

# Process Mining and Intelligence Project

## Emotion Based Music Selection

Ettore Ricci

Francesco Boldrini

Paolo Palumbo

Zahra Omrani

January 22, 2025

## Contents

<b>1</b>	<b>BPMN modeling</b>	<b>2</b>
1.1	Process landscape . . . . .	2
1.2	Process model . . . . .	2
1.2.1	Prepare session . . . . .	2
1.2.2	Generate learning sets . . . . .	2
1.2.3	Develop classifier . . . . .	2
1.2.4	Classify session . . . . .	2
1.2.5	Evaluate classifier performance . . . . .	2
1.2.6	Configure systems . . . . .	2
<b>2</b>	<b>Task level modeling</b>	<b>2</b>
2.1	Segregation system . . . . .	2
2.1.1	Check data balancing . . . . .	2
2.1.2	Check input coverage . . . . .	3
2.2	Development system . . . . .	4
2.2.1	Set iteration number . . . . .	4
2.2.2	Check learning report . . . . .	5
2.2.3	Check validation report . . . . .	6
2.2.4	Check test results . . . . .	8
2.3	Evaluation system . . . . .	9
2.3.1	Evaluate classifier performance . . . . .	9

# 1 BPMN modeling

## 1.1 Process landscape

## 1.2 Process model

### 1.2.1 Prepare session

### 1.2.2 Generate learning sets

### 1.2.3 Develop classifier

### 1.2.4 Classify session

### 1.2.5 Evaluate classifier performance

### 1.2.6 Configure systems

# 2 Task level modeling

Position	Description	Salary	Normalized Salary
Clerk		\$52,000.00	1.00
Data analyst		\$60,000.00	1.15
ML engineer		\$130,000.00	2.50
Data scientist		\$123,000.00	2.37
Domain expert (Neurologist)		\$267,000.00	5.13
Minimum		\$52,000.00	1.00

Table 1: Salary and normalized salary for each position

## 2.1 Segregation system

### 2.1.1 Check data balancing

The task is performed by a Data scientist.

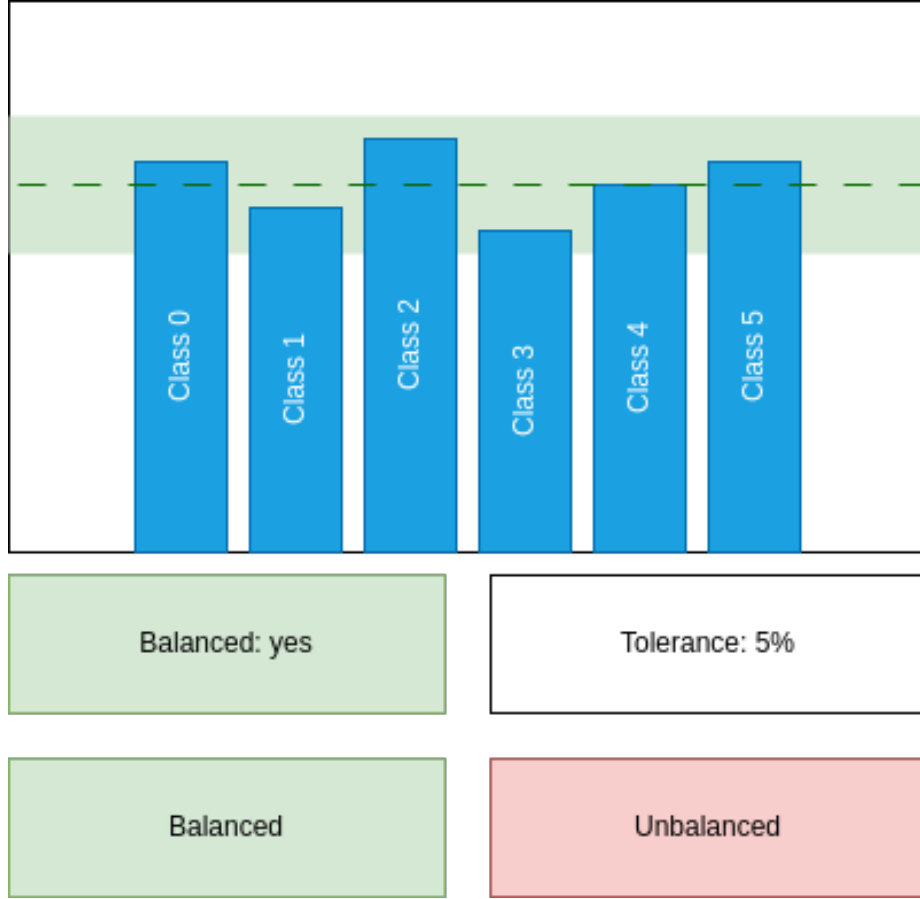


Figure 1: "Check data balancing" mock-up form

Step	O	CL	S	SC
1 <b>ACTOR</b> opens "Check data balancing" form.	1	1	2.37	2.37
2 <b>SYSTEM</b> shows the report.				
3 <b>SYSTEM</b> shows a hint whether the data is balanced or not.				
4 <b>ACTOR</b> checks threshold in the UI.	1	2	2.37	4.74
5 <b>FOR EACH</b> column in the report:	5			
5.1 <b>IF</b> the column is not within the displayed threshold.	4			
5.1.1 <b>THEN</b> the data is not balanced.	4			
6.1 <b>IF</b> the data is balanced.	0.2			
6.1.1 <b>ACTOR</b> clicks "Balanced" button.	0.2	1	2.37	0.47
6.2 <b>ELSE</b>	0.8			
6.2.1 <b>ACTOR</b> clicks "Unbalanced" button.	0.8	1	2.37	1.90
7 <b>SYSTEM</b> shows a confirmation dialog.				
8 <b>ACTOR</b> closes the form.	1	1	2.37	2.37
Human task cost				11.85

Table 2: Detailed use case for "Check data balancing" task

O - Occurrence, CL - Cognitive Level, S - Normalized Salary, SC - Step Cost

### 2.1.2 Check input coverage

The task is performed by a Data scientist.

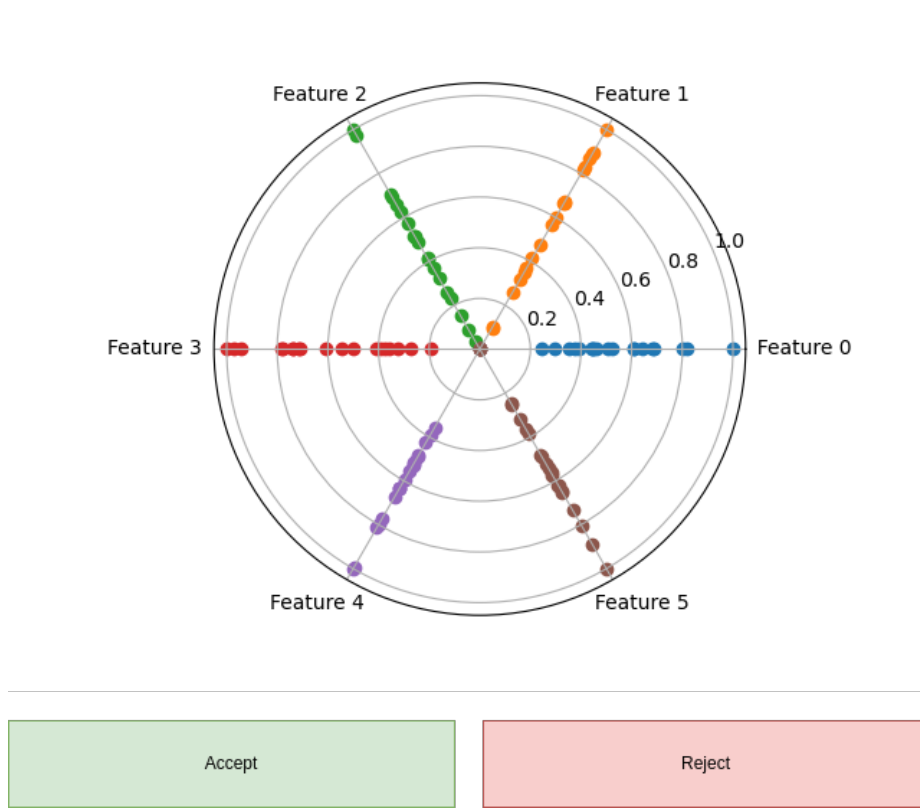


Figure 2: "Check input coverage" mock-up form

Step	O	CL	S	SC
1 <b>ACTOR</b> opens "Check input coverage" form.	1	1	2.37	2.37
2 <b>SYSTEM</b> shows a radar scatter plot of the input distribution.				
3 <b>FOR EACH</b> radius in the radar scatter plot:	5			
3.1 <b>IF</b> the distribution is not uniform as expected.	3.33	4	2.37	31.57
3.1.1 <b>THEN</b> the input coverage is not satisfied.	3.33			
4.1 <b>IF</b> the input coverage is satisfied.	0.33			
4.1.1 <b>ACTOR</b> clicks "Accept" button.	0.33	1	2.37	0.79
4.2 <b>ELSE</b>	0.66			
4.2.1 <b>ACTOR</b> clicks "Reject" button.	0.66	1	2.37	1.57
5 <b>SYSTEM</b> shows a confirmation dialog.				
6 <b>ACTOR</b> closes the form.	1	1	2.37	2.37
Human task cost				38.68

Table 3: Detailed use case for "Check input coverage" task

O - Occurrence, CL - Cognitive Level, S - Normalized Salary, SC - Step Cost

## 2.2 Development system

### 2.2.1 Set iteration number

The task is performed by a ML engineer.

Set Iteration Number

Current Iteration Number 10

Enter New Iteration Number:

Submit

Iteration number updated successfully!

Figure 3: "Set iteration number" mock-up form

Step	O	CL	S	SC
1 <b>ACTOR</b> opens "Set Iteration Number" form.	1	1	2.5	2.5
2 <b>SYSTEM</b> displays the current iteration number.				
3 <b>ACTOR</b> inputs the desired number of iterations.	1	3	2.5	7.5
4 <b>ACTOR</b> clicks "Submit" button to confirm the iteration number.	1	1	2.5	2.5
5 <b>SYSTEM</b> shows a confirmation dialog.				
6 <b>ACTOR</b> closes the form.	1	1	2.5	2.5
Human task cost				15

Table 4: Detailed use case for "Set iteration number" task

O - Occurrence, CL - Cognitive Level, S - Normalized Salary, SC - Step Cost

### 2.2.2 Check learning report

The task is performed by a ML engineer.



Figure 4: "Check learning report" mock-up form

Step	O	CL	S	SC
1 <b>ACTOR</b> opens "Check training report" form.	1	1	2.50	2.50
2 <b>SYSTEM</b> shows the training loss curve.				
3.1 <b>IF</b> the loss is flat for at least half of the iterations:	0.4	3	2.50	3.00
3.1.1 <b>THEN ACTOR</b> clicks "Overfit" button.	0.4	1	2.50	1.00
3.2 <b>IF</b> the loss is not flat at the end of the iterations:	0.4	3	2.50	3.00
3.2.1 <b>THEN ACTOR</b> clicks "Underfit" button.	0.4	1	2.50	1.00
3.3 <b>ELSE</b>	0.2	3	2.50	1.50
3.3.1 <b>ACTOR</b> clicks "Approved" button.	0.2	1	2.50	0.50
4 <b>SYSTEM</b> shows a confirmation dialog.				
5 <b>ACTOR</b> closes the form.	1	1	2.50	2.50
Human task cost				15.00

Table 5: Detailed use case for "Check training report" task

O - Occurrence, CL - Cognitive Level, S - Normalized Salary, SC - Step Cost

### 2.2.3 Check validation report

This task is performed by a ML engineer.

	ID	Depth	Neurons	Train MSE	Valid MSE	Delta MSE
1	954	3	4000	0.13	0.14	0.01
2	321	4	3000	0.23	0.24	0.01
3	5	3	1000	0.35	0.35	0.00
4	764	2	2000	0.24	0.45	0.21
5	202	3	2500	0.20	0.47	0.27
Reject						
Overfitting Tolerance:						0.10

Figure 5: "Check validation report" mock-up form

Step	O	CL	S	SC
1 <b>ACTOR</b> opens "Check validation report" form.	1	1	2.5	2.5
2 <b>SYSTEM</b> shows the best 5 models sorted by increasing Validation Loss.				
3 <b>FOR EACH</b> model in the list:	5			
3.1 <b>IF</b> the model Validation Loss minus the Training Loss is less than the Overfitting Tolerance and the Best Model is not selected.	1	2	2.5	5
3.1.1 <b>THEN</b> select the model as the Best Model.	1	1	2.5	2.5
4 <b>FOR EACH</b> model in the list:	4			
4.1 <b>IF</b> the model is not the Best Model and the Validation Loss minus the Training Loss is less than the Overfitting Tolerance and the Second Best Model is not selected.	1	2	2.5	5
4.1.1 <b>THEN</b> select the model as the Second Best Model.	1	1	2.5	2.5
5.1 <b>IF</b> the Best Model is not selected.	0.05	1	2.5	0.125
5.1.1 <b>ACTOR</b> clicks "Reject" button.	0.05	1	2.5	0.125
5.2 <b>ELSE IF</b> the Second Best Model is not selected or the Validation Loss of the Second Best Model is one order of magnitude greater than the Validation Loss of the Best Model.	0.3	3	2.5	2.25
5.2.1 <b>ACTOR</b> clicks on the Best Model.	0.3	1	2.5	0.75
5.3 <b>ELSE</b>	0.65	3	2.5	4.875
5.3.1 <b>ACTOR</b> clicks on the least complex model among the Best Model and the Second Best Model.	0.65	3	2.5	4.875
6 <b>SYSTEM</b> shows a confirmation dialog.				
7 <b>ACTOR</b> closes the form.	1	1	2.5	2.5
Human task cost				32.91

Table 6: Detailed use case for "Check validation report" task

O - Occurrence, CL - Cognitive Level, S - Normalized Salary, SC - Step Cost

## 2.2.4 Check test results

This task is performed by a ML engineer.

ID	Depth	Neurons	Valid MSE	Test MSE	Delta MSE
954	3	4000	0.14	0.15	0.01
Overfitting Tolerance:					0.10
Accept			Reject		

Figure 6: "Check test results" mock-up form



Step	O	CL	S	SC
1 <b>ACTOR</b> opens "Check test results" form.	1	1	2.5	2.5
2 <b>SYSTEM</b> shows the test results.				
3 <b>ACTOR</b> checks if the difference between the test results and the validation results is within overfitting tolerance.	1	2	2.5	5
4.1 <b>IF</b> the test results is not satisfactory.	0.01			
4.1.1 <b>ACTOR</b> clicks "Reject" button.	0.01	1	2.5	0.025
4.2 <b>ELSE</b>	0.99			
4.2.1 <b>ACTOR</b> clicks "Approve" button.	0.99	1	2.5	2.475
5 <b>SYSTEM</b> shows a confirmation dialog.				
6 <b>ACTOR</b> closes the form.	1	1	2.5	2.5
Human task cost				12.5

Table 7: Detailed use case for "Check test results" task

O - Occurrence, CL - Cognitive Level, S - Normalized Salary, SC - Step Cost

## 2.3 Evaluation system

### 2.3.1 Evaluate classifier performance

×

Evaluate Classifier Performance

Session ID	Expert Label	Classifier Label	Error
0	1	2	Yes
1	1	3	Yes
2	2	1	Yes
3	3	3	No

Max number of errors tolerated (th1): 4

Max number of consecutive error tolerated (th 2) :2

th1 satisfied

3 < 4

Pass

th 2 exceeded

3 > 2

Fail

Figure 7: "Evaluate Classifier Performance" mock-up form

Step	O	CL	S	SC
<b>1 ACTOR</b> opens the "Evaluate Classifier Performance" form.				
<b>2 SYSTEM</b> displays a table of sessions with Expert Label (ground truth) and Classifier Label (predicted label). The difference between the labels (if any) represents an error.				
<b>3 ACTOR</b> reviews the table.				
<b>3.1 IF</b> the total errors or consecutive errors exceed their respective thresholds:				
<b>3.1.1 ACTOR</b> clicks the "Fail" button.				
<b>3.2 ELSE</b>				
<b>3.2.1 ACTOR</b> clicks the "Pass" button.				
<b>4 SYSTEM</b> shows a confirmation dialog.				
<b>5 ACTOR</b> closes the form.				
Human task cost				

Table 8: Detailed use case for "Evaluate Classifier Performance" task

O - Occurrence, CL - Cognitive Level, S - Normalized Salary, SC - Step Cost