Kometa

Knowledge Community for Efficient Training through Virtual Technologies

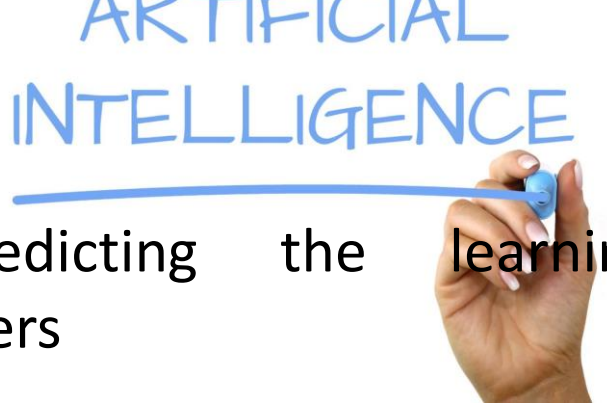
Training in a Virtual Learning Environment: A Process Mining Approach

Annalisa Appice, Pasquale Ardimento, Donato Malerba, Giuseppe Modugno, Diego Marra, Marco Mottola

Presented at the Special Session in Evolving Business Process Management (EvoBPM) @ IEEE EAIS 2020

Scenario

- Workers' training
  - ✓ 360 Video
  - ✓ Virtual Reality
  - ✓ Artificial Intelligence → predicting the learning achievements of the trained workers



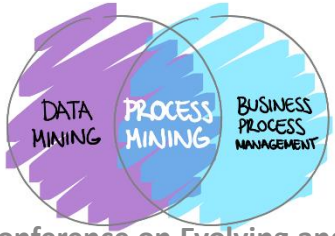
Novelties

- The scenario is different from the one that is commonly investigated in the education studies …
  - ✓ Students are workers instead of under-graduate or post-graduate students
  - ✓ The learning environment involves Virtual Technologies (360 Video+VR) instead of classrooms, traditional e-learning tools (e.g. MOOCs)
  - ✓ Artificial Intelligence to replace multiple-choice tests to determine the learning outcome achievements

Novelties

- Tracking the worker behaviour in the virtual learning environment and logging the tracked data

- Representing the logged data of the training session of each worker as a trace of a training process and designing a predictive process mining approach



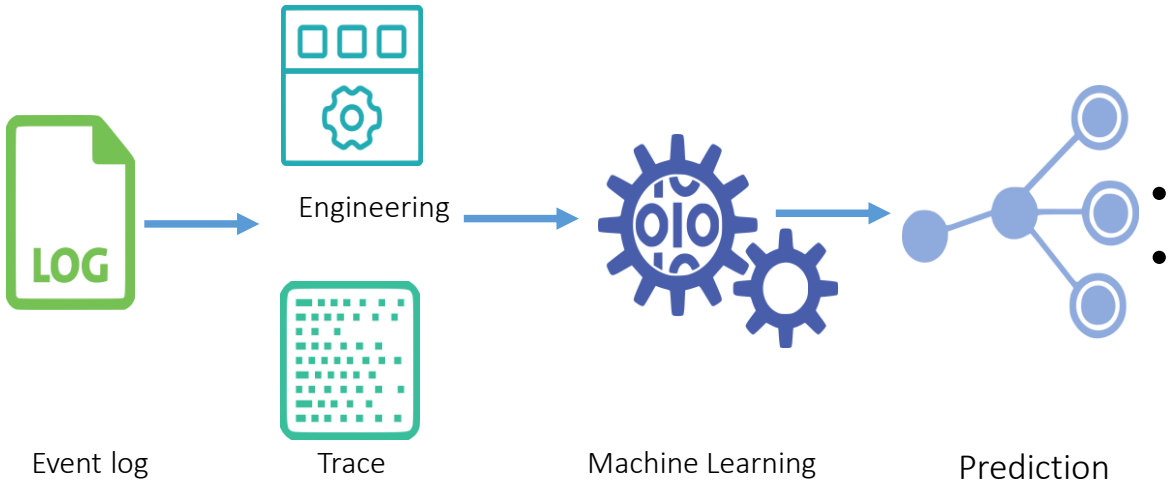
Trace-defined training sessions

Event	Worker	Scene	Time	Completion	Elapsed time	Effort
E1	W1	Scene1	2019-1-10,10:01:32	Completed	17442	18
E2	W1	Scene2	2019-1-10,10:03:02	Completed	27413	27
E3	W1	Scene3	2019-1-10,10:10:02	Completed	122709	48
E4	W1	Scene4	2019-1-10,10:13:22	Completed	100567	21
E5	W1	Scene5	2019-1-10,10:15:01	Uncompleted	458000	45
E6	W1	Scene5	2019-1-10,10:15:10	Uncompleted	3770223	37
E7	W1	Scene5	2019-1-10,10:15:19	Completed	557000	61
	W1			PASSED		
E8	W2	Scene1	2019-1-10,10:30:31	Completed	18244	21
…	…	…	…	…	…	…

Kometa

Knowledge Community for Efficient Training through Virtual Technologies

ViTE- Virtual Training Exam prediction



- Trace-based data engineering
- Predictive pattern learning

Data engineering

- Multi-perspective data engineering
  - ✓ Scene perspective
  - ✓ Control-flow perspective
  - ✓ Performance perspective

Kometa
Knowledge Community for Efficient Training through Virtual Technologies

Data engineering – scene

Event	Worker	Scene	Time	Completion	Elapsed time	Effort
E1	W1	**Scene1**	2019-1-10,10:01:32	**Completed**	17442	18
E2	W1	**Scene2**	2019-1-10,10:03:02	**Completed**	27413	27
E3	W1	**Scene3**	2019-1-10,10:10:02	**Completed**	122709	48
E4	W1	**Scene4**	2019-1-10,10:13:22	**Completed**	100567	21
E5	W1	**Scene5**	2019-1-10,10:15:01	**Uncompleted**	458000	45
E6	W1	**Scene5**	2019-1-10,10:15:10	**Uncompleted**	3770223	37
E7	W1	**Scene5**	2019-1-10,10:15:19	**Completed**	557000	61
PASSED						

Scene1	Scene2	Scene3	Scene4	Scene5
1	1	1	1	3

Scene counter

Scene1C	Scene2C	Scene3C	Scene4C	Scene5C
100%	100%	100%	100%	33.3%

Completion scene

Kometa
Knowledge Community for Efficient Training through Virtual Technologies

Data engineering – control flow

Event	Worker	Scene	Time	Completion	Elapsed time	Effort
E1	W1	**Scene1**	2019-1-10,10:01:32	**Completed**	17442	18
E2	W1	**Scene2**	2019-1-10,10:03:02	**Completed**	27413	27
E3	W1	**Scene3**	2019-1-10,10:10:02	**Completed**	122709	48
E4	W1	**Scene4**	2019-1-10,10:13:22	**Completed**	100567	21
E5	W1	**Scene5**	2019-1-10,10:15:01	**Uncompleted**	458000	45
E6	W1	**Scene5**	2019-1-10,10:15:10	**Uncompleted**	3770223	37
E7	W1	**Scene5**	2019-1-10,10:15:19	**Completed**	557000	61
PASSED						

Control-flow counter

Scene1C→Scene1C	Scene1C→Scene1U	Scene1C→Scene2C	…	Scene5U→Scene5U
0	0	1		1

Data engineering – performance

Event	Worker	Scene	Time	Completion	Elapsed time	Effort
E1	W1	**Scene1**	2019-1-10,10:01:32	Completed	**17442**	18
E2	W1	**Scene2**	2019-1-10,10:03:02	Completed	**27413**	27
E3	W1	**Scene3**	2019-1-10,10:10:02	Completed	**122709**	48
E4	W1	**Scene4**	2019-1-10,10:13:22	Completed	**100567**	21
E5	W1	**Scene5**	2019-1-10,10:15:01	Uncompleted	**458000**	45
E6	W1	**Scene5**	2019-1-10,10:15:10	Uncompleted	**3770223**	37
E7	W1	**Scene5**	2019-1-10,10:15:19	Completed	**557000**	61

PASSED

Trace-level time

count	sum	min	max	median	mean
7	1660153	17442	557000	122709	237164.71

Scene-level time

Scene1Sum	Scene1Min	Scene1Max	...	Scene5Mean
17442	17442	17442		1595074.3

Kometa
Knowledge Community for Efficient Training through Virtual Technologies

Data engineering – performance

Event	Worker	Scene	Time	Completion	Elapsed time	Effort
E1	W1	**Scene1**	2019-1-10,10:01:32	Completed	17442	**18**
E2	W1	**Scene2**	2019-1-10,10:03:02	Completed	27413	**27**
E3	W1	**Scene3**	2019-1-10,10:10:02	Completed	122709	**48**
E4	W1	**Scene4**	2019-1-10,10:13:22	Completed	100567	**21**
E5	W1	**Scene5**	2019-1-10,10:15:01	Uncompleted	458000	**45**
E6	W1	**Scene5**	2019-1-10,10:15:10	Uncompleted	3770223	**37**
E7	W1	**Scene5**	2019-1-10,10:15:19	Completed	557000	**61**

PASSED

Trace-level effort

sum	min	max	median	mean
257	18	61	37	36.71

Scene-level effort

Scene1Sum	Scene1Min	Scene1Max	Scene1Median	...	Scene5Mean
18	18	18	18		47.6

Predictive pattern learning

- Any supervised classification algorithm can be selected, in order to learn a pattern to predict the training outcome based on the features constructed in the data engineering process
- The Principal Component Analysis is performed to deal with the curse of dimensionality

Empirical validation – data

- Virtual training platform developed by MTM Project S.R.L with the content provided by MTM PROJECT S.R.L and Cinemagica S.R.L.

- 97 maintenance technicians have performed a virtual training session with 5 scenes

Empirical validation – data

- A training team has manually evaluated the learning achievements of workers after the virtual training test with a multi-choice test
  - ✓ 71 technicians → passed
  - ✓ 26 technicians → rejected

Empirical validation – methodology

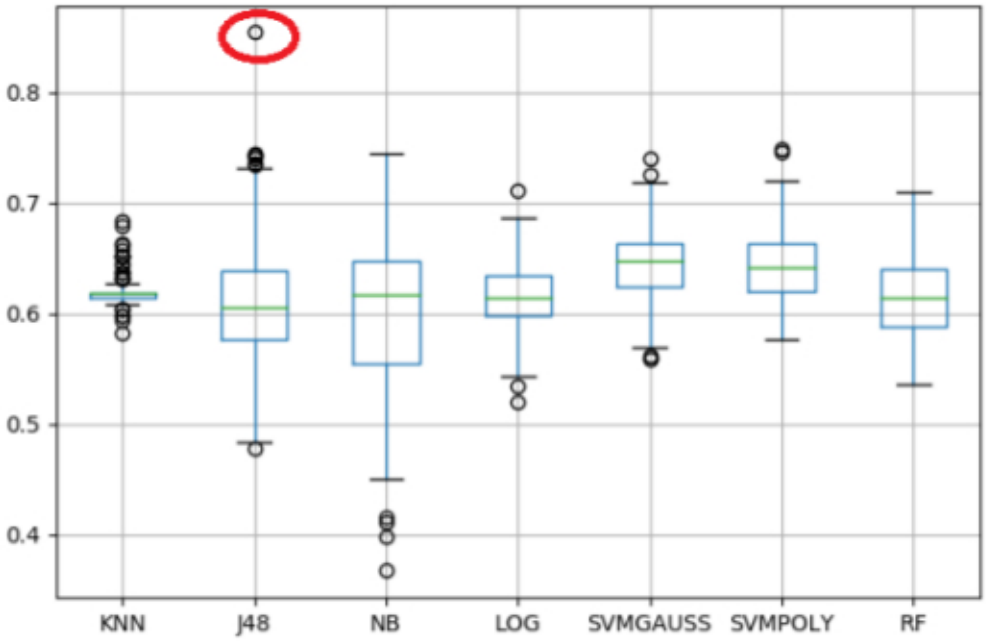
- Leave-one-out validation
- Classification algorithm
  - ✓ KNN, J48, Naive Bayes – NB, Logistic Regression – Log, SVM (Gaussian kernel – GAUSS, polynomial kernel – POLY), Random Forest – RF
- Accuracy performance evaluation
  - ✓ F-score

Kometa

Knowledge Community for Efficient Training through Virtual Technologies

Empirical validation – results by classification algorithm

F-score



Empirical validation – results by data engineering schema

J48

configuration	F-score
scene	0.6186
transition	0.7181
performance	0.6337
scene, transition	0.5830
scene, performance	0.6701
transition, performance	0.7075
scene, transition, performance	**0.8537**

Kometa
Knowledge Community for Efficient Training through Virtual Technologies

Final remarks

Kometa + ViTE	Accuracy	F-score
	0.8556	0.8537

- The accuracy of ViTE is in step with accuracy values in the recent literature for task of outcome prediction in education ranging between 0.7 and 0.81

Future work

- Extending the validation by considering data collected through new training sessions
- Framing the approach in a streaming environment, in order to learn a predictive pattern that may change over time as new training sessions are completed
- Exploring the use of transfer learning to transfer the predictive pattern learned with a training session to a session having a new topic

Thank you for your attention

annalisa.appice@uniba.it

Kometa
Knowledge Community for Efficient Training through Virtual Technologies

This work fulfils the research objectives of the POR Puglia FESR-FSE 2014-2020 - Asse prioritario 1 - Ricerca, sviluppo tecnologico, Innovazione - Sub Azione 1.4.b bando Innolabs - Research project ***KOMETA (Knowledge Community for Efficient Training through Virtual Technologies)*** *funded by Regione Puglia*

UNIONE EUROPEA
FESR "Fondo Europeo di Sviluppo Regionale"

Ministero dello Sviluppo Economico

Regione Puglia
Dipartimento Sviluppo Economico, Innovazione, Istruzione, Formazione e Lavoro
Sezione Ricerca, Innovazione e Capacità Istituzionale

PUGLIA FESR·FSE 2014/2020
Il futuro alla portata di tutti