Kiel University
Christian-Albrechts-Universität zu Kiel
Faculty of Engineering

Department of Computer Science

Process Analytics

Prof. Dr. Agnes Koschmider
Dominik Janssen

Advanced Process Mining

Summer term 2020

Exercise sheet 7

Predictive Process Monitoring

Exercise 1: Sequence encoding

Case ID	Timestamp	Activity	Resource	Pulse	Medication	Gender	Age
1	01.05.20	Admission	ER	40	-	Male	86
1	01.05.20	Surgery	OR	30	Propofol	Male	86
1	01.05.20	CPR	OR	0	Epinephrine	Male	86
1	01.05.20	ToD	OR	0	-	Male	86
2	05.05.20	Admission	ER	60	-	Female	80
2	05.05.20	Surgery	OR	70	Bupivacaine	Female	80
2	05.05.20	Recovery	PACU	55	Saline	Female	80
3	06.05.20	Admission	ER	80	-	Female	75
3	06.05.20	IV Fluids	ER	70	Saline	Female	75
3	06.05.20	Release	Geriatrics	65	-	Female	75
2	10.05.20	Release	Geriatrics	60	Ibuprofen	Female	80

a) In the event log above, identify case and event attributes of the three shown cases.

b) Transform the following event log into an index-based encoded table.

Case ID	Timestamp	Activity	Pulse	Gender	Age
1	01.05.20	Admission	40	Male	86
1	01.05.20	Surgery	30	Male	86
1	01.05.20	CPR	0	Male	86
1	01.05.20	ToD	0	Male	86
2	05.05.20	Admission	60	Female	80
2	05.05.20	Surgery	70	Female	80
2	05.05.20	Recovery	55	Female	80
3	06.05.20	Admission	80	Female	75
3	06.05.20	IV Fluids	70	Female	75
3	06.05.20	Release	65	Female	75
2	10.05.20	Release	60	Female	80

c) Transform the first event log into an aggregation encoded table.

d) Why could it be useful to apply predictive process monitoring in this scenario?

Solution

a) Every attribute that is time invariant (over the recorded time period) is called a *case attribute*. Attributes than are variable and can change based on the event are called *event attributes*.

- Event attributes:

  Timestamp
  Activity
  Resource
  Pulse
  Medication

- Case attributes:

  Case ID
  Gender
  Age

b) Index based encoding:

ID	Gender	Age	$Time_1$	$Time_2$	$Time_3$	$Time_4$	Act_1	Act_2	Act_3	Act_4	$Pulse_1$	$Pulse_2$	$Pulse_3$	$Pulse_4$
1	Male	86	01.05.	01.05.	01.05.	01.05.	Admission	Surgery	CPR	ToD	40	30	0	0
2	Female	80	05.05.	05.05.	05.05.	05.05.	Admission	Surgery	Recovery	Release	60	70	55	60
3	Female	75	06.05.	06.05.	06.05.	-	Admission	IV Fluids	Release	-	80	70	65	-

c) A possible aggregation of the original event log could look like the following:

ID	Gender	Age	*count*(OR)	*min*(Pulse)	*max*(Pulse)	*count*(Medication)	*count*(ToD)
1	Male	86	3	0	40	2	1
2	Female	80	1	55	70	3	0
3	Female	75	0	65	80	2	0

d) Predictive process monitoring could be useful in this scenario to predict the outcome of a hospital treatment. In this case the prediction if a patient dies or survives will be something to observe.