Data Analytics on Machine Learning performance

INDEX

1

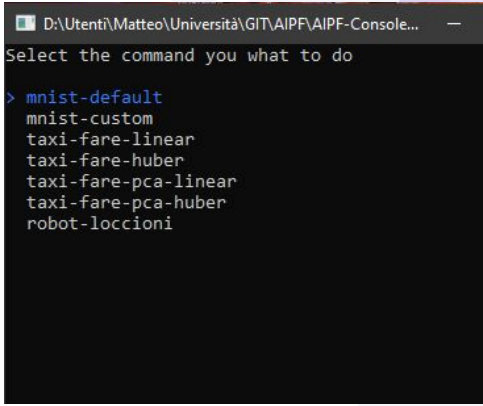
Introduction

Project description & Objective

AIPF System

Project description

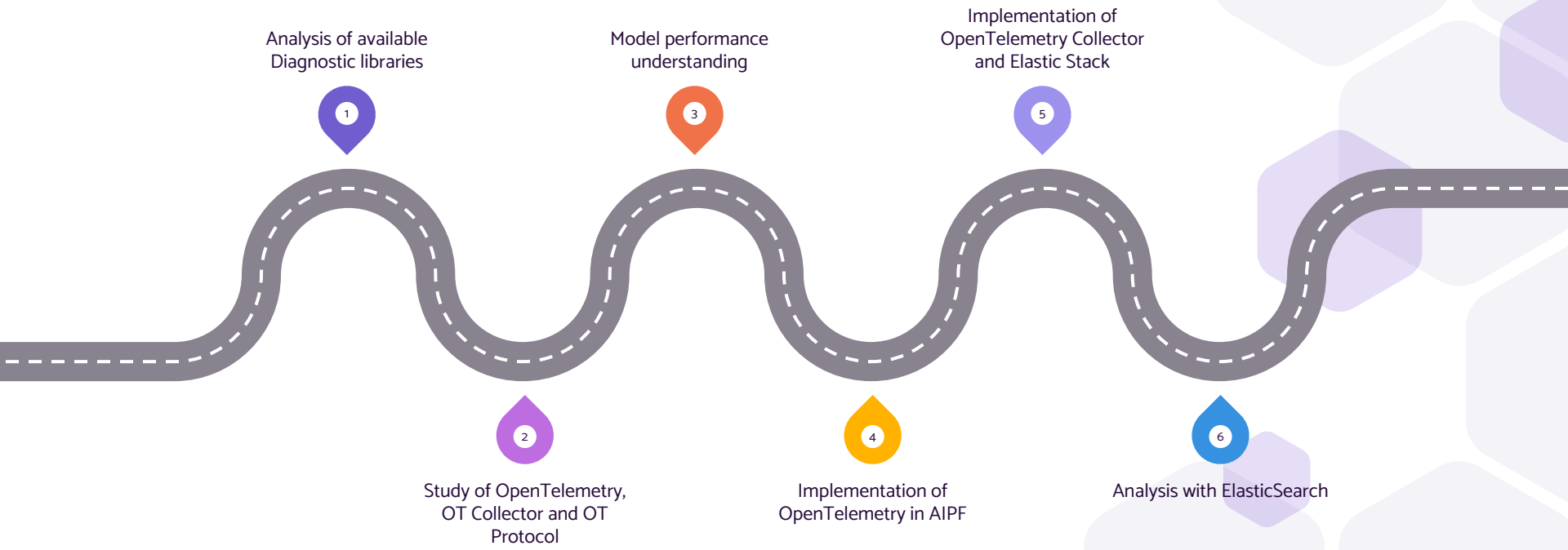
A system that allows the company to integrate Machine Learning Algorithms, for each industrial machine used, in a simple way.



OBJECTIVE

Analyze in real time the diagnostic of the AIPF system, extending it, in order to understand what is the best model.

ROADMAP


Analysis of available Diagnostic libraries
1
2
Study of OpenTelemetry, OT Collector and OT Protocol
Model performance understanding
3
4
Implementation of OpenTelemetry in AIPF
Implementation of OpenTelemetry Collector and Elastic Stack
5
6
Analysis with ElasticSearch

2 Methodology and Technology

System architecture & frameworks

System division

Core

The core is used to define some ML pipeline and models

Diagnostic

Is related to collecting telemetries about the system

FRAMEWORKS

Elastic Search

The fast and scalable search and analytics engine at the heart of the Elastic Stack

Kibana

User interface for Elasticsearch data visualization and navigation

APM Server

Receives data from APM Agents and transforms them into ElasticSearch document

Open Telemetry SDK

.NET library integrated into the AIPF System

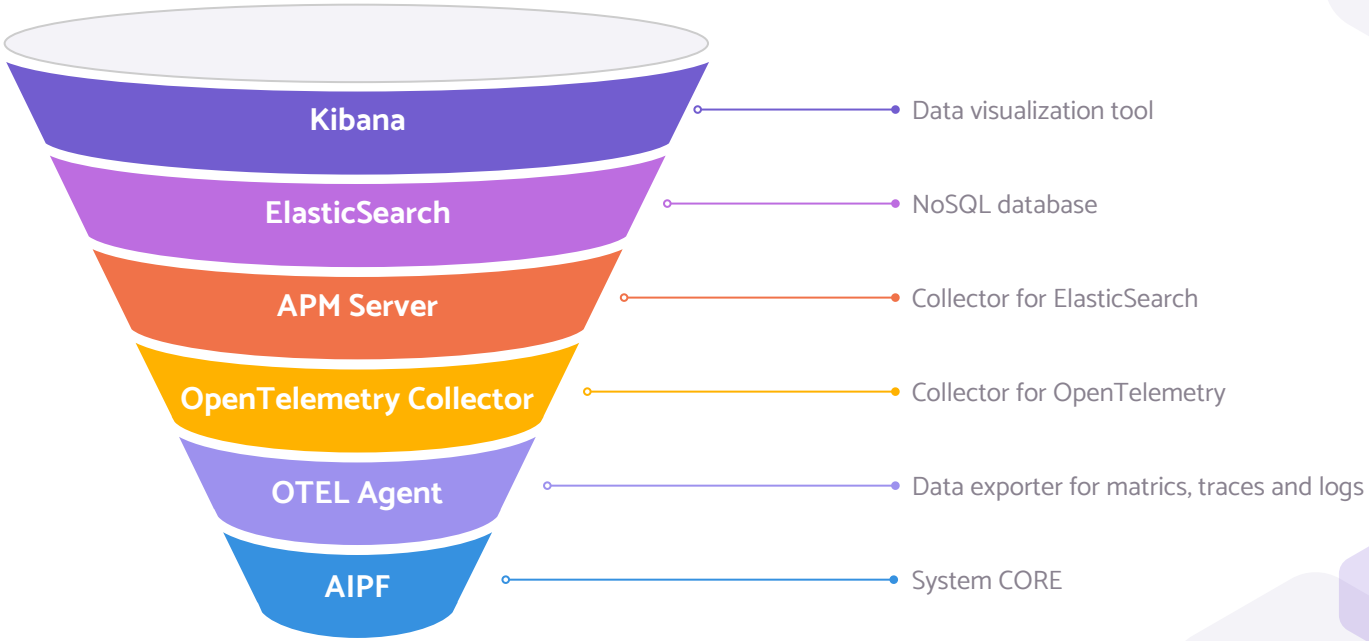
Open Telemetry Collector

The OT collector used to receive, process and export telemetries data

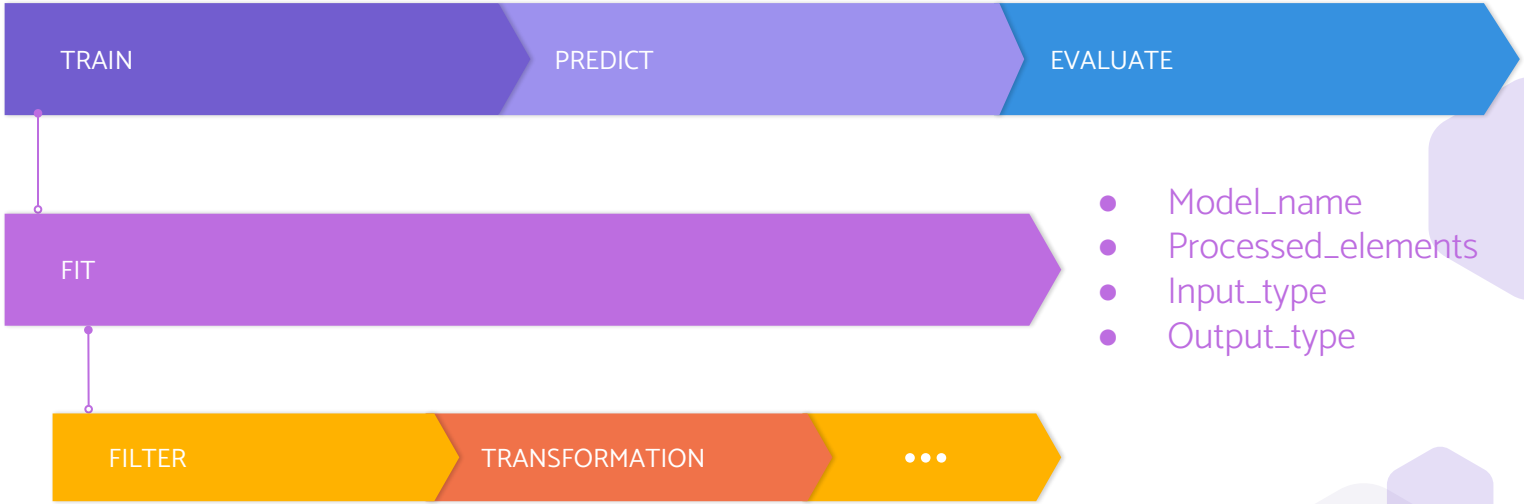
3 Technical implementation

Workflow

Workflow


Kibana
ElasticSearch
APM Server
OpenTelemetry Collector
OTEL Agent
AIPF
Data visualization tool
NoSQL database
Collector for ElasticSearch
Collector for OpenTelemetry
Data exporter for matrics, traces and logs
System CORE

AIPF



- Model_name
- Processed_elements
- Input_type
- Output_type

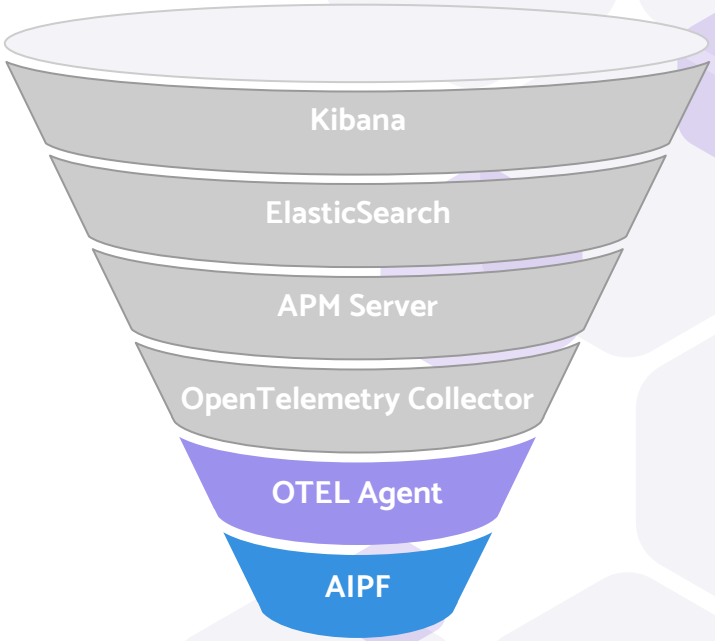
AIPF & OTEL Agent

AIPF

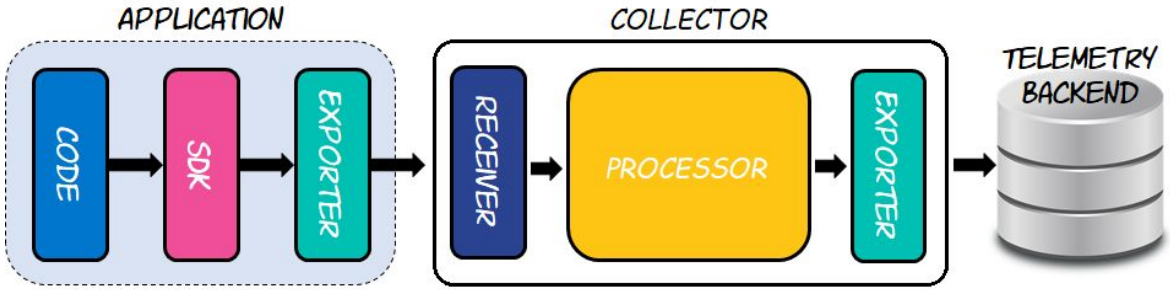
The system workflow begins with the data production by the Core.

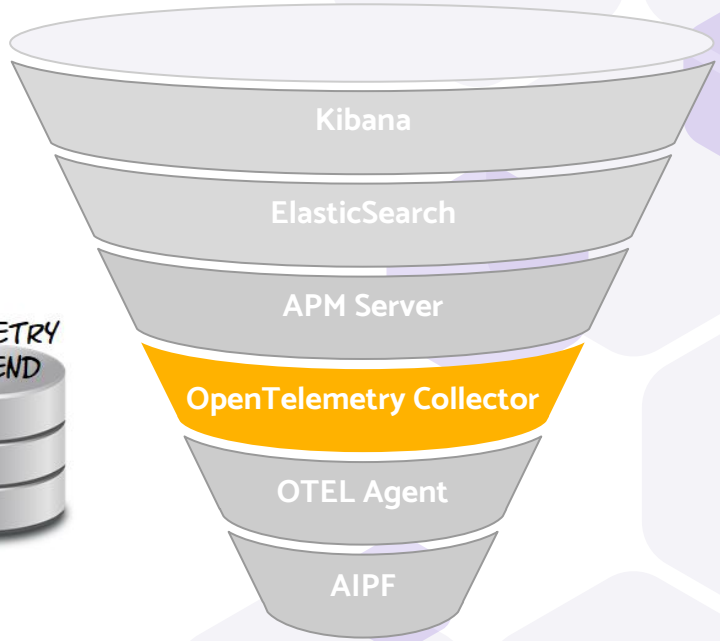
OTEL Agent

We collect the data and then send to the Collector.



Open Telemetry Collector

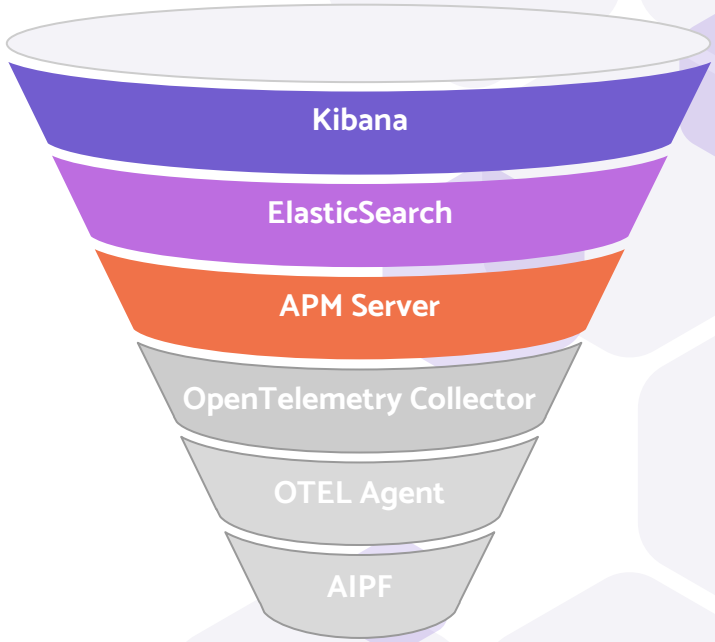




Elastic Stack (APM Server + ElasticSearch + Kibana)

Elastic Stack

Receive data from Open Telemetry Collector and saved inside Elastic; then we analyze them using kibana.



4 Achieved results

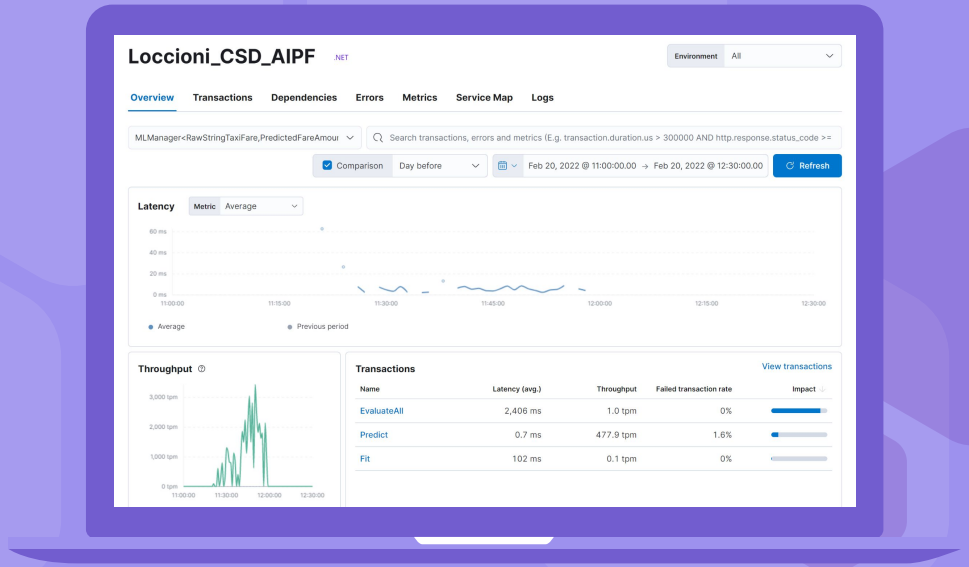
*We are going to analyze the model's performances in typical activities of a ML model (**Train**, **Predict** and **Evaluate**)*

Main Dashboard

Shows the main performance of the system

APM Dashboard

Overview of APM service



TaxiFare Dashboard

Shows the comparison between the TaxiFare models

5 Possible future implementations

Conclusion

CONCLUSION

The integration of OpenTelemetry with our system is limited.

Some further improvements are:

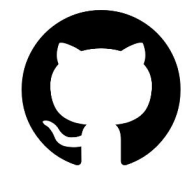
- Add logs inside the AIPF system;
- Add metrics about the states of the pipeline (ex: progress indicator);
- Add more specific metrics about CPU, Memory, Network...

TEAM

Damiano Cacchiarelli

Roberto Cesetti

Matteo Romagnoli



GitHub

THANKS!

You can find us at:

- damiano.cacchiarelli@studenti.unicam.it
- matteo02.romagnoli@studenti.unicam.it
- roberto.cesetti@studenti.unicam.it