UNIVERSITÀ DI PISA

Department of Information Engineering

MSc in Computer Engineering

Software System Engineering Course Project

SMART APPLIANCE MANAGER

Students:

MAZZANTI CARLO

MOHAMMAD TESFAYE YIMAM

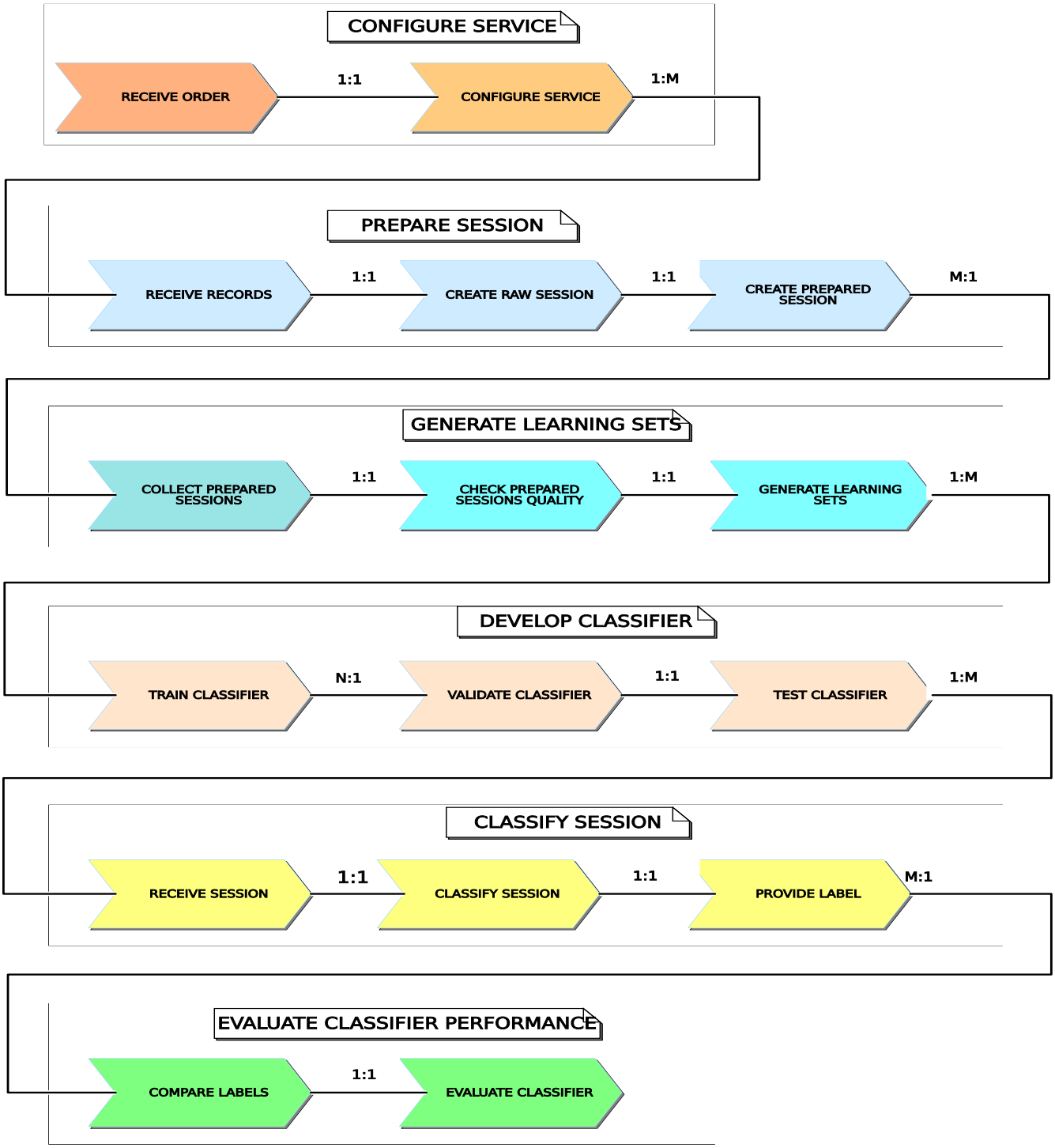
SIMONCINI ALESSIO

VEZZANI LORENZO

ACADEMIC YEAR 2025/26

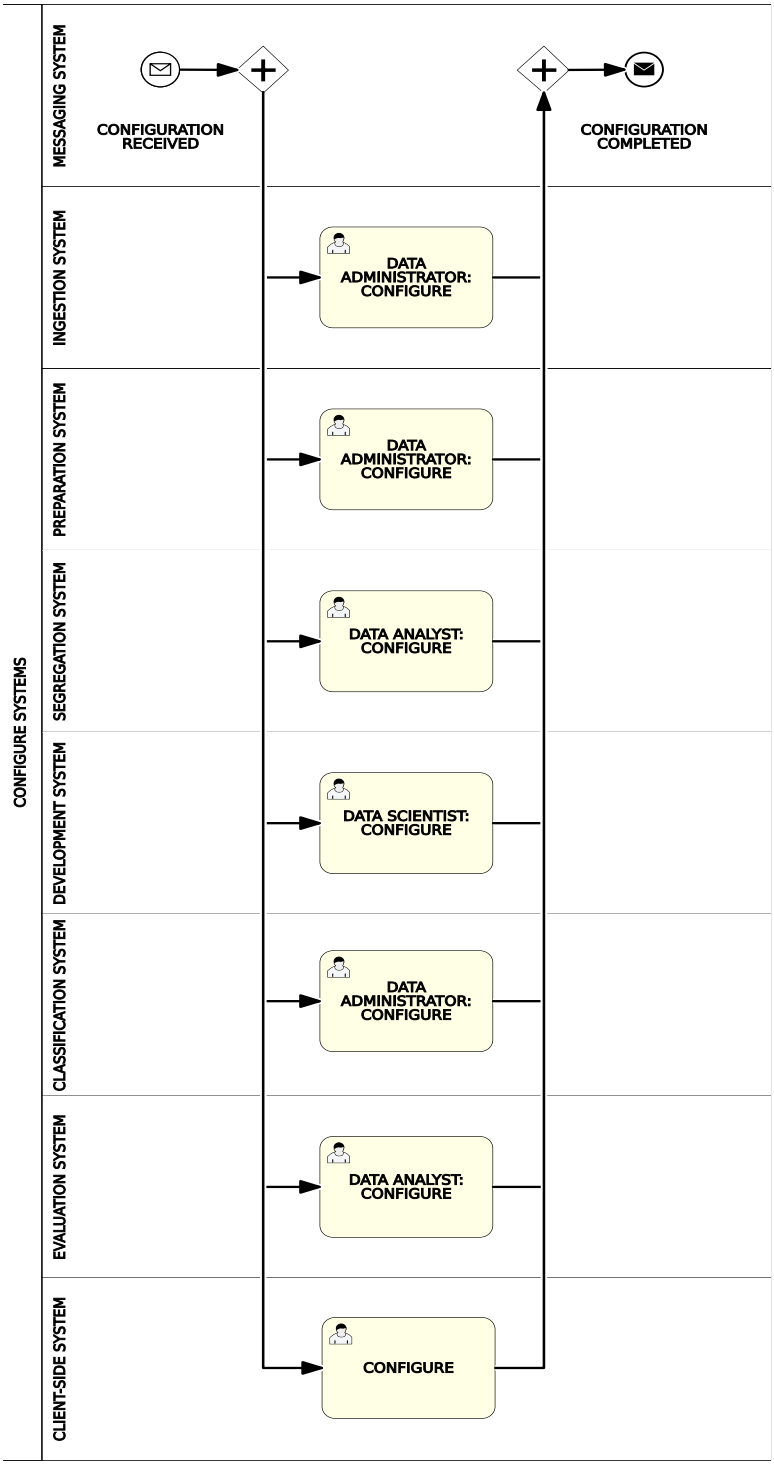
Table of Contents

Workflow Requirements: Process Landscape Diagram (Everyone)

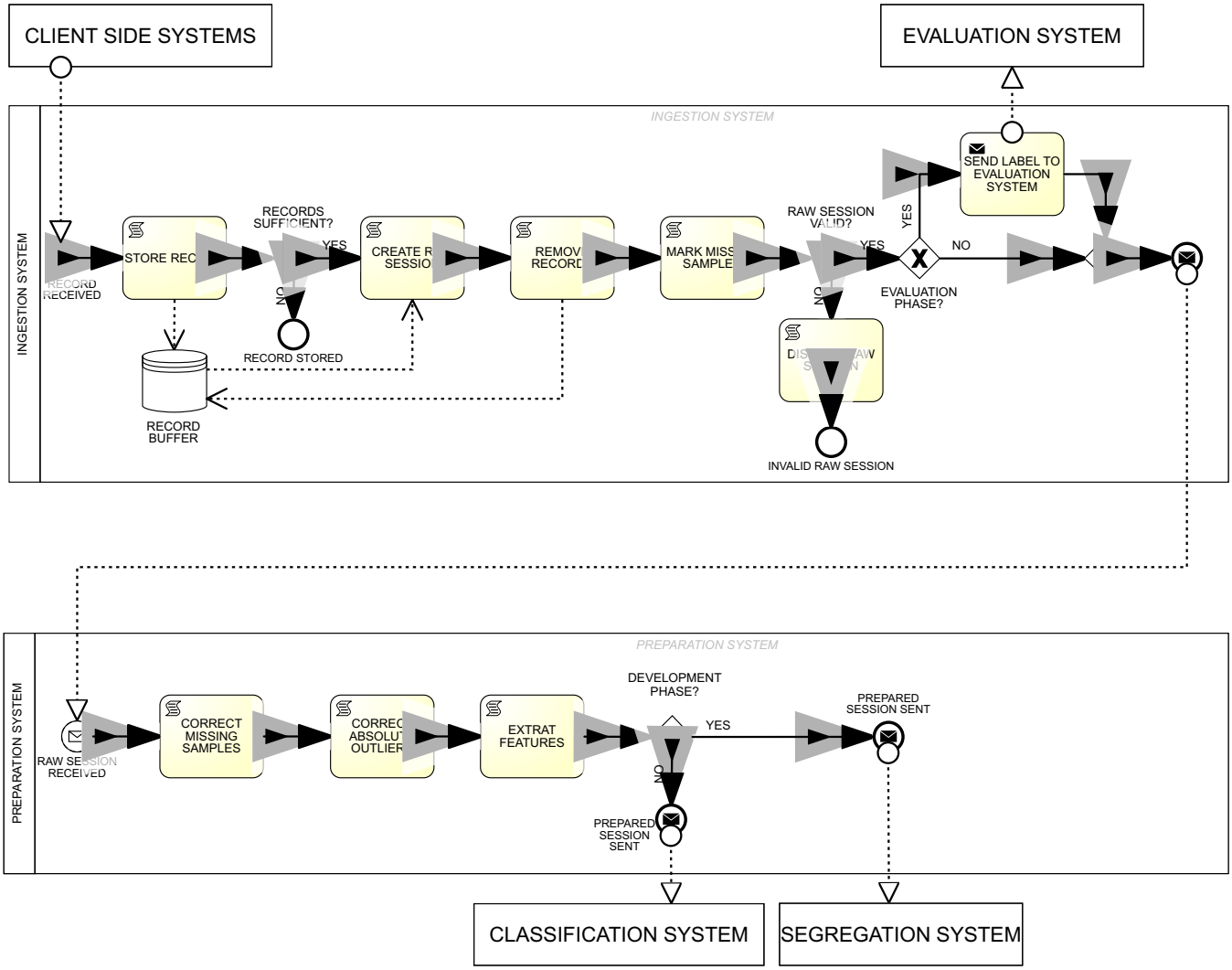


Workflow Requirements: BPMN Modelling

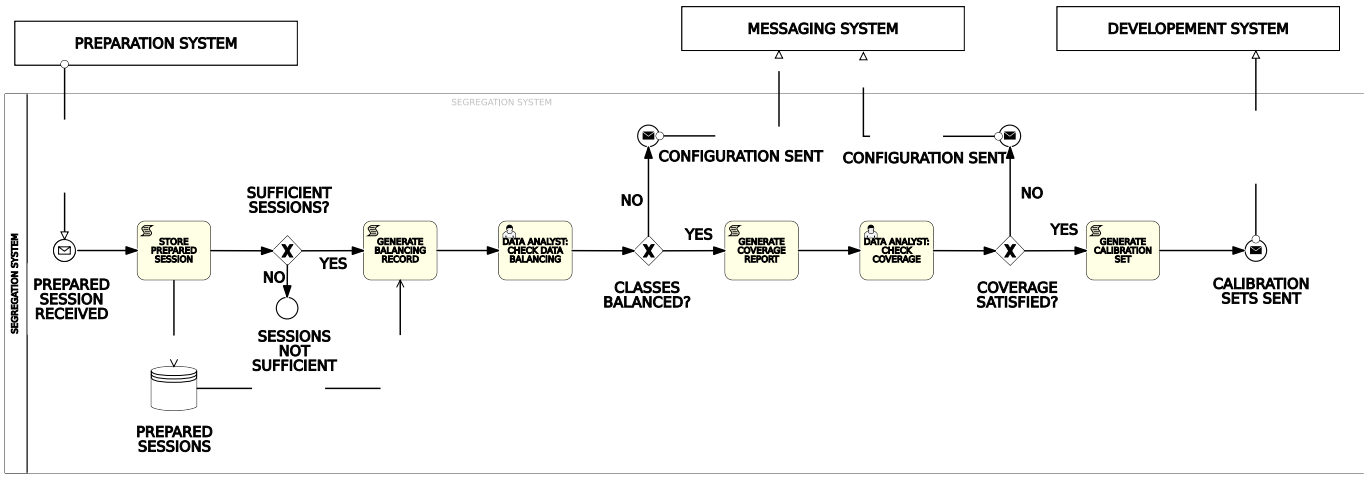
1 - CONFIGURATION PROCESS(Everyone)



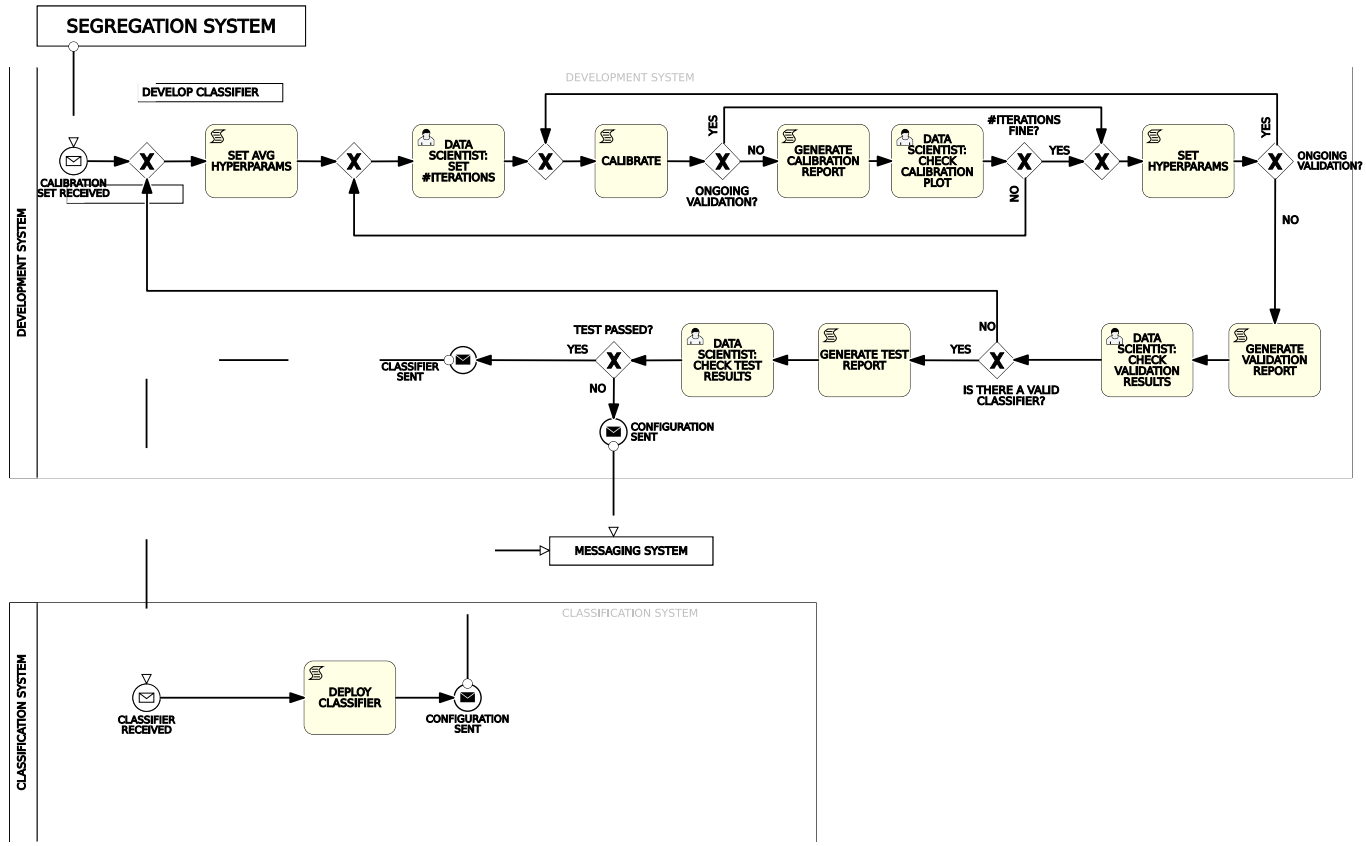
2 - PREPARE SESSION PROCESS (Carlo)



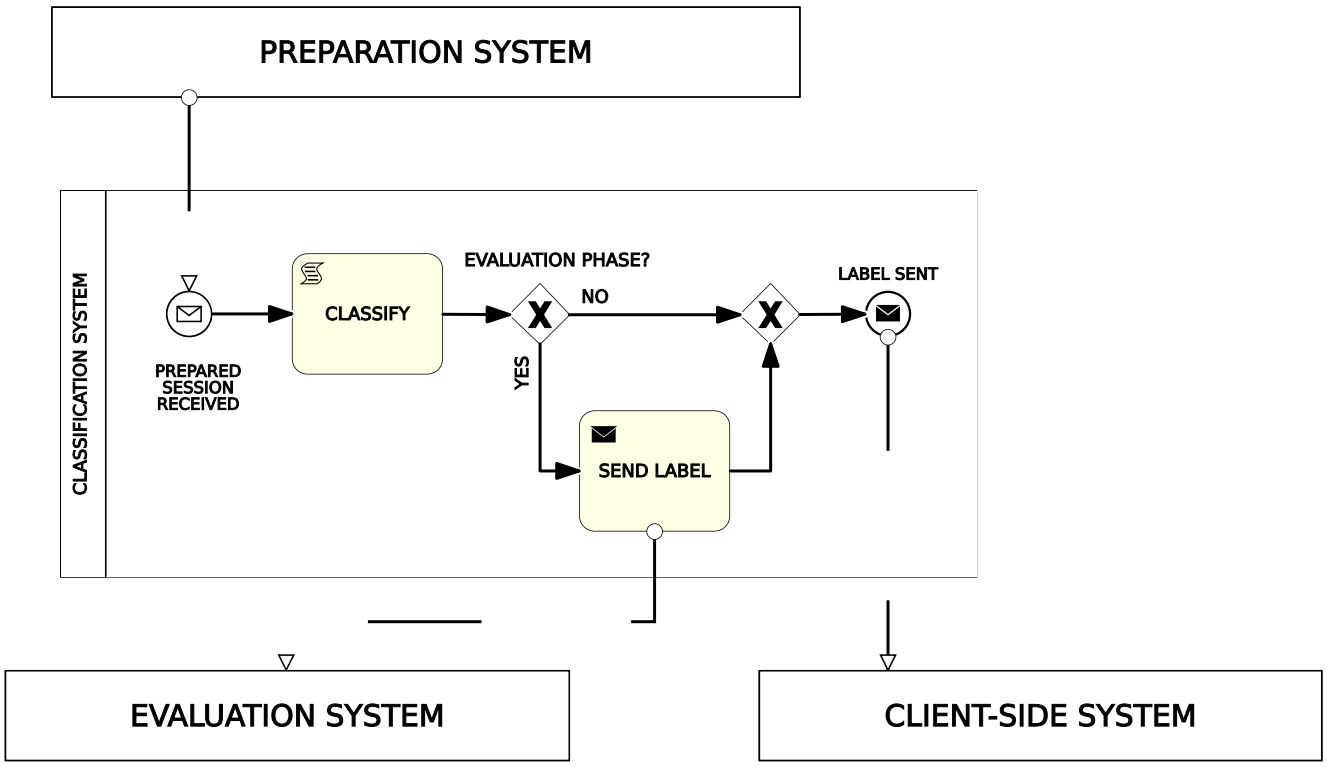
3 - GENERATE LEARNING-SET PROCESS (Lorenzo)



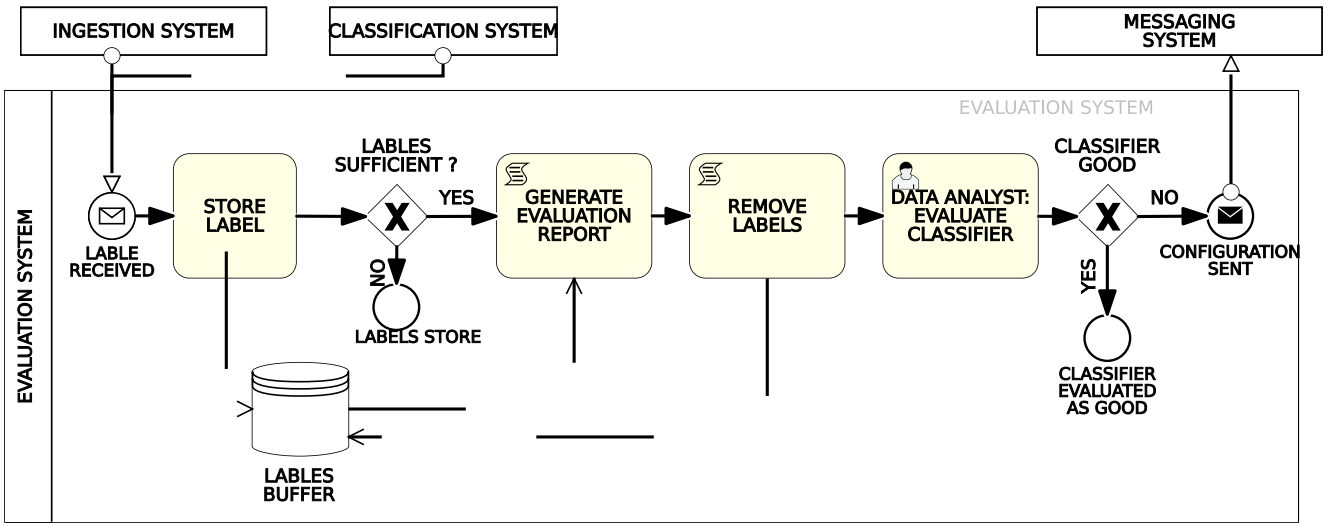
4 - DEVELOP CLASSIFIER PROCESS(Tesfaye)



5 - CLASSIFY SESSION PROCESS (Tesfaye)



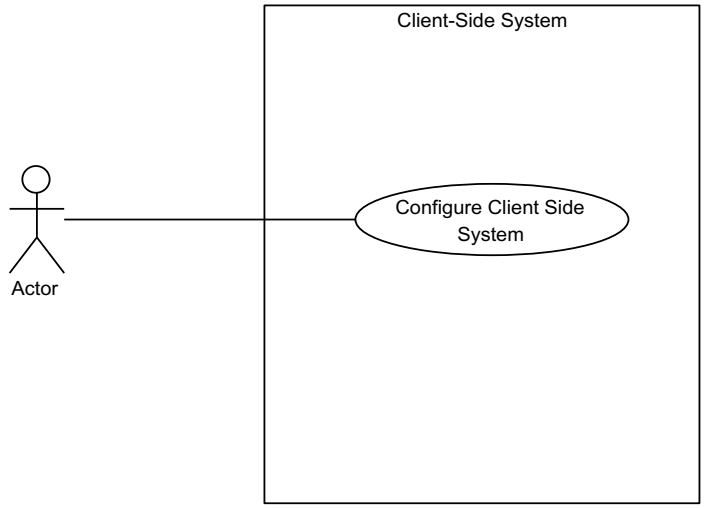
6 - EVALUATE CLASSIFIER PERFORMANCE PROCESS (Alessio)



Workflow ANALYSIS: Human Tasks and Use-Case Diagrams

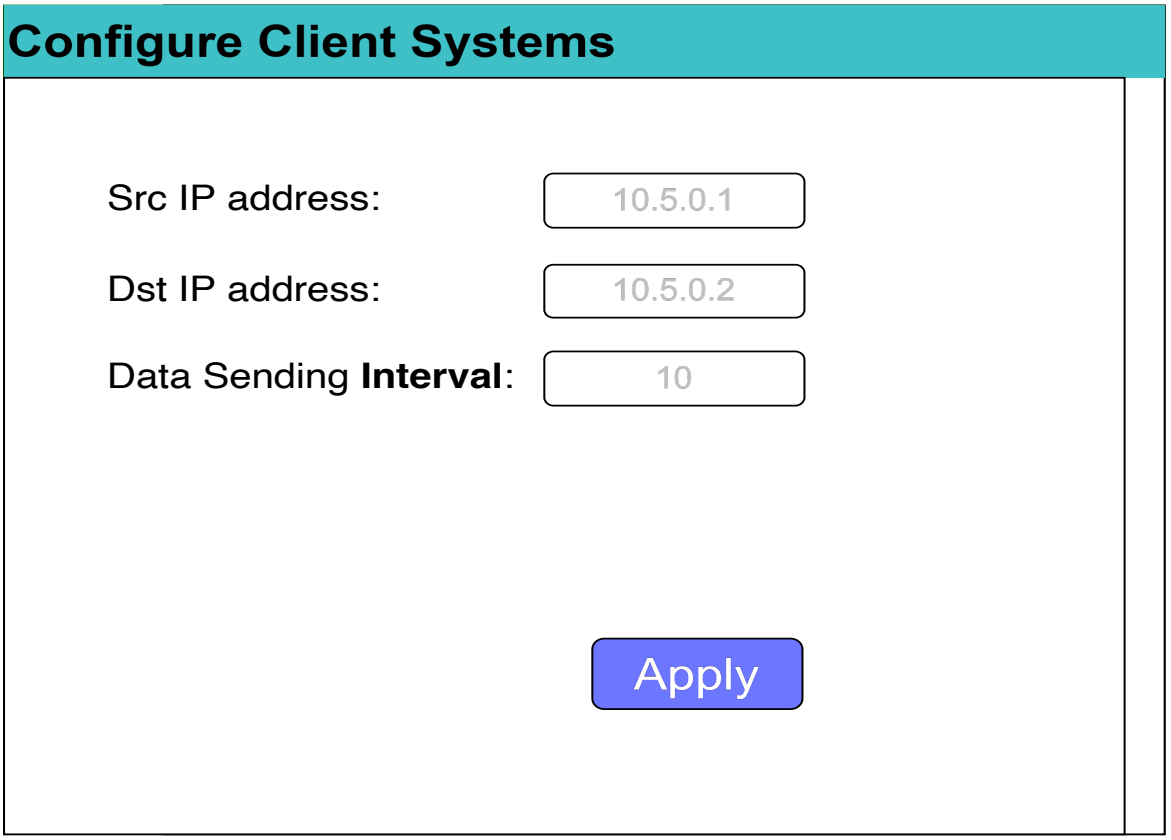
1. CLIENT-SIDE SYSTEM(Tesfaye)

Use-case Diagram



Human Task: Configure Client-Side System

Mock-up



Scenario

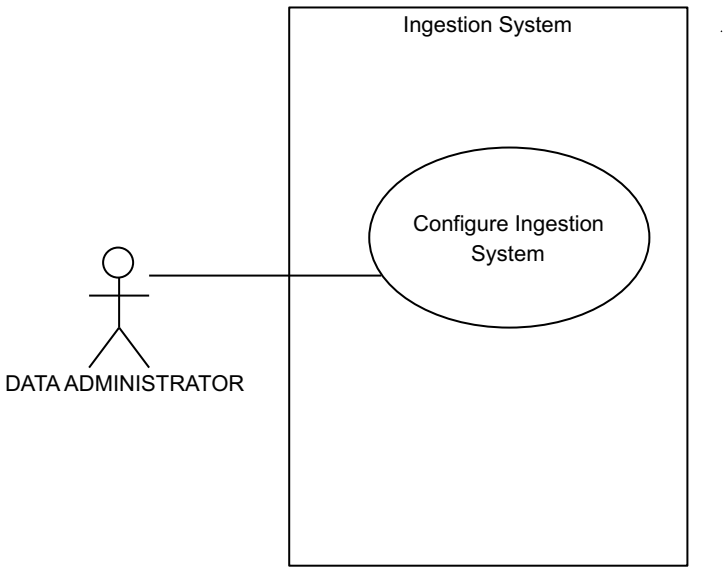
1.The use case starts when the System shows the **Client-Side Configuration Panel**
2. The Actor digits the Source IP address in the <Src IP address> input field
3. The Actor digits the Destination IP address in the <Dst IP address> input field.
4. The Actor digits the Data feeding Interval in the <Data Sending Interval> input field.
5. The Actor presses on the <Apply> button.

Details

Name	Value
Post-conditions	The client system is successfully configured.
	The configuration pop-up is closed.
	The system is now operating with the updated configuration.

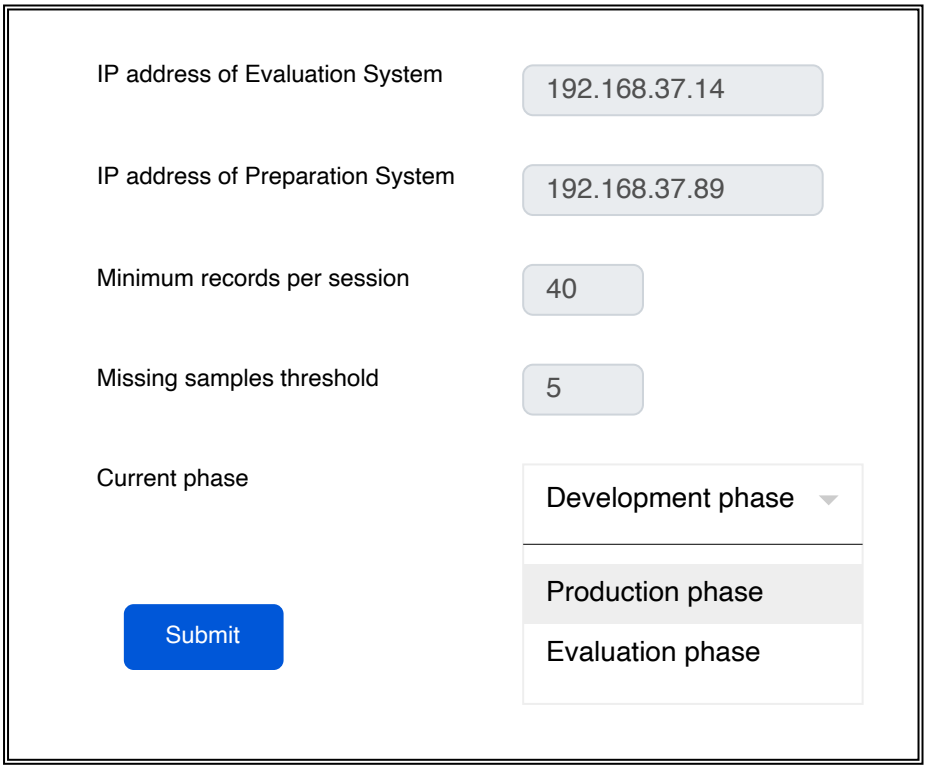
2. INGESTION SYSTEM (Carlo)

Use-case Diagram



Human task: Configure Ingestion System

Mock-up



Scenario

1. The use case starts when the DATA ADMINISTRATOR opens the configuration interface
2. SYSTEM shows inputs for IP address of Evaluation system, IP address of Ingestion system, configuring minimum records per session, missing samples threshold, current phase
3. DATA ADMINISTRATOR enters IP addresses of the Evaluation system and Ingestion system
4. DATA ADMINISTRATOR enters the minimum records per session
5. DATA ADMINISTRATOR enters the missing samples thresholds
6. if the current phase is Development phase
    6.1. DATA ADMINISTRATOR selects Development phase
7. else if the current phase is Production phase
    7.1. DATA ADMINISTRATOR selects Production phase
8. else if the current phase is Evaluation phase
    8.1. DATA ADMINISTRATOR selects Evaluation phase

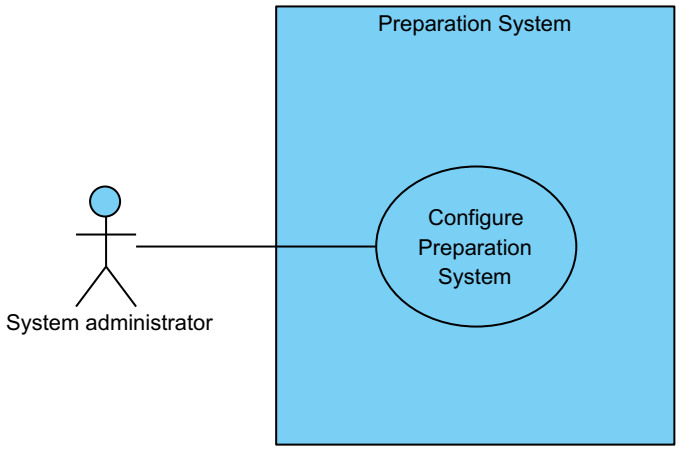
   end if

Details

Post-conditions	The Ingestion System is configured

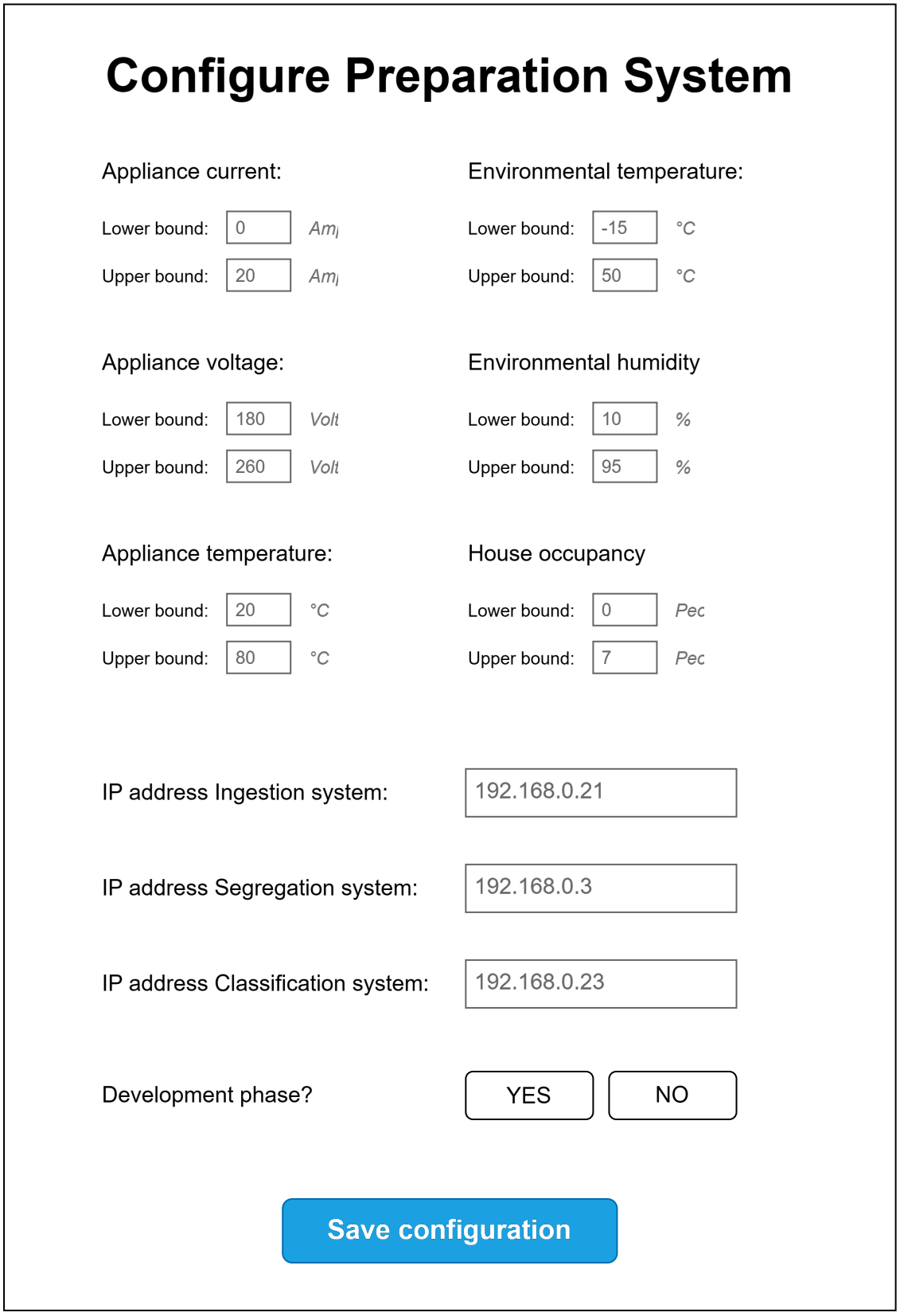
3. PREPARATION SYSTEM(Alessio)

Use-case Diagram



Human Task: Configure Preparation System

Mock-up



Scenario

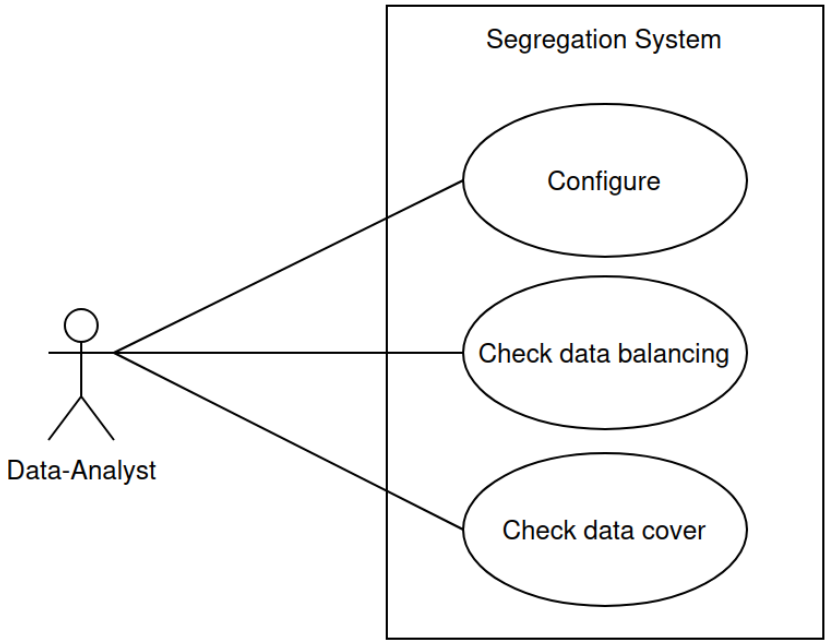
1.Use case starts when the System administrator opens the **Configure Peparation System** interface
2. SYSTEM Displays all the parameters and fields to be set
3.The System administrator sets the *Appliance current* lower bound
4.The System administrator sets the *Appliance current* upper bound
5.The System administrator sets the *Appliance voltage* lower bound
6.The System administrator sets the *Appliance voltage* upper bound
7.The System administrator sets the *Appliance temperature* lower bound
8.The System administrator sets the *Appliance temperature* upper bound
9.The System administrator sets the *Environmental temperature* lower bound
10.The System administrator sets the *Environmental temperature upper bound*
11.The System administrator sets the *Environmental humidity* lower bound
12.The System administrator sets the *Environmental humidity upper bound*
13.The System administrator sets the *House occupancy* lower bound
14.The System administrator sets the *House occupancy* upper bound
15.The System administrator sets the *'IP address Ingestion system'* field
16.The System administrator sets the *'IP address Segregation system'* field
17.The System administrator sets the *'IP address Classification system'* field
18.The System administrator sets the configuration by pressing the **YES** button or the **NO** button
19.The System administrator saves the configuration by pressing the **Save configuration** button

Details

Name	Value
Post-conditions	The phase of the preparation system is set The outlier parameters are set The connected systems' IP addresses are set

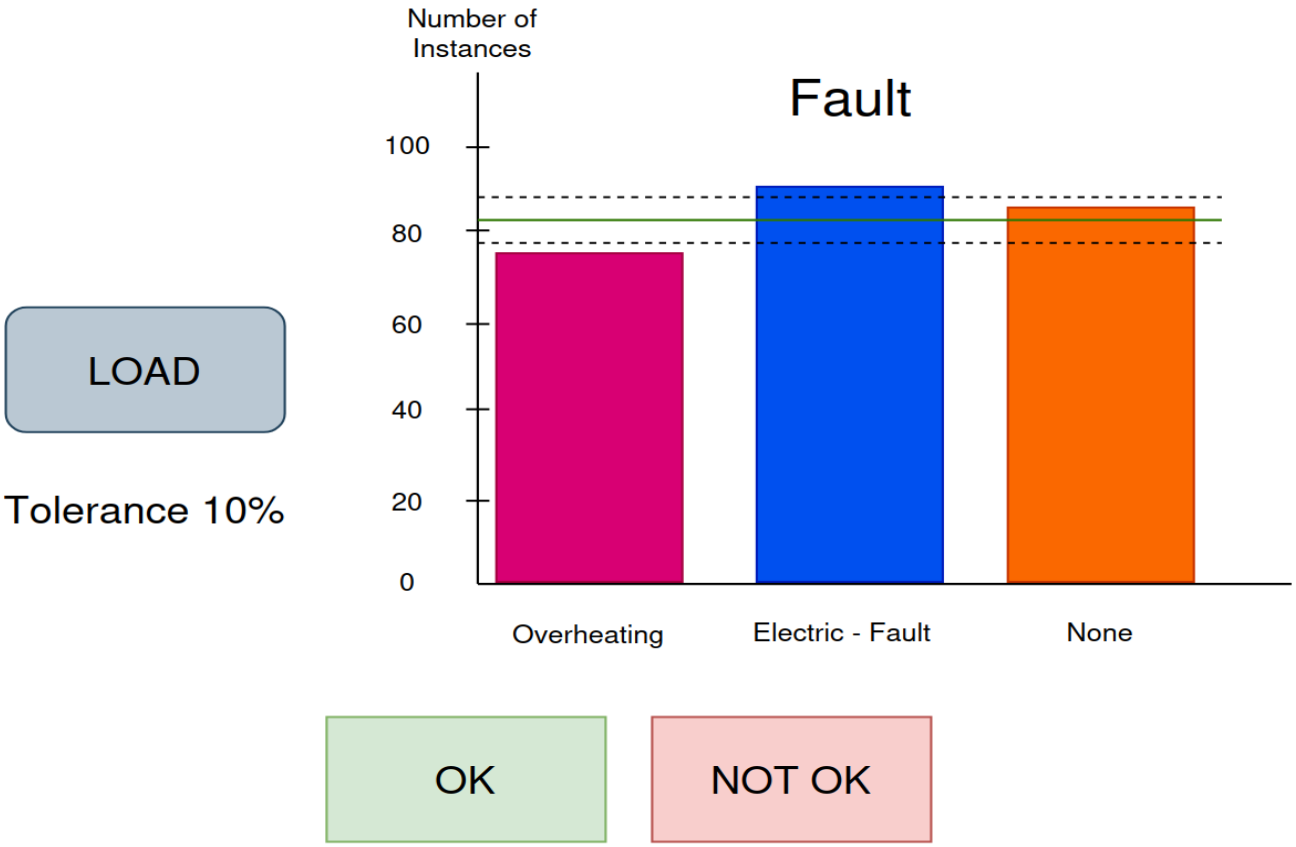
4. SEGREGATION SYSTEM (Lorenzo)

Use-case Diagram



Human Task: Check Data Balancing

Mock-up

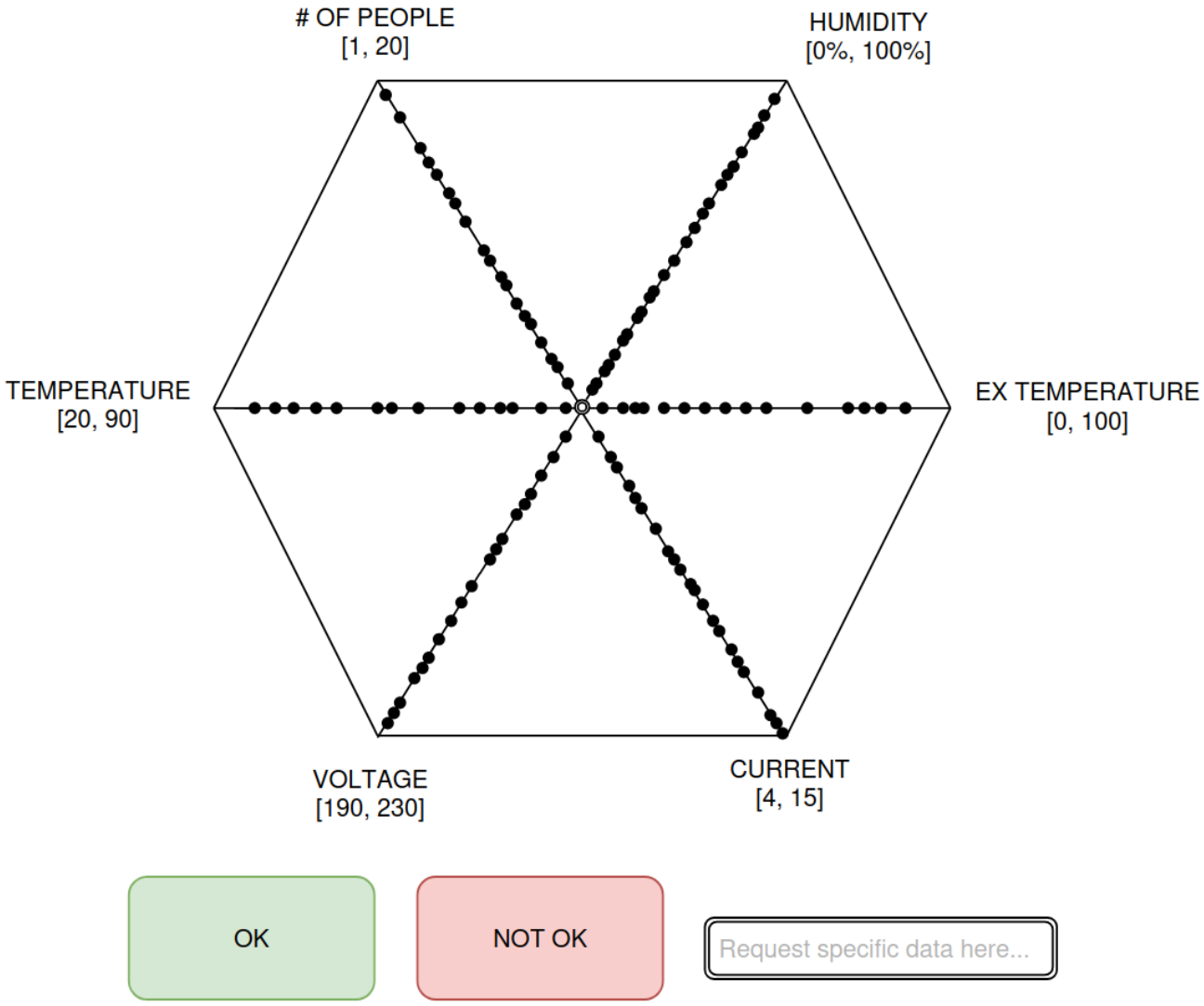


Scenario

1.The Data Analyst launches the software
2. SYSTEM Displays GUI for the Data Analyst
3.The Data Analyst presses the LOAD button
4. for each class
4.1. if the bar is not inside the interval
4.1.2. The Data Analyst selects NOT OK
4.1.3. The use case terminates.
end if
end for
5.The Data Analyst selects OK.

Human Task: Check Data Coverage

Mock-up

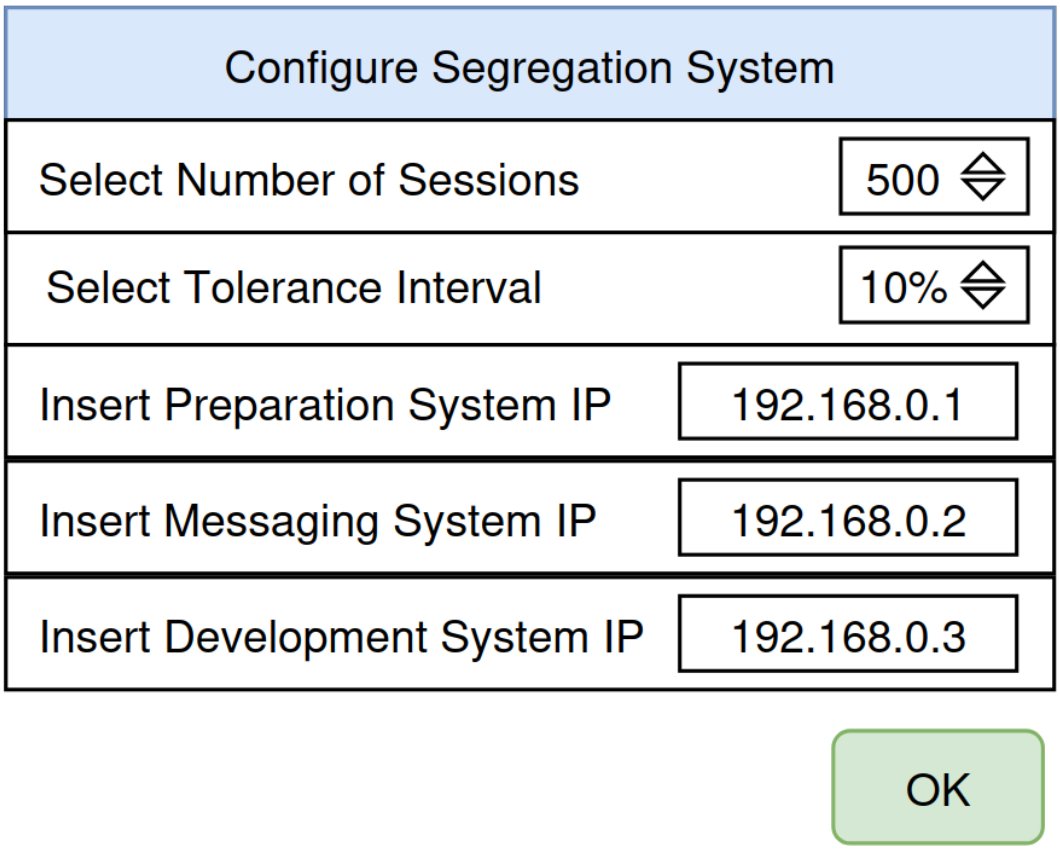


Scenario

1.The Data Analyst launches the software
2. SYSTEM Displays GUI for the Data Analyst
3. The Data Analyst presses the LOAD button
4. if the radar scatterplot is not sufficiently well distributed according to Data Analyst
4.1 The Data Analyst requires for specific data using the textbox
4.2. The Data Analyst selects "NOT OK"
4.3. The use case terminates.
end if
5. The Data Analyst selects "OK".

Human Task: Configure the Segregation System

Mock-up

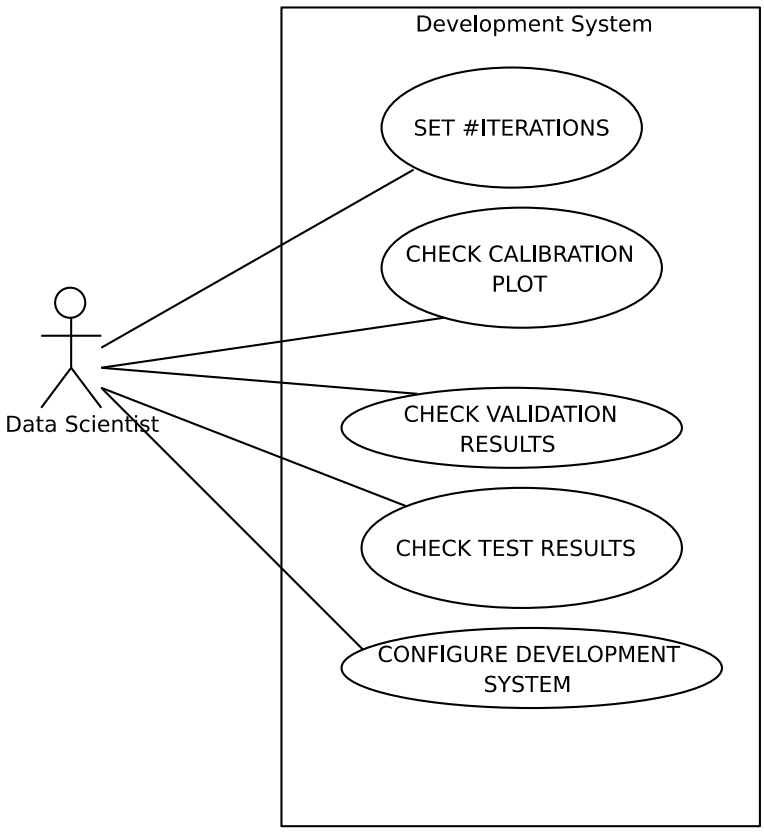


Scenario

1. The Data Analyst launches the software.
2. The Data Analyst selects the number of sessions.
3. The Data Analyst selects a tolerance percentage.
4. The Data Analyst sets Preparation System IP address.
5. The Data Analyst sets Messaging System IP address.
6. The Data Analyst sets Development System IP address.
7. The Data Analyst selets OK.

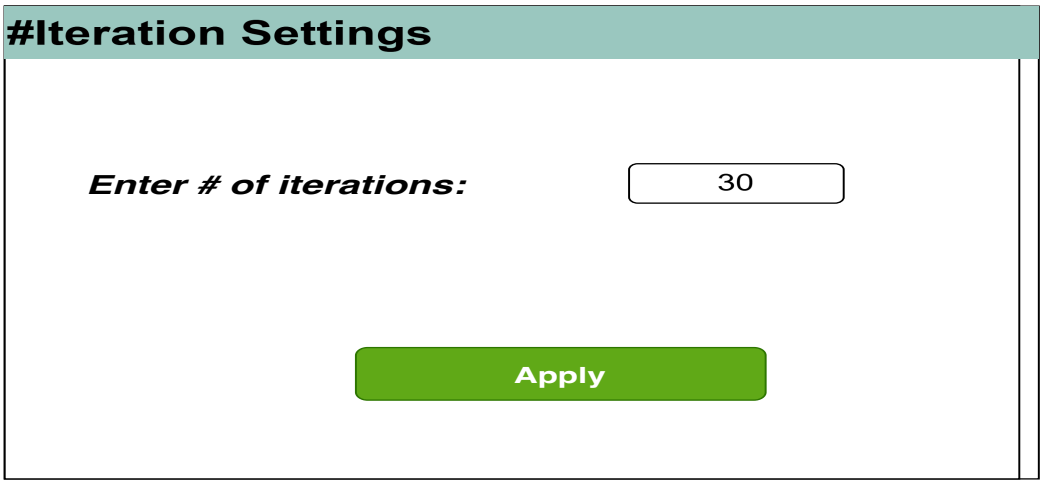
5. DEVELOPMENT SYSTEM (Tesfaye, Carlo)

Use-case Diagram



Human Task: Set #Iterations

Mock-up



Scenario

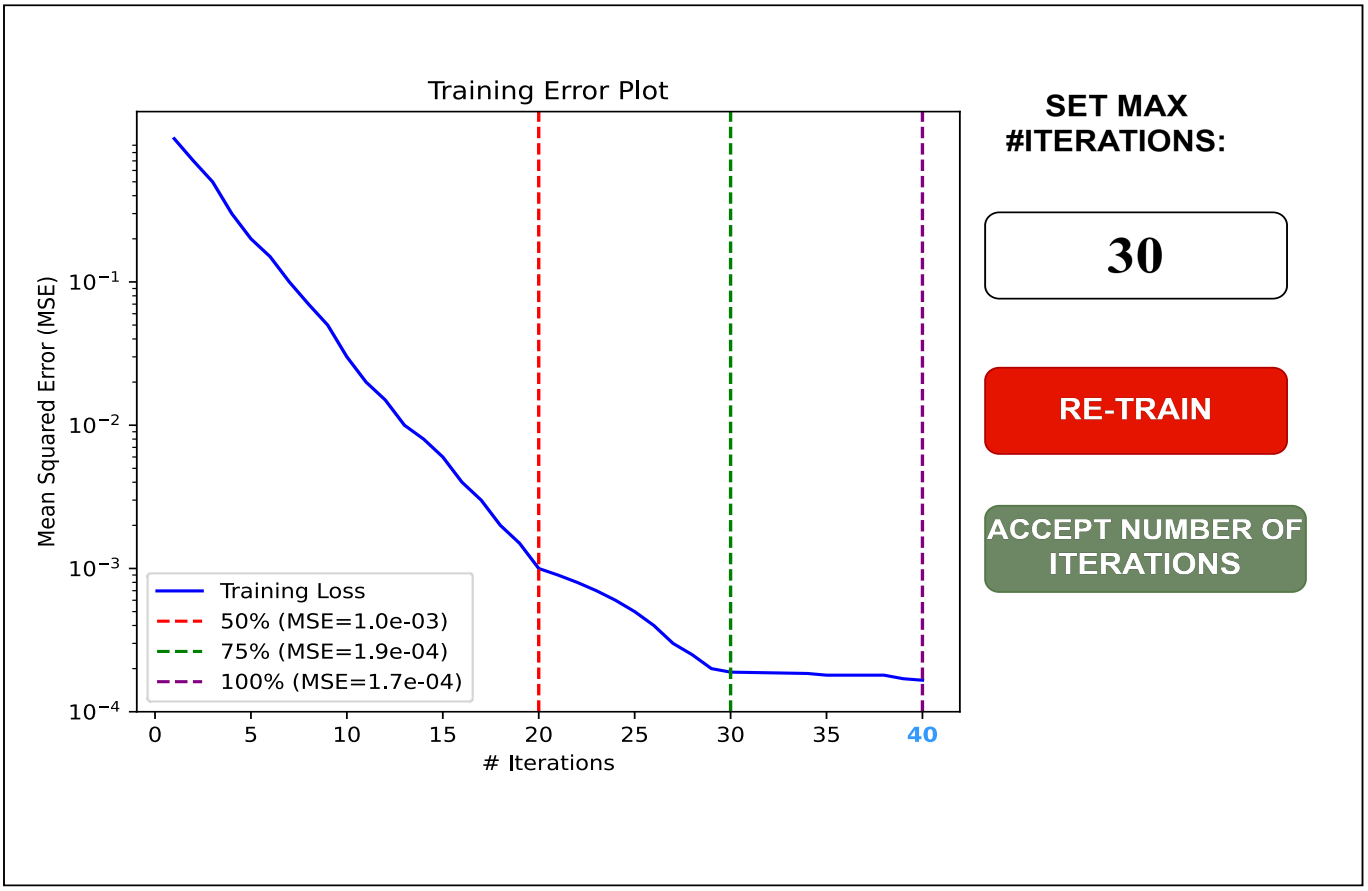
1.The use case starts when the Data Scientist opens the UI to set the # of iterations.
2.The Data Scientist inserts the #iterations from the keyboard in the <**Enter # of iterations**> field.
3.The Data Scientist presses the <**Apply**> button.

Details

Name	Value
Post-conditions	The development system updates the number of calibration/learning iterations.

Human Task: Check Calibration Plot

Mock-up



Scenario

1. The use case starts when the Data Scientist opens on <**Check Calibration Plot**>
2.SYSTEM shows a pop-up window with the Mean Squared Error (MSE) plot against the number of iterations.
3. if the Data Scientist observes that the MSE curve becomes almost flat well before the last iterations then
3.1. The Data Scientist reduces the number of iterations via the provided field to manage overfitting.
3.2. The Data Scientist press the <**RE-TRAIN**> button
3.3. SYSTEM updates the plot with the new number of iterations
4. else if the Data Scientist sees that the MSE curve is still clearly decreasing near the final iterations
4.1. The Data Scientist increases the number of iterations based on the provided fields and the button on the pop-up window.
4.2. The Data Scientist press the <**RE-TRAIN**> button
4.3. SYSTEM updates the plot with the new number of iterations
5. else if the Data Scientist considers the current number of iterations are adequate
5.1. The Data Scientist presses the <**ACCEPT NUMBER OF ITERATIONS**> button.
6. end if

Details

Name	Value
Post-conditions	**Alternative Flow 1:** 1.The Training Error Report has been analysed. 2.The system continued with its execution by setting the hyperparameters according to the grid search.
	Alternative Flow 2: 1.The number of **iterations** has been tuned. 2.A new neural network training has started.

Human Task: Check Validation Results

Mock-up

Select the best network

	Validation Results						
	ID	**Validation Error**	**Training Error**	**Difference**	**Number of Layers**	**Number of Neurons**	**Complexity**
☑	1	0.12	0.10	0.02	7	20	140
☐	2	0.15	0.12	0.03	3	40	120
☐	3	0.20	0.15	0.05	7	30	210
☐	4	0.30	0.20	0.10	8	50	400
☐	5	0.40	0.21	0.19	8	100	800

OVERFITTING TOLERANCE: 0.05 | CONFIRM | REFUSE ALL NETWORKS

Scenario

1. the use case starts when SYSTEM visualizes the test results for the top 5 classifiers
2. for each classifier in the report
   2.1. SYSTEM shows network ID, Validation Error, Training Error, Difference, Number of Layers, Number of Neurons, Complexity

   end for each
3. SYSTEM shows Overfitting Tolerance
4. for each classifier in the report
   4.1. DATA SCIENTIST checks the difference

   end for each
5. if DATA SCIENTIST sees no classifiers with difference lower than the overfitting tolerance
   5.1. DATA SCIENTIST clicks on "REFUSE ALL NETWORKS" button
6. else
   6.1. if there are multiple networks with difference lower than the overfitting tolerance
      6.1.1. DATA SCIENTIST selects the network with lowest difference

   end if
   6.2. if there are multiple networks with similar difference w.r.t. the lowest difference
      6.2.1. DATA SCIENTIST selects the network with the lowest complexity

   end if
   6.3. DATA SCIENTIST clicks on "CONFIRM" button

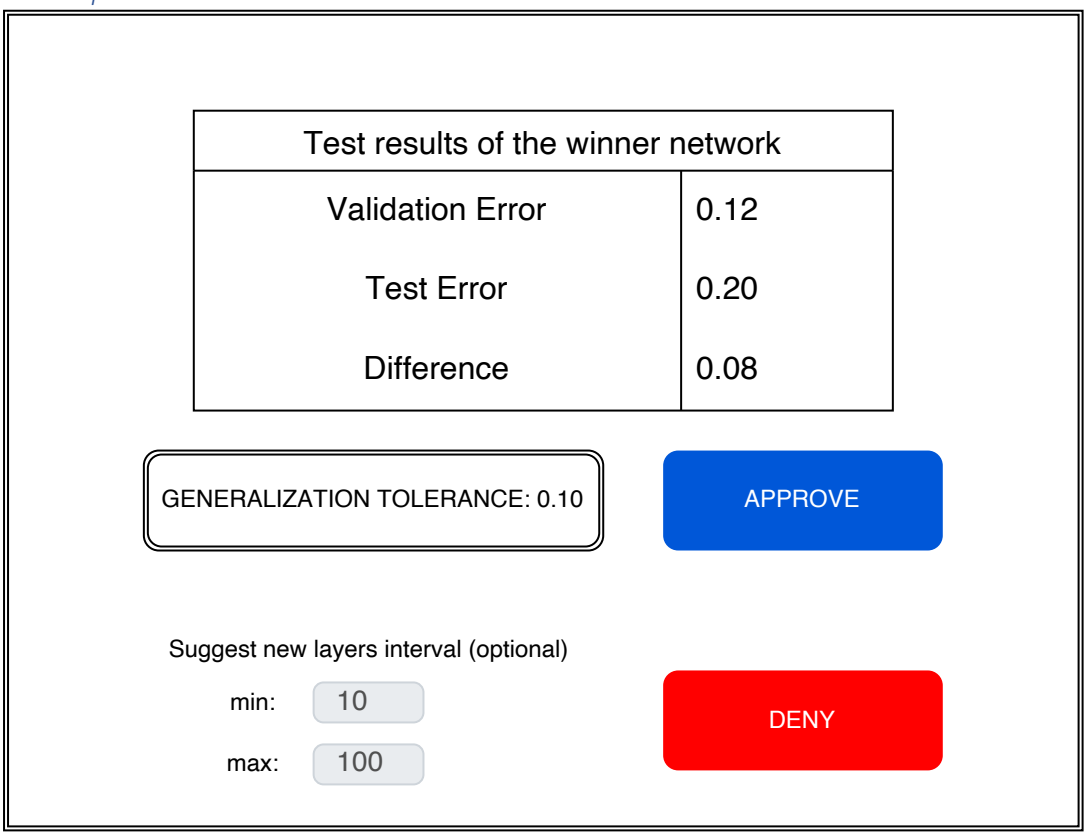
end if

Details

Preconditions	The validation report has been generated
Post-conditions	the best classifier is selected or a new training is started

Human Task: Check Test Results

Mock-up



Scenario

1. the use case starts when the DATA SCIENTIST opens the test results
2. SYSTEM shows validation error, test error and difference of the winner network
3. SYSTEM shows the generalization tolerance
4. DATA SCIENTIST checks the results
5. if DATA SCIENTIST sees difference lower than the generalization tolerance
    5.1. DATA SCIENTIST clicks on the "APPROVE" button
6. else
    6.1. DATA SCIENTIST optionally enters a suggestion on the new interval of number of layers
    6.2. DATA SCIENTIST clicks on the "DENY" button

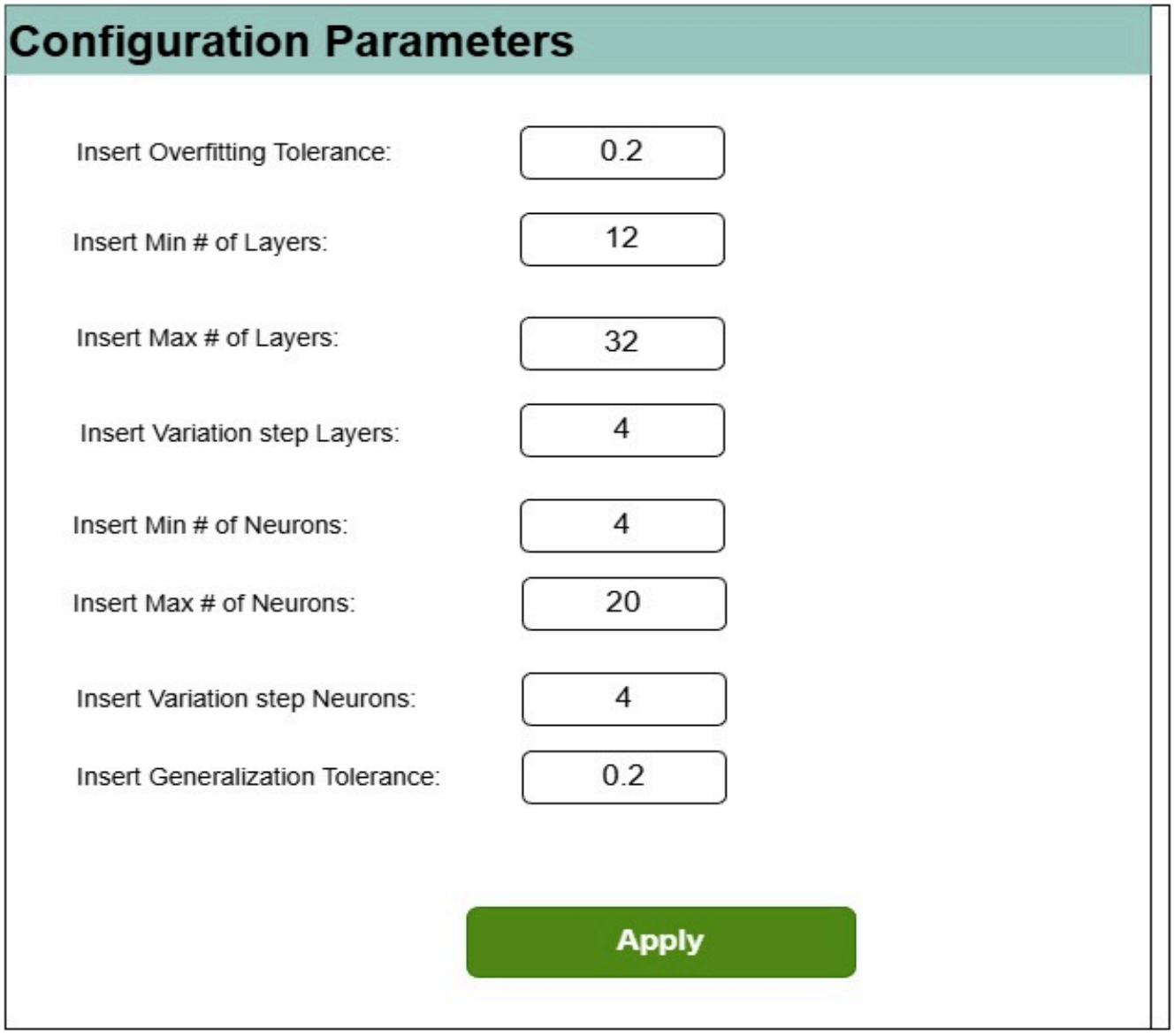
end if

Details

Preconditions	The test report has been generated
Post-conditions	the test result is approved or not

Human Task: Configure Development System

Mock-up



Scenario

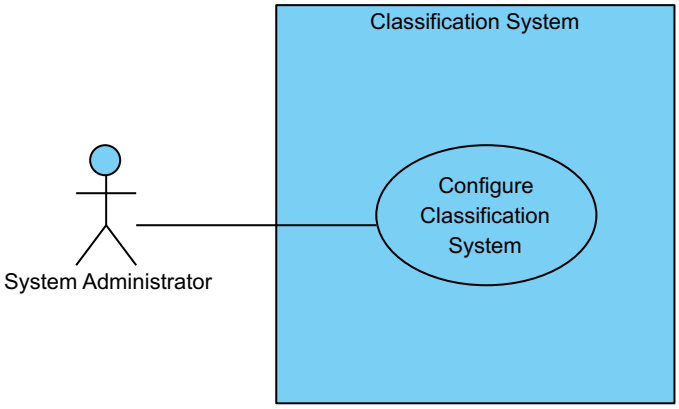
1.The use case starts when the Data Scientist opens the interface to configure the development system.
2.SYSTEM shows Development System configuration parameter interface.
3.The Data Scientist inserts from the keyboard the Overfitting tolerance in the <Insert Overfitting tolerance> field.
4.The Data Scientist digits the Min # of Layers in the <Insert Min # of Layers> field.
5.The Data Scientist digits the Max # of Layers in the <Insert Max # of Layers> field.
6.The Data Scientist digits the Variation step Layers in the “Insert Variation step Layers” field.
7.The Data Scientist digits the Min # of Neurons in the “Insert Min # of Neurons” field.
8.The Data Scientist digits the Max # of Neurons in the “Insert Max # of Neurons” field.
9.The Data Scientist digits the Variation step Neurons in the “Insert Variation step Neurons” field.
10.The Data Scientist digits the Generalization tolerance in the “Select Generalization tolerance” field.
11.The Data Scientist presses the "Apply" button.

Details

Name	Value
Post-conditions	The configuration of the development system has been updated.

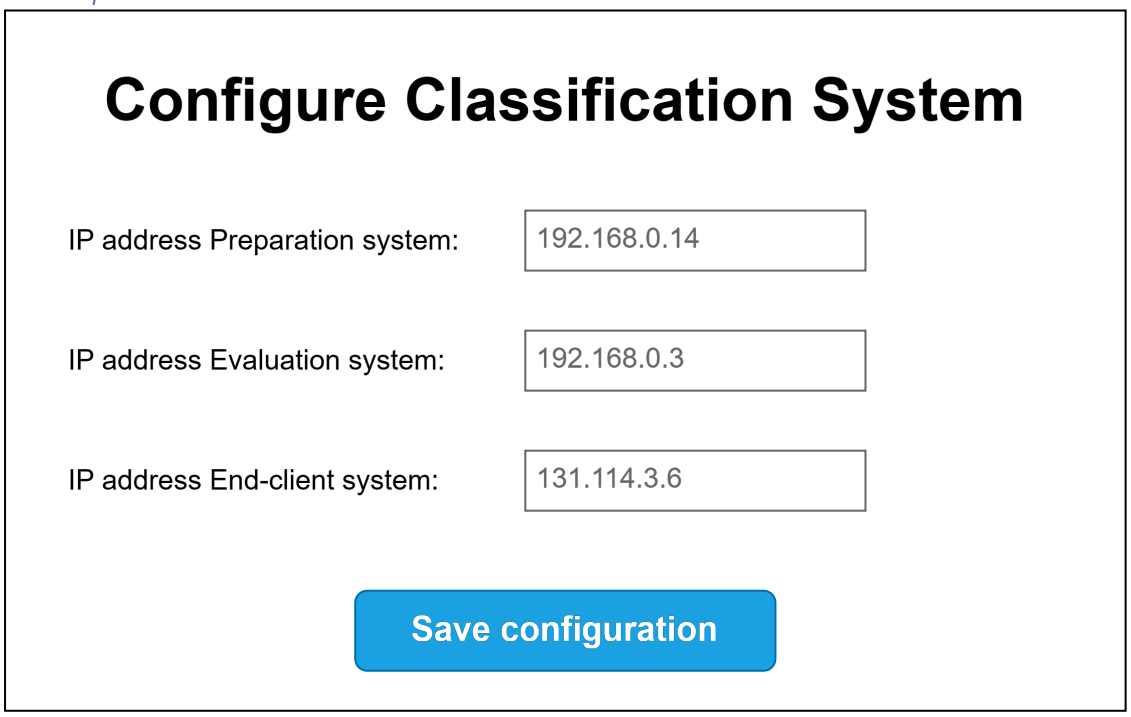
6. CLASSIFICATION SYSTEM(Alessio)

Use-case Diagram



Human Task: Configure Classification System

Mock-up



Scenario

1.Use case starts when the System Administrator opens the **Configure Classification System** interface
2. SYSTEM Displays the interface with the IP addresses to be set
3.The System Administrator sets the *'IP address Preparation system'* field
4.The System Administrator sets the *'IP address Evaluation system'* field
5.The System Administrator sets the *'IP address End-Client system'* field

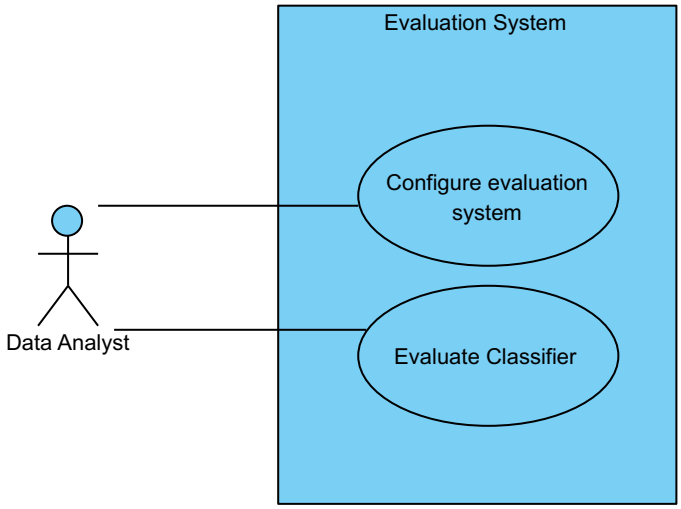
6.The System Administrator saves the configuration by pressing the **Save configuration** button

Details

Name	Value
Post-conditions	The phase of the classification system is set The connected systems' IP addresses are set

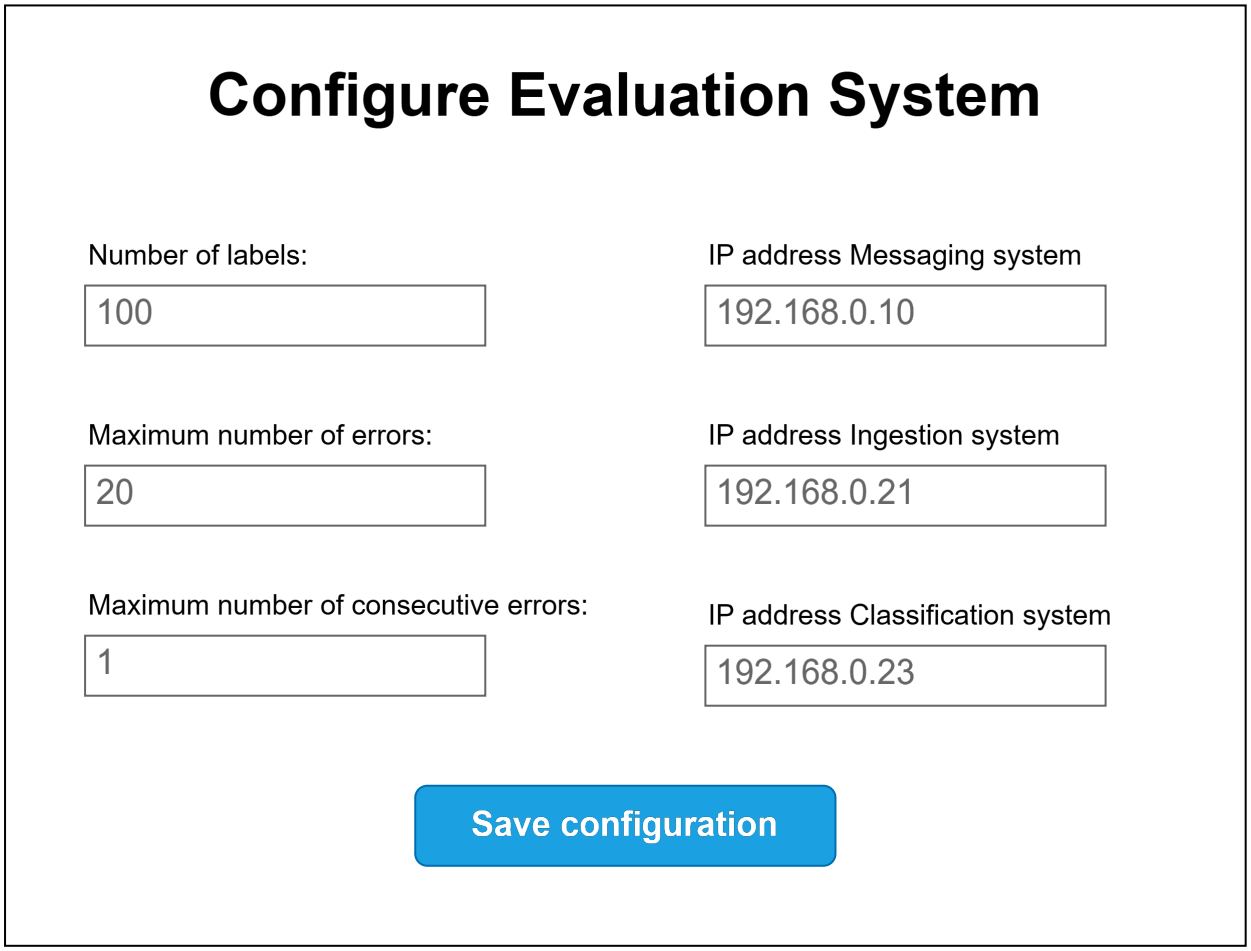
7. EVALUATION SYSTEM (Alessio)

Use-case Diagram



Human Taks: Configure Evaluation System

Mock-up



Scenario

1.Use case starts when the Data Analyst opens the **Configure Evaluation System** interface
2. SYSTEM Displays all the parameters to be set and a writable textbox for each one
3.The Data Analyst sets the *'Number of labels'* parameter
4.The Data Analyst sets the *'Maximum number of errors'* parameter
5.The Data Analyst sets the *'Maximum number of consecutive errors'* parameter
6.The Data Analyst sets the *'IP address Messaging system'* field
7.The Data Analyst sets the *'IP address Ingestion system'* field
8.The Data Analyst sets the *'IP address Classification system'* field
9.The Data Analyst saves the configuration by pressing the **Save configuration** button

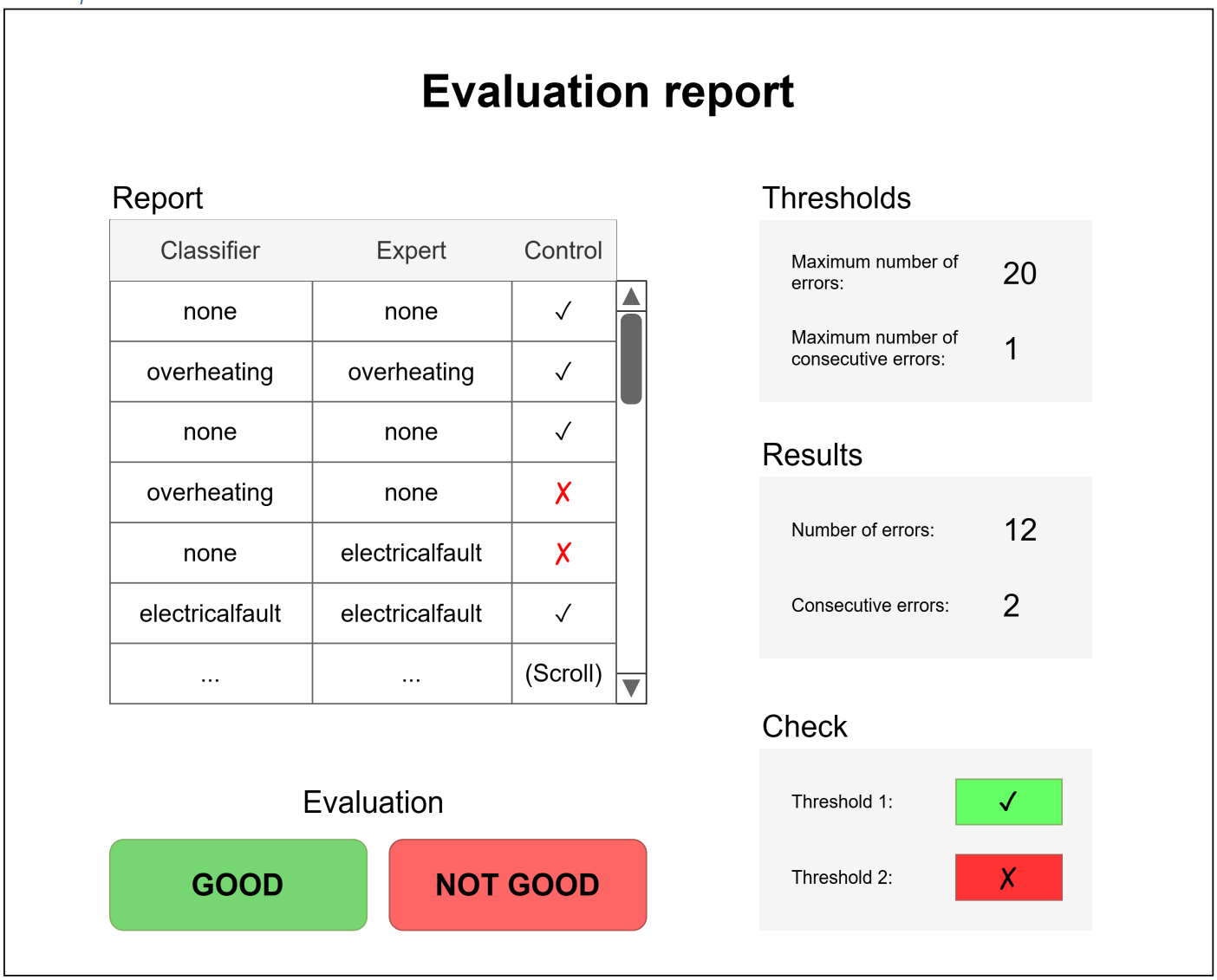
Details

Name	Value
Post-conditions	The Evaluation System is configured

	The connected systems' IP addresses are set

Human Task: Evaluate Classifier

Mock-up



Scenario

1.Use case starts when the Data Analyst opens the **Evaluate Classifier** interface
2. SYSTEM Displays the interface with all the relative information
3. if The Data Analyst sees that the threshold 1 (*Number of errors*) is respected:
3.1. if The Data Analyst sees that the threshold 2 (*Consecutive errors*) is respected:
3.1.1.The Data Analyst selects the **GOOD** button
3.2. else
3.2.1.The Data Analyst selects the **NOT GOOD** button

end if
4. else
4.1. The Data Analyst selects the **NOT GOOD** button
end if

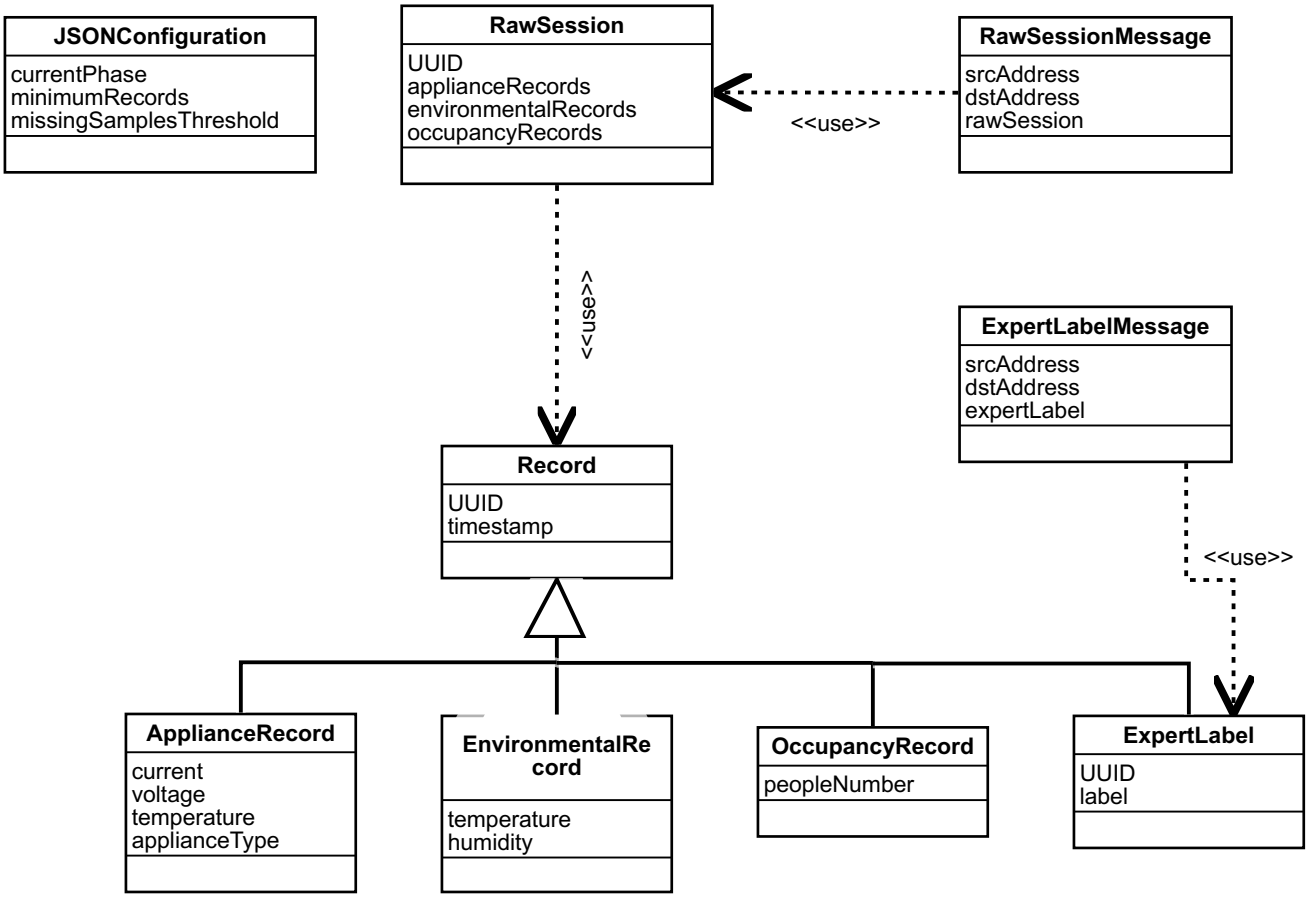
Details

Name	Value
Preconditions	A sufficient number of labels (specified in the configuration) has been collected
Post-conditions	The classifier is evaluated

Workflow ANALYSIS: Data and Application Logic Models

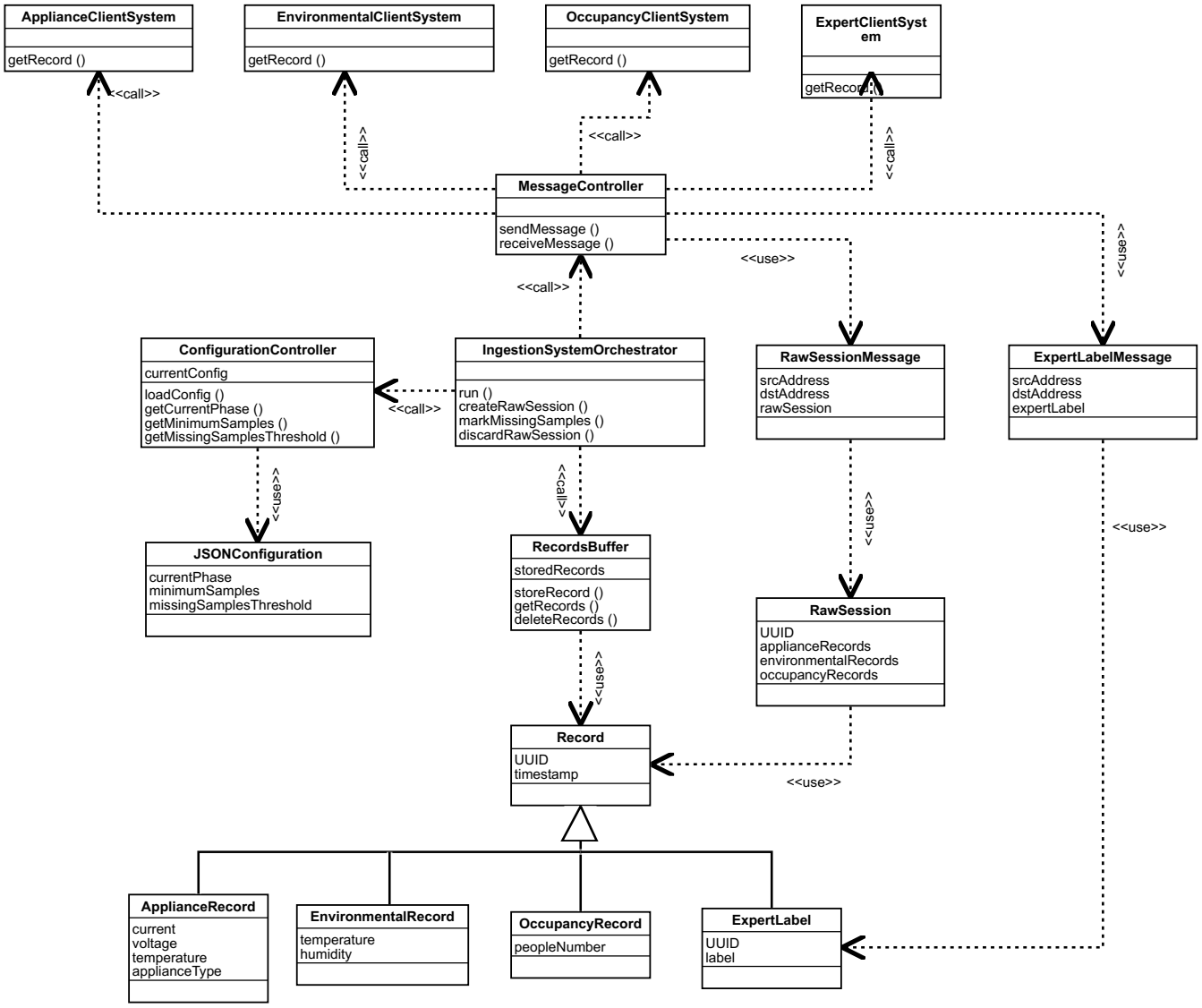
1. INGESTION SYSTEM (Carlo)

Data model

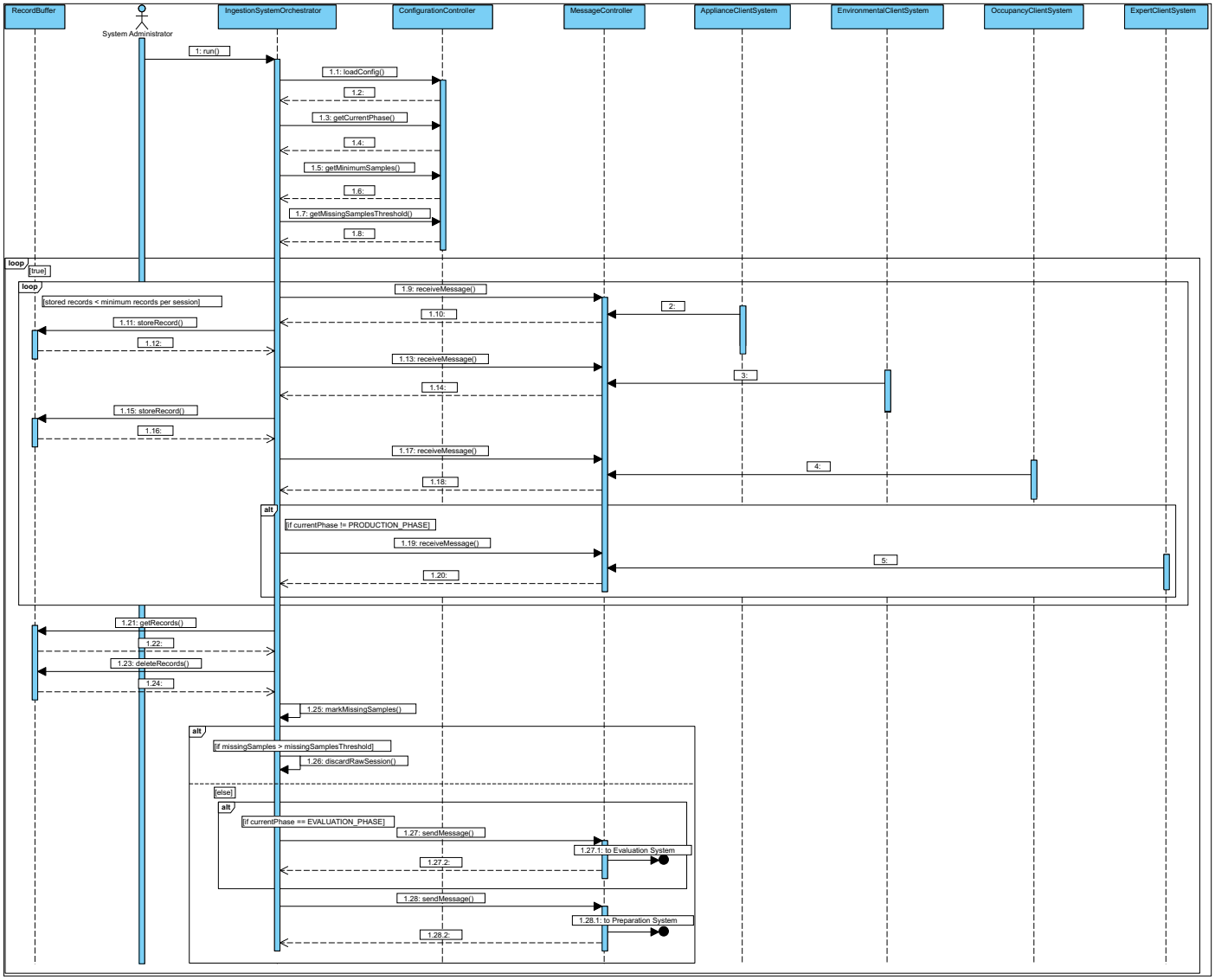


Application logic model

UML Class Diagram

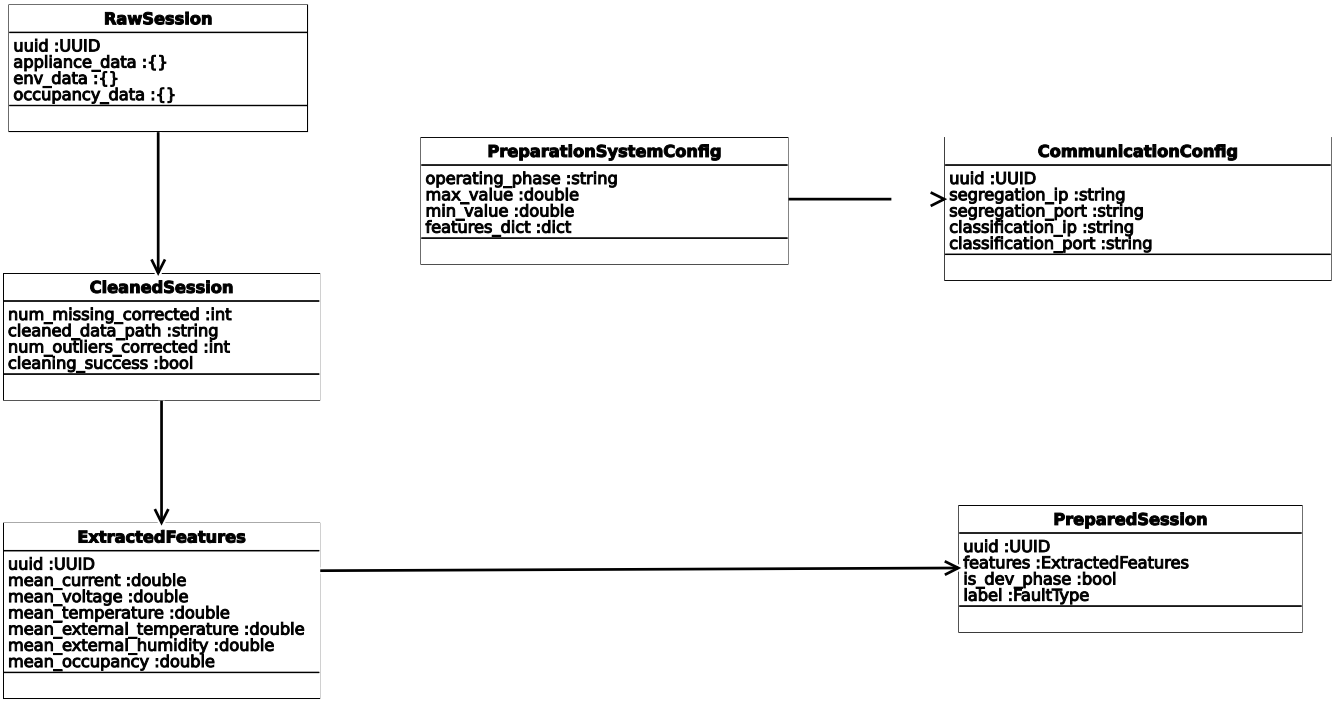


Workflow Realization (Sequence Diagram)



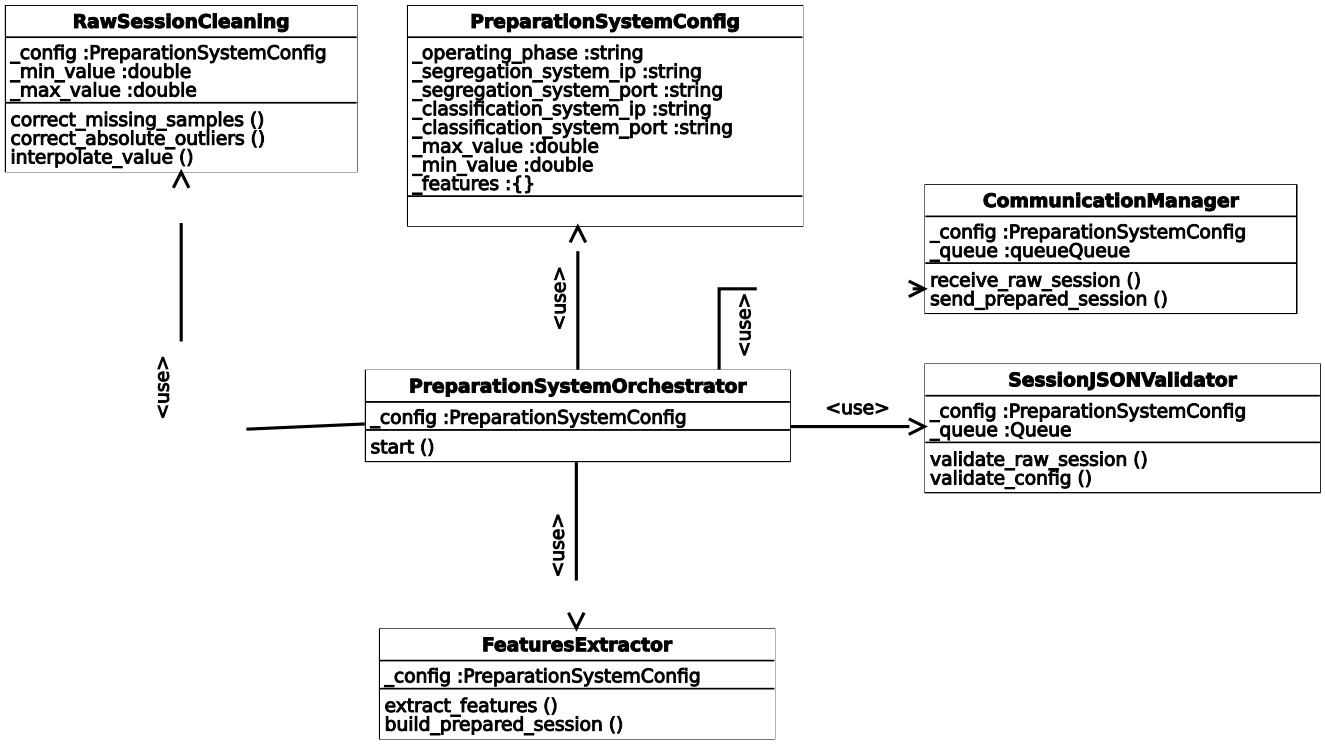
2. PREPARATION SYSTEM(Tesfaye)

Data Model

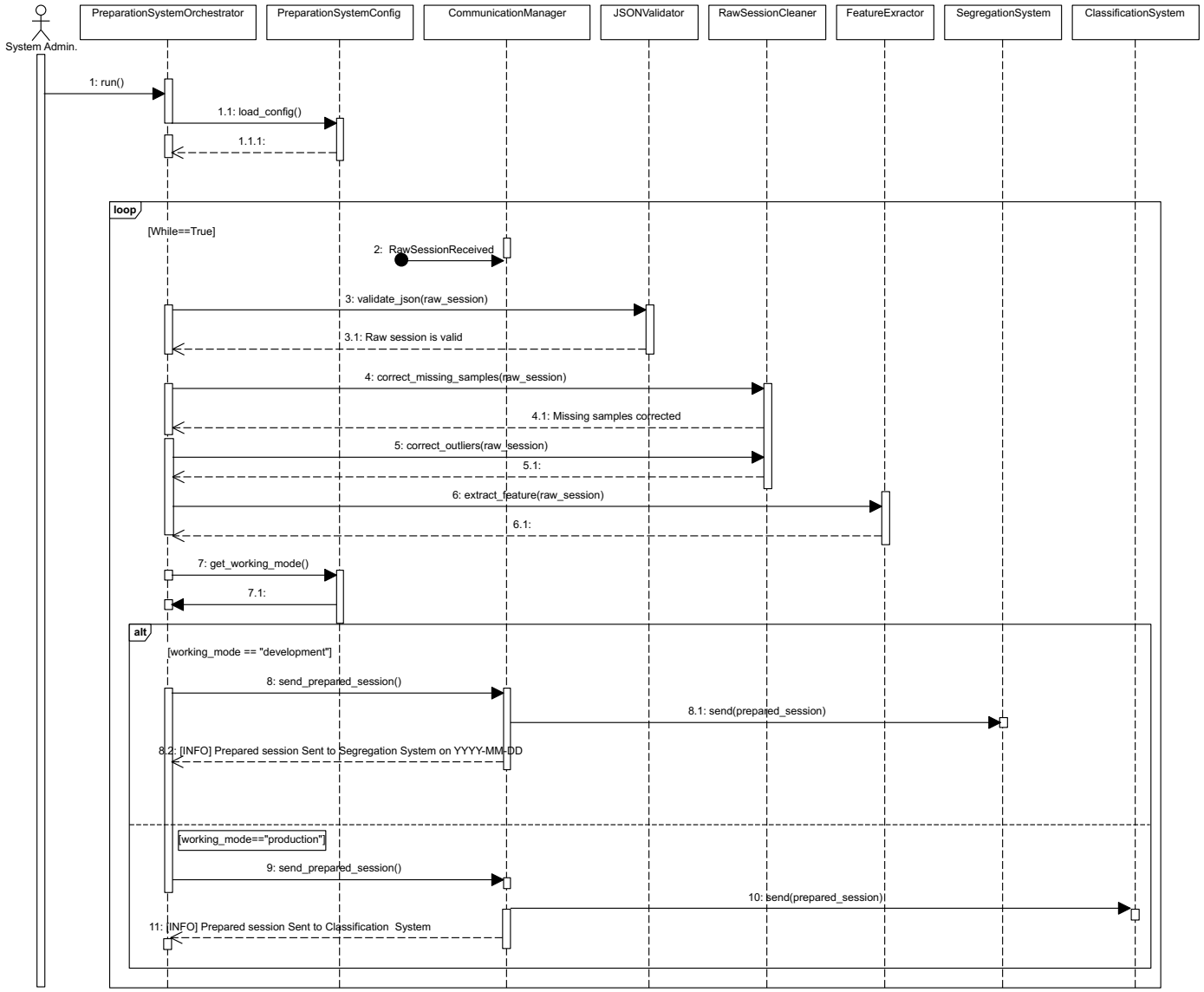


Application Logic Model

UML Class Diagram

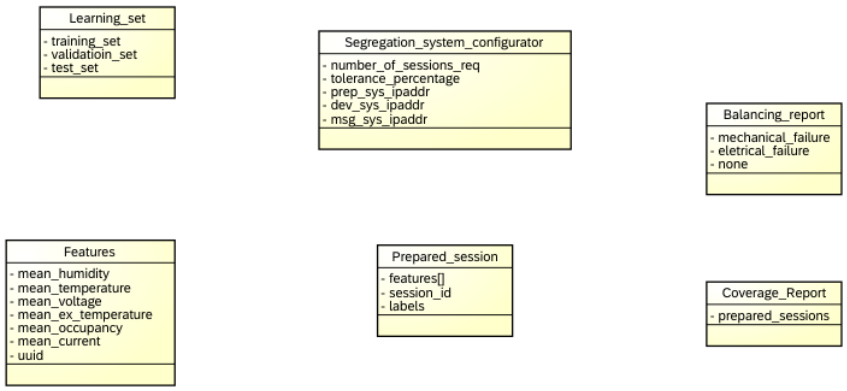


Workflow Realization (Sequence Diagram)



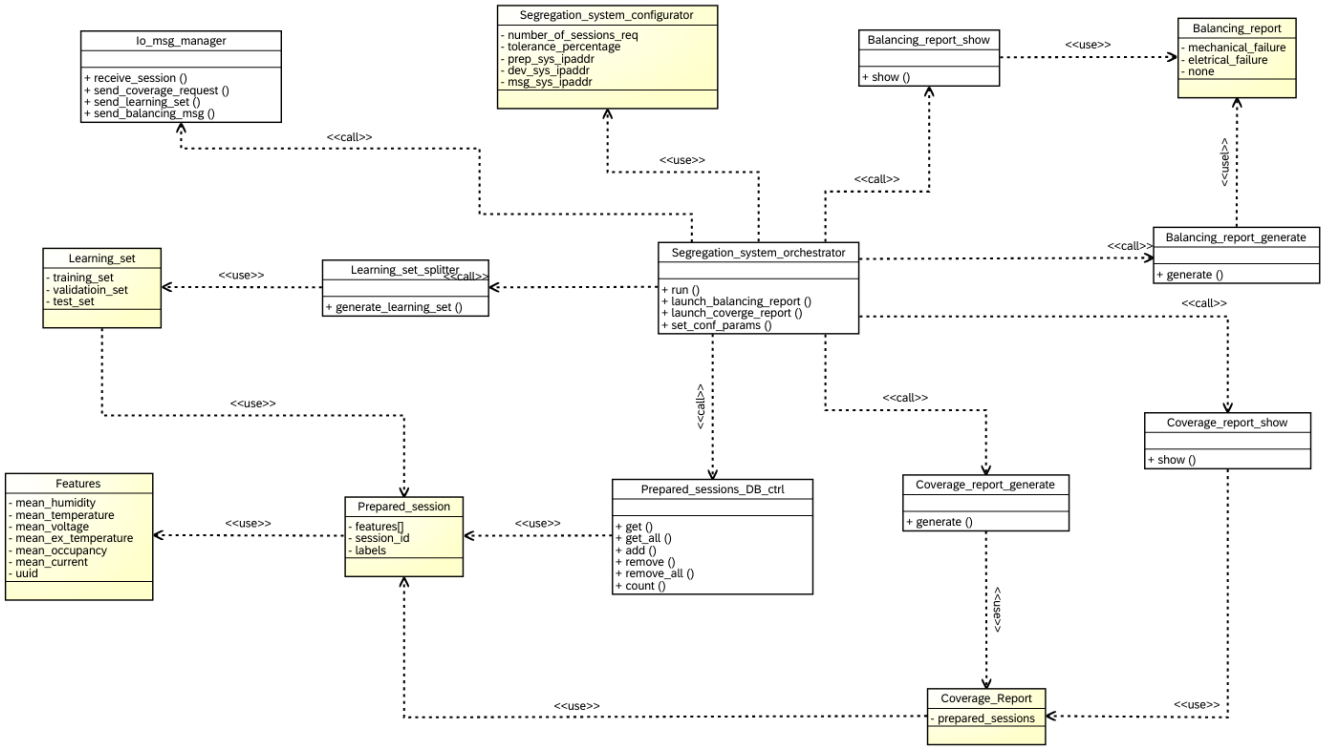
3. SEGREGATION SYSTEM (Lorenzo V)

Data Model

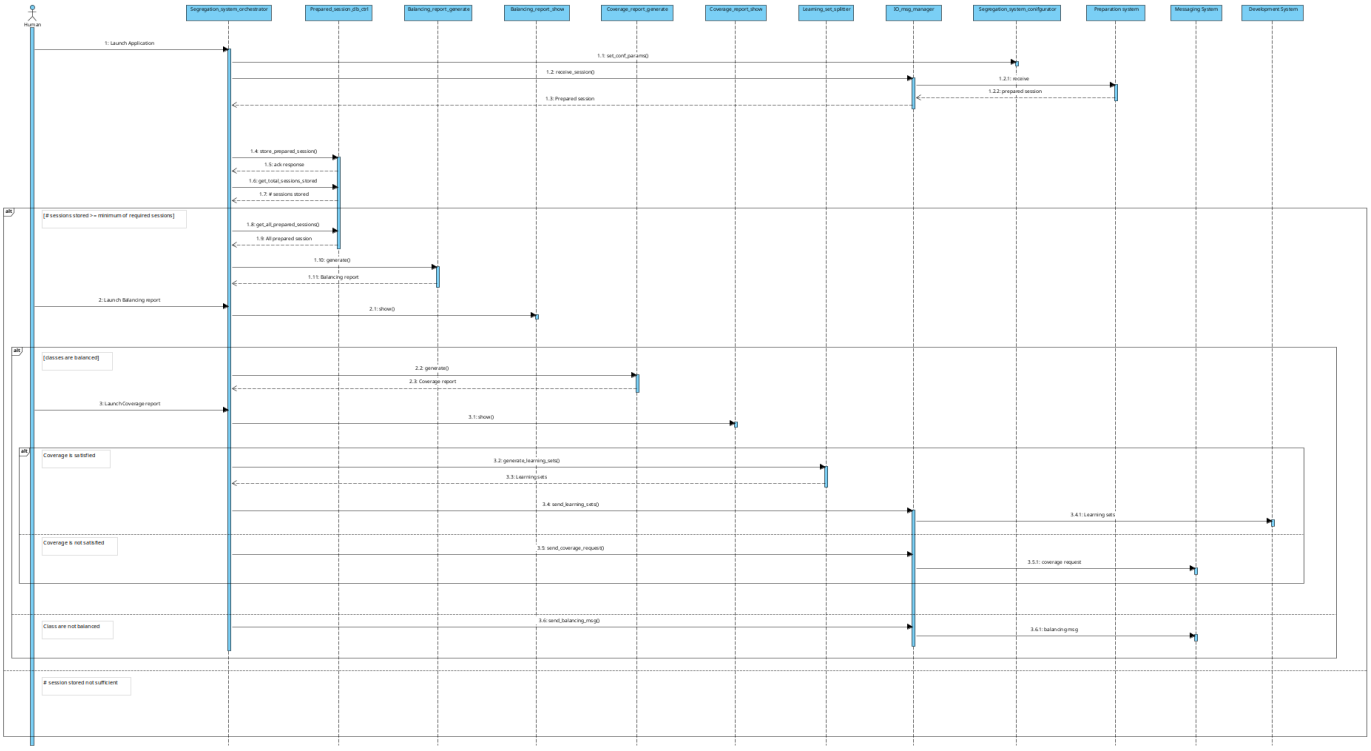


Application Logic Model

UML class diagram

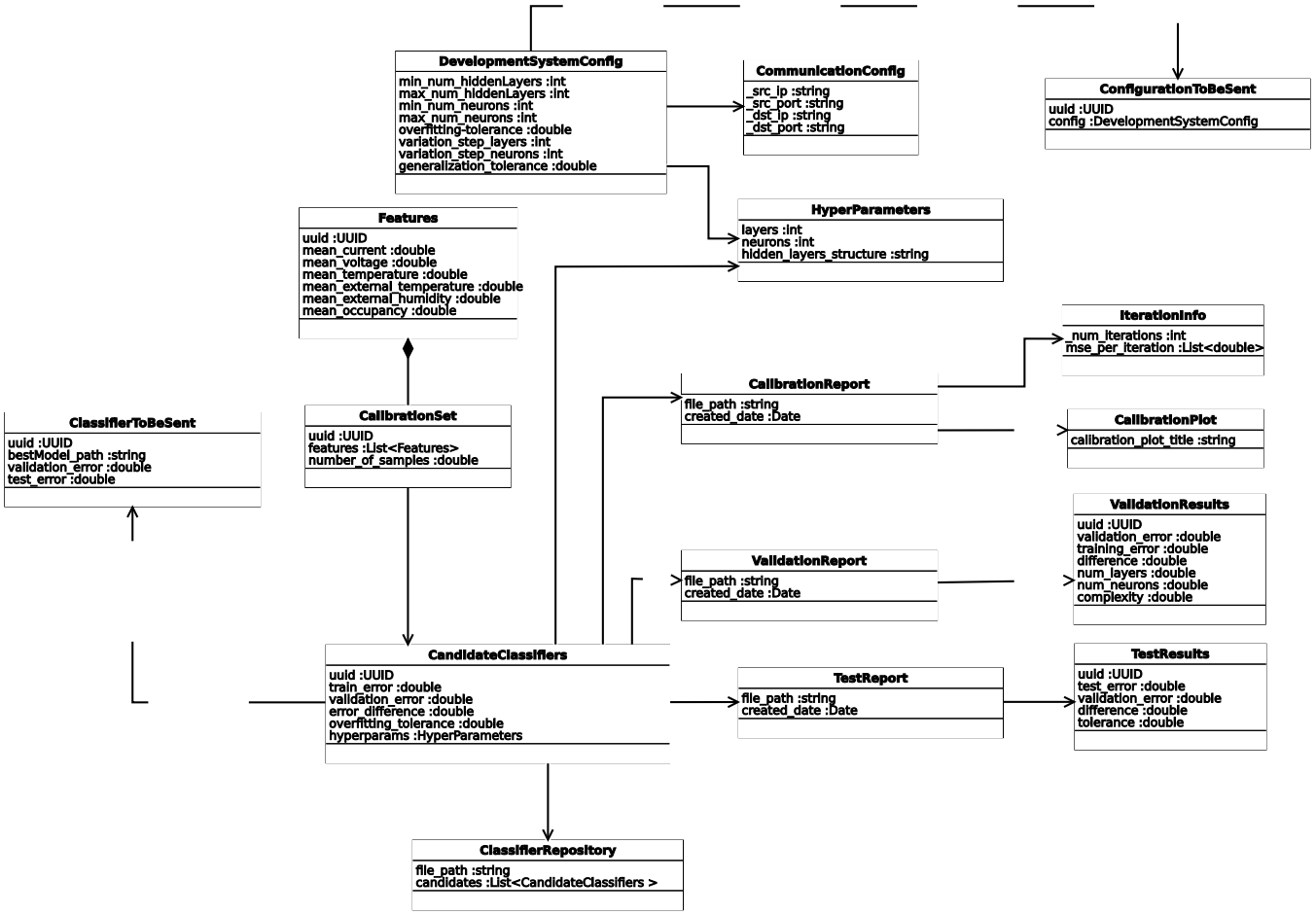


Workflow Realization (Sequence Diagram)



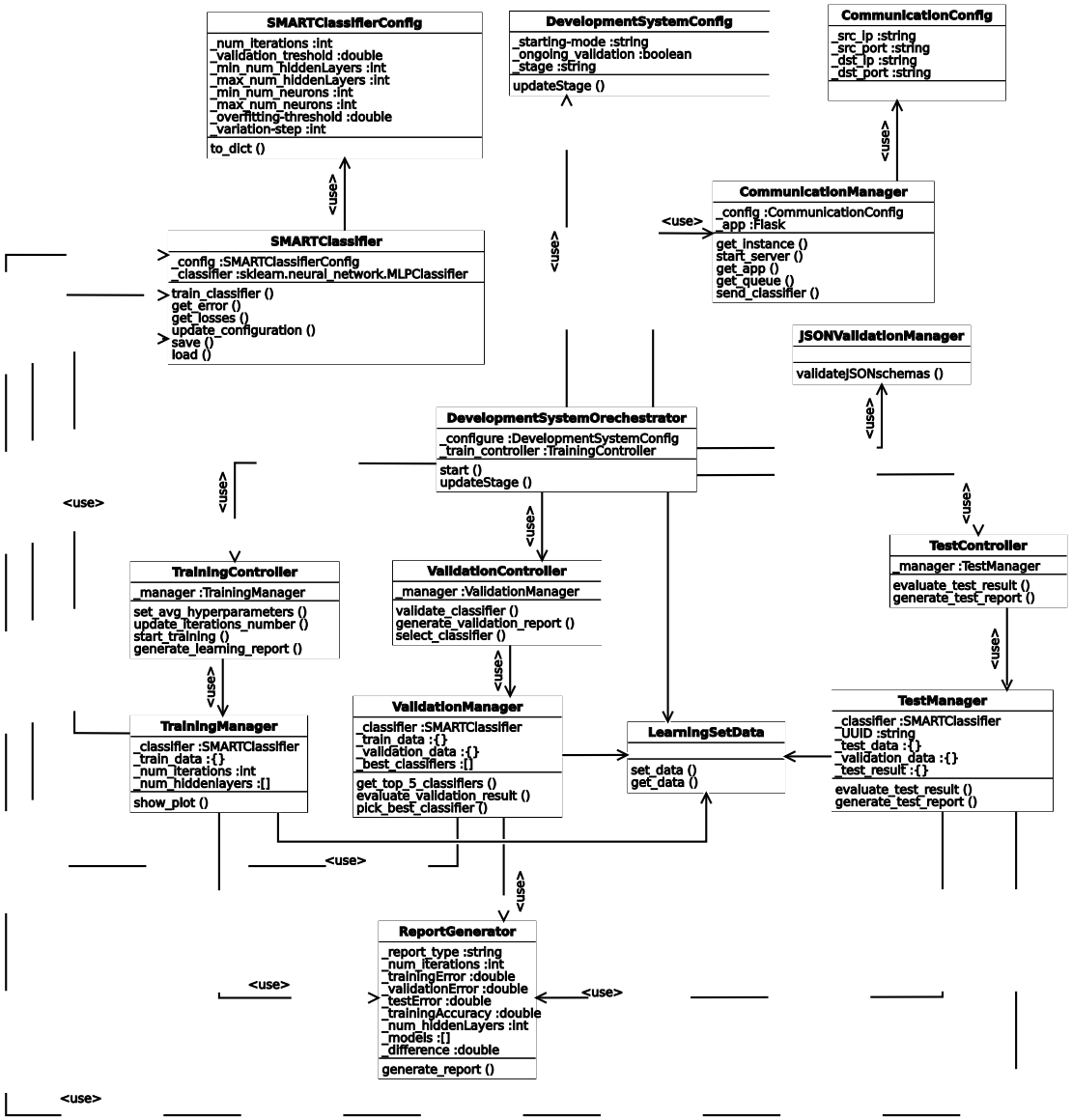
4. DEVELOPMENT SYSTEM(Tesfaye)

Data Model

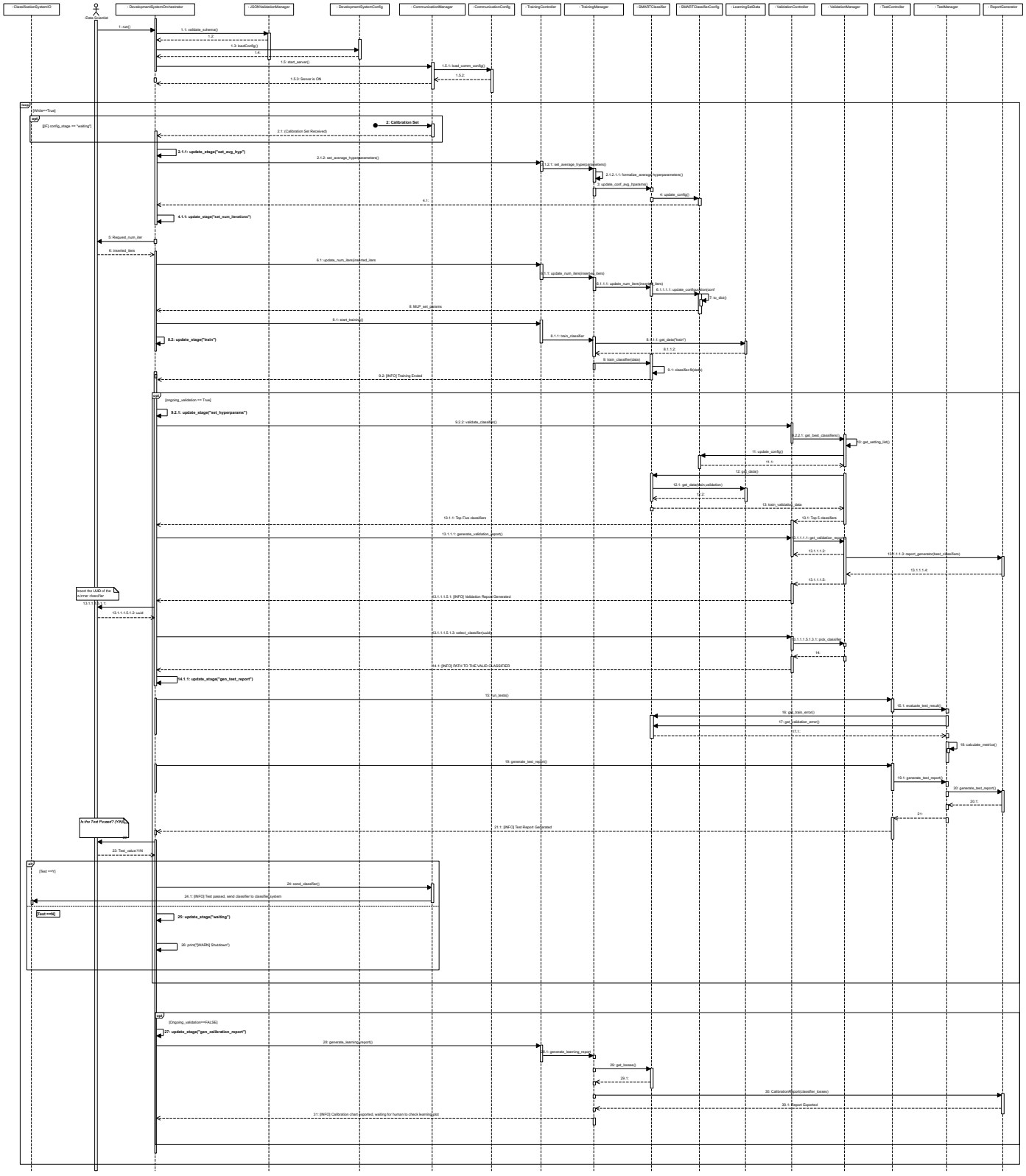


Application Logic Model

UML Class Diagram

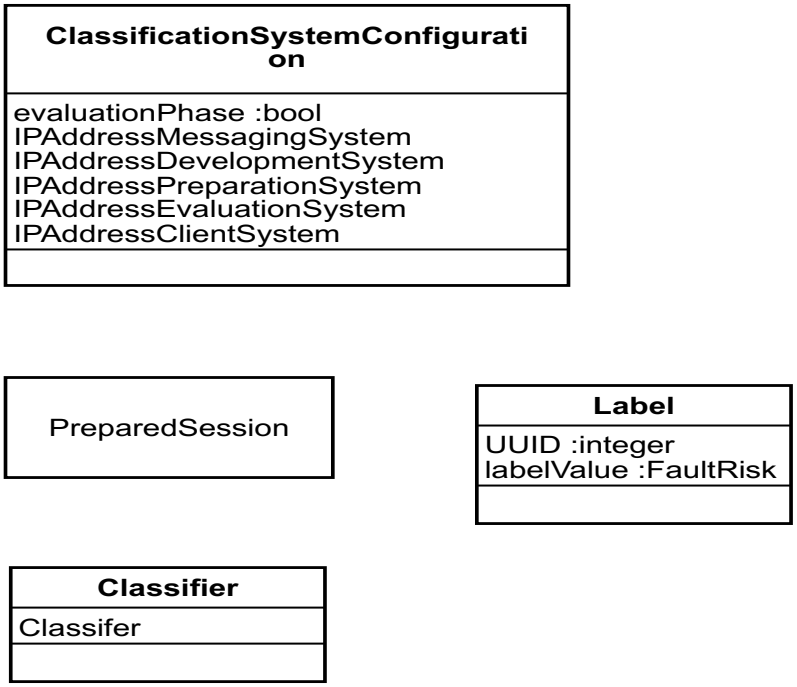


Workflow Realization (Sequence Diagram)



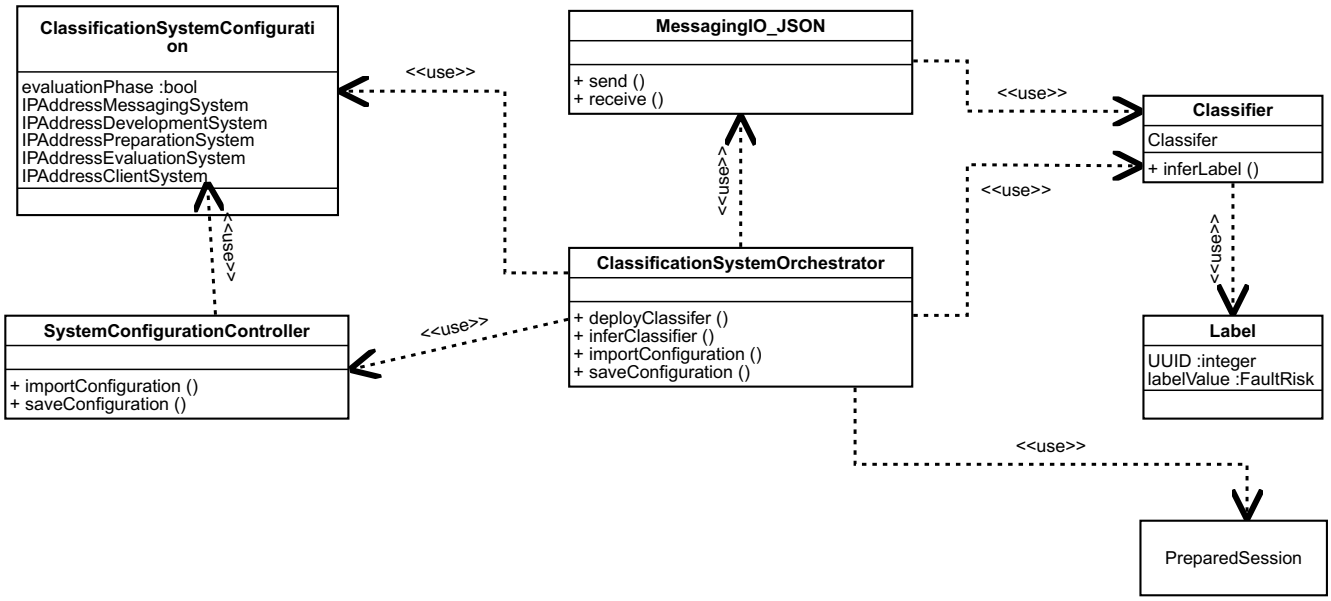
5. CLASSIFICATION SYSTEM (Alessio)

Data model

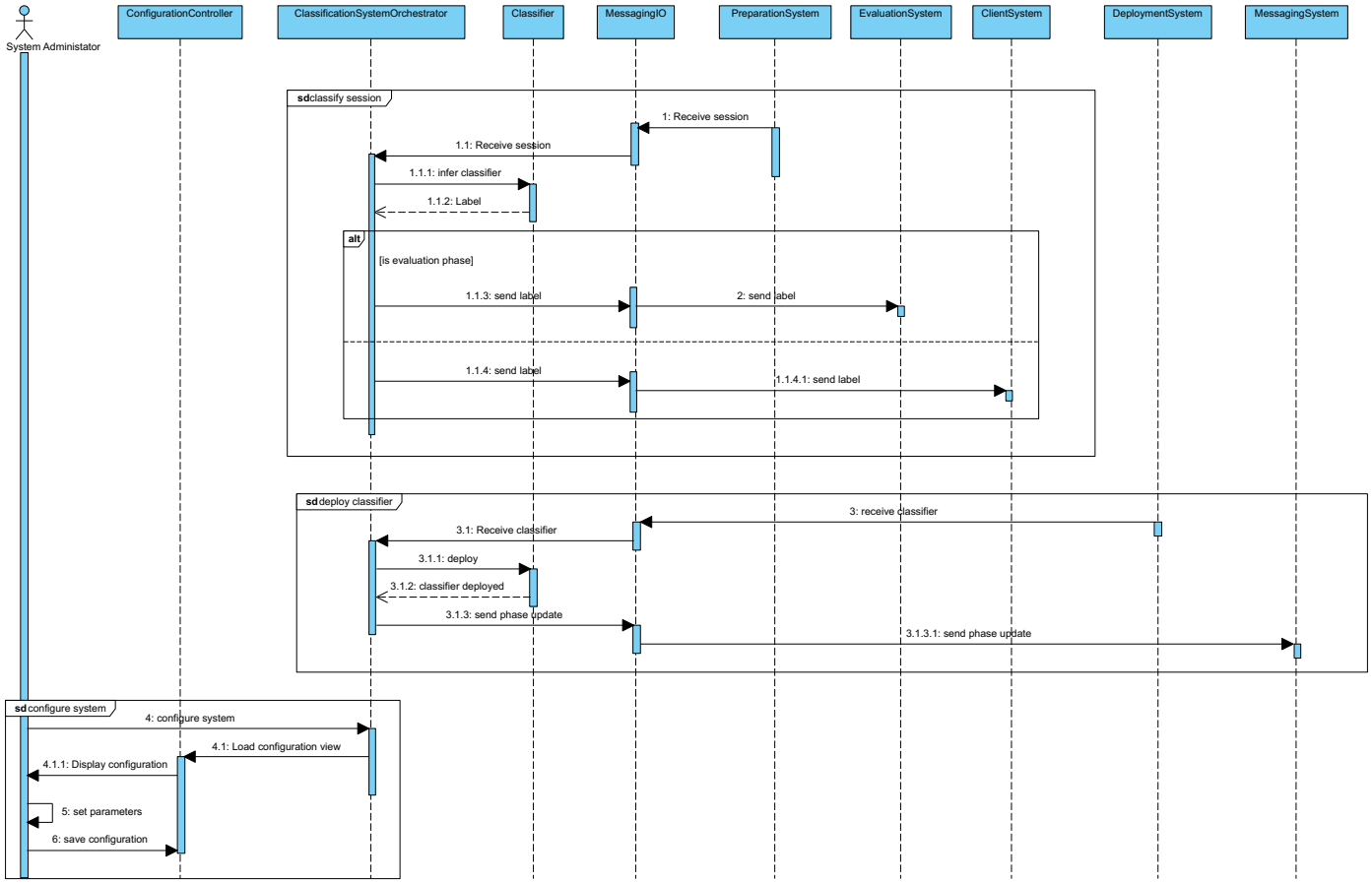


Application logic model

UML Class Diagram

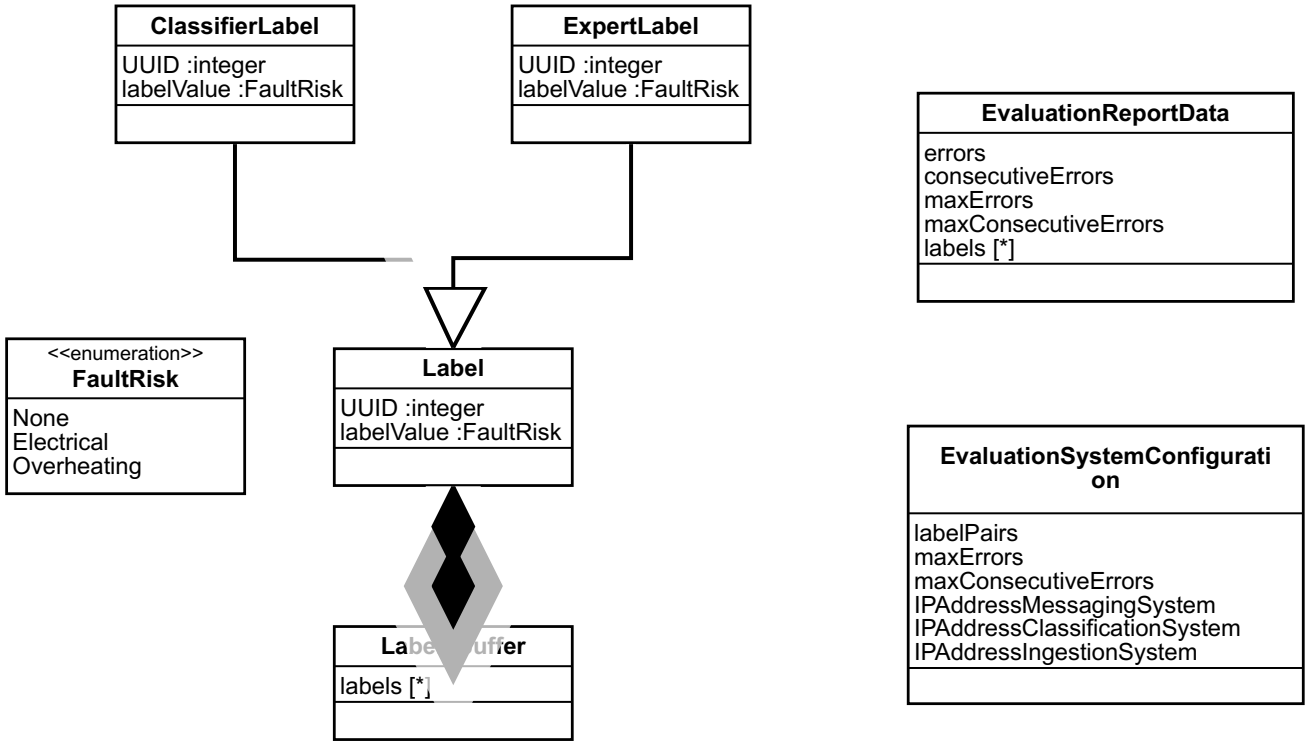


Workflow Realization (Sequence Diagram)



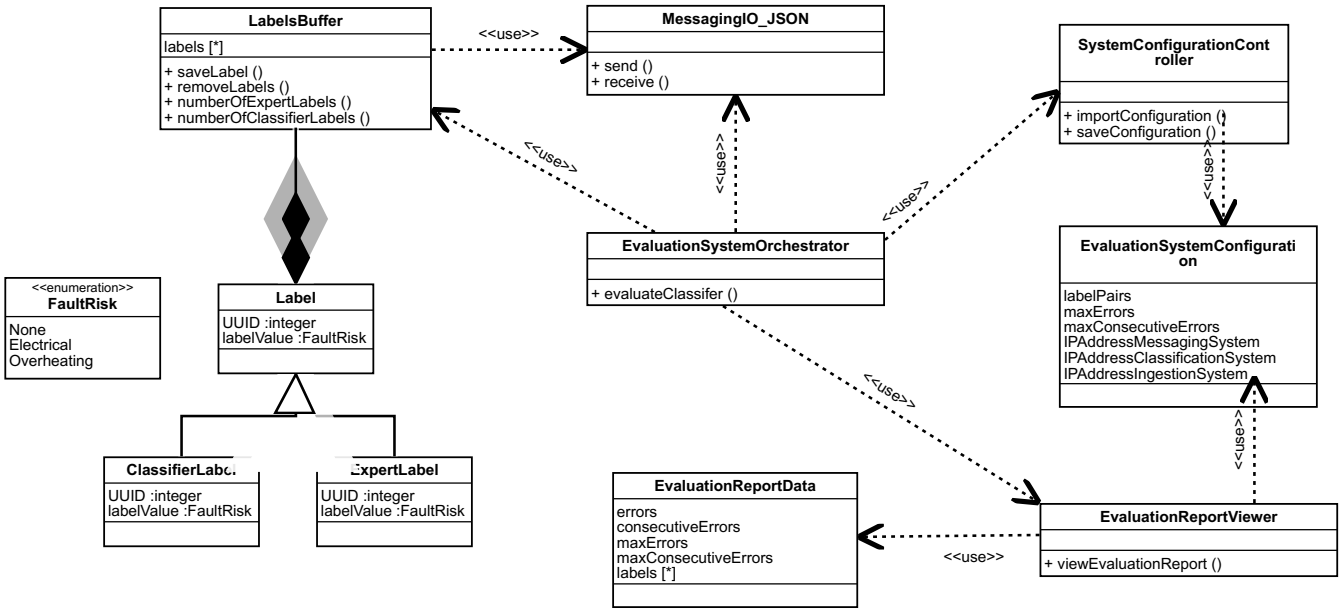
6. EVALUATION SYSTEM (Alessio)

Data model

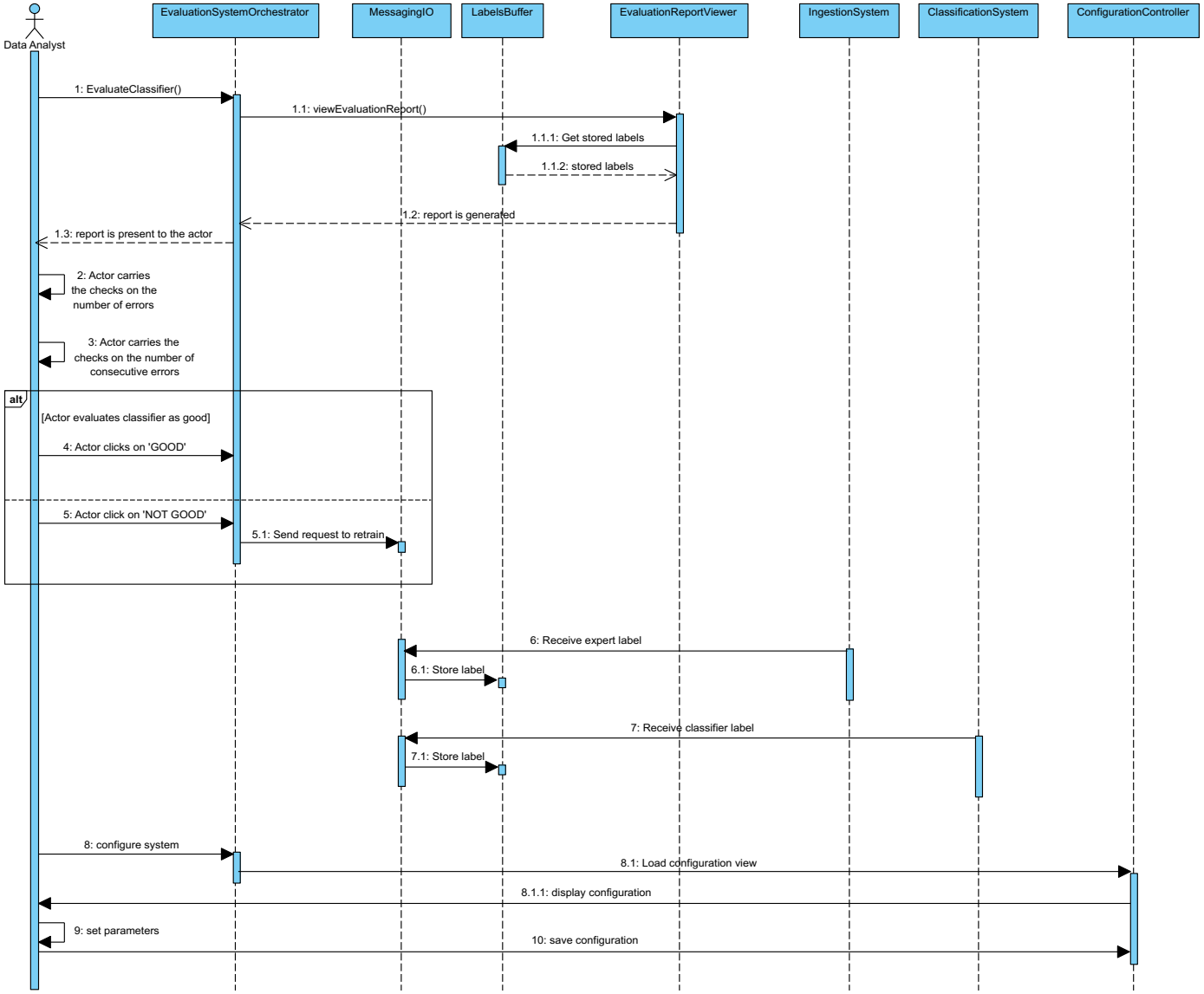


Application logic model

UML Class Diagram



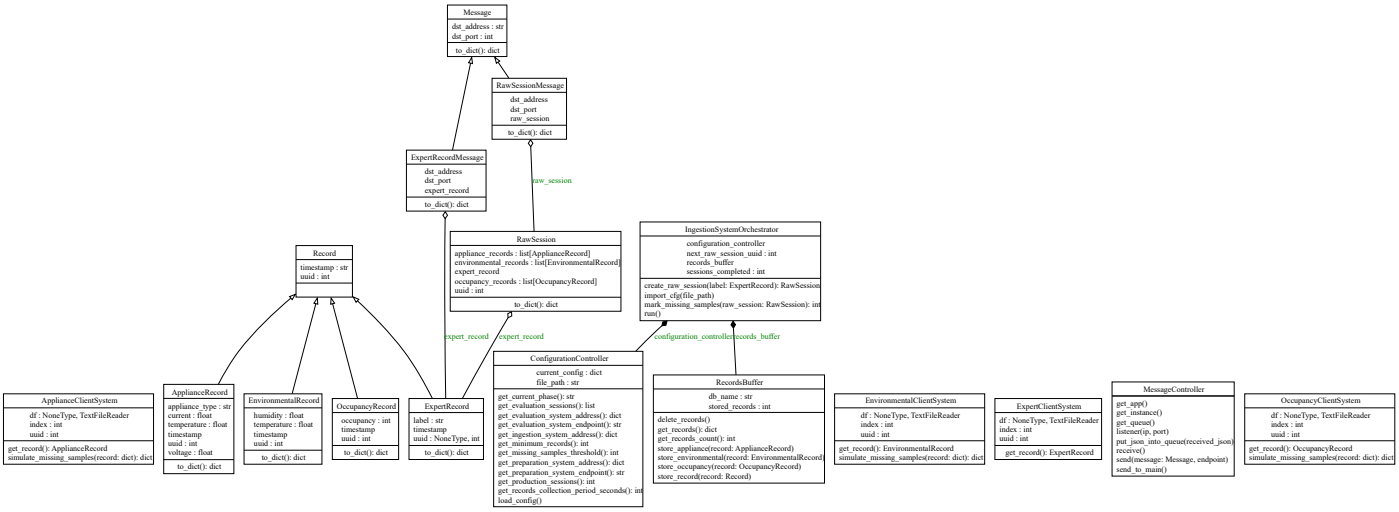
Workflow Realization (Sequence Diagram)



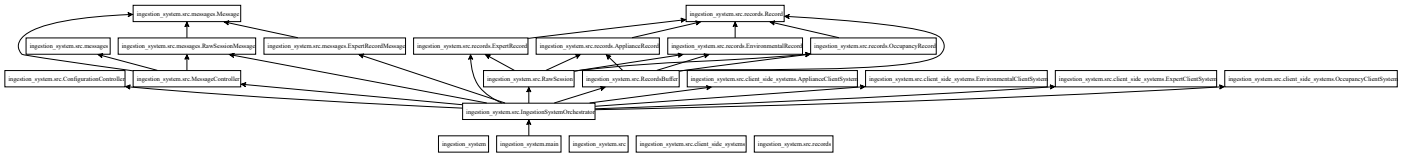
Workflow Design: UML class and Sequence diagrams

1. PREPARE SESSION PROCESS (Carlo)

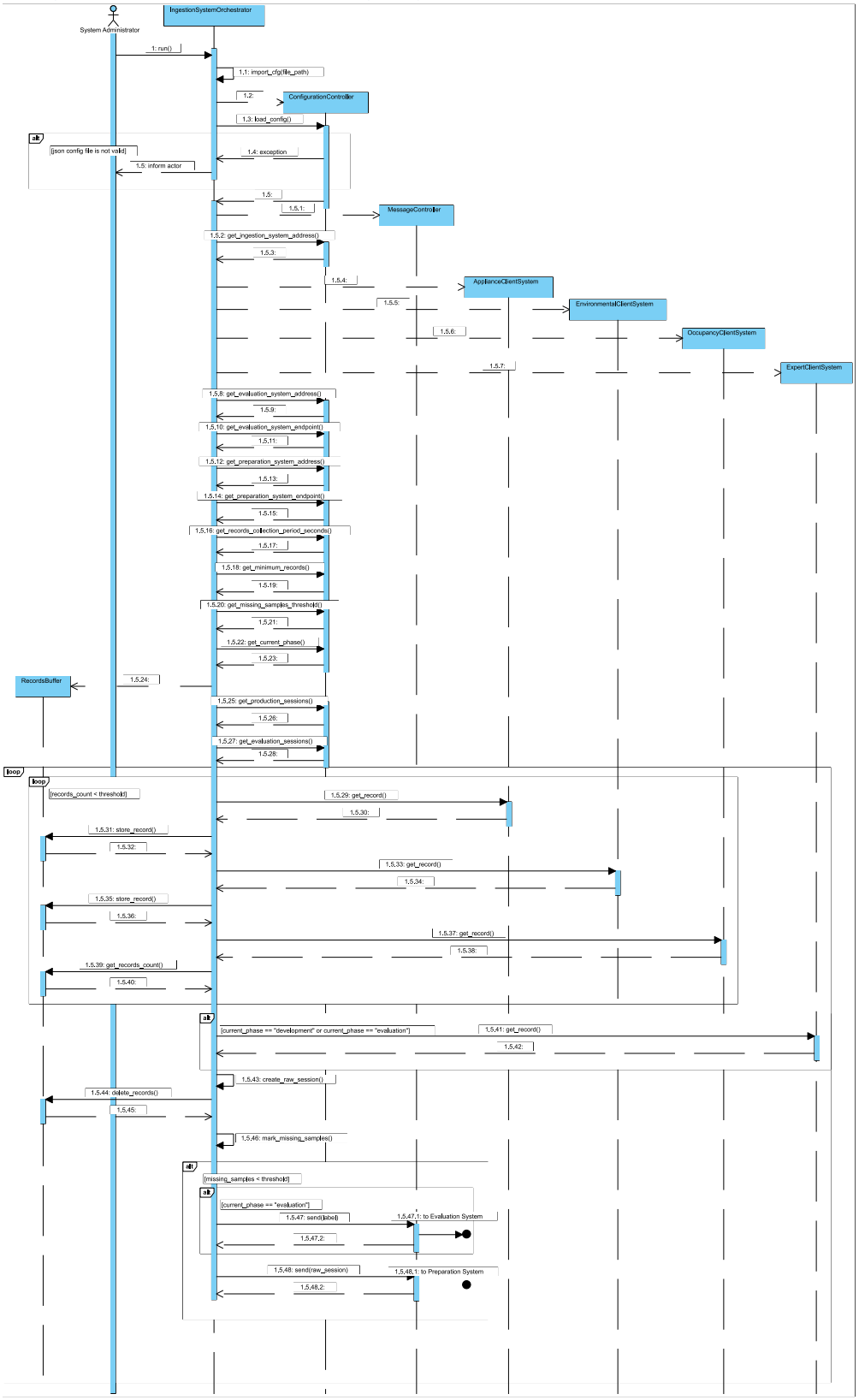
UML Class Diagram



Package diagram

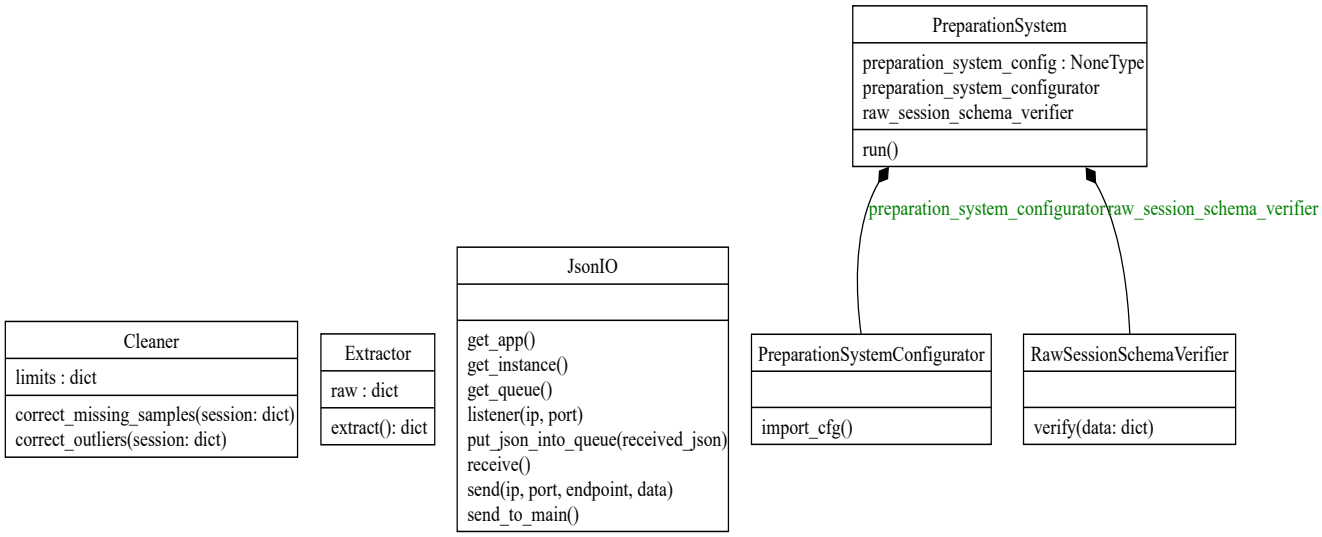


Workflow Realization (Sequence Diagram)

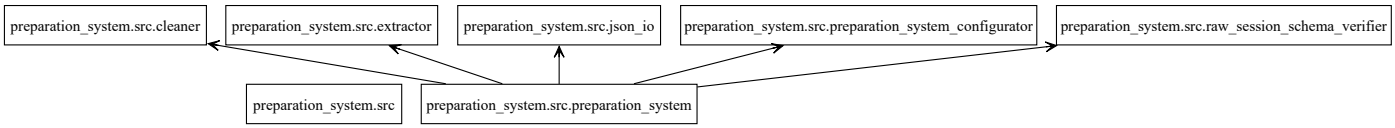


2. PREPARATION SYSTEM

UML Class Diagram

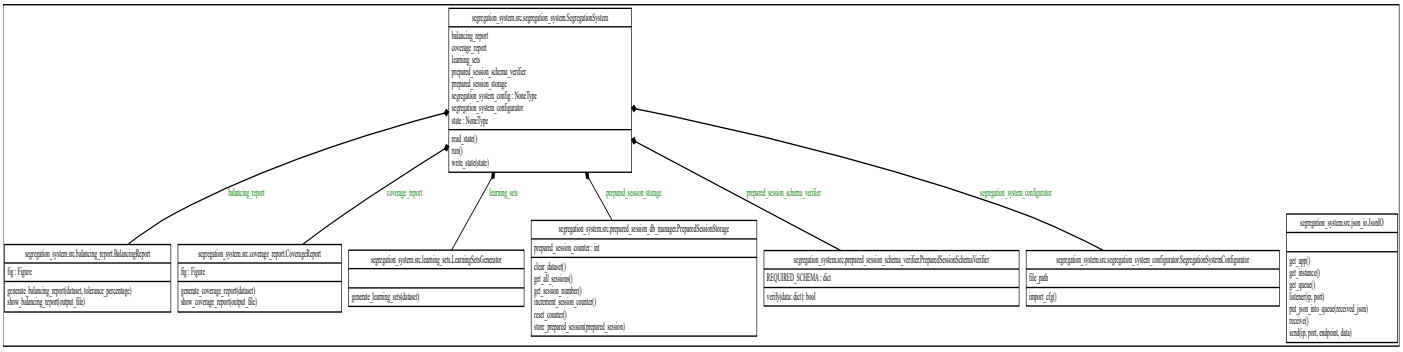


Package Diagram

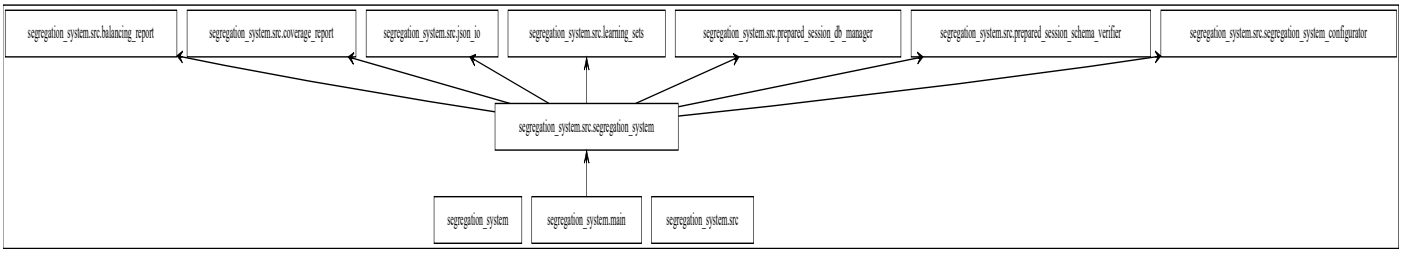


3. SEGREGATION SYSTEM

UML Class Diagram



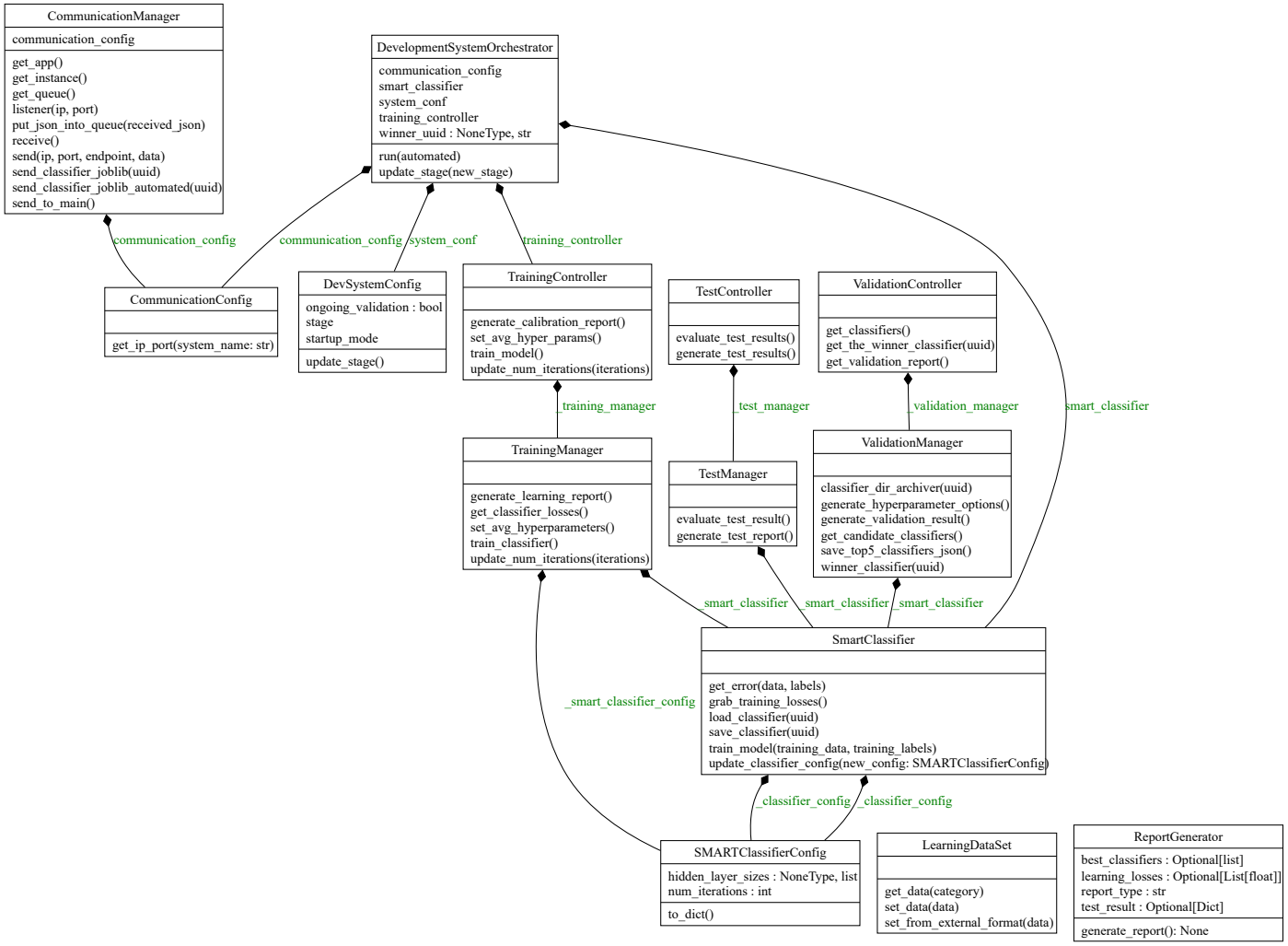
Package Diagram



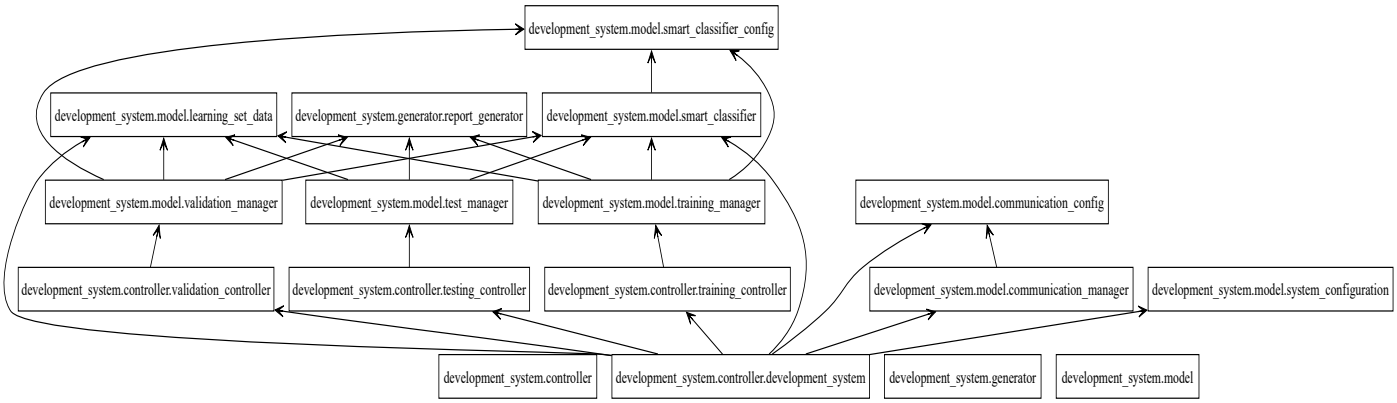
4. DEVELOPMENT SYSTEM (Tesfaye)

The UML class and package diagrams for the Development System were automatically derived using the Pyreverse module from the Pylint suite in Python.

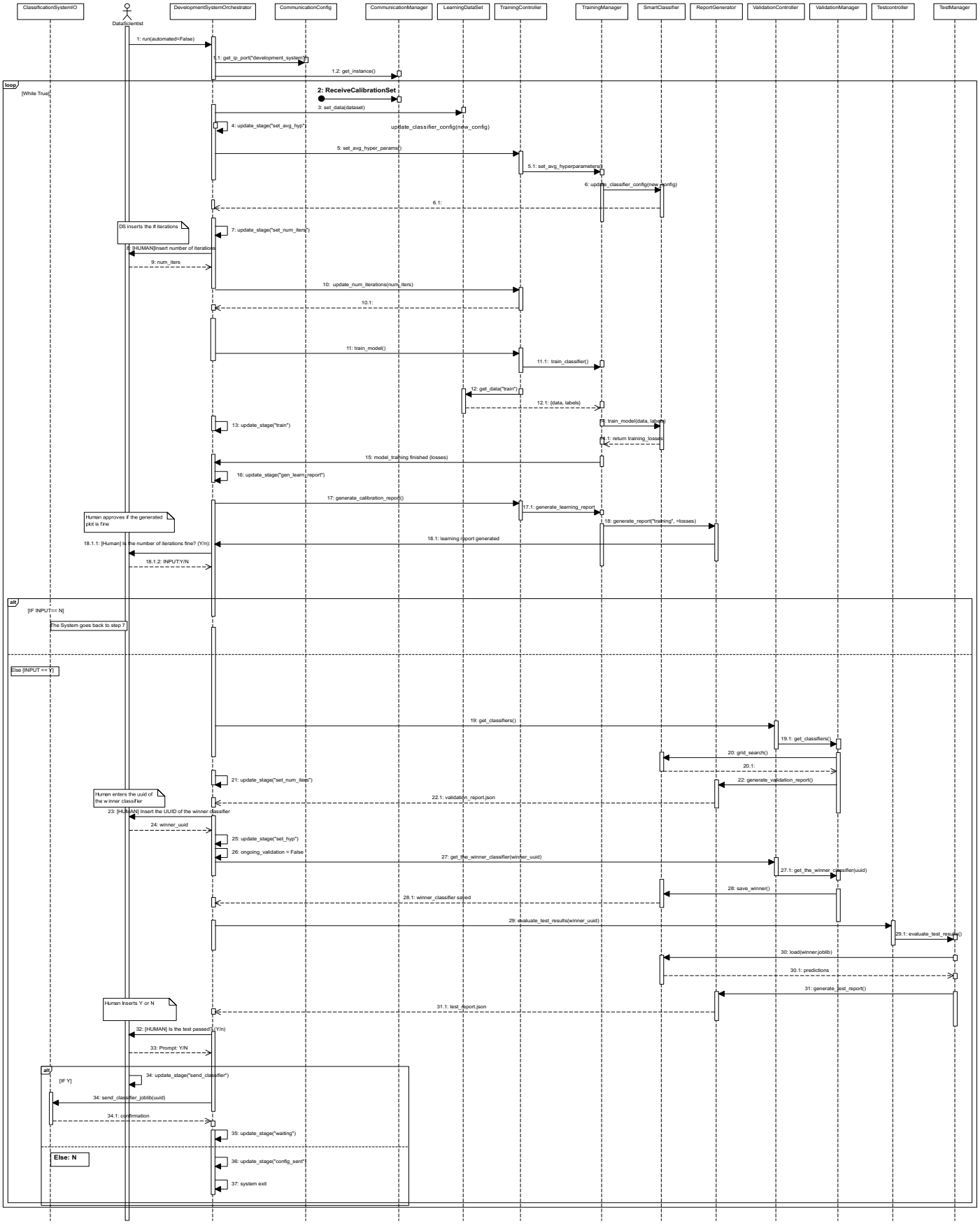
UML Class Diagram



Package Diagram

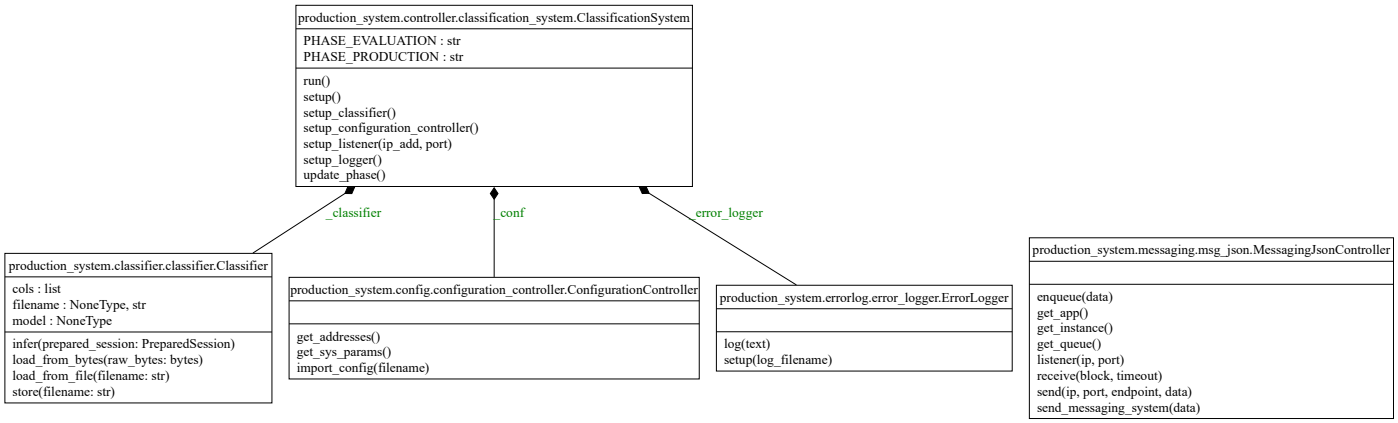


Workflow Realization (Sequence diagram)

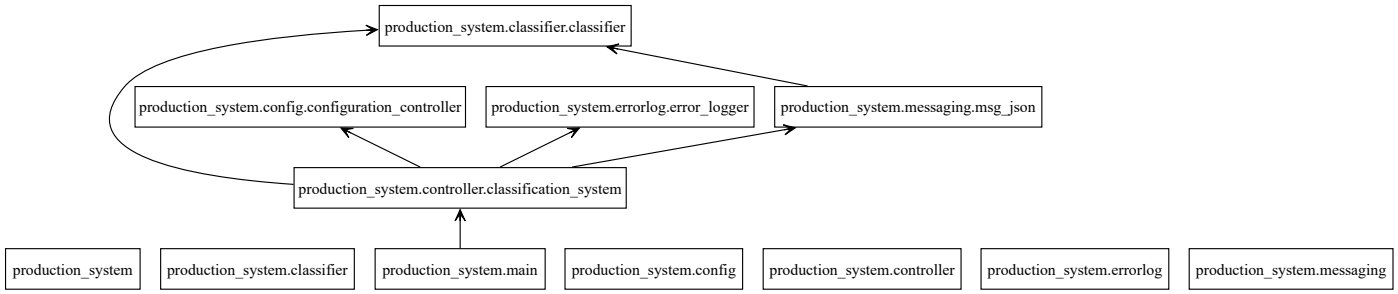


5. CLASSIFICATION SYSTEM (Alessio)

UML Class Diagram

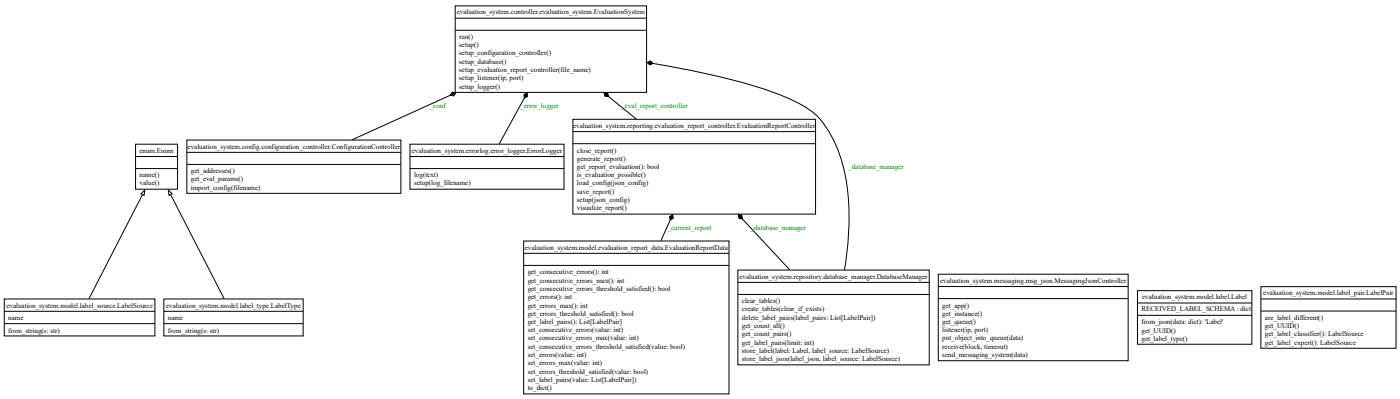


Packages Diagram

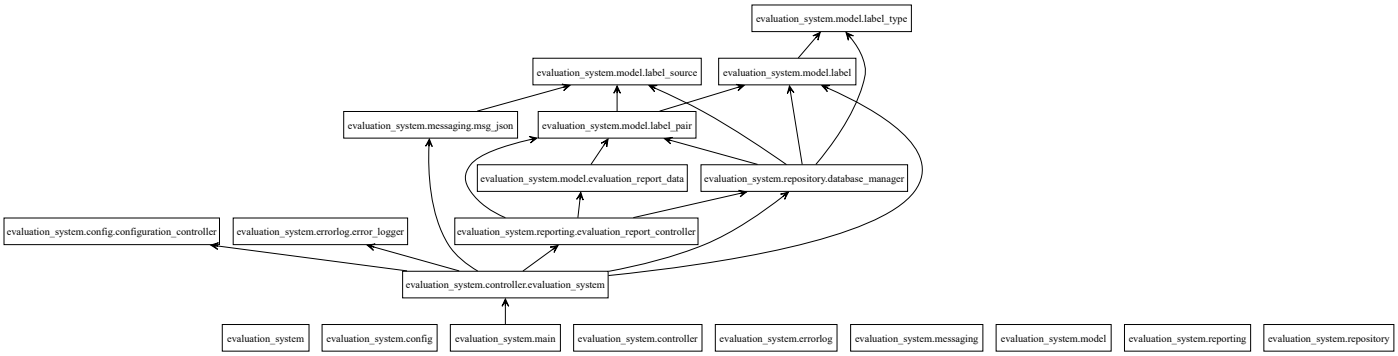


6. EVALUATION SYSTEM (Alessio)

UML Class Diagram

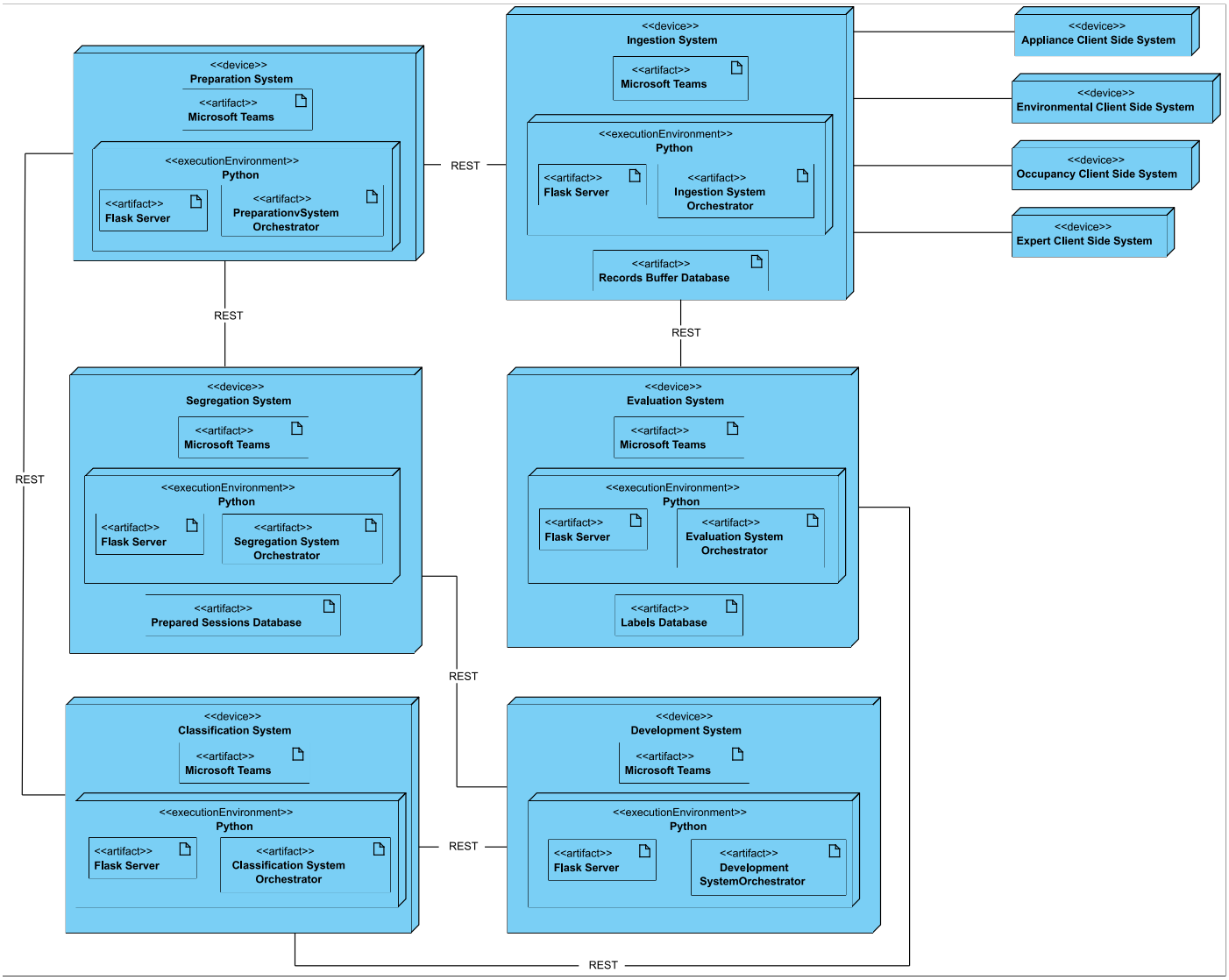


Package Diagram



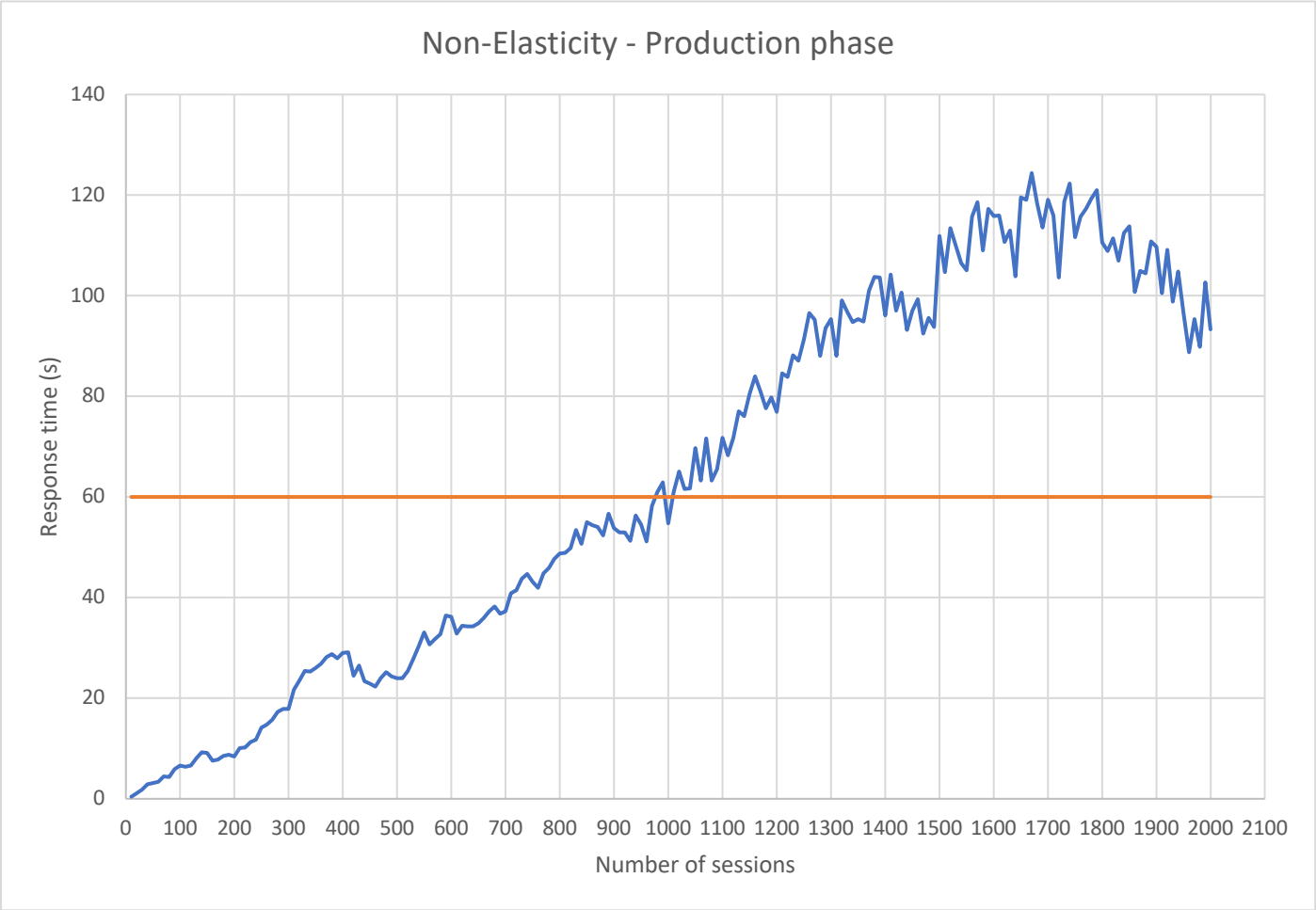
Workflow Implementation:

Deployment Diagram



Workflow Testing:

1. RESPONSIVENESS-ELASTICITY of the production phase



After all systems startup, the response time for each session begins to gradually increase as the number of classifications requests arrive. This linear-like behaviour is because the queues of the preparation system and the classification system begin to fill and cache memories are empty, so many misses are generated.

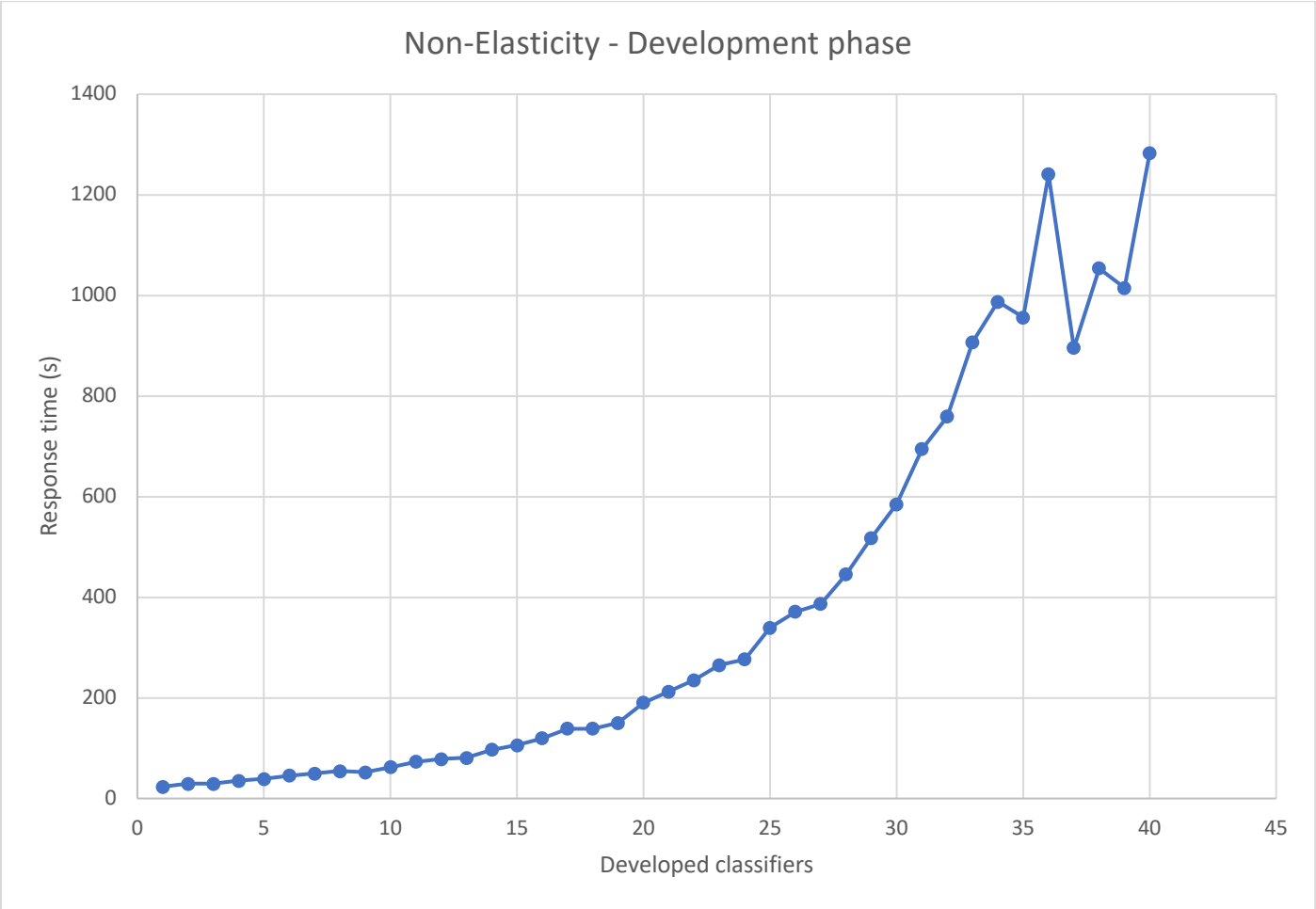
This queues buildup is caused by the important rate of arrival of records from the client-side systems and the performance bottleneck of classification system during inferences.

After a certain threshold, the caches of the systems will contain most of the useful data necessary for the computations, so the processing will be faster, hence the final stabilization of the response time.

In this phase the Ingestion system will be the bottleneck because it won't be able to increase the sending rate of the raw sessions.

For our purposes, we would like to provide a classification for each 60 seconds, so our responsiveness threshold lays at around 1000 sessions.

2. AUTOMATION RESPONSIVENESS-ELASTICITY of the development phase



Increasing the number of classifiers to develop, the response time of the development phase will increase exponentially due to the heavy computational load sustained by the development system during the training process.

For developing a new classifier, we would have the possibility to create it during the night, where clients won't be bothered from the temporary service discontinuation, so we would have about 5 hours available. Tests demonstrate that we would be able to withstand a considerable number of developed classifiers and still be below the time constraint of 5 hours.

3. RESILIENCY

Ingestion System

ID	INPUT	CONSEQUENCE	SCORE
I1	Missing records	If below the threshold, it is accepted (correction is assigned to the preparation system). If over the threshold, the raw session is discarded	1
I2	Missing labels	Error detected, raw session is discarded	4
I3	Duplicated records	Error raised by the unique identifier constraint	4
I4	Record containing an outlier	Error undetected	5
I5	Malformed configuration file	Error detected, the system will stop providing the details	2
I6	Late records	Error undetected	5
I7	Wrong record data type	Error undetected	5
TOTAL			**26**

Preparation System

ID	INPUT	CONSEQUENCE	SCORE
P1	Wrong structured Raw Session	Error detected and Raw Session discarded	4
P2	Missing samples	Error detected and corrected exploiting interpolation	1
P3	Late Samples	Error non detected	5
P4	Absolute Outliers	Error detected and corrected using predefined min / max values for each physical quantity	1
TOTAL			11

Segregation System

ID	INPUT	CONSEQUENCE	SCORE
S1	Wrong structured Prepared Session	Error detected and Prepared Session discarded	4
S2	Unbalanced Class	Error not detected. A user must analyse the graph and decided if classes are well balanced.	5
S3	Not uniform Data Distribution	Error not detected. A user must analyse the graph and decided if data are sufficiently well plotted.	5
TOTAL			14

Development System

ID	Input	Consequence	Score
D1	Late learning sets	System remains in the waiting stage.	4
D2	Hyperparameters file corrupted.	Issue detected. Trace back displayed and System stops.	3
D3	Wrong input of the Human task of inserting the UUID of the winner classifier.	The development system goes back to where the **average hyper parameters** are set.	1
D4	Wrong calibration set **values**	The model is trained on the wrong data.	5
D5	Wrong request body from the **Segregation** system.	Returns 404	1
TOTAL			**14**

Classification System

ID	INPUT	CONSEQUENCE	SCORE
CS1	Wrong classifier structure received	No detection, the classifier is deployed	**5**
CS2	Wrong prepared session structure	Validation error: the error is detected and noted in a log file for future analysis	**3**
CS3	Wrong prepared session values	No detection	**5**
CS4	Receiving a prepared session while in development phase	Validation error: the error is detected and noted in a log file for future analysis	**3**
CS5	Receiving a classifier while in production phase	Validation error: the error is detected and noted in a log file for future analysis	**3**
TOTAL			**19**

Evaluation System

ID	SCORE	CONSEQUENCE	SCORE
ES1	Invalid Label received	Validation error: the error is detected and noted in a log file for future analysis	3

ES2	Duplicated Label received	Duplicated label is automatically detected and is not added to the database	1
ES3	Missing labels from either expert or classifier	Presence of unpaired labels are detected and noted. Unpaired labels are exclude from the reporting process	3
TOTAL			7