TrueLearn: A Python Library for Personalised Informational Recommendations with (Implicit) Feedback

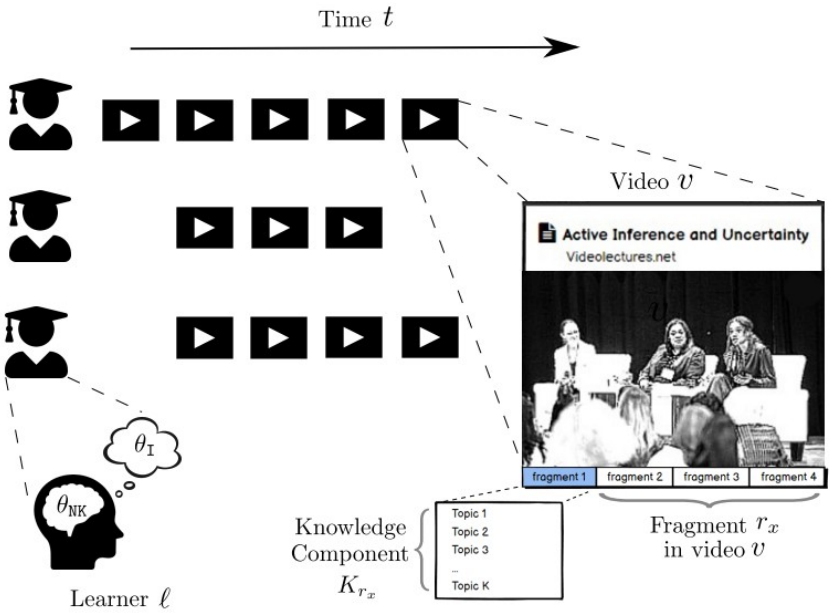
YUXIANG QIU, KARIM DJEMILI, DENIS ELEZI, AANEEL SHALMAN, MARÍA PÉREZ-ORTIZ AND SAHAN BULATHWELA

UCL Centre for Artificial Intelligence

Educational Recommendation

- Matching relevant educational materials to the right learner at the right time
- Goes beyond prepared curriculum
- Can leverage a large collection of educational materials to personalise
- Need to account for informational novelty that goes beyond content similarity ***[Bulathwela et al., AAAI'20]***
- Need transparency to build trust ***[Bull & Kay, Adv. In ITS'2010]***

Problem Setting



[Bulathwela et al., MDPI Sustainability 2022].

Related Work

Main task of focus, personalise exercises. Python library **PyBKT** ***[Bulut et al., MDPI Psych 2023]*** available.

$$P(correct\ answer|\theta_l, d_r)$$

Two main approaches for personalising education

1) Bayesian Knowledge Tracing (HMM of Learner State)
2) Item Response Theory (A function of Learner State and Exercise Difficulty)

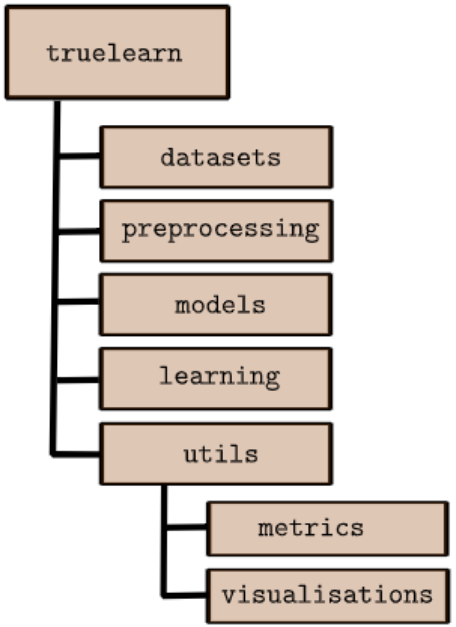
Training is expensive and only focuses on a small number of exercises and learners

Related Work

- TrueLearn, one of the first models to predict video engagement with educational videos ***[Bulathwela et al., AAAI'20]***
- Has multiple models to recover a knowledge and interest representation ***[Bulathwela et al., MDPI Sustainability 2022]***
- Also, can formulate a transparent learner state representation
- There is no software that allows evaluation and implementation of educational recommenders

TrueLearn Library

- Can be used to access datasets, use personalisation models, visualise user states and for offline evaluation
- Inspired by API design of scikit-learn, a well-tested ML library
- API is designed for usability and extensibility

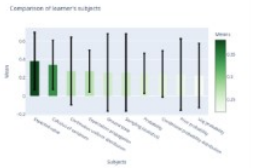


Experimental Results

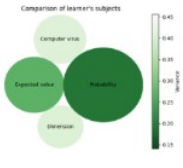
- Validate the implementation accuracy by aligning our results with prior work results
- The PEEK dataset was used to fit the models ***[Bulathwela et al., ORSUM'21]***
- Classification metrics are used to evaluate performance

Model	Acc.	Prec.	Rec.	F1
TrueLearn Interest	58.13	52.08	*78.61*	63.00
TrueLearn Novelty	*64.78*	*58.52*	**80.91**	**65.53**
TrueLearn INK	**78.32**	**64.32**	64.03	*64.00*

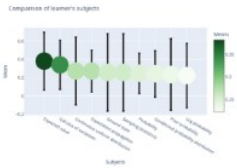
Visualization Capabilities



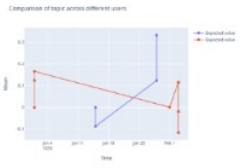
BarPlotter Example



BubblePlotter Example



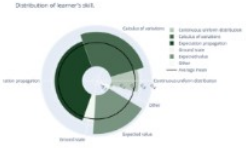
DotPlotter Example



LinePlotter Example

PiePlotter Example

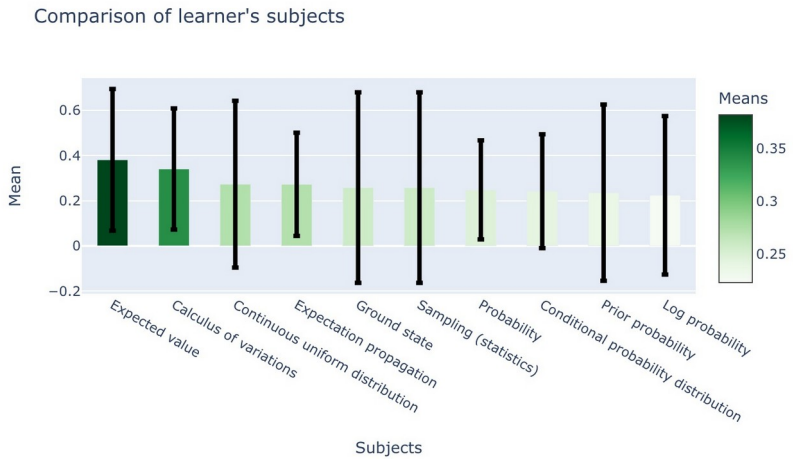
RadarPlotter Example

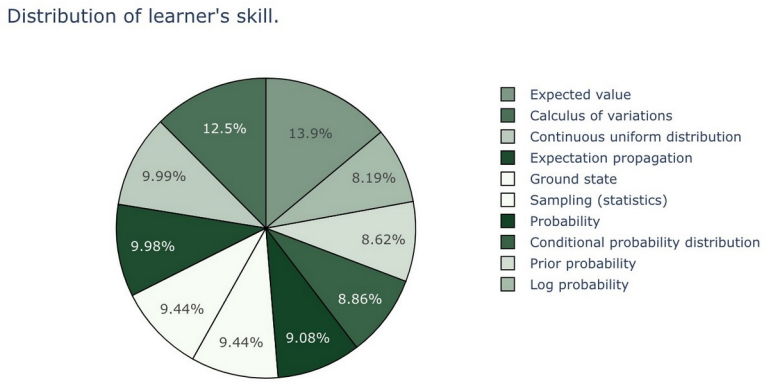


RosePlotter Example

TreePlotter Example

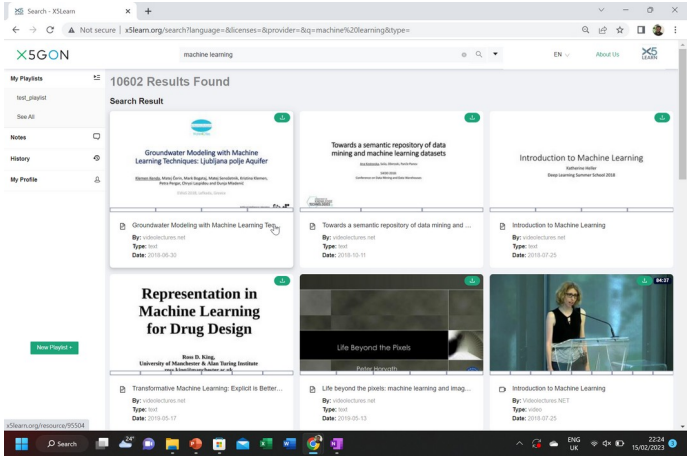
Visualization Capabilities



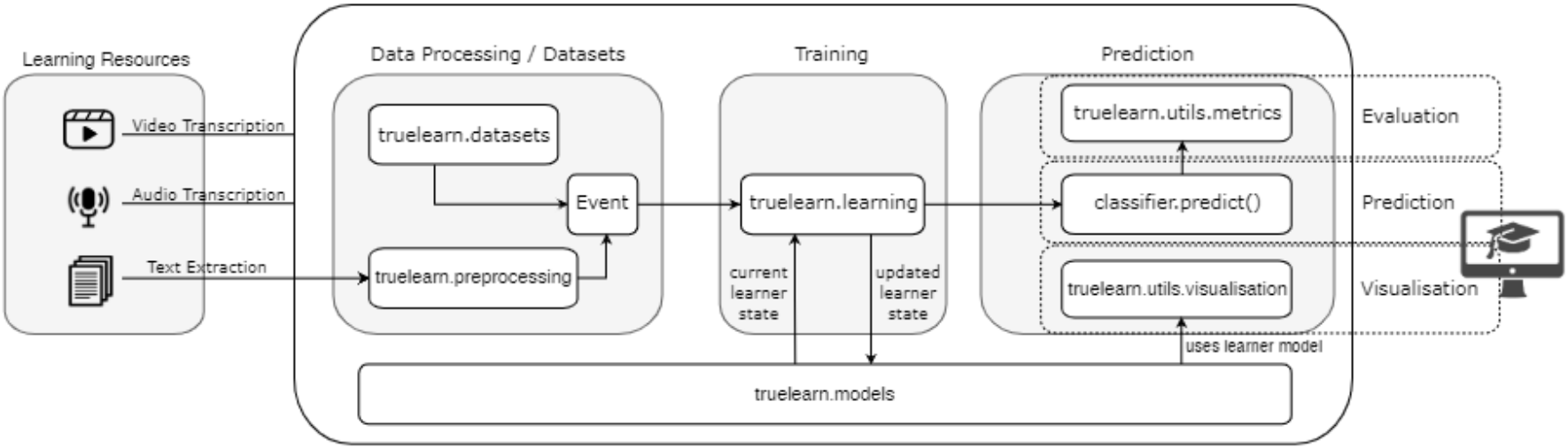


Impact

- A scalable library to model learner engagement for real-world educational videos and to carry out offline evaluation of learner modelling algorithms with datasets and metrics
- Extensive documentation, examples and specific instructions for contributors
- Currently being integrated to https://x5learn.org/



Impact


Learning Resources
Video Transcription
Audio Transcription
Text Extraction
Data Processing / Datasets
truelearn.datasets
Event
truelearn.preprocessing
Training
truelearn.learning
current learner state
updated learner state
Prediction
truelearn.utils.metrics
classifier.predict()
truelearn.utils.visualisation
Evaluation
Prediction
Visualisation
uses learner model
truelearn.models

Future Work

- Incompatibility
- Extend Classifier
- Empirical Study on Visualizations
- Generalisability

Department of Computer Science

TrueLearn: A Python Library for Personalised Informational Recommendations with (Implicit) Feedback

https://github.com/TrueLearnAI/truelearn

https://truelearn.readthedocs.io/en/latest/

YUXIANG QIU, KARIM DJEMILI, DENIS ELEZI, AANEEL SHALMAN, MARÍA PÉREZ-ORTIZ AND SAHAN BULATHWELA

UCL Centre for Artificial Intelligence

m.bulathwela@ucl.ac.uk

References

Lars Buitinck, Gilles Louppe, Mathieu Blondel, Fabian Pedregosa, Andreas Mueller, Olivier Grisel, Vlad Niculae, Peter Prettenhofer, Alexandre Gramfort, Jaques Grobler, Robert Layton, Jake VanderPlas, Arnaud Joly, Brian Holt, and Gaël Varoquaux. 2013. API design for machine learning software: experiences from the scikit-learn project. CoRR abs/1309.0238 (2013). arXiv:1309.0238 http://arxiv.org/abs/1309.0238

Sahan Bulathwela, María Pérez-Ortiz, Emine Yilmaz, and John Shawe-Taylor 2020. TrueLearn: A Family of Bayesian Algorithms to Match Lifelong Learners to Open Educational Resources. In AAAI Conf. on Artificial Intelligence,

Sahan Bulathwela, Maria Perez-Ortiz, Erik Novak, Emine Yilmaz, and John Shawe-Taylor. 2021. PEEK: A Large Dataset of Learner Engagement with Educational Videos. In Proc. of RecSys Workshop on Online Recommender Systems and User Modeling (ORSUM'21). https://arxiv.org/abs/2109.03154

Bulathwela, Sahan and Pérez-Ortiz, María and Yilmaz, Emine and Shawe-Taylor, John. 2022. Power to the Learner: Towards Human-Intuitive and Integrative Recommendations with Open Educational Resources. Sustainability 14, 18 (2022). https://doi.org/10.3390/su141811682

Susan Bull and Judy Kay. 2010. Open Learner Models. Vol. 308. 301–322. https://doi.org/10.1007/978-3-642-14363-2_15

Okan Bulut, Jinnie Shin, Seyma N. Yildirim-Erbasli, Guher Gorgun, and Zachary A. Pardos. 2023. An Introduction to Bayesian Knowledge Tracing with pyBKT. Psych 5, 3 (2023), 770–786.