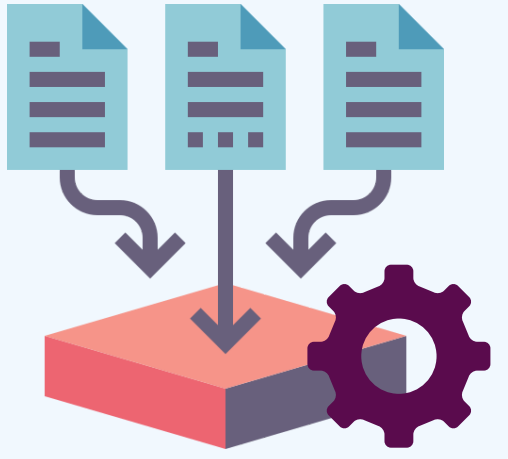


EEN EMPIRISCHE ANALYSE

INDUCTIVE MINER

1 INTRODUCTIE

Competitive testing van Process Discovery algoritmes biedt waardevolle inzichten door aan te tonen welke algoritmes superieur zijn ten opzichte van andere. Het heeft ons geleerd dat de **Inductive Miner (IM)** een goed algoritme is om processen te ontdekken via event logs (Augusto, et al. 2019). De vraag blijft **waarom** de Inductive Miner goed werkt.



Q1: "Wat is het effect van Design Decisions in het algoritme op de prestatie van de Inductive Miner?"



Q2: "Is er interactie tussen de Task Properties en de Design Decisions van de Inductive Miner?"

2 INDUCTIVE MINER

- Divide-en-Conquer werkwijze
- Splitst logs op in sublogs volgens bepaalde **splitsingsvolgorde**

Design Choice
CHOICE - SEQUENCE - PARALLEL - LOOP
→ **24 algoritme varianten**

4 QUALITY MEASURES



1. Fitness
2. Precision
3. Generalisation
4. Simplicity

3 TASK PROPERTIES



<u>Sequence</u>	<u>Parallel</u>	#Activities 10-30
10%-65%	10%-65%	
<u>Choice</u>	<u>Loop</u>	
10%-65%	10%-65%	

5 METHODE

Event log generatie

- Task Properties
- 75 processen => 75 logs
- 1000 traces

Inductive Miner

- Repeated measures
- Q1: 4 experimenten
- Q2: 32 experimenten
- => 1800 obs/measure

Experimenten

- Repeated measures
- Q1: 4 experimenten
- Q2: 32 experimenten

Statistische analyse

- Friedman-test (Demsar, et al. 2006)



6 RESULTATEN EXP 1-4

H0: De splitsingsvolgorde van de Inductive Miner heeft geen invloed op een kwaliteitsmaatstaf van het ontdekte model ten opzichte van de event log.

$SN = 95\%, \alpha = 0.05$

F-stat > FF => H0 kon niet verworpen worden op 95%

Exp 1-4	FF	F-Stat	F-stat < FF
Fitness	-38.65	1.54	X
Precision	-38.52	1.54	X
Generalisation	-39.40	1.54	X
Simplicity	-38.99	1.54	X

7 RESULTATEN EXP 5-36

H0: De locatie in de splitsingsvolgorde van een splitsing heeft geen impact op een kwaliteitsmaatstaf voor procesmodellen van processen met een hoge kans op een bepaalde operator.

$SN = 95\%, \alpha = 0.05$

F-stat > FF => H0 kon niet verworpen worden op 95%

8 CONCLUSIE



- 1 Een verandering in splitsingsvolgorde heeft geen invloed op de qualitymeasures
- 2 IM = robuust algoritme dat consistent presteert onder verschillende contexten