UNIVERSITÀ DI PISA

Artificial Intelligence and Data Engineering

Process Mining and Intelligence

Personal Protective Equipment Detection

Project report

Luana Bussu

Luca Caprioli

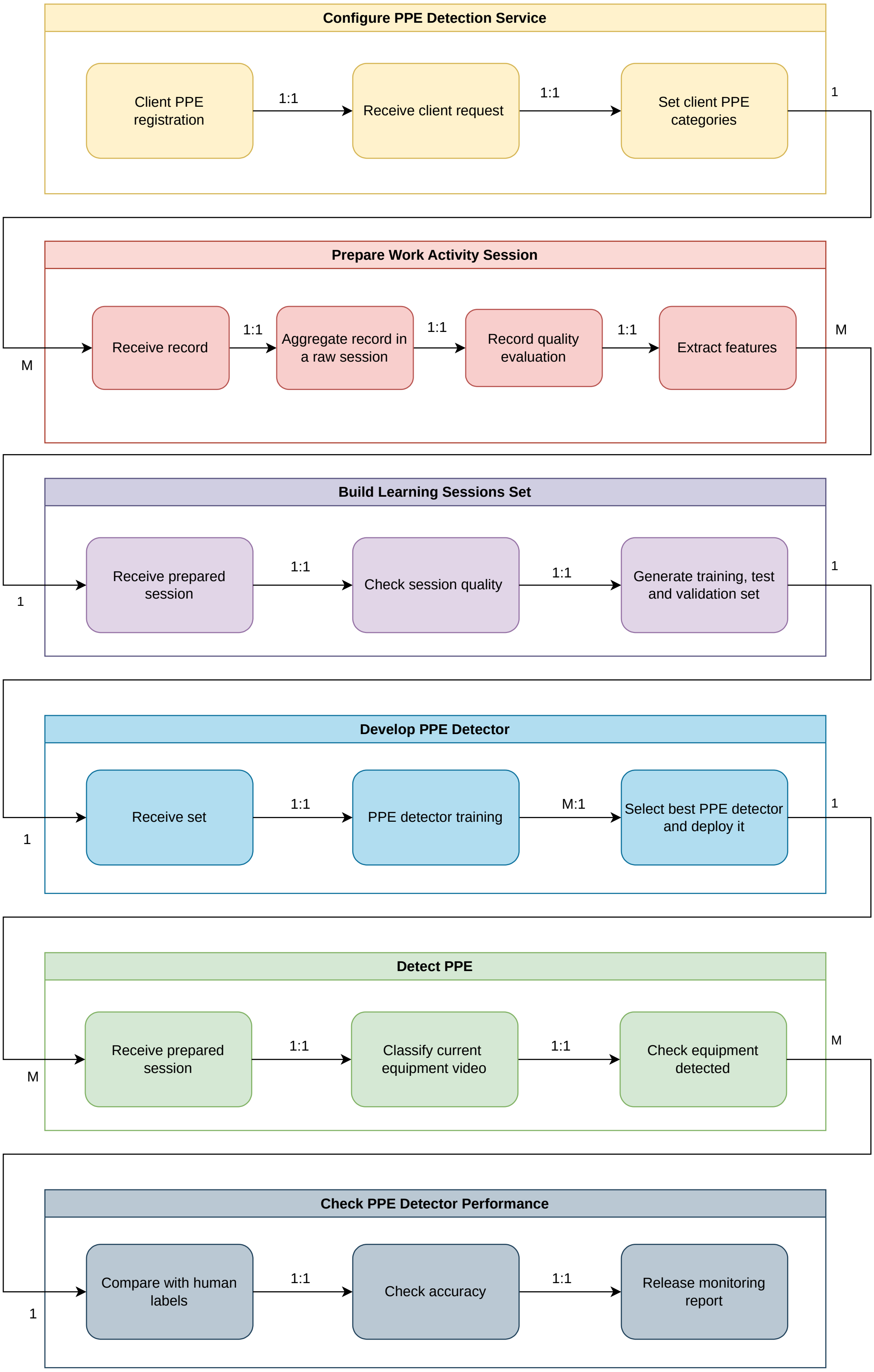
Martina Marino

Roberta Matrella

Academic Year 2021/2022

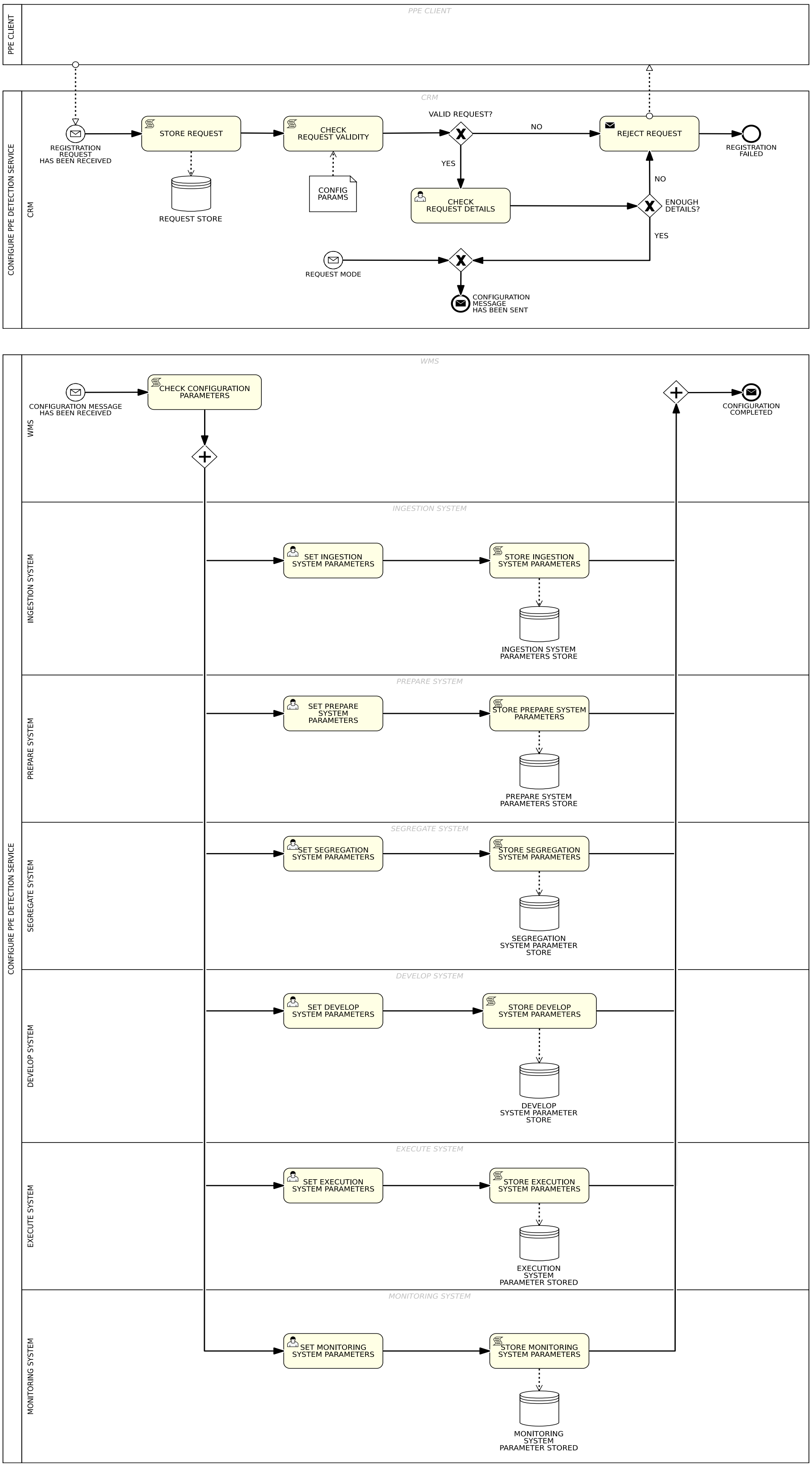
Summary

1 Process Landscape (All)

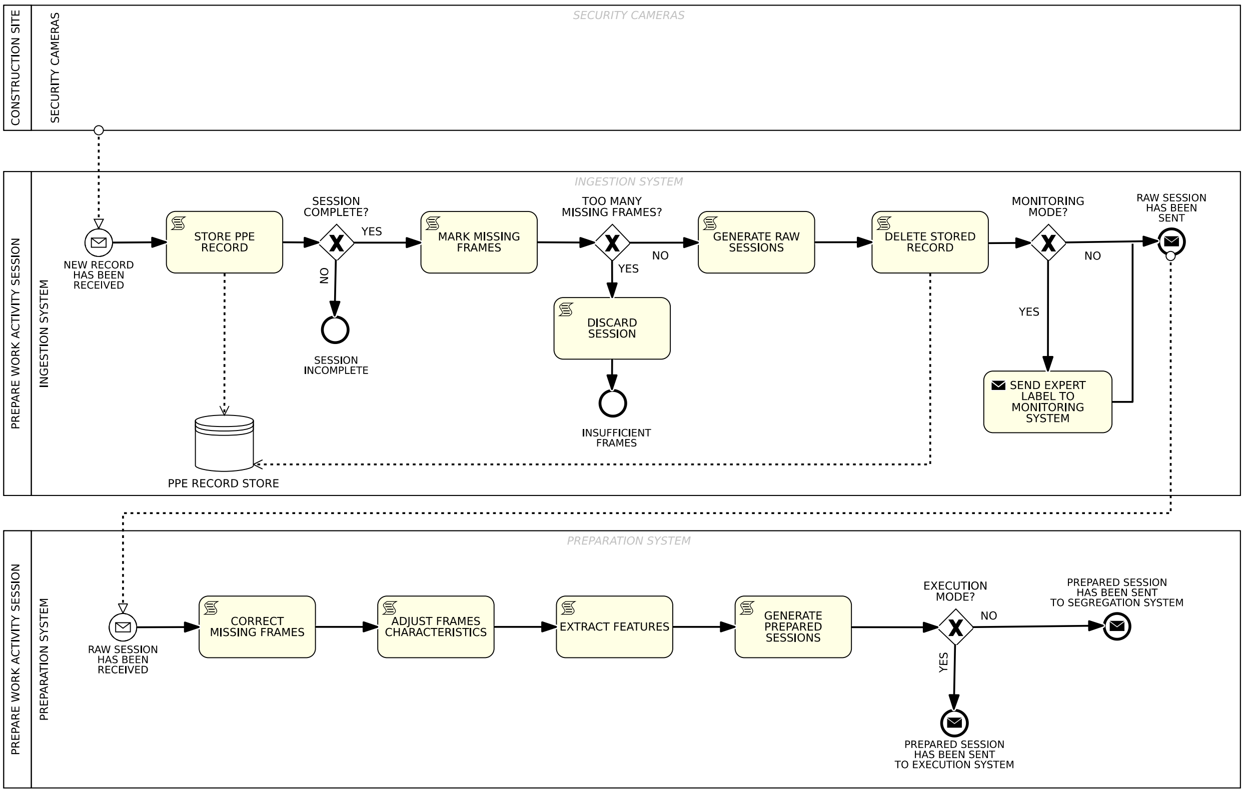


Set in Develop PPE Detector: set of training, test and validation data

2 Configure PPE Detection Service (ALL)



3 Prepare Work Activity Session (Martina Marino)



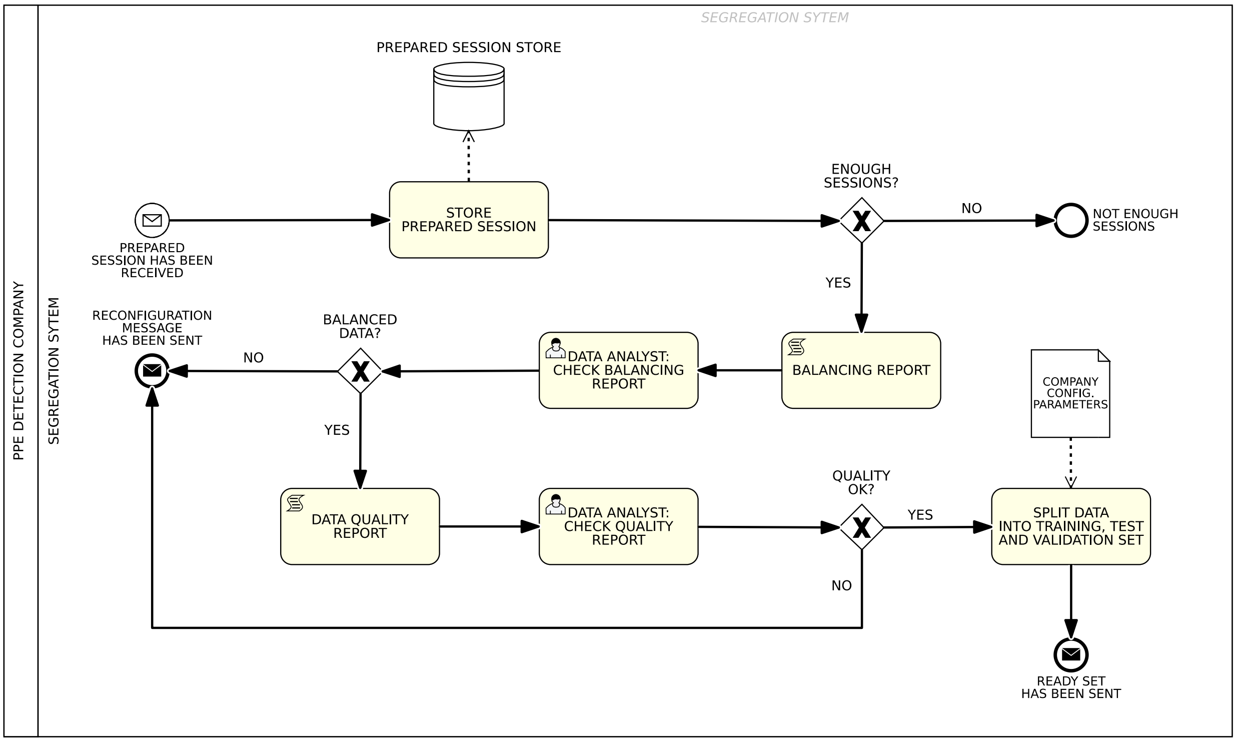
A session is a **set of records** synchronized by time.

A record is composed by:

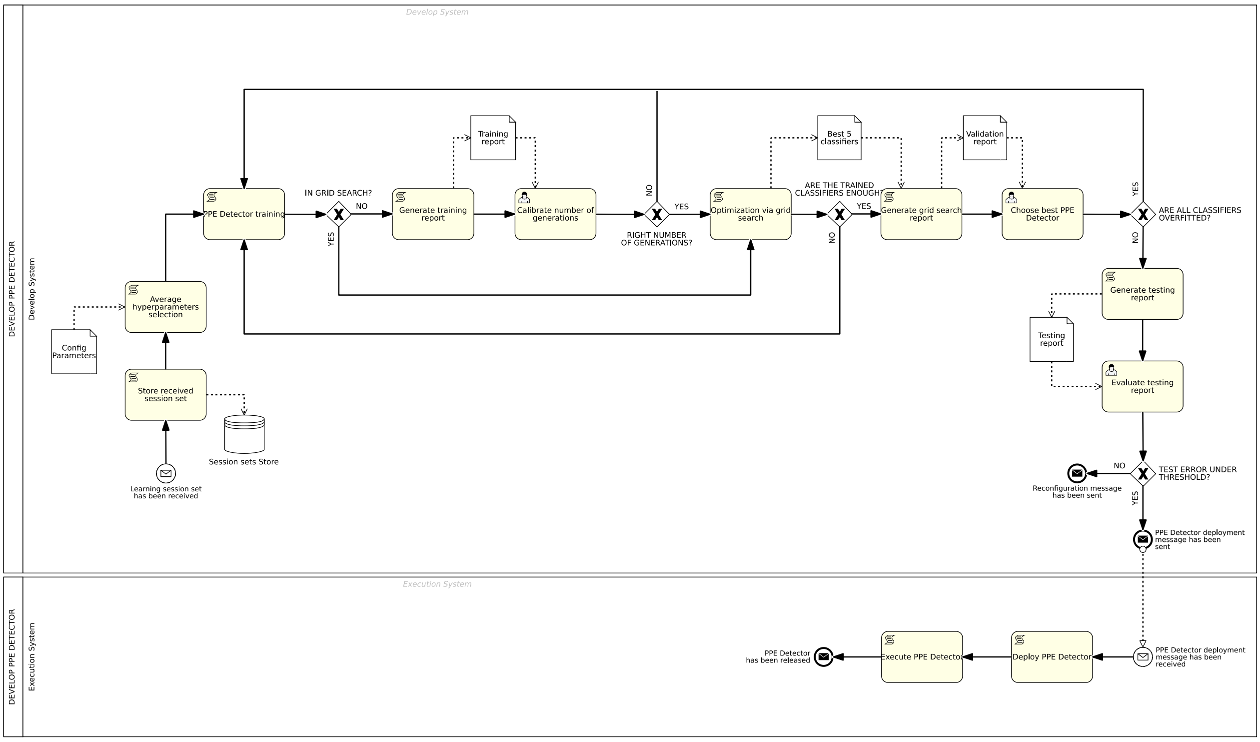
- Extract of a video
- Type of activity (role)
- Characteristics of the construction site (Indoor/outdoor)

A session is composed by an **expert label** (objects identified) if the next step is the develop system.

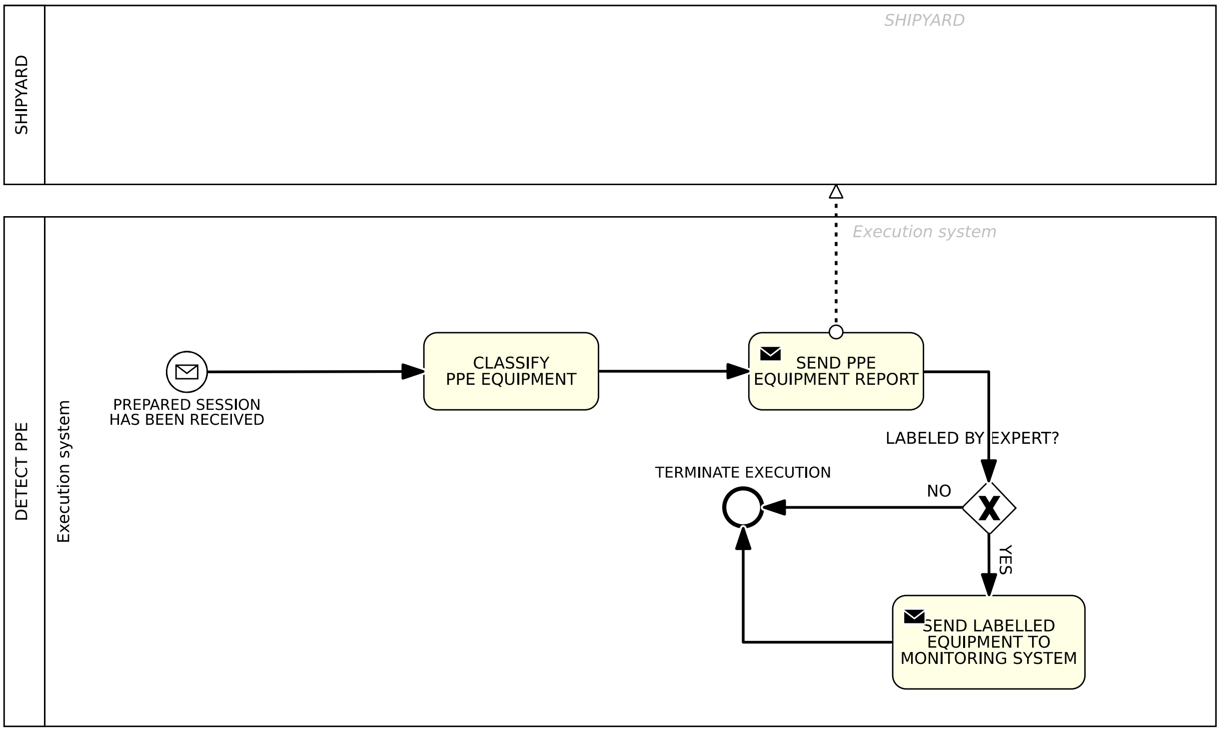
4 Build Learning Sessions Set (Luana Bussu)



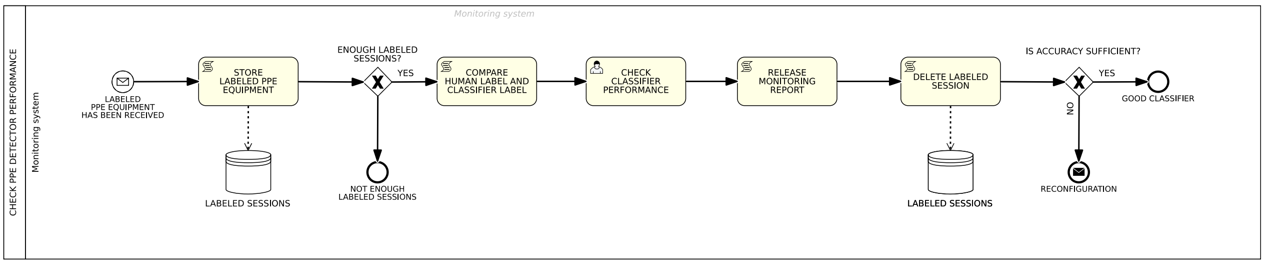
5 Develop PPE Detector (Luca Caprioli)



6 Detect PPE (Roberta Matrella)



7 Check PPE Detector Performance (Roberta Matrella)



8 Salaries

Role	Reference	Description	Cost per hour	Min cost per hour	Normalized Cost
Data Analyst	Data analyst job profile | Prospects.ac.uk	Data analysts translate data into information that can be used to solve problems or track business. They also need to use data analysis to produce graphs or charts.	£23.57	£12.50	1.89
ML Engineer	Machine learning engineer job profile | Prospects.ac.uk	They are responsible for creating programs and algorithms that enable machines to take actions without being directed.	£28.41	£19.88	1.43
Administrative	https://ca.indeed.com/career-advice/finding-a-job/what-is-administrative-work	Administrative helps you to complete tasks associated with business management. This can include filing paperwork, meeting with internal and external stakeholders, providing critical information, implementing procedures, or responding to employee or customer inquiries.	£17.23	£15.59	1.11
Quality engineer	https://www.energy.gov/eere/wind/career-map-quality-engineer#:~:text=They%20work%20with%20quality%20assurance,regulations%2C%20and%20satisfy%20client%20expectations	They work with quality assurance and quality control teams to develop processes, test procedures and implement systems that ensure the products and processes fulfill quality standard. They evaluate the quality of the final product.	£21.02	£18.18	1.14

For the cost per hour, 220 weekdays per year and 8 hours per day considered.

9 Cognitive effort

- **Remember (1)**: cognitive effort consists in selecting an option (i.e., open a window or push a button)
- **Understand (2)**: move from one form of representation to another, find a specific example of a concept, determine that something belongs to a category, abstract a general theme or important points, draw a logical conclusion from the information presented, detect correspondences between two concepts, construct a cause-and-effect model of a system.
- **Apply (3)**: perform a procedure for a familiar and unfamiliar task, producing models, presentations, interviews, simulations.
- **Analyze (4)**: to distinguish the relevant parts from the irrelevant ones of the material presented, to determine how the elements adapt to the function within a structure, to determine the point of view, the prejudices, the values.

10 Use case Diagrams

In determining the cognitive effort of each activity carried out in the various processes, the type of construction site is of the utmost importance, which differs based on some fundamental characteristics:

- Number of objects to recognize
- Environmental brightness of the workplace
- Presence of dust/fumes in the workplace
- Crowding of objects to be recognized (in which and how many activities each device is used)

10.1 CRM System (Roberta Matrella)

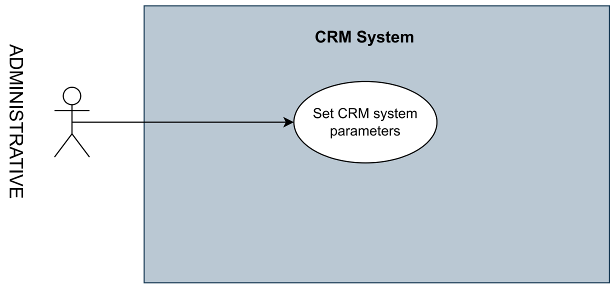


Figure 1 - Use case of the CRM system

10.1.1 Check request details

Task	Subtask	%	Role	Cognitive Effort	Cost
Check request details	1. Open requests interface		Administrative	Remember (1)	1*1.11
	2. **FOR EACH** request parameter				**4.44**
	2.1 Check parameter		Administrative	Remember (1)	1*1.11
	3. **IF** enough details	2			
	3.1 Approve request		Administrative	Remember (1)	1*1.11
	4. **ELSE**	98			
	4.1 Reject request		Administrative	Remember (1)	1*1.11
	5. Close requests interface		Administrative	Remember (1)	1*1.11

Total Cost: 7.77

Requests Parameters

- Construction site ID (1)
- Objects to detect (1)
- Activities of the construction site (1)
- Roles of the construction site (1)

All these parameters are used to save the request and check the similar sites

10.2 Ingestion System (Martina Marino)

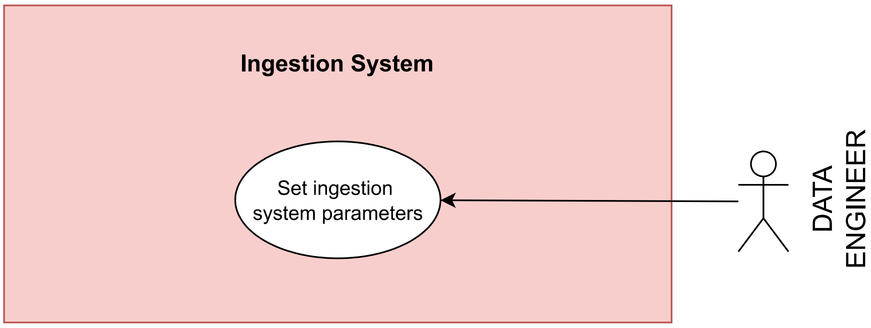


Figure 2 - Use case of the ingestion system

10.2.1 Set ingestion system parameters

Task	Subtask	%	Role	Cognitive Effort	Cost
Configure ingestion system parameters	1. Open parameters interface		Data Engineer	Remember (1)	1*1.52
	2. Set missing frame threshold		Data Engineer	Analyse (4)	4*1.52
	3. **FOR EACH** parameter of cost (1)		Data Engineer		**4.56**
	3.1. Set value retrieved from message		Data Engineer	Remember (1)	1*1.52
	4. Save configuration into file		Data Engineer	Remember (1)	1*1.52
	5. Close parameters interface		Data Engineer	Remember (1)	1*1.52

Total Cost: 15.20

In step 2 the threshold for the missing frames is multicriterial and done by experience. Depending on the activity, there are different sources of danger, and the dynamics of harmful events are of various type, from a low-level of risk (earplugs not worn) to a high risk (falling from scaffold).

Ingestion System Parameters

- Missing frame threshold for the session (4)
- Number of records per session to be complete (1)
- Mode (execution or monitoring) (1)
- Preparation system address (1)

10.3 Preparation System (Martina Marino)

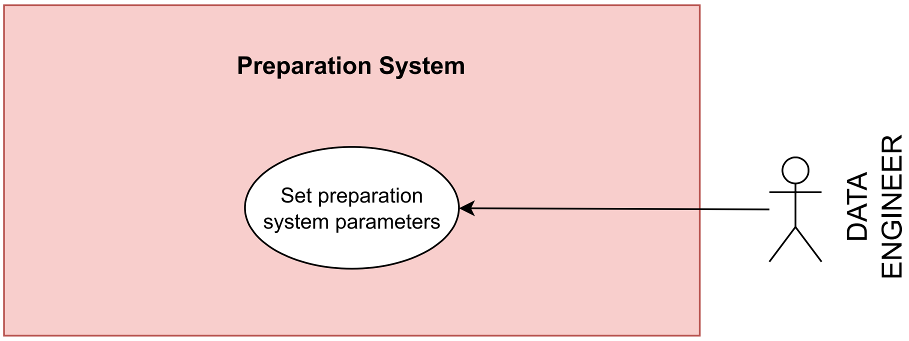


Figure 3 - Use case of the preparation system

10.3.1 Set preparation system parameters

Task	Subtask	%	Role	Cognitive Effort	Cost
Configure preparation system parameters	1. Open parameters interface		Data Analyst	Remember (1)	1*1.89
	2. Set features to extract value		Data Analyst	Analyze (4)	4*1.89
	3. Set mode value		Data Analyst	Remember (1)	1*1.89
	4. Set segregation system address		Data Analyst	Remember (1)	1*1.89
	5. Save configuration into file		Data Analyst	Remember (1)	1*1.89
	6. Close parameters interface		Data Analyst	Remember (1)	1*1.89
					Total Cost: 17.01

The step 2 needs the operator to choose the best configuration for extracting the most representative features according to number of objects, ambient brightness, crowding of objects, presence of dust, smoke or fog to name a few. For this reason it is multicriterial.

The Role can also be a ML Engineer: the skills needed for this task are related to an hybrid of the two roles.

Preparation System Parameters
• Features to extract (4) • Mode (execution or monitoring) (1) • Segregation system address (1)

10.4 Segregation System (Luana Bussu)

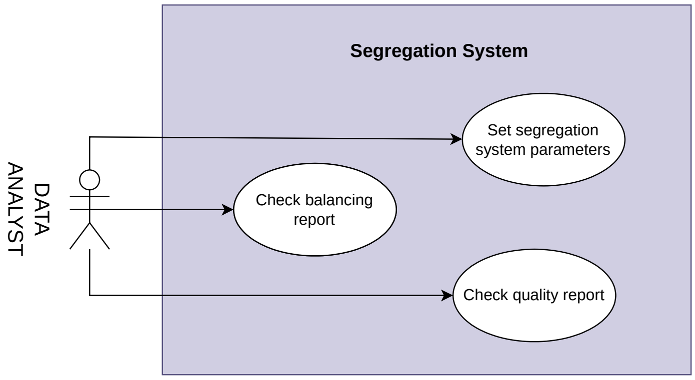


Figure 4 - Use case of the segregation system

10.4.1 Set Segregation System parameters

Task	Subtask	%	Role	Cognitive Effort	Cost
Configure segregation system parameters	1. Open parameters interface		Data Analyst	Remember (1)	1*1.89
	2. Set number of sessions to collect		Data Analyst	Analyze (4)	4*1.89
	3. Set minimum requirement per class		Data Analyst	Analyze (4)	4*1.89
	4. Set percentage of tolerance		Data Analyst	Analyze (4)	4*1.89
	5. Set configuration system address		Data Analyst	Remember (1)	1*1.89
	6. Set develops system address		Data Analyst	Remember (1)	1*1.89
	7. Close parameters interface		Data Analyst	Remember (1)	1*1.89

Total Cost: 30.24

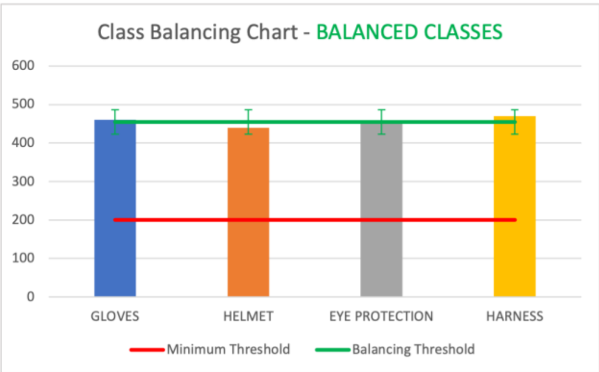
Segregation System Parameters

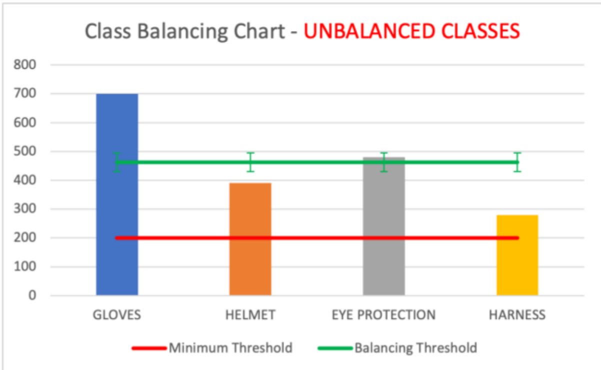
- Number of sessions to collect (4: the number of data to be collected depends on the site characteristics mentioned at the beginning of the session;)
- Minimum requirement per class (4: the choice is made by experience since it is necessary to consider how many activities each object appears in, therefore the minimum per class is chosen based on the characteristics of the activity in which it is used)
- Percentage of tolerance within the average value (4: based on the number of activities in which each object appears and the characteristics of the construction site)
- Configuration System address (1)
- Develop System address (1)

10.4.2 Check balancing report

Task	*Subtask*	*%*	*Role*	*Cognitive Effort*	*Cost*
Check balancing report	1. Open Reports Interface		Data Analyst	Remember (1)	1*1.89
	2. Show Balancing Report		System		
	3. **FOR EACH** class				
	3.1 Check tolerance threshold		System		
	4. Check the Balancing Report		Data Analyst	Analyze (2)	2*1.89
	IF classes are unbalanced				
	4.1 Request new data	20	Data Analyst	Remember (1)	0.2*1*1.89
	ELSE				
	4.2 Close Reports interface	80	Data Analyst	Remember (1)	0.8*1*1.89
	5. **END**				

Total Cost: 7.56





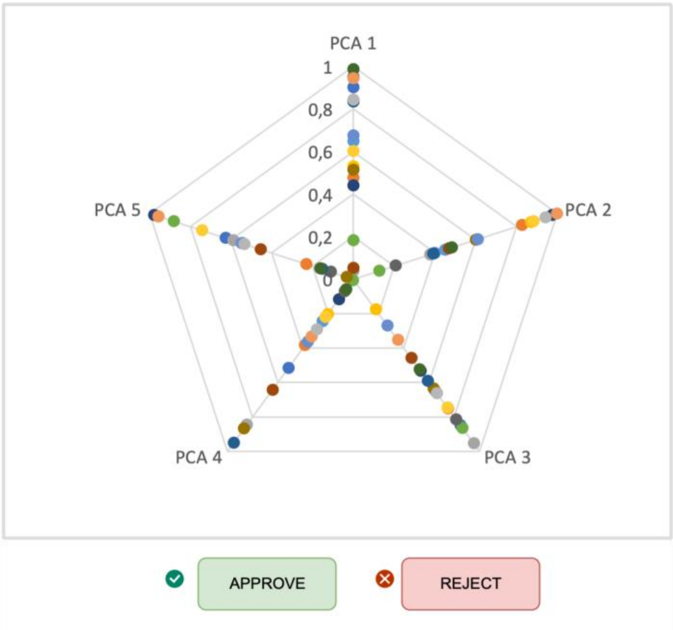
Probability of having balanced (unbalanced) data = 80% (20%)

The percentage of unbalanced data depends on:

- QUANTITY OF DATA COLLECTED: higher the number of data collected, lower will be the probability to have unbalanced data (but higher the cost) and vice versa, so it is important to choose the right trade off.
- NUMBER OF CLASSES: with the same amount of data collected, the higher the number of classes to detect, the higher the probability of unbalanced data will be.

10.4.3 Check quality report

Task	Subtask	%	Role	Cognitive Effort	Cost
Check quality report	1. Open Reports Interface		Data Analyst	Remember (1)	1*1.89
	2. Show Quality Report		System		
	3. **FOR EACH** feature (5 times)				
	3.1 Check distribution		Data Analyst	Analyze (4)	5*4*1.89
	4. **IF** all distributions are good	95			
	4.1 Approve quality		Data Analyst	Remember (1)	1*1.89
	5. **ELSE**	5			
	5.1 Reject quality		Data Analyst	Remember (1)	1*1.89
	6. Close Report Interface		Data Analyst	Remember (1)	1*1.89
				Total Cost:	**45.36**



The distribution on the components of the PCA depends on the input data (the features are not standard but are decided for each construction site) and the change of these distributions as the input data varies is decided by the analyst through experience, for this reason the cognitive effort is 4.

10.5 Develop System (Luca Caprioli, Martina Marino)

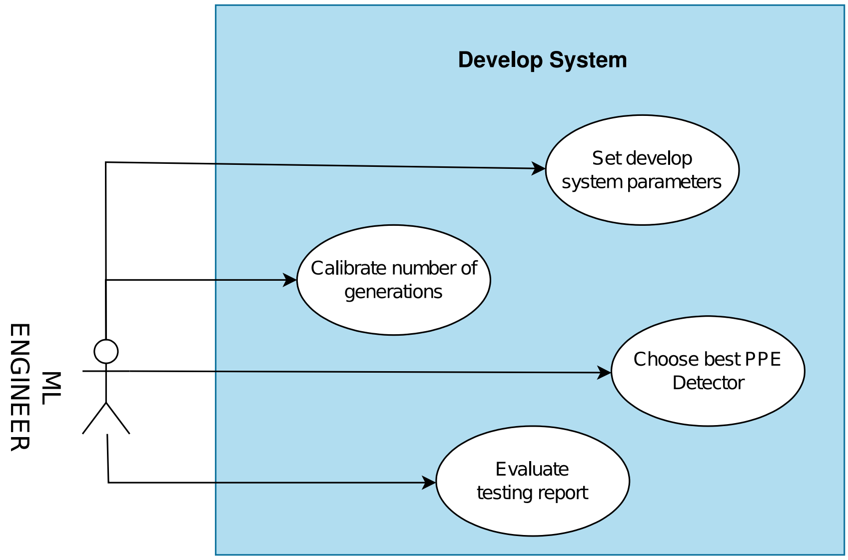


Figure 5 - Use case of the develop system

10.5.1 Set develop system parameters (Luca Caprioli)

Task	*Subtask*	*%*	*Role*	*Cognitive Effort*	*Total Cost*
Configure develop system parameters	1. Open parameters interface		M.L. Engineer	Remember (1)	1*1.89
	2. Set Number of classes		M.L. Engineer	Remember (1)	1*1.89
	3. **FOR EACH** parameter of cost (4)		M.L. Engineer		**4*4*1.89**
	3.2. Set develop system parameter		M.L. Engineer	Analyze (4)	4*1.89
	4. Close parameters interface		M.L. Engineer	Remember (1)	1*1.89

Total: 35,91

Develop System Parameters

- Min, Max, step hyperparameters (4)
- Training error threshold (4)
- Validation error threshold (4)
- Test error threshold (4)
- Number of classes (1)

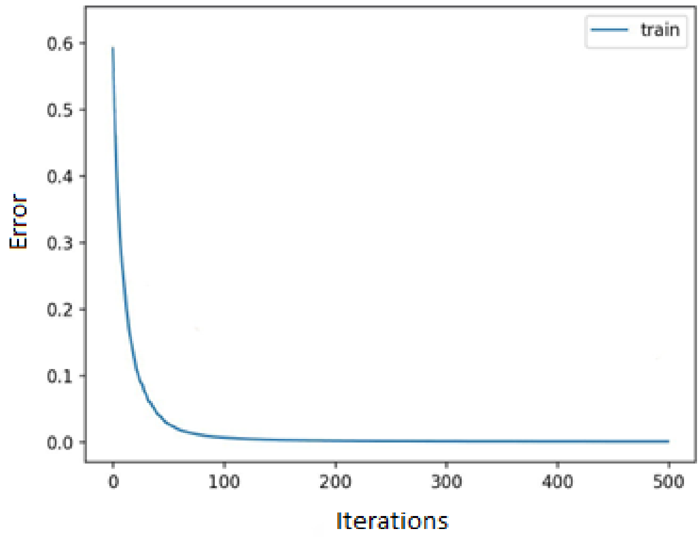
All the parameters of cost (4) are multicriterial and depend on:

Data quality, Data numerosity, Data heterogeneity, Classifier performance on data, Number of features, Number of classes.

10.5.2 Calibrate number of generations (Martina Marino)

Task	Subtask	%	Role	Cognitive Effort	Cost
Set number of generations	1. Open user interface		M.L. Engineer	Remember (1)	1*1.43
	2. Check training report		M.L. Engineer	Apply (3)	3*1.43
	3. **IF** the error is flat for at least half of the iterations	75			**4.29**
	4.1. Reduce of one third the number of iterations		M.L. Engineer	Apply (3)	3*1.43
	4.2. Submit current configuration		M.L. Engineer	Remember (1)	1*1.43
	4.3. → 3				
	4. **IF** the loss is not flat at the end of the iterations	25			**1.07**
	5.1. Enlarge by one third the number of iterations		M.L. Engineer	Apply (3)	3*1.43
	6. Close user interface		M.L. Engineer	Remember (1)	1*1.43

Total: 12.51



10.5.3 Choose best PPE Detector (Luca Caprioli)

Task	*Subtask*	*%*	*Role*	*Cognitive Effort*	*Cost*
Choose best PPE detector	1. Open user interface		M.L. Engineer	Remember (1)	1*1.43
	2. Show user interface		System		
	3. Select grid search		M.L. Engineer	Remember (1)	1*1.43
	4. Execute grid search		System		
	5. **FOR EACH** network				**6*5*1.43**
	5.1. Check $\lvert val_{err} - train_{err} \rvert \leq threshold$			Apply (3)	3*1.43
	5.2. Check complexity			Apply (3)	3*1.43
	6. **IF** all errors over threshold	10			**0.1*1*1.43**
	6.1. → 8				
	7. **ELSE**	90			**0.9*3*1.43**
	7.1. Choose best network with parameters		M.L. Engineer	Apply (3)	3*1.43
	7.2. → 8				
	8. Close user interface		M.L. Engineer	Remember (1)	1*1.43
	9. **END**				

Total: 55.48

$$|val_{err} - train_{err}| \leq 0.02$$

Network	Complexity		Train error	Validation error
	#of layers	#of neurons		
MLP 1	4	90	0,1023	0,1182
MLP 2	9	70	0,1134	0,1469
MLP 3	10	70	0,1234	0,1365
MLP 4	5	70	0,1123	0,1184
MLP 5	7	50	0,1345	0,1435

1. Choose network with lowest error.
2. Check if within threshold.
    a. If not go back to step 1.
3. Choose network with second lowest error.
4. Check if within threshold.
    a. If not go back step 3.
5. Choose network with lower complexity between the two selected.

10.5.4 Evaluate testing report (Luca Caprioli)

Task	***Subtask***	***%***	***Role***	***Cognitive Effort***	***Cost***
Evaluate testing report	1. Open testing report		M.L. Engineer	Remember (1)	1*1.43
	2. Check testing report		M.L. Engineer	Analyze (3)	3*1.43
	3. **IF** $\lvert \text{test}_{\text{err}} - \text{val}_{\text{err}} \rvert \le$ threshold?	80			**0.8*1*1.43**
	1.1. Submit deployment message			Remember (1)	1*1.43
	1.2. → 4				
	ELSE	20			**0.2*1*1.43**
	1.3. Submit reconfiguration message		M.L. Engineer	Remember (1)	1*1.43
	1.4. → 4				
	4. Close user interface		M.L. Engineer	Remember (1)	1*1.43
	5. **END**				
					Total: 8.58

$$|\boldsymbol{test_{err}} - \boldsymbol{val_{err}}| \le \mathbf{0.02}$$

Network	**Complexity**		**Validation error**	**Test error**
	#of layers	#of neurons		
MLP 1	4	90	0,1023	0,1182

10.6 Execution system (Roberta Matrella)

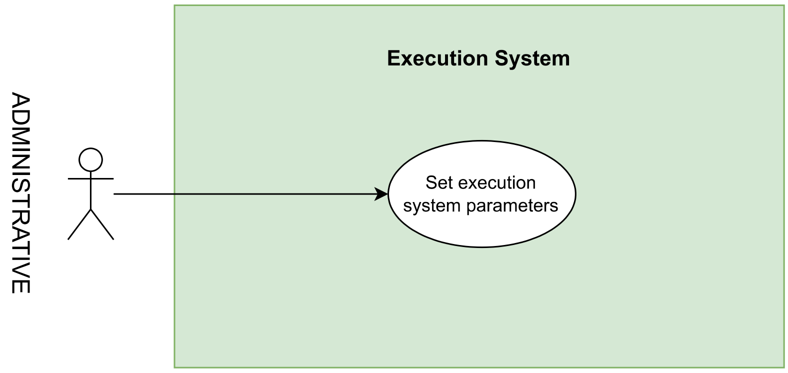


Figure 6 Use case of execution system

10.6.1 Execution system parameters

Task	*Subtask*	*%*	*Role*	*Cognitive Effort*	*Cost*
Configure execute system	1. Open parameters interface		Administrative	Remember (1)	1*1.11
	1.1 Set address to send result		Administrative	Remember (1)	1*1.11
	1.2 Save configuration into file		Administrative	Remember (1)	1*1.11
	2. Close parameters interface		Administrative	Remember (1)	1*1.11
					Total Cost: 4.44

Execute System Parameters

- Address to send result of classifier execution (1)
- Address of configuration to set execution mode (1)
- Address of monitoring system (1)

10.7 Monitoring system (Roberta Matrella)

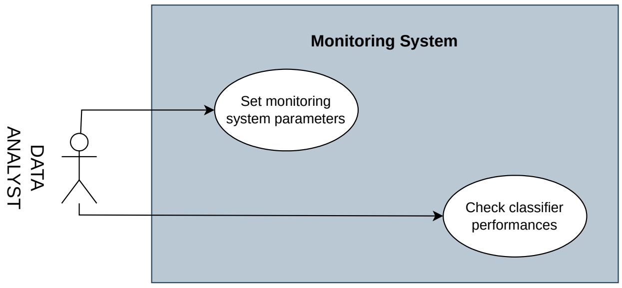


Figure 7 - Use case of the monitoring system

10.7.1 Set monitoring system parameters

Task	*Subtask*	*%*	*Role*	*Cognitive Effort*	*Cost*
Configure monitoring system	1. Open parameters interface		Quality Engineer	Remember (1)	1*1.15
	2. Set number of labeled sessions		Quality Engineer	Analyze (4)	4*1.15
	3. Set accuracy threshold		Quality Engineer	Analyze (4)	4*1.15
	4. Set monitoring address		Quality Engineer	Remember (1)	1*1.15
	5. Save configuration into file		Quality Engineer	Remember (1)	1*1.15
	6. Close parameters interface		Quality Engineer	Remember (1)	1*1. 15

Total Cost: 13.8

Monitoring System Parameters

- Number of labeled sessions (4: the expert must decide it basing on different criteria such as if the construction site is outdoor or indoor or the different kind of PPE to detect; so, it depends on the risk of the activity, and so the risk of not use the equipment and the conditions of the construction site).
- Accuracy threshold (4: the accuracy depends on the previous value of accuracy that can vary basing on different reasons; if there are more risks not wearing an equipment maybe a control more frequent is needed)
- Address for message configuration to be sent (1)

10.7.2 Check classifier performances

Task	Subtask	%	Role	Cognitive Effort	Cost
Check classifier performances	1. Open monitoring interface		Quality Engineer	Remember (1)	1*1.15
	2. Show monitoring report		System		
	3. Assess accuracy		Quality Engineer	Understand (2)	2*1.15
	IF accuracy > accuracy threshold?	97			
	4. Open reconfiguration mode		Quality Engineer	Remember (1)	1*1.15
	2. Show reconfiguration modes		System		
	3. Set reconfiguration to development		Quality Engineer	Remember (1)	1*1.15
	4. Send reconfiguration request		Quality Engineer	Remember (1)	1*1.15
	ELSE	3			
	5. Close monitoring interface		Quality Engineer	Remember (1)	1*1.15
	6. **END**				

Total Cost: 6,39

10.7.3 Monitoring report

Labeled by expert	Labeled by classifier	Correctly labeled?
helmet	helmet	YES
helmet	gloves	NO
gloves	gloves	NO
herness	eye protection	NO
harness	gloves	NO
harness	harness	YES
eye protection	helmet	NO
harness	helmet	YES

The computation of the accuracy is made by the system, the Data Analyst needs only to compare the value of accuracy with the accuracy threshold set in the configuration system.

Accuracy < Accuracy threshold	62% < 75%

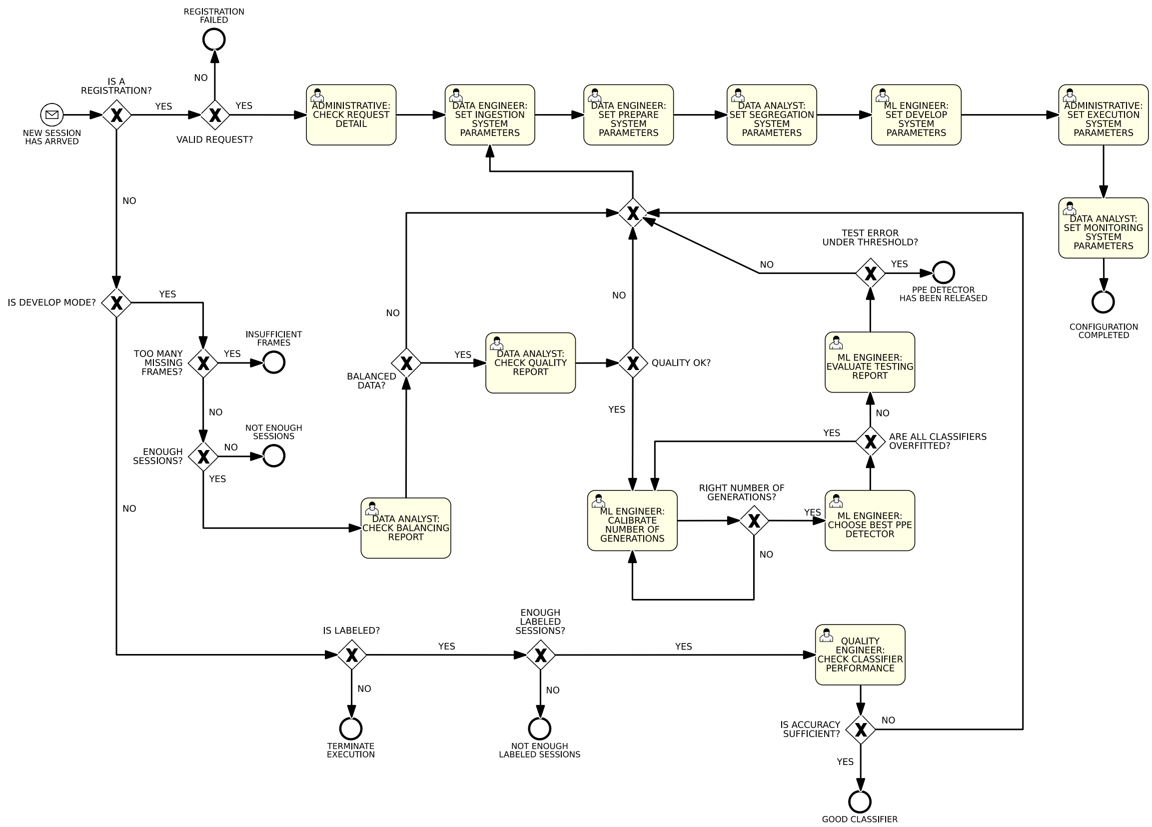
11 “As-is” (Luana Bussu, Roberta Matrella)

The “As-is” model has been realized grouping all the human tasks in a single BPMN. Then, a simulation of it was performed.

In order to evaluate the cost of all the human processes, a uniform distribution of ±15% has been taken. The BIMP simulation has been performed using the following configuration:

Instances	10000
Arrival time	90 Seconds
Duration	Uniform

Number of classifiers	2
Trainings per classifier	3
Number of monitoring (per classifier)	12
Execution sessions (per classifier)	2500
Training sessions (per classifier)	600
Sessions per monitoring	20



Considering the hypotheses written above, we have considered a total of:

- Total number of training sessions: 600x3x2 = 3600 (to branch ‘yes’ of gateway ‘IS DEVELOP MODE’)
- Total number of sessions to run: 2500x2 = 5000 (to the end event ‘TERMINATE EXECUTION’).
- Total number of monitoring sessions: 20x12x2 = 480 (to branch ‘yes’ of gateway ‘IS LABELED?’)
- Number of new requests: 3 (1 failed) (to branch ‘yes’ of gateway ‘IS A REGISTRTION?’)

We obtained a total of 9083 token that we approximated to 10000 because of the discarded tokens (for example sessions discarded for a high number of missing frames).

When the gateway 'ENOUGH SESSIONS?' is achieved, we considered that 3594 tokens go to the end event 'NOT ENOUGH SESSIONS' because, since 600 sessions are needed for training, the tokens relating to the first 599 sessions of each training (6 in total) will not be enough and for this reason they will be directed towards this end event (599x6 = 3594).

The same consideration has been made at the gateway 'ENOUGH LABELED SESSIONS?' where 20 sessions per monitoring are needed so the first 19 of each monitoring (12x2 = 24 total monitoring) go to the end event 'NOT ENOUGH SESSIONS' (24*19 = 456 tokens).

Starting from these assumptions we compute the percentage of the gateways for execute the BIMP simulation (for the gateways like 'BALANCED DATA' we used the percentages used in the relative table of human tasks in the previous section).

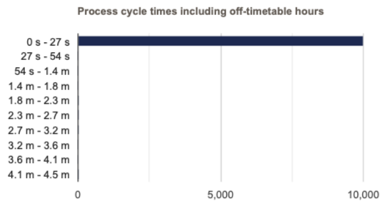
Gateway	**Yes**	**No**
It is a registration?	0,02%	99,98%
Valid request?	98%	2%
Is develop mode?	36%	64%
Session complete?	98%	2%
Too many missing frames?	5%	95%
Enough sessions?	0,2%	99,8%
Is balanced data?	95%	5%
Quality ok?	95%	5%
Are all classifier overfitted?	10%	90%
Right number of generations?	75%	25%
Test error under threshold?	80%	20%
Is labeled?	5%	95%
Enough labeled sessions?	5%	95%
Is accuracy sufficient?	97%	3%

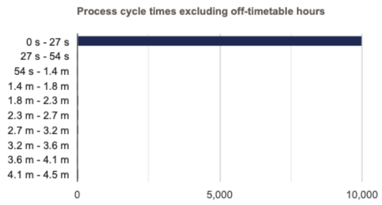
11.1 Simulation results

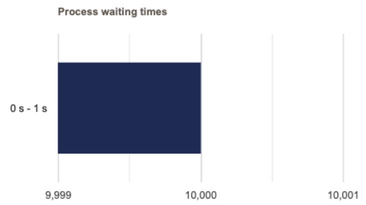
General information

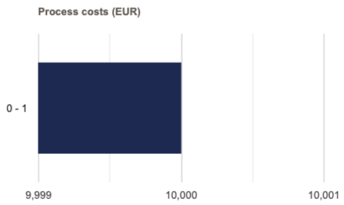
Completed process instances 10000	
Total cost 0 EUR	
Total simulation time 6.4 weeks	

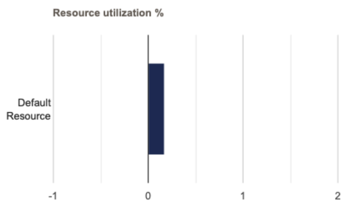
Charts











Scenario Statistics

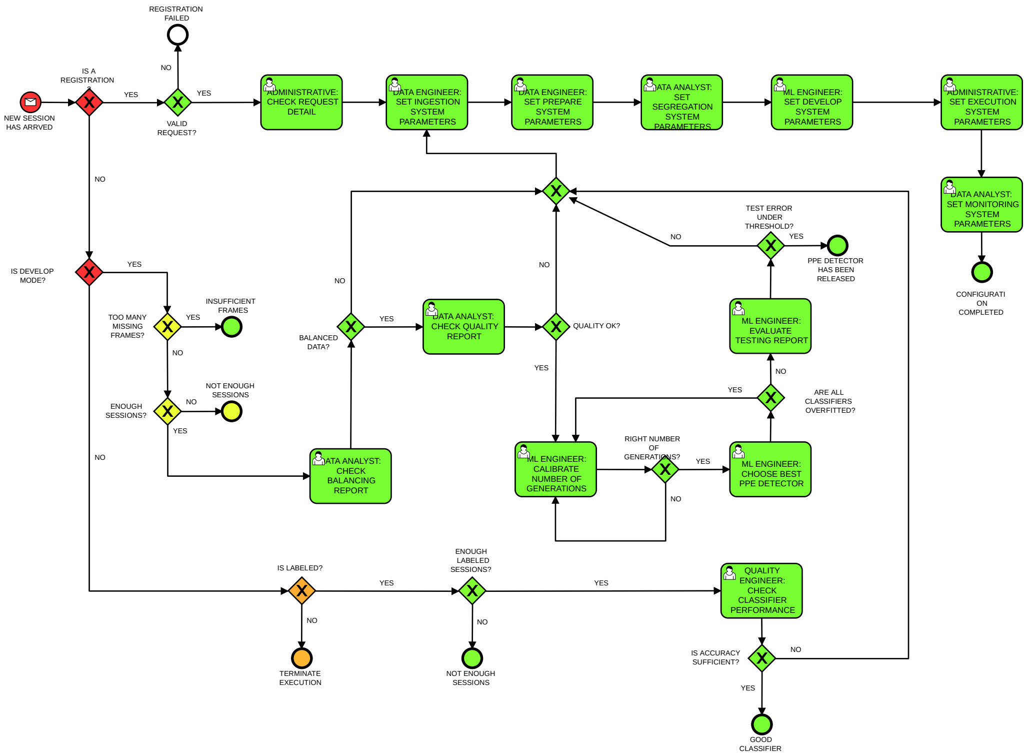
	Minimum	Maximum	Average
Process instance cycle times including off-timetable hours	0 seconds	4.4 minutes	0.2 seconds
Process instance cycle times excluding off-timetable hours	0 seconds	4.4 minutes	0.2 seconds
Process instance costs	0 EUR	0 EUR	0 EUR

Activity Durations, Costs, Waiting times, Deviations from Thresholds

Name		**Waiting time**			**Duration**			**Duration over threshold**			**Cost**			**Cost over threshold**		
	Count	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max
ADMINISTRATIVE:
SET EXECUTION SYSTEM PARAMETERS	6	0 s	0 s	0 s	3.9 s	4.4 s	5.1 s	0 s	0 s	0 s	0	0	0	0	0	0
DATA ANALYST:
SET MONITORING SYSTEM PARAMETERS	6	0 s	0 s	0 s	12.4 s	14 s	15.3 s	0 s	0 s	0 s	0	0	0	0	0	0
DATA ANALYST:
SET SEGREGATION SYSTEM PARAMETERS	6	0 s	0 s	0 s	28.3 s	30.3 s	34.1 s	0 s	0 s	0 s	0	0	0	0	0	0
DATA ENGINEER:
SET INGESTION SYSTEM PARAMETERS	6	0 s	0 s	0 s	13.3 s	15.8 s	17.4 s	0 s	0 s	0 s	0	0	0	0	0	0
DATA ENGINEER:
SET PREPARE SYSTEM PARAMETERS	6	0 s	0 s	0 s	15.4 s	17.8 s	19.6 s	0 s	0 s	0 s	0	0	0	0	0	0
ML ENGINEER:
SET DEVELOP SYSTEM PARAMETERS	6	0 s	0 s	0 s	35.7 s	37.9 s	40.8 s	0 s	0 s	0 s	0	0	0	0	0	0
ML ENGINEER:
CALIBRATE NUMBER OF GENERATIONS	7	0 s	0 s	0 s	10.7 s	11.5 s	13.3 s	0 s	0 s	0 s	0	0	0	0	0	0
ML ENGINEER:
CHOOSE BEST PPE DETECTOR	5	0 s	0 s	0 s	47.7 s	53.2 s	1 m	0 s	0 s	0 s	0	0	0	0	0	0
ML ENGINEER:
EVALUATE TESTING REPORT	5	0 s	0 s	0 s	8.2 s	8.8 s	9.5 s	0 s	0 s	0 s	0	0	0	0	0	0
ADMINISTRATIVE:
CHECK REQUEST DETAIL	2	0 s	0 s	0 s	8.4 s	8.5 s	8.6 s	0 s	0 s	0 s	0	0	0	0	0	0
DATA ANALYST:
CHECK BALANCING REPORT	7	0 s	0 s	0 s	6.4 s	7.8 s	8.6 s	0 s	0 s	0 s	0	0	0	0	0	0
DATA ANALYST:
CHECK QUALITY REPORT	6	0 s	0 s	0 s	38.8 s	44.5 s	47 s	0 s	0 s	0 s	0	0	0	0	0	0
QUALITY ENGINEER:
CHECK CLASSIFIER PERFORMANCE	16	0 s	0 s	0 s	5.5 s	6.2 s	7.3 s	0 s	0 s	0 s	0	0	0	0	0	0

Heatmap

Heatmap based on Counts

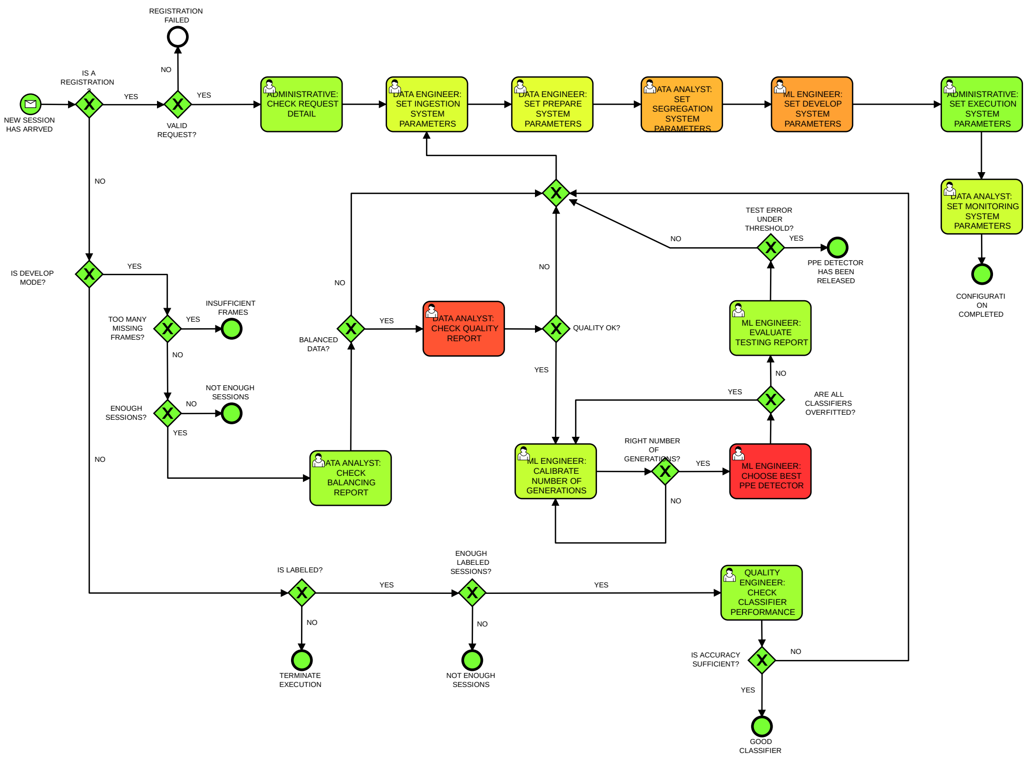


Legend

Color	Value
	0
	1111
	2222
	3333
	4444
	5556
	6667
	7778
	8889
	10000

Heatmap

Heatmap based on Durations



Legend

Color	Value
	0 s
	6 s
	12 s
	18 s
	23 s
	29 s
	35 s
	41 s
	47 s
	53 s

In the “Duration” plot, where the duration represents the cost, it is possible to observe that the human tasks with the highest values are “ML Engineer: Choose best PPE detector” and “Data Analyst: Check quality report”. It was expected because of the main role of these two tasks.

In the “Counts” plot there are no red human tasks. Only three gateways are red or orange, but it is expectable because they discriminate the main branches with high number of tokens.

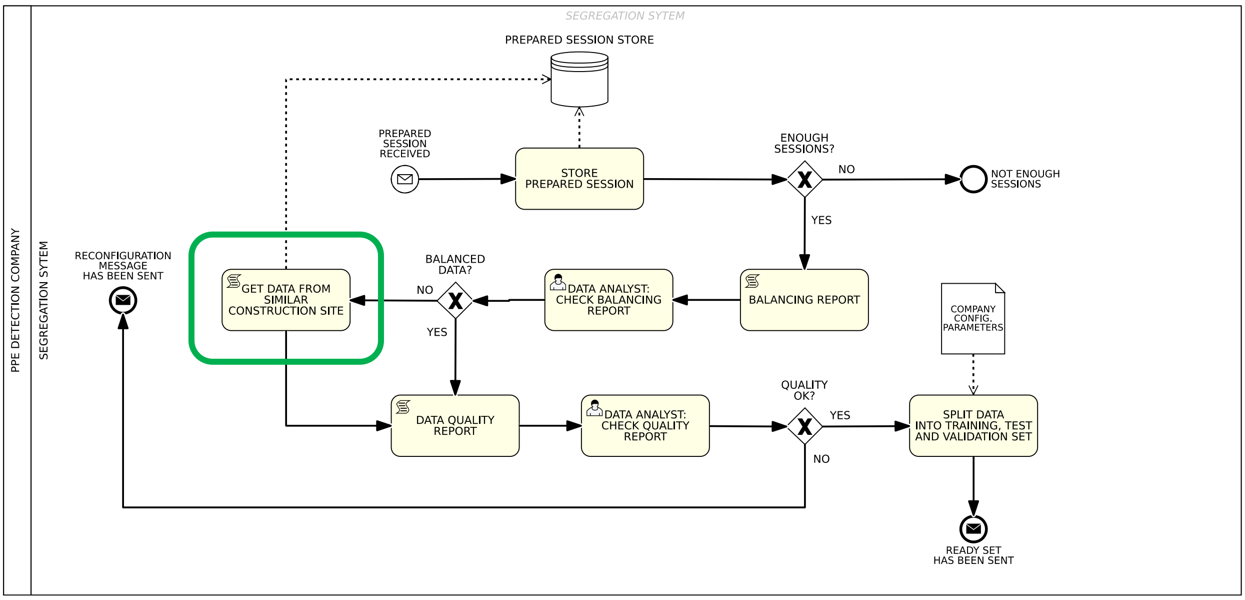
12 "To-Be” (Luca Caprioli, Martina Marino)

There are 3 changes executed to the “As Is” model to monitor the token workflow. Given the construction site under examination **A** and already established construction site **B** most like A, under the hypothesis of a clustering executed on the subscription parameters given by the other clients:

- Hand-Off: If the Data is Unbalanced, data from B is used to balance data of A instead of issuing a reconfiguration.
- Service: With the same reasoning of the Hand-Off changes, instead of conducting a full grid search optimization, hyperparameters from the chosen detector of B are used, skipping the selection of the best classifier of the grid search.
- Task: As the hyperparameters are now taken from the detector of B, the cognitive cost of configuring them drops significantly. The same is true for the validation error. **Min/Max/step hyperparameters** and **Validation error threshold** become cognitive effort **Remember (1)**.

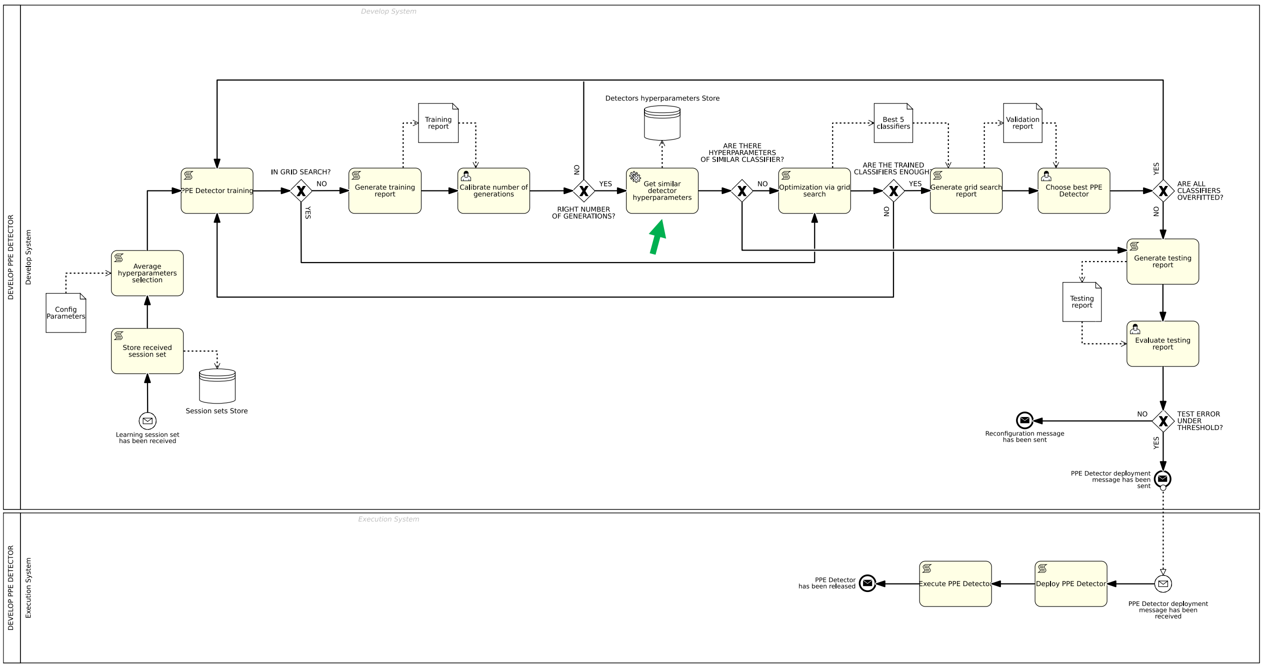
12.1 Hand-off level improvement

The improvement can be seen in the “BALANCED DATA?” gateway where in the “NO” option the next human task is the same of the “YES” option. In the middle there is a script task that recovers the information needed to go on.



12.2 Service level improvement

To introduce this type of improvement another gateway has been added after the “RIGHT NUMBER OF GENERATIONS?” one. In the “YES” case, if there is a classifier like the actual one, its hyperparameters are retrieved and used. The next step will be directly evaluating the testing report. Supposing 5 new registrations a month and considering that there are 7 main types of construction sites (residential, institutional/commercial, power generation, waste, heavy construction/transport, special industrial construction sites, gas & oil, as reported in [Construction Sites: Types, Safety & Tips | ProEst](#)) we considered that 5 categories over 7 have been already processed and their hyperparameters can be taken.



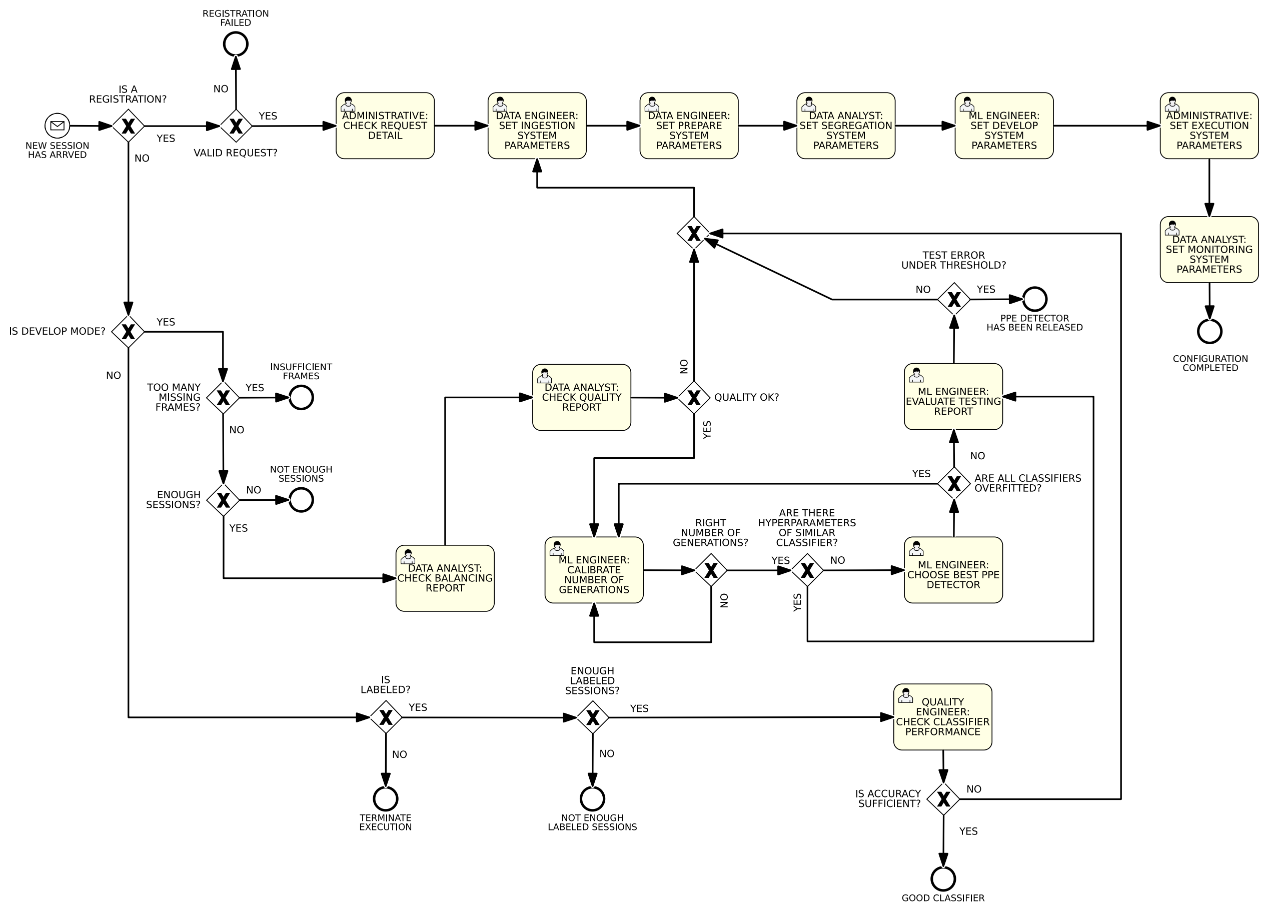
12.3 Task level improvement

The diagram is the same as the AS-IS and the improvement can be seen in the human task costs.

Task	Subtask	%	Role	Cognitive Effort	Total Cost
Configure develop system parameters	1. Open parameters interface		M.L. Engineer	Remember (1)	1*1.89
	2. Set Number of classes		M.L. Engineer	Remember (1)	1*1.89
	3. **FOR EACH** parameter		M.L. Engineer		**1*4*1.89**
	3.1. Get parameter of similar detector		M.L. Engineer	Remember (1)	1*1.89
	4. Close parameters interface		**M.L. Engineer**	**Remember (1)**	**1*1.89**

12.4 Model

The final model with the three improvements applied is the following.



12.5 Simulation results

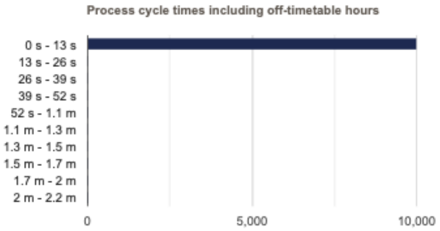
General information

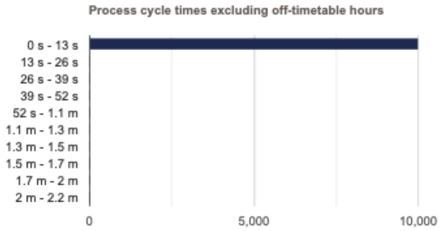
Completed process instances 10000

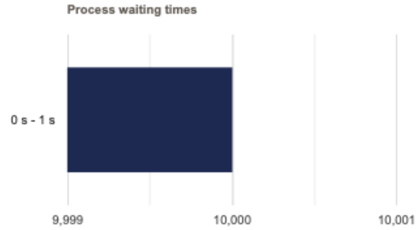
Total cost 0 EUR

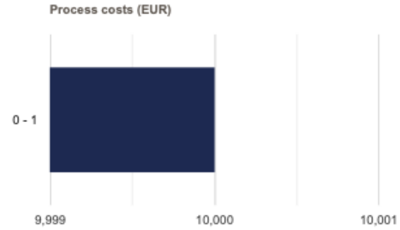
Total simulation time 6.2 weeks

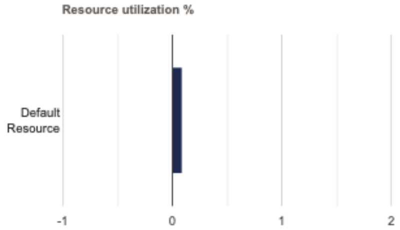
Charts











Scenario Statistics

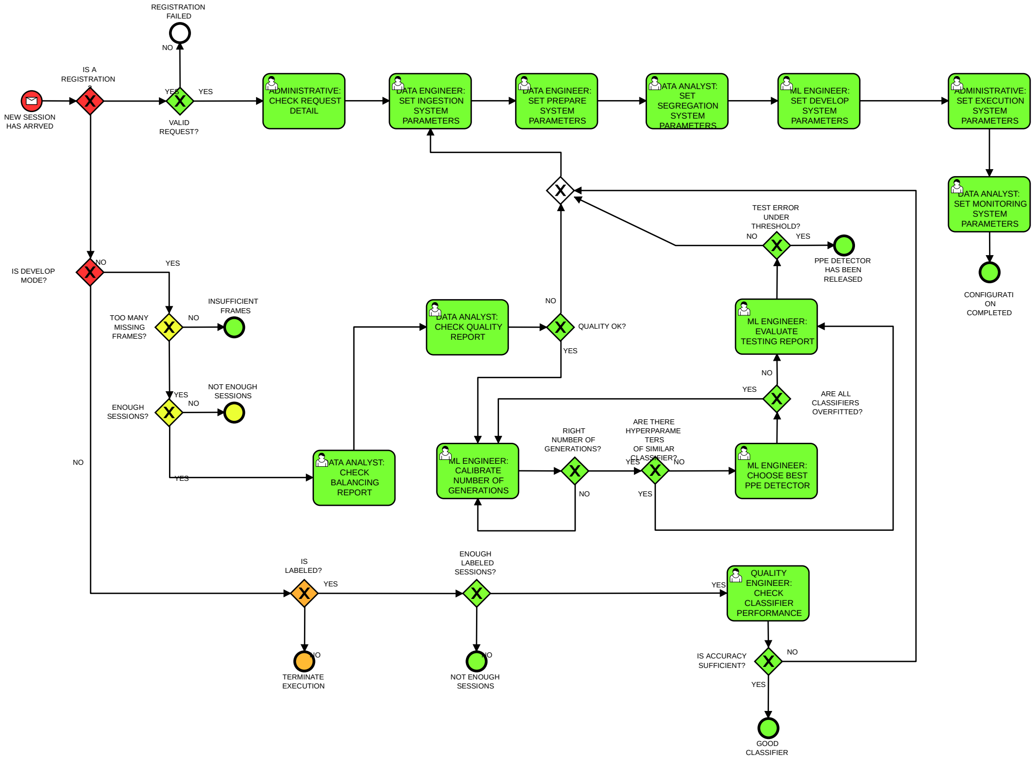
	Minimum	Maximum	Average
Process instance cycle times including off-timetable hours	0 seconds	2.2 minutes	0.1 seconds
Process instance cycle times excluding off-timetable hours	0 seconds	2.2 minutes	0.1 seconds
Process instance costs	0 EUR	0 EUR	0 EUR

Activity Durations, Costs, Waiting times, Deviations from Thresholds

Name		**Waiting time**			**Duration**			**Duration over threshold**			**Cost**			**Cost over threshold**		
	Count	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max
ADMINISTRATIVE:
SET EXECUTION SYSTEM PARAMETERS	4	0 s	0 s	0 s	3.8 s	4.4 s	4.8 s	0 s	0 s	0 s	0	0	0	0	0	0
DATA ANALYST:
SET MONITORING SYSTEM PARAMETERS	4	0 s	0 s	0 s	14.1 s	14.9 s	15.9 s	0 s	0 s	0 s	0	0	0	0	0	0
DATA ANALYST:
SET SEGREGATION SYSTEM PARAMETERS	4	0 s	0 s	0 s	27.2 s	29.8 s	32.5 s	0 s	0 s	0 s	0	0	0	0	0	0
DATA ENGINEER:
SET INGESTION SYSTEM PARAMETERS	4	0 s	0 s	0 s	13.6 s	14 s	14.3 s	0 s	0 s	0 s	0	0	0	0	0	0
DATA ENGINEER:
SET PREPARE SYSTEM PARAMETERS	4	0 s	0 s	0 s	15 s	16.5 s	18.5 s	0 s	0 s	0 s	0	0	0	0	0	0
ML ENGINEER:
SET DEVELOP SYSTEM PARAMETERS	4	0 s	0 s	0 s	11.5 s	13 s	14.4 s	0 s	0 s	0 s	0	0	0	0	0	0
ML ENGINEER:
CALIBRATE NUMBER OF GENERATIONS	3	0 s	0 s	0 s	10.4 s	10.8 s	11.2 s	0 s	0 s	0 s	0	0	0	0	0	0
ML ENGINEER:
CHOOSE BEST PPE DETECTOR	2	0 s	0 s	0 s	48.9 s	52.3 s	55.7 s	0 s	0 s	0 s	0	0	0	0	0	0
ML ENGINEER:
EVALUATE TESTING REPORT	3	0 s	0 s	0 s	7.8 s	8.6 s	9 s	0 s	0 s	0 s	0	0	0	0	0	0
ADMINISTRATIVE:
CHECK REQUEST DETAIL	3	0 s	0 s	0 s	6.7 s	6.8 s	7 s	0 s	0 s	0 s	0	0	0	0	0	0
DATA ANALYST:
CHECK BALANCING REPORT	3	0 s	0 s	0 s	7.9 s	8.3 s	8.6 s	0 s	0 s	0 s	0	0	0	0	0	0
DATA ANALYST:
CHECK QUALITY REPORT	3	0 s	0 s	0 s	39.8 s	42 s	45.7 s	0 s	0 s	0 s	0	0	0	0	0	0
QUALITY ENGINEER:
CHECK CLASSIFIER PERFORMANCE	13	0 s	0 s	0 s	5.7 s	6.5 s	7.3 s	0 s	0 s	0 s	0	0	0	0	0	0

Heatmap

Heatmap based on Counts

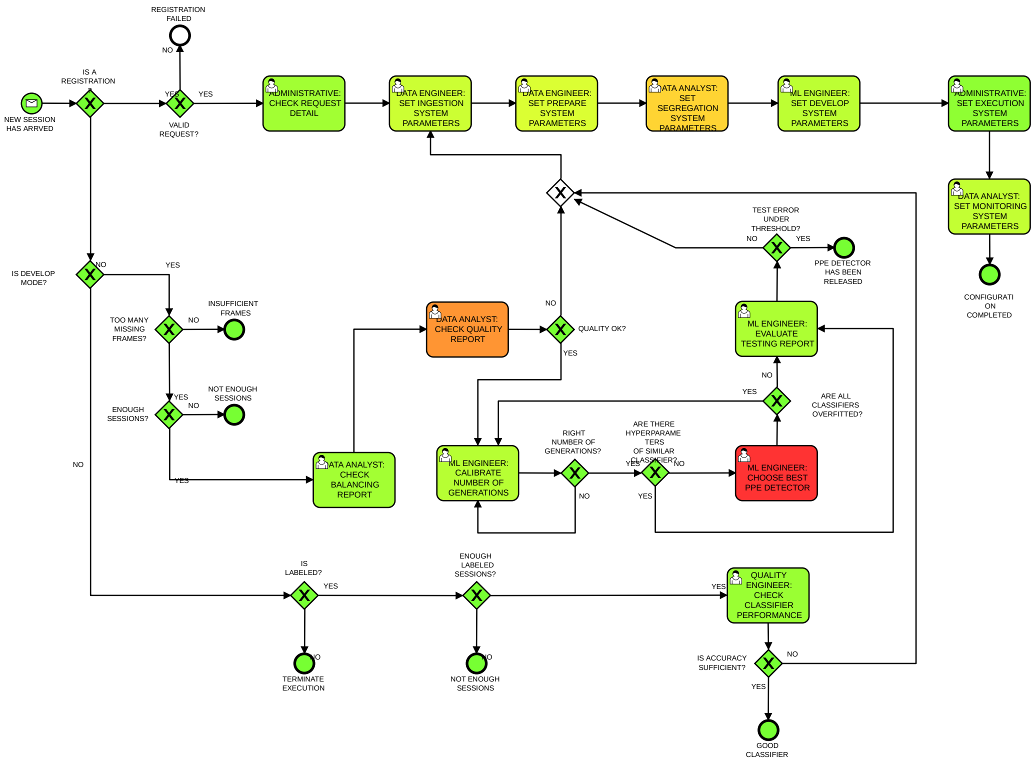


Legend

Color	Value
	0
	1111
	2222
	3333
	4444
	5556
	6667
	7778
	8889
	10000

Heatmap

Heatmap based on Durations



Legend

Color	Value
	0 s
	7 s
	13 s
	20 s
	27 s
	33 s
	40 s
	47 s
	54 s
	1 m

12.6 Comparison

Comparing the two models, the average process instance cycle time decreases from 0.2 sec to 0.1 sec. Moreover, the cost of the "set develop parameters" has been reduced and the cognitive effort of the M.L. Engineer.

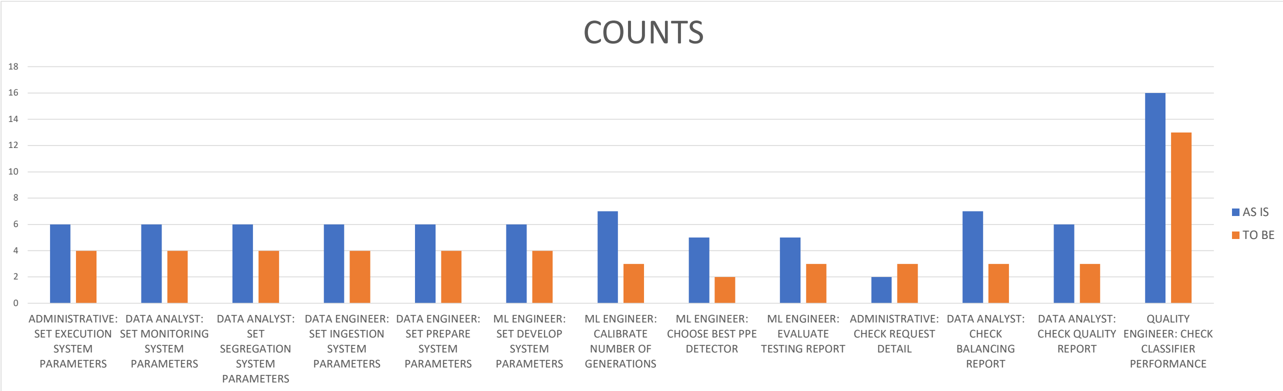


Figure 8 - Counts comparison of AS-IS and TO-BE models

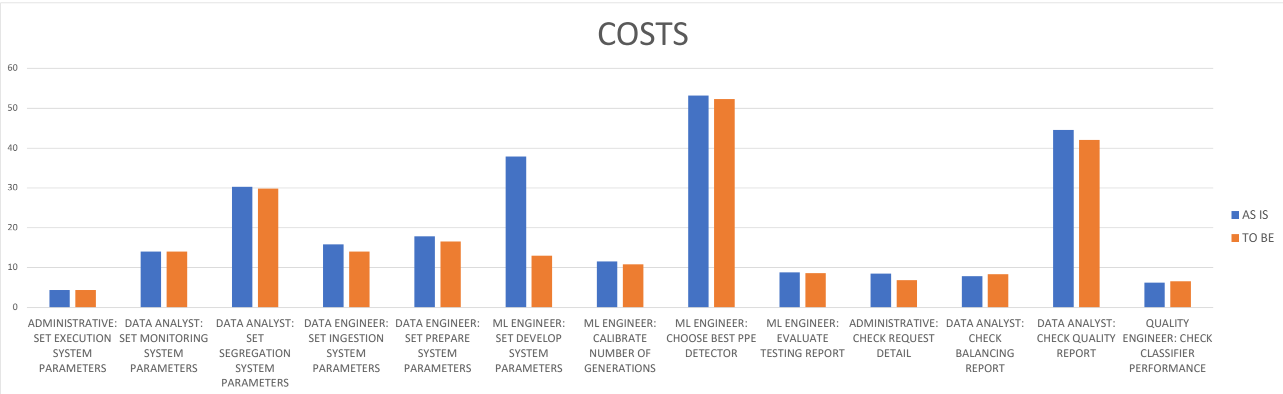


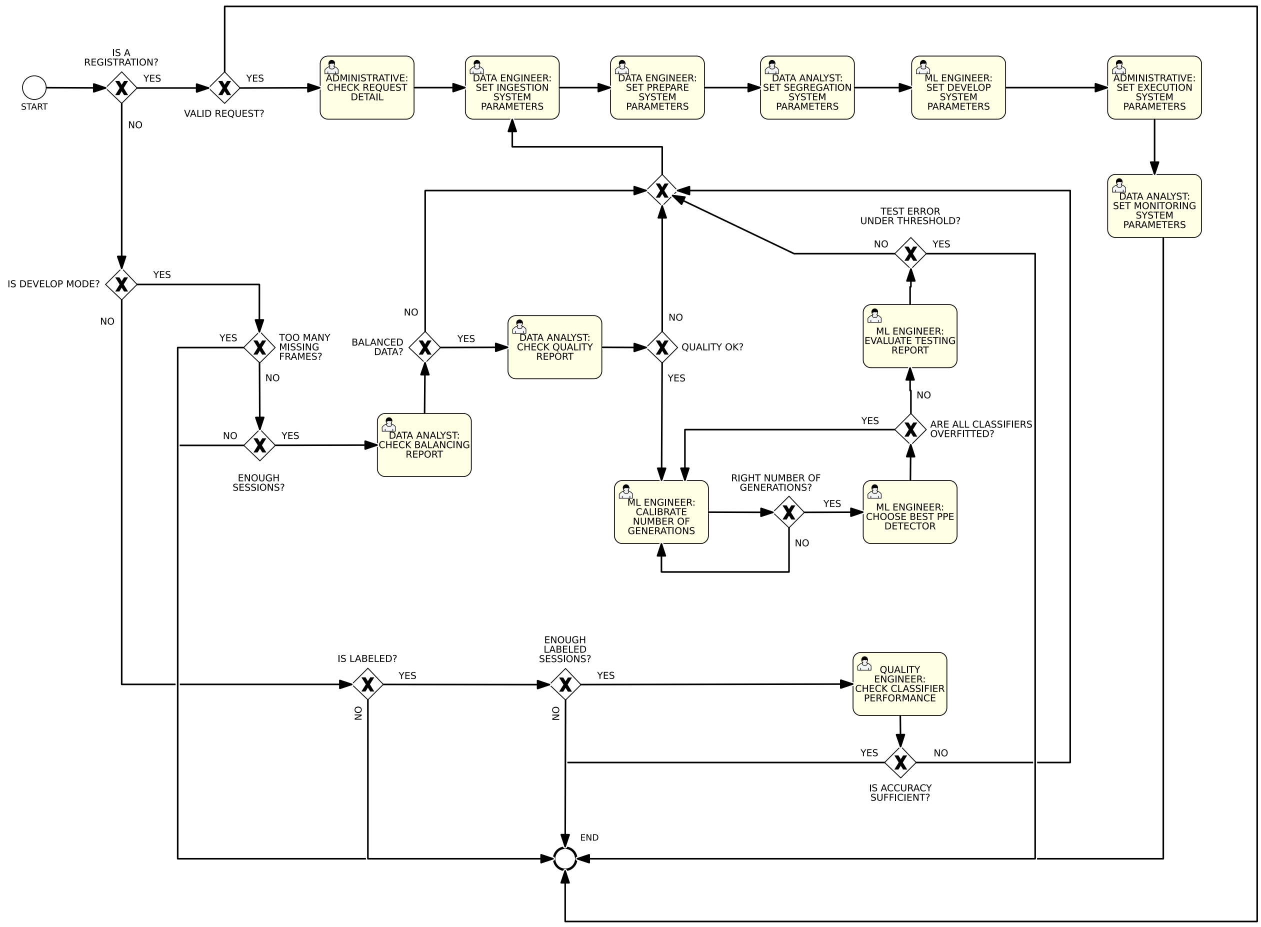
Figure 9 - Costs comparison of AS-IS and TO-BE models

13 Process mining

13.1 Normative Model (As-is)

To obtain the **original log**, normative model is simulated using BIMP considering the following configuration:

Instances	100
Resources per lane	10
Gateway probability	50%
Cost per task	1 €



Original model: 13 tasks, 14 gateways, 41 control Flows. Complexity = 68.

13.2 Transition map of the As-Is normative model

13.2.1 Transition map using Disco (Martina Marino)

The following transition map has been obtained importing on Disco the original AS-IS model using the MXML logs generated by Bimp.

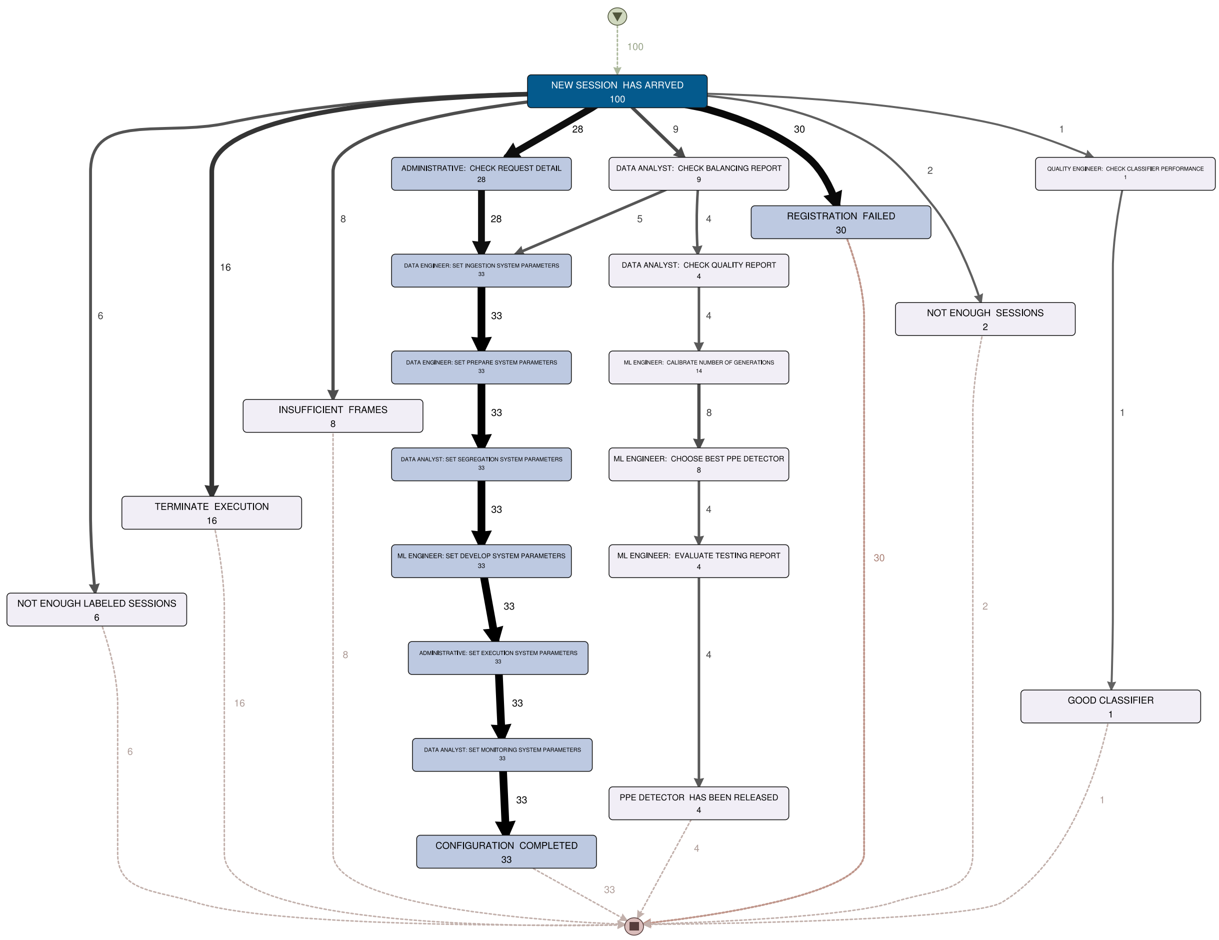


Figure 10 - Transition Map of the Normative Model on original log

In the transition map the dashed lines represent all the end events. BPMN allows process models to have multiple start and multiple end events, but such process models can be re-written as process models with a single start and a single end event, hence process models can be restricted with a single start and a single end event without loss of generality. Also, the models produced by the automated process discovery techniques from this point have a single start and a single end event.

Since it is better to have a single end point during the phase of process mining, the original As-Is model has been modified: the previous start event receiving a message has become a simple START event and all the end event now conveys to a single END event. So, simulating on Bimp the new model, the transition map obtained is the following.

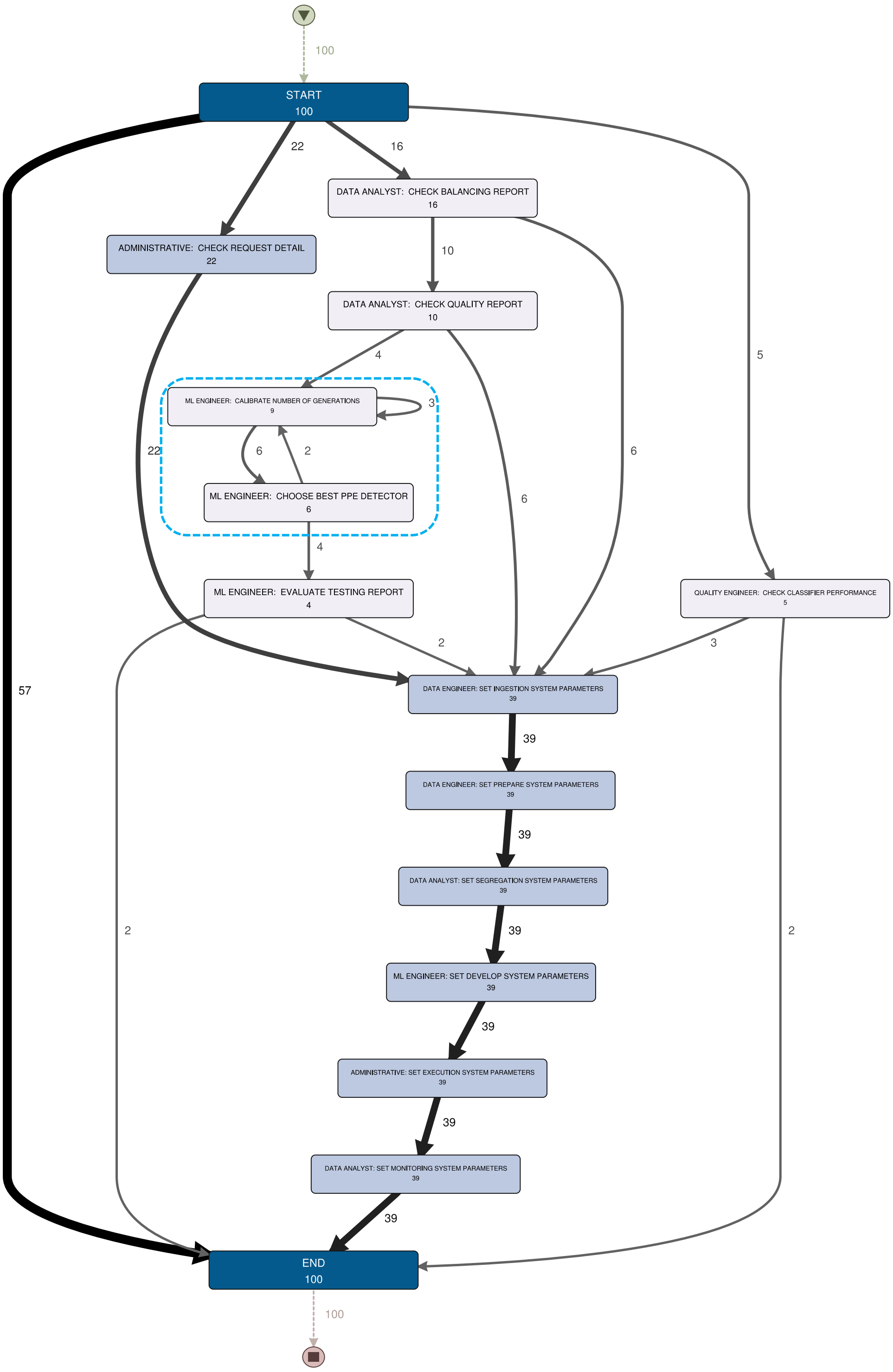


Figure 11 - Transition Map of the Normative Model with one end event

The transition map shows how many times a task is executed, or a path is run. Given 100 tokens in input, the number on the initial paths also represent the percentage of that scenario.

Variants (9)	
Complete log	All cases (100)
Variant 1	57 cases (57%)
Variant 2	22 cases (22%)
Variant 3	6 cases (6%)
Variant 4	6 cases (6%)
Variant 5	3 cases (3%)
Variant 6	2 cases (2%)
Variant 7	2 cases (2%)
Variant 8	1 case (1%)
Variant 9	1 case (1%)

Focusing on the loop between “ML ENGINEER: CALIBRATE NUMBER OF GENERATIONS” and “ML ENGINEER: CHOOSE BEST PPE DETECTOR”, the variants involved with the tasks are the following:

	Activity	Resource
1	START	
2	DATA ANALYST: CHECK BALANCING REPORT	Default Resource-000004
3	DATA ANALYST: CHECK QUALITY REPORT	Default Resource-000005
4	ML ENGINEER: CALIBRATE NUMBER OF GENERATIONS	Default Resource-000006
5	ML ENGINEER: CHOOSE BEST PPE DETECTOR	Default Resource-000007
6	ML ENGINEER: EVALUATE TESTING REPORT	Default Resource-000008
7	DATA ENGINEER: SET INGESTION SYSTEM PARAMETERS	Default Resource-000008
8	DATA ENGINEER: SET PREPARE SYSTEM PARAMETERS	Default Resource-000008
9	DATA ANALYST: SET SEGREGATION SYSTEM PARAMETERS	Default Resource-000004
10	ML ENGINEER: SET DEVELOP SYSTEM PARAMETERS	Default Resource-000009
11	ADMINISTRATIVE: SET EXECUTION SYSTEM PARAMETERS	Default Resource-000002
12	DATA ANALYST: SET MONITORING SYSTEM PARAMETERS	Default Resource-000005
13	END	

Figure 12 - Variant 6 (2 cases)

	Activity	Resource
1	START	
2	DATA ANALYST: CHECK BALANCING REPORT	Default Resource-000006
3	DATA ANALYST: CHECK QUALITY REPORT	Default Resource-000009
4	ML ENGINEER: CALIBRATE NUMBER OF GENERATIONS	Default Resource-000010
5	ML ENGINEER: CALIBRATE NUMBER OF GENERATIONS	Default Resource-000001
6	ML ENGINEER: CHOOSE BEST PPE DETECTOR	Default Resource-000002
7	ML ENGINEER: CALIBRATE NUMBER OF GENERATIONS	Default Resource-000003
8	ML ENGINEER: CHOOSE BEST PPE DETECTOR	Default Resource-000003
9	ML ENGINEER: CALIBRATE NUMBER OF GENERATIONS	Default Resource-000008
10	ML ENGINEER: CALIBRATE NUMBER OF GENERATIONS	Default Resource-000007
11	ML ENGINEER: CALIBRATE NUMBER OF GENERATIONS	Default Resource-000010
12	ML ENGINEER: CHOOSE BEST PPE DETECTOR	Default Resource-000003
13	ML ENGINEER: EVALUATE TESTING REPORT	Default Resource-000006
14	END	

Figure 13 - Variant 8 (1 case)

	Activity	Resource
1	START	
2	DATA ANALYST: CHECK BALANCING REPORT	Default Resource-000001
3	DATA ANALYST: CHECK QUALITY REPORT	Default Resource-000002
4	ML ENGINEER: CALIBRATE NUMBER OF GENERATIONS	Default Resource-000003
5	ML ENGINEER: CHOOSE BEST PPE DETECTOR	Default Resource-000004
6	ML ENGINEER: EVALUATE TESTING REPORT	Default Resource-000005
7	END	

Figure 14 - Variant 9 (1 case)

In variants 6 and 9 token do not loop but go straight to the end with a linear path while in variant 8 the token loop in the “CALIBRATE NUMBER OF GENERATIONS” and for 2 times comes back from the “CHOOSE PPE DETECTOR” because of the overfitting.

13.2.2 Transition map using Apromore (Roberta Matrella)

A process map (also called dependency graph) shows the activities of the process and the transitions between consecutive activities.

The main difference between Disco and Apromore is that, while Apromore, using case frequency, highlights how many different tokens have accessed a task, Disco shows the number of times the task has been accessed.

So, using both Apromore and Disco, it is possible to offer detailed information on what happen during simulation.

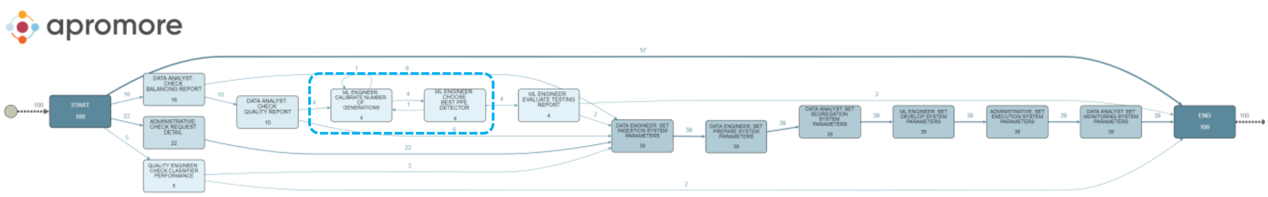


Figure 15 Transition map of the normative model with case frequency

For example, with Apromore it is possible to observe that the “Calibration number of generations” and “Choose best PPE detector” (highlighted with the light blue marker) have been accessed by four tokens and, observing the transition map provided by Disco, it is possible also to know how many times the tokens have looped between the two. The same token has loop in the “Calibrate number of generations” three times while the same token has looped two times between "Calibrate number of generations" and "Choose best PPE detector".

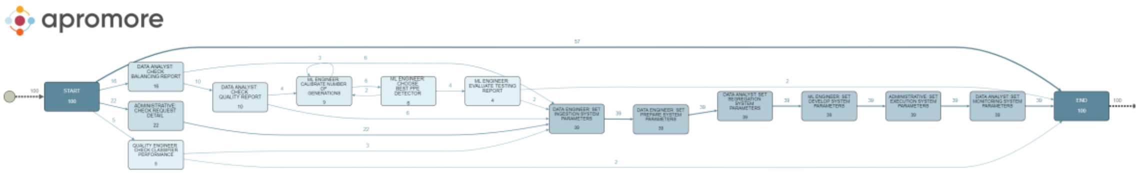
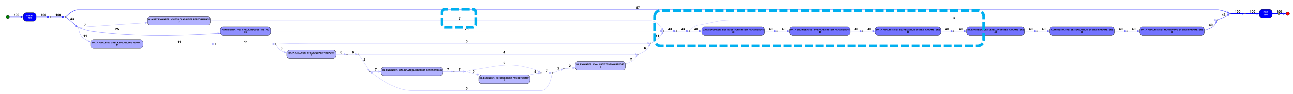


Figure 16 Transition map of the normative model with total frequency

As it is possible to observe, the transition map generated using total frequency, that shows the number of times the tokens pass through the tasks. In this case, the transition map generated is the same as the one generated with Disco.

13.2.3 Transition map using ProM (Luca Caprioli)

Using the ProM Inductive Visual Miner, a transition map of the log derived from the normative model was generated.

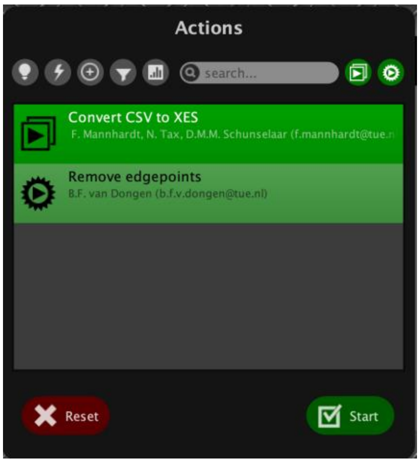


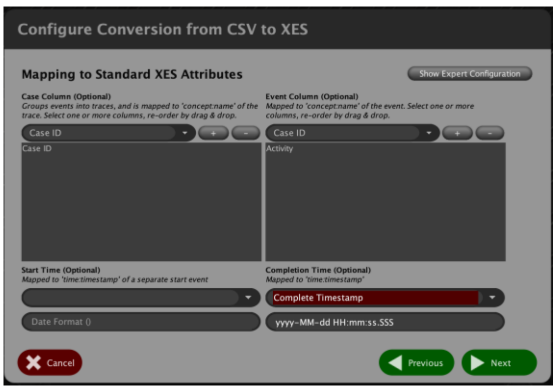
The results shown are almost in line with the already discussed maps of both Disco and the Apromore one with total case frequencies. The only real difference of the ProM map is that there are 7 tokens in input to check classifier performance instead of 5. These 2 more tokens are then equally divided before the system parameters task chain. This could be given by the tendency of the Inductive Miner to over-generalized the behavior of the log, especially in the bipartition before the configuration. On the modeling side the only difference of the ProM Inductive Visual Miner is that aggregates the token in input to a process coming from different flows.

13.3 BPMN model of the normative model (Luca Caprioli, Luana Bussu)

13.3.1 BPMN model using ProM (Luca Caprioli, Luana Bussu)

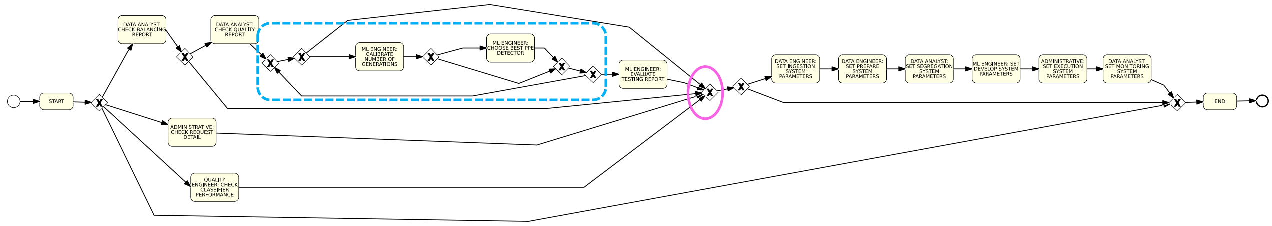
The MXML log generated by BIMP contains additional start/end events separated from the process which cause issues for conformance. For this reason, the log has been exported as CSV adding the end points using DISCO and imported in ProM. Using ProM the log has been converted into XES using the 'Convert CSV





to XES function', where, in order to select only the time end process, the Complete Timestamp has been selected as Completition time. The resulting XES have been used as log in ProM to generate a new BPMN.

To generate the mined model, after selecting the log, the 'BPMN miner' filter has been applied and, subsequently, the BPMN model has been obtained using the 'Inductive miner'. With the 'Inductive miner' with a noise value set to 0 to guarantee the maximum fitness.



ProM: 15 tasks, 10 gateways, 32 control flows between start and end. Complexity = 57.

13.3.2 BPMN model using Apromore (Roberta Matrella)

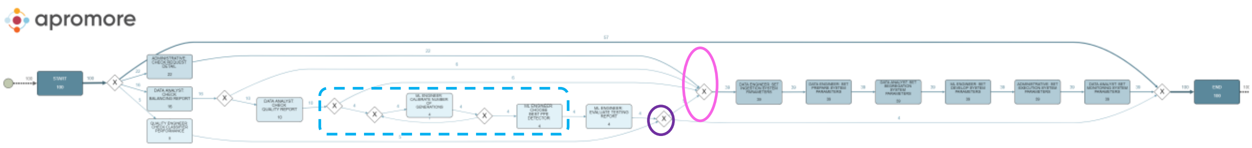


Figure 17 BPMN model original log

Apromore: 15 tasks, 8 gateways, 29 control flows between start and end. Complexity = 52

In the BPMN model mined by ProM, after the "Calibrate Number of Generations", it is possible to do the "Choose best PPE detector" or pass directly to the "Evaluate testing report". In the BPMN model mined by Apromore(blue box), the "Choose best PPE detector" cannot be avoided before the "Evaluate testing report" as in the original behavior. Indeed, as in the original model, there is a gateway that allows to come back to "Calibrate number of generations" or go to "Choose best PPE detector".

In Apromore, it is not possible to perform the "Calibrate Number of Generation" again after the "Choose best PPE detector" and this not reflects the original behavior. In the original model, there is a gateway that allows to go back to "Calibrate number of generations" or through "Evaluate testing report".

In ProM, "Evaluate testing report", "Check request detail" and "Check classifier performances" converges all to the same gateway (pink marker), while, in Apromore, "Evaluate testing report" and "Check classifier performances" converge to a gateway (purple maker) that is directed to the gateway in which converges "Check request detail" too (pink marker). In the original model this additional gateway is not present.

In ProM it is possible to skip the configuration tasks after the "Check request details", thanks to the (pink marker) gateway to which all the previous tasks converge. In Apromore, after the "Check request detail" the configuration tasks must be performed, as in the original model.

13.4 Conformance checking

In this section we discuss the different metrics utilized to check the original log against the mined model produce with ProM and the mined model generated with Apromore. We want to understand how much the original log fits the mined models.

13.4.1 Metrics

The four metrics taken in consideration are:

- **Trace-Fitness**: the fitness of a mined model measures how much the observed behaviour in the log is captured by the process model. It is computed as

$$fitness = \frac{1}{2}\left(1 - \frac{m}{c}\right) + \frac{1}{2}\left(1 - \frac{r}{p}\right)$$

  Where m is the number of missing (unproduced) output tokens, c is the number of input tokens correctly consumed, r is the number of input tokens remaining unconsumed, p is the number of output tokens correctly produced. Good value for the fitness is 1.

- **Complexity:** follows the Occam's Razor principle: the simplest model that can explain the behaviour seen in the log is the best model. We compute the complexity of the model summing up the number of gateways, activities, and sequence flows.

$$complexity = \#gateways + \#sequence\ flow + \#activities$$

  There is not a universal formula, but it depends on the criteria of the analysis. In our case, we have a fixed number of activities and the numbers of gateways and sequence flows that can change according to the mining algorithm and the used event log.

Fitness and complexity are not enough to judge the quality of a discovered process model. For this reason, we select other two metrics.

- **Precision:** should <u>avoid underfitting</u> the model. The model should not allow more behaviour than that occurred in the event logs. It's defined as:

$$Precision(L, M) = \frac{1}{|\epsilon|} \sum_{e \in \varepsilon} \frac{|en_L(e)|}{|en_M(e)|}$$

  Where ε is the collection of unique events in a context of the log, $enM(e)$ represents the enabled activities in the model M, $enL(e)$ represents the observed activities actually executed in a similar context in L.

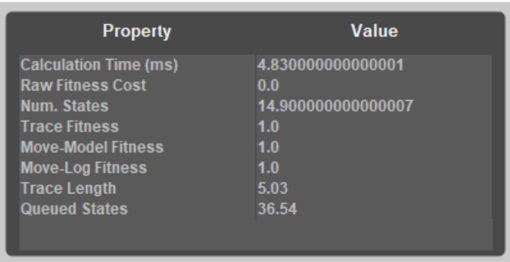
- **Generalization:** should help to <u>avoid overfitting</u>. All cases in the log where the process model is based should fit into the model. It refers to the capability of the model to abstract.

13.4.2 Results on normative model

ProM has been used also to evaluate the metrics previously discussed. In order to evaluate the fitness, the following steps were performed:

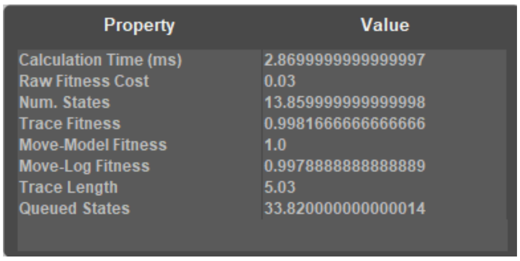
1. The BPMN mined model generated from the original log and the original log has been uploaded on ProM using 'Select BPMN Diagram'.
2. Thanks to the 'Convert BPMN Diagram to Petri Net (Control-flow)' the Petri Net has been generated.
3. The 'Replay a log on Petri Net for conformance analysis" has been applied to the Petri Net and the original log (only the resources with t_act have been set).

The fitness obtained from the BPMN model generated with ProM is shown below:



Property	Value
Calculation Time (ms)	4.830000000000001
Raw Fitness Cost	0.0
Num. States	14.900000000000007
Trace Fitness	1.0
Move-Model Fitness	1.0
Move-Log Fitness	1.0
Trace Length	5.03
Queued States	36.54

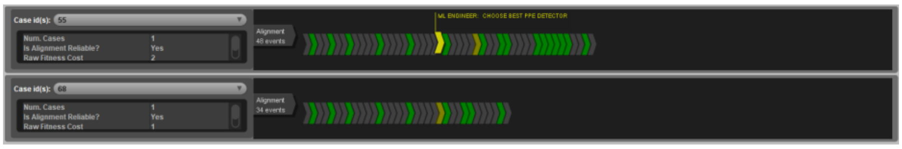
Figure 18: Fitness model ProM



Property	Value
Calculation Time (ms)	2.8699999999999997
Raw Fitness Cost	0.03
Num. States	13.859999999999998
Trace Fitness	0.9981666666666666
Move-Model Fitness	1.0
Move-Log Fitness	0.9978888888888889
Trace Length	5.03
Queued States	33.820000000000014

Figure 19: Fitness model Apromore

The alignment has been evaluated to understand the non-maximum value of the fitness. As it is possible to observe, only three cases don't respect the model. On the previous considerations, it doesn't surprise that the problem arises with the 'Choose best PPE detector" and "Calibrate number of generations" tasks. Indeed, as previously observed, Apromore doesn't allow to come back directly to the 'Calibrate number of generations' after the selection of the best PPE detector.



Model	Fitness	Complexity	Precision	Generalization
Model mined by ProM	1	57	0,68952	0,82341
Model mined by Apromore	0,998	52	0,80945	0,79311

We can see a higher value of Generalization for the ProM model and analyzing the model we can see that after the "Calibrate Number of Generations", it is possible to pass directly to the "Evaluate testing report" and this was not possible in the original model.

14 Modified Log experiments

For this experiment we considered three violations described and explained below:

- **Check Request Details**: this task has been skipped under the assumption that only valid requests are received (there is an external validation or is the building site that takes care of validates it following predefined guidelines).
  This violation has been made deleting from the log the task 'Check request detail' of case IDs 9-45-79-74.
- **Check Balancing Report:** this task has been skipped under the assumption that are used data of similar building sites and that these data are balanced.
  This violation has been made deleting from the log the task 'Check balancing report' of case IDs 13-64.
- **Choose best PPE Detector**: this task has been skipped under the assumption that are used hyperparameters of similar building sites making the grid search and the selection of the best classifier superfluous.
  This violation has been made deleting from the log the task 'Choose best PPE detector' of case IDs 68-55-68.

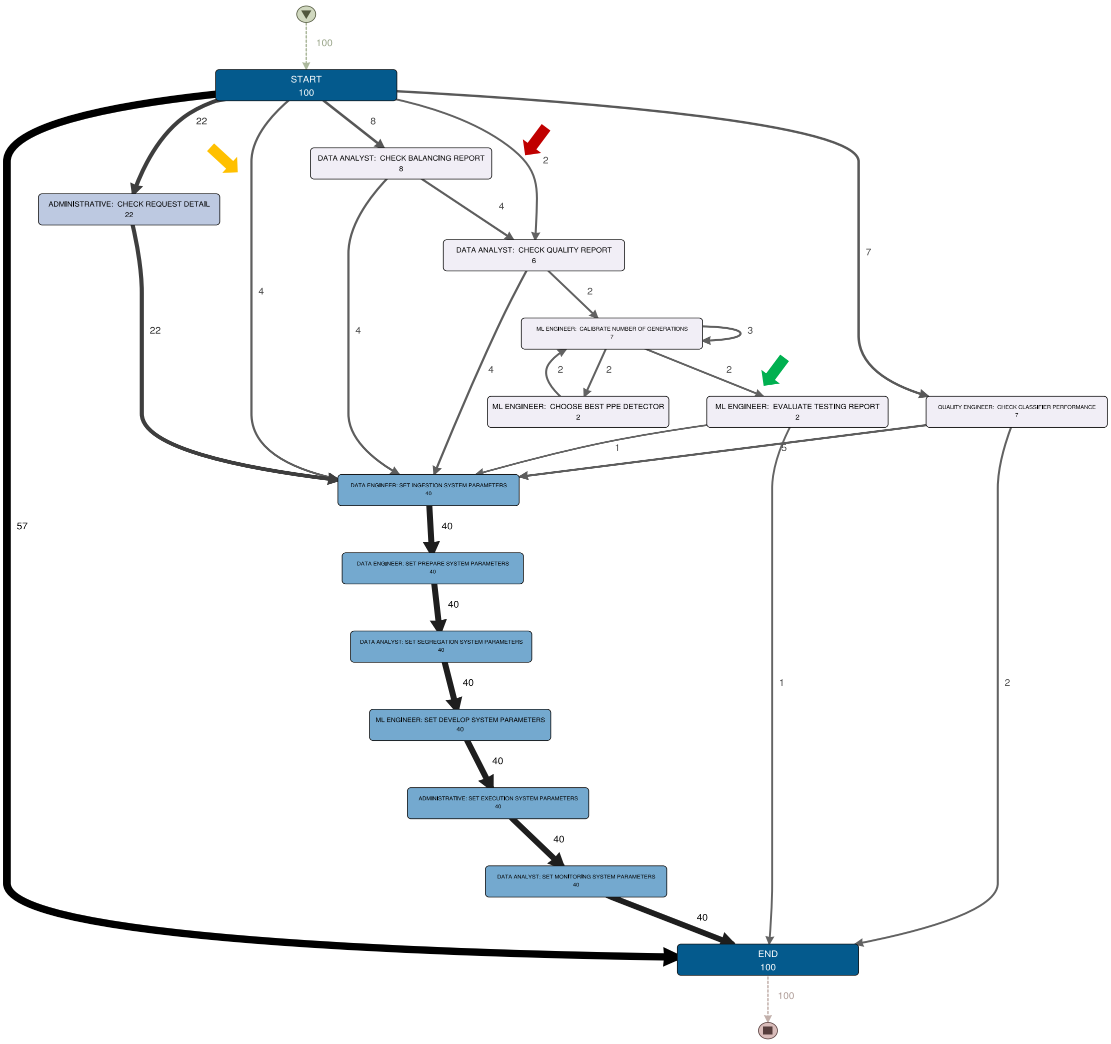
14.1 Transition map generated from modified log

14.1.1 Transition map using Disco (Martina Marino)

In the transition map generated using the modified log we can highlight the differences related to the records changed in the .csv file.

- **Check Request Details skipped**: there is a new arrow from START to DATA ENGINEER: SET INGESTION SYSTEM PARAMETERS under the assumption that only valid requests are received.
- **Check Balancing Report skipped:** the edge pointed in red is the one corresponding to the case in which data of similar building sites are used and they are already balanced.
- **Choose best PPE Detector skipped**: under the assumption that grid search and the consequent selection of the best classifier can be jumped up if hyperparameters of similar building sites are available, a new arrow now connects ML ENGINNER: CALIBRATE NUMBER OF GENERATION and ML ENGINEER: EVALUATE TESTING REPORT.

Furthermore, it can be noticed that there is a missing arrow connecting "ML ENGINEER: CHOOSE BEST PPE DETECTOR" and "ML ENGINEER: EVALUATE TESTING REPORT". The reason is that in the modified log there is not a case in which the model is good enough to go directly to that task.



14.1.2 Transition map using Apromore (Roberta Matrella)

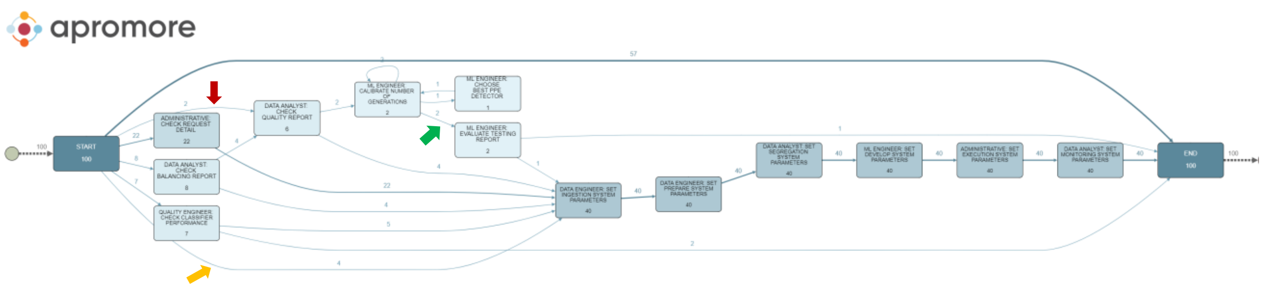


Figure 20 Transition map on Apromore with case frequency

The considerations for the transition map of Disco are also valid for the transition map generated with Apromore.

- The control flow highlighted by the red arrow, that goes from Start to “Check quality report”, represent the case in which the data is taken, balanced, from another site.
- The yellow arrow highlights the new control flow that connects directly Start with “Set injestion parameters”. It represents the case in which the request is checked by an external validator.

- The new control flow highlighted by the green arrow represents the case in which the grid search and the consequent choice of the best classifier can be skipped because the hyperparameters have been taken by another site.

In the violation of the “Choose best PPE detector” only two tokens pass through the “Evaluate testing report” because two times the same caseID has been modified. There are 4 tokens from Start to the “Set injection system parameters” and only 2 to “Check quality report” because of a different number of CaseID modified in the log.

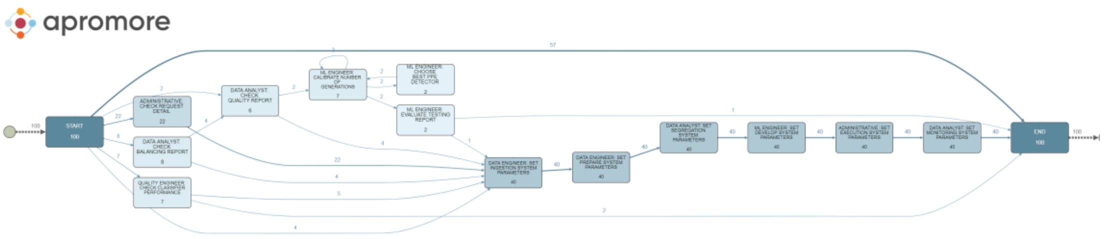
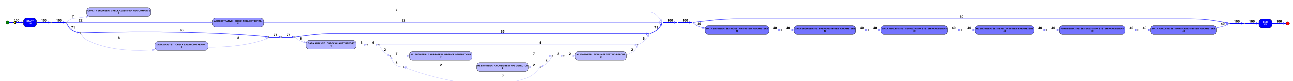


Figure 21 Transition map on Apromore with total frequency

Using the total frequency, as it is possible to observe, as in the normative model, the model generated with Apromore and with Disco are the same.

14.1.3 Transition map using ProM (Luca Caprioli)

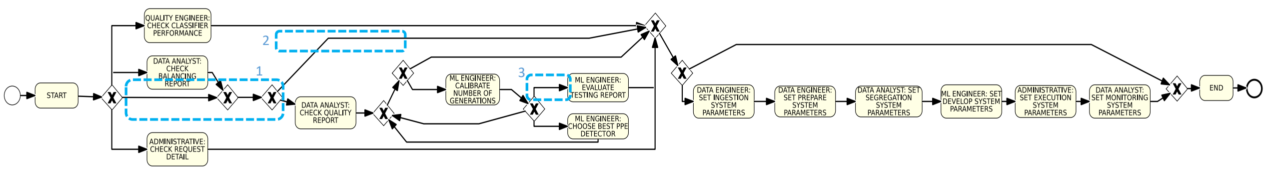
A transition map using ProM Inductive Visual Miner was created for comparison with the others just examined.



The results in this map report that 6 tokens pass in the “Check Quality Report” task, 2 tokens pass in the “Choose Best PPE Detector” task and 40 tokens pass directly from the START event to the “Set Ingestion System Parameter” task. All these results are in line with the Apromore total frequency and Disco transition map, confirming that this kind of algorithm also evaluate the total frequency of the tokens.

14.2 BPMN model generation from the modified log

14.2.1 BPMN model using ProM (Luana Bussu, Luca Caprioli)



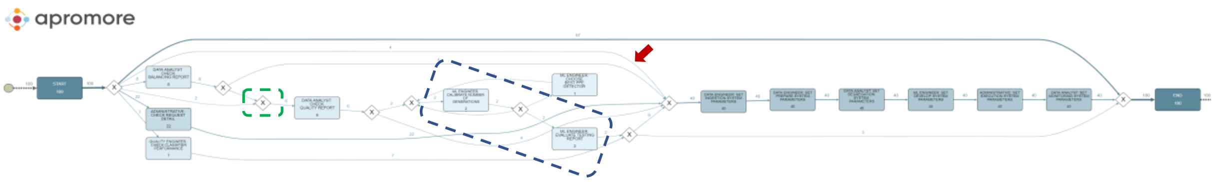
ProM: 15 tasks, 9 gateways, 31 control flow. Complexity = 55

Differences with the normative model:

1. New flow from the first gateway to the task ‘CHECK QUALITY REPORT’ without passing through the task ‘CHECK BALANGING REPORT’. This was expected because with the change on the log based on

the assumption that balanced data are received from similar shipping sites it is not necessary to have a balancing report.
2. New flow to the task 'SET INGESTION SYSTEM PARAMETERS' without passing through the task 'CHECK REQUEST DETAILS'. This was expected because of the changes on the log based on the assumption that only valid requests are received.
3. New path that allows to move to the task 'EVALUATE TESTING REPORT' without passing through the task 'CHOOSE BEST PPE DETECTOR', also expected based on the assumption the hyperparameters of a similar site are used so it is not necessary to make a grid search.

14.2.2 BPMN model using Apromore



Apromore: 15 tasks, 9 gateways, 33 control flow. Complexity = 57.

Apromore differences:

1. There is one direct control flow from the start to the first gateway (red arrow). This has been caused by the "Check request details" violation that assumes that the validity of the requests can be performed outside.
2. There is a control flow that is directed, thanks to a new gateway, (green marker) to "Check quality report" without pass through "Check balancing report". This is because of the "Check balancing report" violation. Indeed, the change in the log assumes that the data arrives already balanced from similar shipping sites and, for this reason, there is no need of balancing report.
3. A new path (blue marker) allows to go directly to "Evaluate testing report" without pass through "Choose best PPE detector" task. This is due to the "Choose best PPE detector" violation based on the assumption that the hyperparameters can be taken from other shipped sites, so it is not necessary to perform grid search.

Now we have a control flow that allows to go from "Choose best PPE detector" to "Calibrate number of generations" because the modification on the logs have been performed on the caseID's that were not respecting the model in the original log. In this way we have the control flow that respects the original behavior.

14.3 Conformance checking

Following the metrics previously exposed in the paragraph 13.4.1, a conformance checking phase has been done.

14.4 Results on normative model (with modified log).

Property	Value
Calculation Time (ms)	3.5600000000000014
Raw Fitness Cost	0.060000000000000001
Num. States	15.119999999999997
Trace Fitness	0.9941818181818182
Move-Model Fitness	0.9935555555555556
Move-Log Fitness	1.0
Trace Length	4.940000000000001
Queued States	36.70000000000001

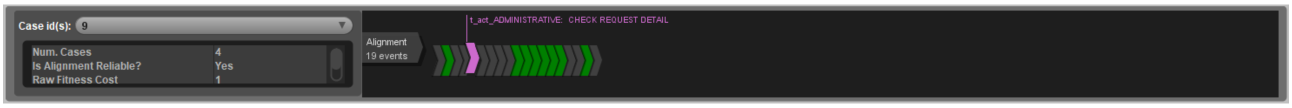
Figure 23 Fitness with ProM

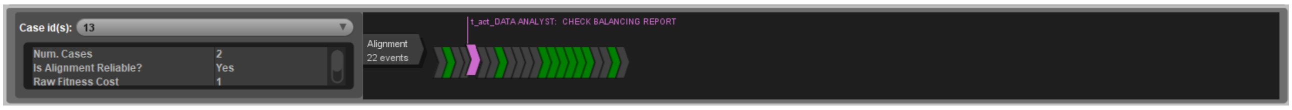
Property	Value
Calculation Time (ms)	4.309999999999999
Raw Fitness Cost	0.090000000000000001
Num. States	13.740000000000006
Trace Fitness	0.9921291866028709
Move-Model Fitness	0.9924444444444446
Move-Log Fitness	0.9988235294117647
Trace Length	4.940000000000001
Queued States	33.519999999999996

Figure 22 Fitness with Apromore

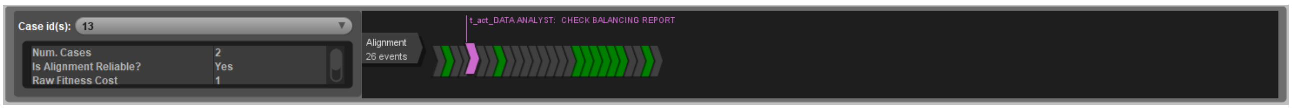
As we can see, the fitness has decreased in both models since the used log contains some violations.

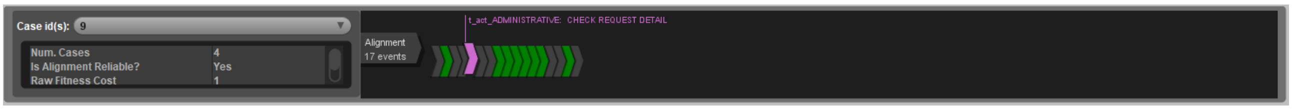
In the ProM model there are two cases that doesn't reach the fitness equal to 1, they correspond to skipped task (purple marker) that as expected correspond to the changes made on the log.
The task 'CHOOSE BEST PPE DETECTOR' is not skipped because the log was modified removing two cases of the same token so, the second time it has been skipped but the first one it is accessed (and it was expected because of the new direct path to the task 'EVALUATE TESTING REPORT').





The same think can be said for Apromore were the task skipped (purple markers) correspond to the tasks skipped in the modified log.





Model	Fitness	Complexity	Precision	Generalization
Model mined by ProM	0,994	57	0,7333	0,865
Model mined by Apromore	0,992	52	0,8196	0,869

14.5 Results on violated model

The log has been modified introducing the previously described violations and imported into ProM. Then, the metrics have been evaluated using the modified log and the model mined (with the modified log) using ProM and Apromore. The following results have been obtained:

Property	Value
Calculation Time (ms)	2.8899999999999997
Raw Fitness Cost	0.0
Num. States	19.419999999999998
Trace Fitness	1.0
Move-Model Fitness	1.0
Move-Log Fitness	1.0
Trace Length	4.940000000000001
Queued States	46.67999999999999

Figure 24 Fitness with ProM

Property	Value
Calculation Time (ms)	1.2099999999999993
Raw Fitness Cost	0.0
Num. States	13.659999999999997
Trace Fitness	1.0
Move-Model Fitness	1.0
Move-Log Fitness	1.0
Trace Length	4.940000000000001
Queued States	35.56999999999999

Figure 25 Fitness with Apromore

Model	Fitness	Complexity	Precision	Generalization
Model mined by ProM	1	55	0,7195	0,8428
Model mined by Apromore	1	57	0,7497	0,8428

As expected, the fitness values are very high (they reach the maximum value): differently from the conformance checking on the normative model, we can see that also Apromore reaches the maximum value of 1 and in fact analyzing the model we can see that now is possible to come back directly to the 'Calibrate number of generations' after the selection of the best PPE detector. The other metrics are similar to the previous ones. Basing on the complexity, it is possible to say that the model is more complex due to the new paths present in the model mined by the modified log. Precision and generalization have both increased w.r.t. the original model.

15 Final Results

Model		Fitness	Complexity	Precision	Generalization
Mined On	**Evaluated on**				
Normative model by ProM	**Original Log**	1	57	0,6895	0,8234
Normative model by Apromore	**Original Log**	0,998	52	0,8094	0,7931
Violated model by ProM	**Modified Log**	1	55	0,7195	0,8428
Violated model by Apromore	**Modified Log**	1	57	0,7497	0,8428
Normative model by ProM	**Modified Log**	0,994	57	0,7333	0,865
Normative model by Apromore	**Modified Log**	0,992	52	0,8196	0,869

In the table above, all the results obtained during the mining phase are reported.

Considering the **original log**, basing on the 4 metrics taken into consideration, the BPMN model generated by Apromore achieves the best results. Indeed, it has a fitness of about 1, it has less complexity, a much better precision and a similar generalization w.r.t. the BPMN model generated by ProM.

Evaluating the 4 metrics on the **model generated with the modified log**, ProM and Apromore do not differ a lot. Fitness and generalization values are the same with the only difference regarding a little more complex but also precise values with Apromore.

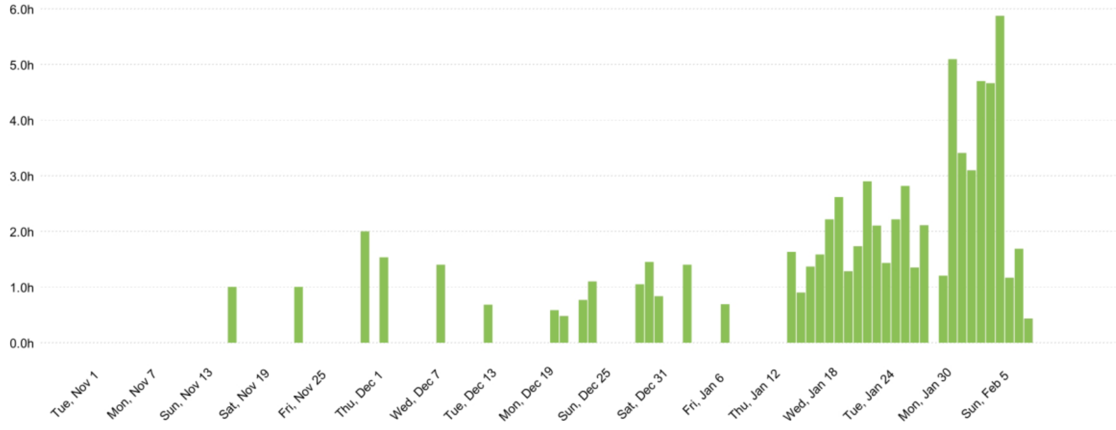
The last comparison is made on the **normative model on the modified log** showing a decrease of fitness given by the violations that are unexpected for this model. Given the same value of generalization, the results of the other two metrics can be discussed. As for the original model and log, the Apromore model is less complex with a precision even more similar. For these reasons, Apromore model is preferrable.

Comparing the different models from a business perspective, we can assert that the modified model adds alternative paths that generate a lower cost for the company (internal prospective). The reason is that, in the modified model, some tasks have been skipped, increasing the velocity of the processes and reducing the costs in terms of human resources.

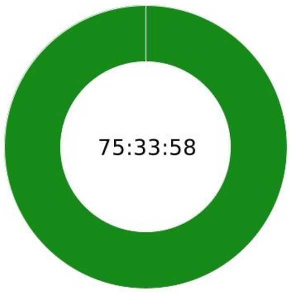
16 Clockify

16.1 Luana Bussu

Total: 75:33:58

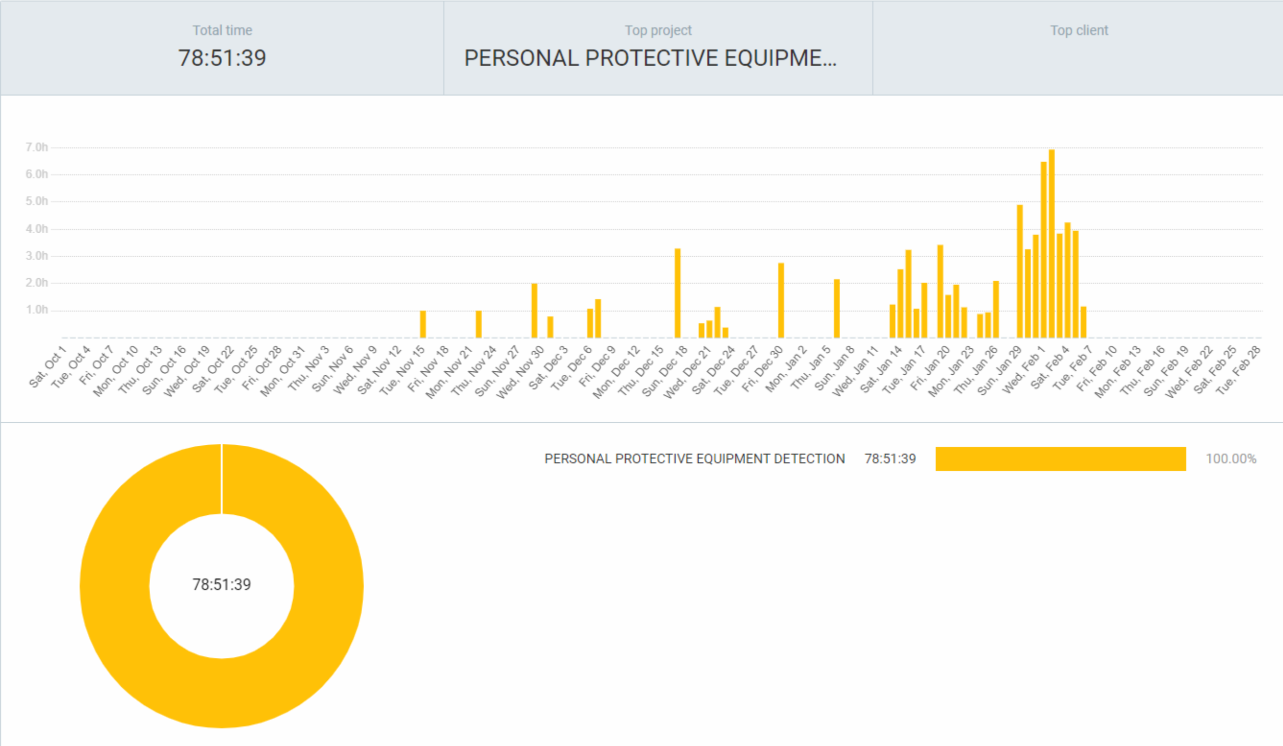


User

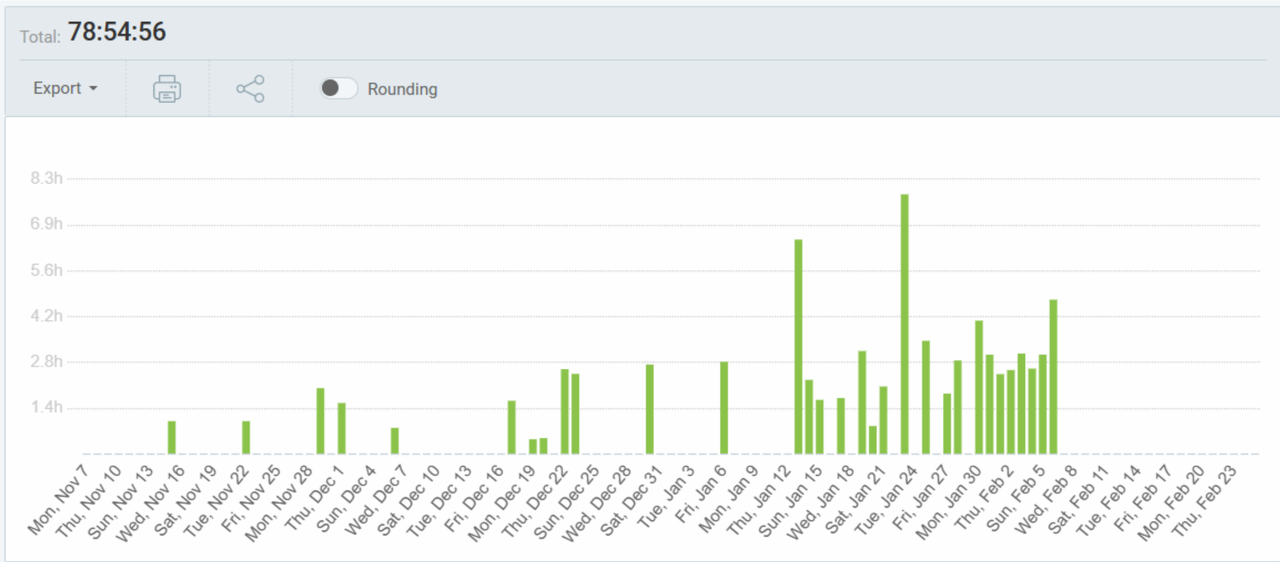


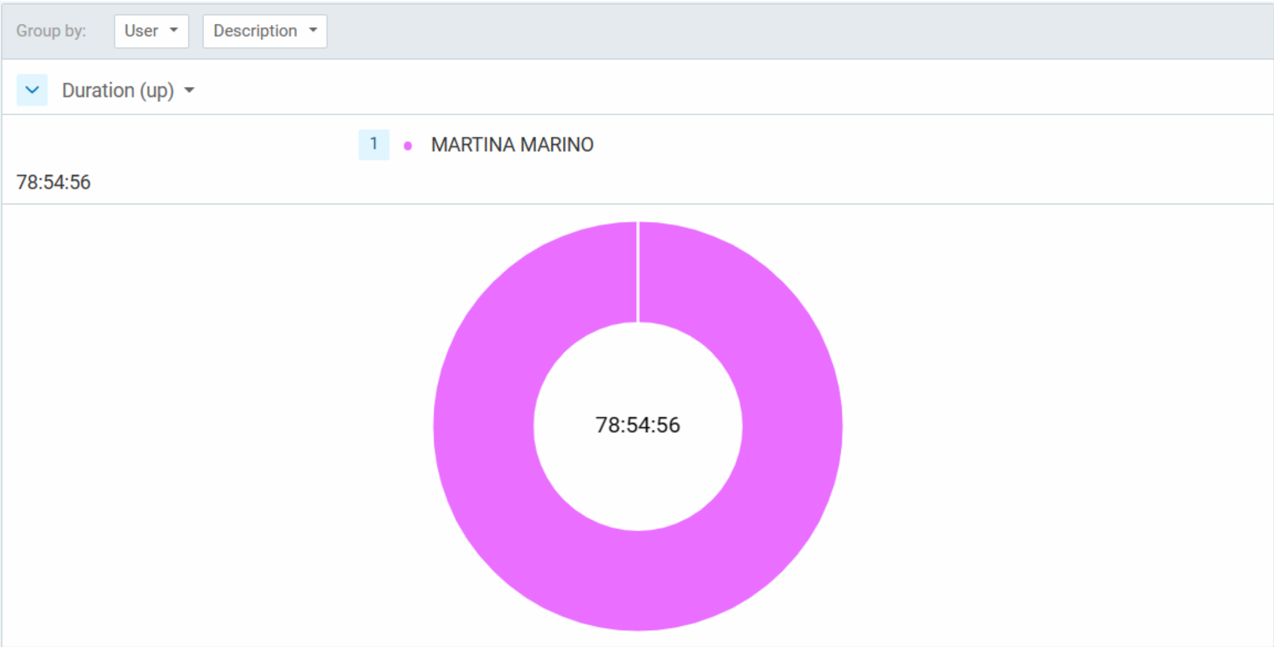
Luana Bussu	75:33:58	100.00%

16.2 Luca Caprioli



16.3 Martina Marino





16.4 Roberta Matrella