

Event Log Sampling for Predictive Monitoring

---- User Manual

Prepared by Malek Alhelwany, Taekeun Jeong, Xiaoyan Jin

<https://github.com/Malekhy/ws2122-lspm>

Overview

The event log sampling for Predictive Monitoring (LSPM) proposes an instance selection procedure that allows sampling training process instances for prediction models.

We implement 3 different sampling algorithms as a training service in the form of web services, combine it with existing predictive models, and produce prediction results.

On our web service page, the user will be able to import (csv/xes) and export (xes) their files. With choosing the suitable sampling method as they want, they will get the prediction results with reliable levels of prediction accuracy instantly.



Getting started!

Steps (using virtual environment)


- Clone the project from <https://github.com/Malekhy/ws2122-lspm>
- Open the command terminal 'cmd'
- Move into the directory folder of the project:
[\(project-path\)\ws2122-lspm](#)
- Move into project environment:
 - For Windows users: `.\Scripts\activate`
 - For Mac users: `source Scripts\activate`
- Install requirements: `pip install -r requirements.txt`
- Run the web app: `python manage.py runserver`
- Open the browser and hit the URL: <http://localhost:8000/>




Steps (using docker image)

- Clone the project from <https://github.com/Malekhy/ws2122-lspm>
- Open the command terminal 'cmd'
- Move into the directory folder of the project:
[\(project-path\)\ws2122-lspm](#)
- Build the docker image using command: `docker build . -t lspm`
- Run the docker container using command:
`docker run -d -p 8000:8000 lspm`
- Open the browser and hit the URL: <http://localhost:8000/>

Then you will get...

 Chair of Process and Data Science



Home Sample Model Help Page

Event Logs

Please upload the event log: No file chosen

CPMExport - Copy.csv
CPMExport - Copy_QjkElkp.csv
CPMExport.csv
FirstAssignmentBPI- concept.csv
FirstAssignmentBPI-DataSet2.csv
ItalianHelpdeskFinal- concept - Copy.csv
ItalianHelpdeskFinal- concept.csv
ItalianHelpdeskFinal- Concept.xes
ItalianHelpdeskFinal.csv
ItalianHelpdeskFinal.xes
exampleSample.xes

is used as input.

Please choose event log characteristics

Please choose the Case ID:

Please choose the Activity Name:

Please choose the Start Time Stamp:

Please choose the sampling method

☒ Unique Selection

☐ Logarithmic distribution

☐ Division



Instruction

1. Click here to choose files

2. Click here to add the file into the working directory

*Click here to delete the chosen files

*Click here to download the chosen files on the list

Home Sample Model Help Page

Event Logs

Please upload the event log: No file chosen

CPMExport - Copy.csv
CPMExport - Copy_QjkEklp.csv
CPMExport.csv
FirstAssignmentBPI- concept.csv
FirstAssignmentBPI-DataSet2.csv
ItalianHelpdeskFinal- concept - Copy.csv
ItalianHelpdeskFinal- concept.csv
ItalianHelpdeskFinal- Concept.xes
ItalianHelpdeskFinal.csv
ItalianHelpdeskFinal.xes

is used as input.

Please choose event log characteristics

Please choose the Case ID:

Please choose the Activity Name:

Please choose the Start Time Stamp:

Please choose the sampling method

☒ Unique Selection
☐ Logarithmic distribution
☐ Division



Instruction

*Click here to see statistics comparison between the original event log and the sampling results

5. Click here to start sampling

3. Click here to adjust the preferred characteristics

4. Click here to choose the sampling methods

Home Sample Model Help Page

Event Logs

Please upload the event log: No file chosen

CPMExport - Copy.csv
CPMExport - Copy (2).csv
CPMExport.csv
FirstAssignmentBPI- concept.csv
FirstAssignmentBPI-DataSet2.csv
ItalianHelpdeskFinal- concept - Copy.csv
ItalianHelpdeskFinal- concept.csv
ItalianHelpdeskFinal- Concept.xes
ItalianHelpdeskFinal- Concept.xes
ItalianHelpdeskFinal.csv
ItalianHelpdeskFinal.xes

ItalianHelpdeskFinal- concept - Copy.csv is used as input.

Please choose event log characteristics

Please choose the Case ID:

Please choose the Activity Name:

Please choose the Start Time Stamp:

Please choose the sampling method

☒ Unique Selection
☐ Logarithmic distribution
☐ Division




Instruction


results page 

In the Bounce Rate column we can see the reduced percentage after sampling process.

e.g. In this row we can see the significant changes after implementing our sampling methods.



Chair of Process
and Data Science



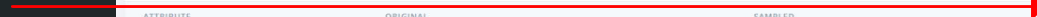
Home Sample Model Help Page

Sample model window

Original eventlog and sampled one some statistics [See all](#)

ATTRIBUTE	ORIGINAL	SAMPLED	BOUNCE RATE
Event log name	Italiani-helpdeskFinal-concept - Copy.csv	SAMPLED.FILE20220107-214243.csv	
Number of cases	4579	2123	↓ 46,53%
Number of events	21342	10366	↓ 36,49%
Average case duration	0d 0h 0m 0s	0d 0h 0m 0s	↓ 50,87%
Number of variants	226	111	↓ 46,53%
Total case duration	0d 0h 0m 0s	0d 0h 0m 0s	↓ 46,53%

Please navigate to Home tab for changing the selected log



Happy process mining :)
