time on various models but explore how to understand, use, and deploy logistic regression models of knowledge tracing. This family of knowledge tracing models that includes logistic knowledge tracing is very broad because the underlying method is simply regression, and the contributions of different factors of learning can be included by simply adding them as predictors in the regression [4, 5, 11, 12, 15, 20, 22]. This leads to several issues to consider, which will be discussed in this tutorial.

1.1 Content and Timeline

1.1.1 Features of components as predictors (30 mins)

The basis of prior LKT models is revealed as features (e.g., a count of the prior instances) for a component of the data (e.g., a column identifier). Each feature is computed for all levels (e.g., KCs) in the identifier column, which occurs per student [16, 17, 23, 24].

1.1.1.1 Components

Components is the term used to describe the factors of the data that are used to “compute” a predictor, like for the student overall or a particular KC or item. Computing a predictor for a component might mean simply fitting an intercept to each of its levels. Still, it could also be complex and involve parameterized curvilinear effects as a function of the component's prior history. The possible computations for a component are described as “features”.

1.1.1.2 Feature types

Features are simple and not-so-simple functions of the component and its history. They can be usefully grouped into 3 types. Static features are simple, including the intercepts of student, item, or knowledge component. Dynamic features are more interesting and are most typified by the “lineafm” feature popularized in the DataShop project [14]. Lineafm is a linear effect of the count of prior learning events for a component. Most dynamic and adaptive features are computed for each component separately for each student since we typically look for changes due to prior learning or adaptation as a function of prior performance within the student. Adaptive features are typified by PFA [20] variants, where the feature uses counts of success and/or failures to adjust the predictions for a KC or item. Adaptive features are quite powerful at the student level and can be used to provide a dynamic estimate of achievement.

1.1.1.3 Non-linear parameters in features

Many curvilinear features are allowed, where the feature computed for the component-based using 1 or more non-linear parameters that are solved for by nesting the logistic regression in a gradient descent optimizer. One useful dynamic feature is recency, typically applied for items or KCs by computing the t^{-d} as the feature where

t is the time since the KC or item was last encountered. More complex temporal features are available.

1.1.2 Data preparation (30 mins Phil)

Describe the requirements for getting models to run, including all preprocessing. This will include discussing time-based features and their requirements for time-formatted data. You may bring data that follows the requirements explained here: https://cran.r-project.org/web/packages/LKT/vignettes/Basic_Operations.html, or use the embedded data built into the package. For those with their data, we will help with various formatting problems and provide advice about special cases of data formats and how conversions can occur.

1.1.3 Previous classes of LKT model (30 mins Luke)

We will review the many varieties of LKT models that have been seen in prior work (e.g., primitive like AFM [3] or PFA[20]). We will examine the function inputs and outputs of LKT and see examples of many features in past work. We will create and fit the AFM and PFA models using your data or the embedded data. Then, we will add the recency feature to illustrate it and how to add non-linear features to the logistic regression model.

1.1.4 Searching for optimal feature & Crossvalidation (30 mins Luke)

LKT is under continuous development as part of the Optimal Learning Lab at the University of Memphis. In 2023, we introduced two methods for model search: stepwise feature search and LASSO feature search [17]. These methods are quite powerful, and we will review the results of this paper, which showed that LKT models can do better than some deep learning approaches. Using your data or the embedded data, we will illustrate how to crossvalidate results from complex models using the LKT functions to compute the test fold feature for held-out data.

1.1.5 Creating new features (30 mins Phil)

We will go under the hood with the LKT codebase and illustrate how to add a feature. We will give a hands-on tutorial where we create a new feature in your data and use it in LKT as part of the model we have been developing throughout the tutorial. This will illustrate how you can do feature engineering in the context of and using LKT. LKT is open-source, and we are glad to add any new stable and useful features to the R package on the CRAN repository.

1.1.6 Application to Optimal Learning (30 mins Luke and Phil)

We will walk through a concrete example of using a fitted LKT model for practice scheduling. In the MoFaCTS LKT model, predictions are used in combination with a decision rule to sequence practice. We will walk through the code flow and examine how it solves the problems: what to learn next, when to increase spacing, and when to introduce new items for learning. We will show how the output can be used in combination with decision rules to guide a student's practice sequence. We will discuss how the model's features influence the interpretation of the model output and its potential use in an ITS.

1.1.7 Application to proficiency reporting (30 mins Luke)

Another primary use is reporting proficiency (to teachers, students, or administrators). For reporting proficiency, we will demonstrate how the output of the knowledge tracing model can be used to

group students according to shared ability or provide proficiency estimates that are superior to cumulative averages. Grouping students requires transforming the model outputs into category labels (e.g., "Needs work", "on track", "proficient") as well as pedagogical theory to guide what other factors should determine student grouping. Finally, practice items are sometimes related to multiple skills simultaneously. We will show how this situation can be accounted for with LKT.

2. FORMAT

The tutorial will be lecture and discussion, but we hope to go at a pace where prepared participants (with R and R-studio on their computers) can follow along executing examples in real-time to obtain any tailored assistance with software and interpretation of results.

3. PRESENTERS

Philip Pavlik is an Associate Professor at the Institute for Intelligent Systems at the University of Memphis (ppavlik@memphis.edu). His research program, which began in January 2001, has focused on the effects of practice and forgetting on performance in declarative memory tasks [18]. The ultimate goals of this research are to create accurate mathematical models of practice and forgetting that can be used to understand the factors governing recall performance and to use these models to improve performance in educational settings. This work applies economics and cognitive psychology principles to look for how to schedule practice for maximal benefit for the student [9, 10, 23, 27].

Pavlik's approach is informed by his training at Carnegie Mellon, where he worked with researchers John Anderson and Ken Koedinger. As a graduate student working with John Anderson, Pavlik developed new learning models and applied them to instructional strategy optimization [19]. This work continued with Ken Koedinger, co-director of the 10-year Pittsburgh Science of Learning Center [20].

At the University of Memphis, Pavlik leads the Optimal Learning Lab, which applies cognitive theory and methods to educational contexts. He and his students have investigated learning in various domains, including mathematics, language learning, statistics, musical learning, and anatomy and physiology. In this work, he develops software and tools for educational applications, such as the Mobile Fact and Concept Training System (MoFaCTS)[21], which serves as a testbed for educational interventions, and the R package logistic knowledge tracing (LKT)[17, 24], which aims to make student modeling more efficient and accurate.

Luke Eglington is a Staff Learning Scientist at Amplify Education (leglington@amplify.com). Luke completed his PhD at Dartmouth College with supervisor Dr. Sean Kang. As a postdoctoral fellow at the University of Memphis, he worked with Dr. Philip Pavlik to develop novel knowledge tracing models and methods of using them to optimally sequence practice. His research initially focused on the effects of retrieval practice [8], spacing [13], and interleaving [28] on student learning and metacognition. More recently, he has investigated how best to build knowledge-tracing models [7, 10, 18, 25] and how to use them to sequence practice optimally [9]. He is also interested in developing methods for better use of the outputs of KT models, such as adaptive student grouping, practice decision rules, and human-interpretable (and actionable) reports.

Meng Cao is a Ph.D. candidate working with Dr. Pavlik in the Optimal Learning Lab at the University of Memphis (mcao@memphis.edu). Her research primarily focuses on

optimizing training sequences for Mandarin tone learning and category learning. In her work, she has extended the Performance Factors Analysis Model (PFA) by integrating prior practice difficulty and categorization theories, aiming to determine the optimal item difficulty [2]. This strategic incorporation facilitates the application of the PFA model in adaptive training systems for personalized item selection during practice sessions. More recently, she has extended to integrating attentional factors into the PFA model [1]. This inquiry seeks to understand the nuanced impact of interleaving and blocking in category learning.

Wei Chu is a fifth-year Ph.D. candidate working with Dr. Pavlik in the Optimal Learning Lab at the University of Memphis (wchu@memphis.edu). Her research focused on applications of the spacing effect and testing effect in the context of second language learning. More recently, she has explored how to track learners' correct response time fluctuations to improve the predictive accuracy of the Performance Factors Analysis model, utilizing the logistic knowledge tracing (LKT) package [6].

4. TARGET AUDIENCE

The target audience is rather broad since we intend to give a detailed explanation of the LKT methodology before discussing how to implement such models. For this reason, we welcome anyone interested in quantitative student models of skill or knowledge learning. We expect no more than 25 participants.

5. ACKNOWLEDGMENTS

The preparation of this manuscript was supported in part by National Science Foundation award #2301130.

6. REFERENCES

- [1] Cao, M. and Pavlik Jr, P.I., 2022. A Variant of Performance Factors Analysis Model for Categorization. In *Proceedings of the 15th International Conference on Educational Data Mining*, A. Mitrovic and N. Bosch Eds. International Educational Data Mining Society, Durham, United Kingdom, 763-766. DOI= <http://dx.doi.org/10.5281/zenodo.6852974>.
- [2] Cao, M., Pavlik Jr, P.I., and Bidelman, G.M., 2019. Incorporating Prior Practice Difficulty into Performance Factor Analysis to Model Mandarin Tone Learning. In *Proceedings of the 12th International Conference on Educational Data Mining*, C. Lynch, A. Merceron, M. Desmarais and R. Nkambou Eds., 516-519.
- [3] Cen, H., Koedinger, K.R., and Junker, B., 2006. Learning Factors Analysis - A general method for cognitive model evaluation and improvement. In *Proceedings of the 8th International Conference on Intelligent Tutoring Systems* Springer Berlin / Heidelberg, 164-175.
- [4] Cen, H., Koedinger, K.R., and Junker, B., 2008. Comparing two IRT models for conjunctive skills. In *Proceedings of the Proceedings of the 9th International Conference on Intelligent Tutoring Systems* (Montreal, Canada2008), ACM, Inc., 796-798.
- [5] Chi, M., Koedinger, K.R., Gordon, G., Jordan, P., and VanLehn, K., 2011. Instructional factors analysis: A cognitive model for multiple instructional interventions. In *Proceedings of the 4th International Conference on Educational Data Mining (EDM 2011)*, M. Pechenizkiy, T. Calders, C. Conati, S. Ventura, C. Romero and J. Stamper Eds. International Educational Data Mining Society, Eindhoven, The Netherlands, 61-70. DOI= <http://dx.doi.org/http://doi:10.1.1.230.9907>.
- [6] Chu, W. and Philip, I.P., Jr., 2023. The Predictiveness of PFA is Improved by Incorporating the Learner's Correct Response Time Fluctuation. In *Proceedings of The 16th International Conference on Educational Data Mining*, F. Mingyu, K. Tanja and T. Partha Eds., 244-250. DOI= <http://dx.doi.org/10.5281/zenodo.8115643>.
- [7] Doroudi, S. and Brunskill, E., 2017. The Misidentified Identifiability Problem of Bayesian Knowledge Tracing. In *Proceedings of 10th International Conference on Educational Data Mining*, X. Hu, T. Barnes, A. Hershkovitz and L. Paquette Eds., 143-149.
- [8] Eglington, L.G. and Kang, S.H.K., 2018. Retrieval Practice Benefits Deductive Inference. *Educational Psychology Review* 30, 1 (2018/03/01), 215-228. DOI= <http://dx.doi.org/10.1007/s10648-016-9386-y>.
- [9] Eglington, L.G. and Pavlik Jr, P.I., 2020. Optimizing practice scheduling requires quantitative tracking of individual item performance. *npj Science of Learning* 5, 1 (2020), 15. DOI= <http://dx.doi.org/10.1038/s41539-020-00074-4>.
- [10] Eglington, L.G. and Pavlik, P.I., 2023. How to Optimize Student Learning Using Student Models That Adapt Rapidly to Individual Differences. *International Journal of Artificial Intelligence in Education* 33, 3 (2023/09/01), 497-518. DOI= <http://dx.doi.org/10.1007/s40593-022-00296-0>.
- [11] Galyardt, A. and Goldin, I., 2015. Move your lamp post: Recent data reflects learner knowledge better than older data. *Journal of Educational Data Mining* 7, 2 (2015), 83-108. DOI= <http://dx.doi.org/https://doi.org/10.5281/zenodo.3554671>.
- [12] Gong, Y., Beck, J.E., and Heffernan, N.T., 2011. How to construct more accurate student models: Comparing and optimizing knowledge tracing and performance factor analysis. *International Journal of Artificial Intelligence in Education* 21, 1 (2011), 27-46. DOI= <http://dx.doi.org/http://doi:10.3233/JAI-2011-016>.
- [13] Kang, S.H.K., Eglington, L.G., Schuetze, B.A., Lu, X., Hinterstoisser, T.M., and Huaco, J., 2023. Using Cognitive Science and Technology to Enhance Financial Education: The Effect of Spaced Retrieval Practice. *Journal of Financial Counseling and Planning*, 1, 20-31. DOI= <http://dx.doi.org/10.1891/JFCP-2021-0032>.
- [14] Koedinger, K.R., Baker, R.S., Cunningham, K., Skogsholm, A., Leber, B., and Stamper, J., 2010. A data repository for the EDM community: The PSLC DataShop. In *Handbook of Educational Data Mining*, C. Romero, S. Ventura and M. Pechenizkiy Eds. CRC Press, CRC Press, Boca Raton, FL, USA, 43-56. DOI= <http://dx.doi.org/http://doi:10.1201/B10274-6>.
- [15] Liu, R. and Koedinger, K.R., 2017. Towards reliable and valid measurement of individualized student parameters. In *International Conference on Educational Data Mining*, X. Hu, T. Barnes, A. Hershkovitz and L. Paquette Eds., Wuhan, China, 135-142.
- [16] Pavlik Jr, P.I. and Eglington, L.G., 2021. LKT: Logistic Knowledge Tracing. In CRAN R package version 1.6.
- [17] Pavlik Jr, P.I. and Eglington, L.G., 2023. Automated Search Improves Logistic Knowledge Tracing, Surpassing Deep Learning in Accuracy and Explainability. *Journal of Educational Data Mining* 15, 3 (12/26), 58-86. DOI= <http://dx.doi.org/10.5281/zenodo.10363337>.
- [18] Pavlik Jr., P.I. and Anderson, J.R., 2005. Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. *Cognitive Science* 29, 4

- (2005), 559–586. DOI=http://dx.doi.org/10.1207/s15516709cog0000_14.
- [19] Pavlik Jr., P.I. and Anderson, J.R., 2008. Using a model to compute the optimal schedule of practice. *Journal of Experimental Psychology: Applied* 14, 2 (2008), 101–117. DOI= <http://dx.doi.org/10.1037/1076-898X.14.2.101>.
- [20] Pavlik Jr., P.I., Cen, H., and Koedinger, K.R., 2009. Performance factors analysis -- A new alternative to knowledge tracing. In *Proceedings of the 14th International Conference on Artificial Intelligence in Education*, V. Dimitrova, R. Mizoguchi, B.D. Boulay and A. Graesser Eds. IOS Press, Brighton, England, 531–538. DOI= <http://dx.doi.org/http://doi:10.3233/978-1-60750-028-5-531>.
- [21] Pavlik Jr., P.I., Kelly, C., and Maass, J.K., 2016. Using the mobile fact and concept training system (MoFaCTS). In *Proceedings of the 13th International Conference on Intelligent Tutoring Systems*, A. Micarelli and J. Stamper Eds. Springer, Switzerland, 247-253.
- [22] Pavlik Jr., P.I., Yudelson, M., and Koedinger, K.R., 2011. Using contextual factors analysis to explain transfer of least common multiple skills. In *15th International Conference Artificial Intelligence in Education*, G. Biswas, S. Bull, J. Kay and A. Mitrovic Eds. Springer, Auckland, New Zealand, 256–263. DOI= http://dx.doi.org/http://doi:10.1007/978-3-642-21869-9_34.
- [23] Pavlik Jr., P.I. and Zhang, L., 2022. Using autoKC and Interactions in Logistic Knowledge Tracing. In *Proceedings of The Third Workshop of the Learner Data Institute, The 15th International Conference on Educational Data Mining (EDM 2022)*, Durham, UK, 1-6.
- [24] Pavlik, P.I., Eglinton, L.G., and Harrell-Williams, L.M., 2021. Logistic Knowledge Tracing: A Constrained Framework for Learner Modeling. *IEEE Transactions on Learning Technologies* 14, 5, 624-639. DOI= <http://dx.doi.org/10.1109/TLT.2021.3128569>.
- [25] Pelánek, R., 2017. Bayesian knowledge tracing, logistic models, and beyond: an overview of learner modeling techniques. *User Modeling and User-Adapted Interaction* 27, 3 (2017), 313-350. DOI= <http://dx.doi.org/http://doi:10.1007/s11257-017-9193-2>.
- [26] Piech, C., Spencer, J., Huang, J., Ganguli, S., Sahami, M., Guibas, L., and Sohl-Dickstein, J., 2015. Deep Knowledge Tracing. *arXiv preprint arXiv:1506.05908*.
- [27] Scruggs, R., Baker, R.S., Pavlik, P.I., McLaren, B.M., and Liu, Z., 2023. How well do contemporary knowledge tracing algorithms predict the knowledge carried out of a digital learning game? *Educational Technology Research and Development* 71, 901-918. DOI= <http://dx.doi.org/10.1007/s11423-023-10218-z>.
- [28] Yan, V.X., Schuetze, B.A., and Eglinton, L.G., 2020. A review of the interleaving effect: Theories and lessons for future research. *PsyArXiv*(2020), 1-39. DOI= <http://dx.doi.org/10.31234/osf.io/ur6g7>.