Classification and Retrieval System for Gas Pipe Repairs
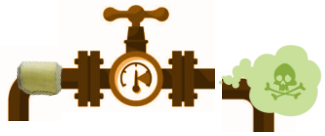
CheckPoint #3

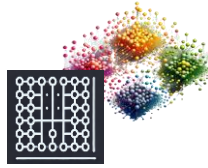*Team 1: Minal Jamshed, Prima Acharjee, Michele Panteleo*
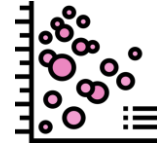
Outline

Recap

Data
Preparation

Methodology

Evaluation

What's next?

Problem Description

This project deals with **gas leaks.** Whenever a fault happens, gasfitters are asked to **fix** it by choosing the **best** strategy (welding, substituting, **patching**).
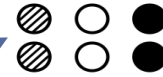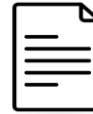
Gas leak

Mission

Our mission is to develop a **chatbot** to help gasfitters about **patching or not,** explaining **why**. Chatbot will answer based on **fault description** and **past interventions**.
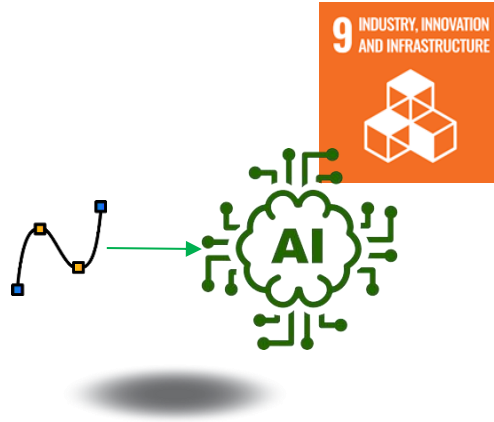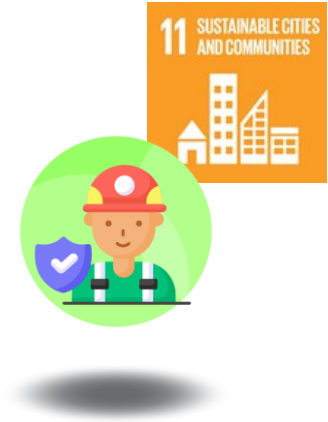

Retrieve

Decide

Explain

Value Proposition

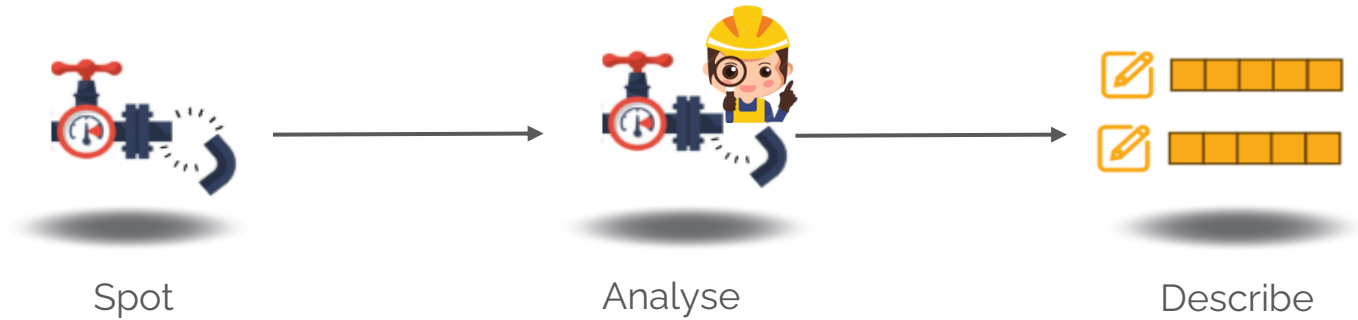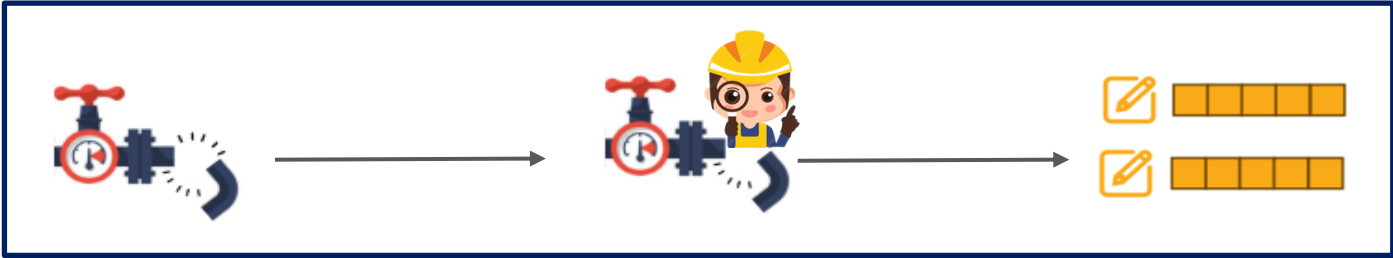**Innovate**

Reduce Gas Leaks Impact

Improve Safety

Data Preparation

Data Collection

Spot

Analyse

Describe

Data Collection

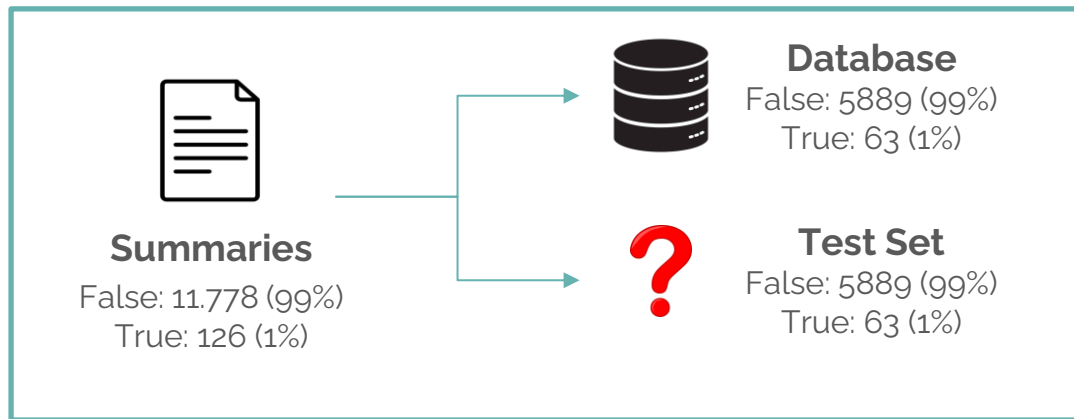| ID	Repair_Code	Kit_Size_num	Label	Kit_Compliance	Damage_Type	Pipe_Exposure	Severe_Corrosion	Pipe_Covered	Damaged_Valve	Summary	Summary_1	Kit_Size_str
ID000001	INT000001	5	FALSE	The size and ty	Non-sheared lin	Aerial pipe	TRUE	TRUE	TRUE	The aerial	The aerial stee	XL
ID000002	INT000002	5	FALSE	The size and ty	Non-sheared lin	Aerial pipe	TRUE	TRUE	TRUE	A non-shea	A localized line	XL
ID000003	INT000003	5	FALSE	The size and ty	Non-sheared lin	Aerial pipe	TRUE	TRUE	FALSE	Aerial stee	Aerial steel pip	XL
ID000004	INT000004	5	FALSE	The size and ty	Non-sheared lin	Aerial pipe	TRUE	TRUE	FALSE	The aerial	The steel aeria	XL

Collect and
Organize in a CSV

Data Collection

ID	Repair_Code	Kit_Size_num	Label	Kit_Compliance	Damage_Type	Pipe_Exposure	Severe_Corrosion	Pipe_Covered	Damaged_Valve	Summary	Summary_1	Kit_Size_str
ID000001	INT000001	5	FALSE	The size and ty	Non-sheared lin	Aerial pipe	TRUE	TRUE	TRUE	The aerial	The aerial stee	XL
ID000002	INT000002	5	FALSE	The size and ty	Non-sheared lin	Aerial pipe	TRUE	TRUE	TRUE	A non-shea	A localized line	XL
ID000003	INT000003	5	FALSE	The size and ty	Non-sheared lin	Aerial pipe	TRUE	TRUE	FALSE	Aerial stee	Aerial steel pip	XL
ID000004	INT000004	5	FALSE	The size and ty	Non-sheared lin	Aerial pipe	TRUE	TRUE	FALSE	The aerial	The steel aeria	XL

Features considered

Two Setup

Half-Half Split

Summaries
False: 11.778 (99%)
True: 126 (1%)

Database
False: 5889 (99%)
True: 63 (1%)

Test Set
False: 5889 (99%)
True: 63 (1%)

Rephrased Split

Summaries and **Rephrased**

Database
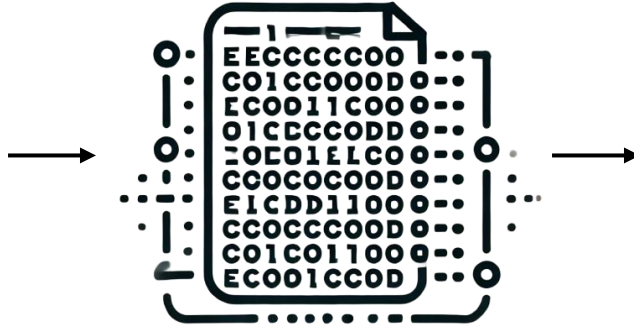False: 11.778
True: 126

Test Set
False: .11.778.
True: 126
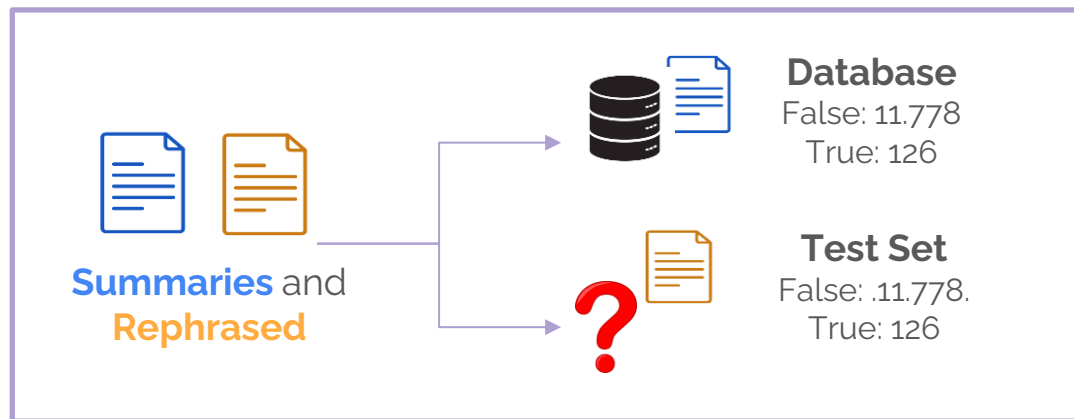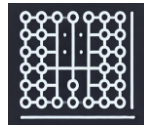
Methodology

Generating Embeddings

The aerial steel pipe has a non-sheared linear…..

Steel aerial pipe suffers from a non-sheared l…..

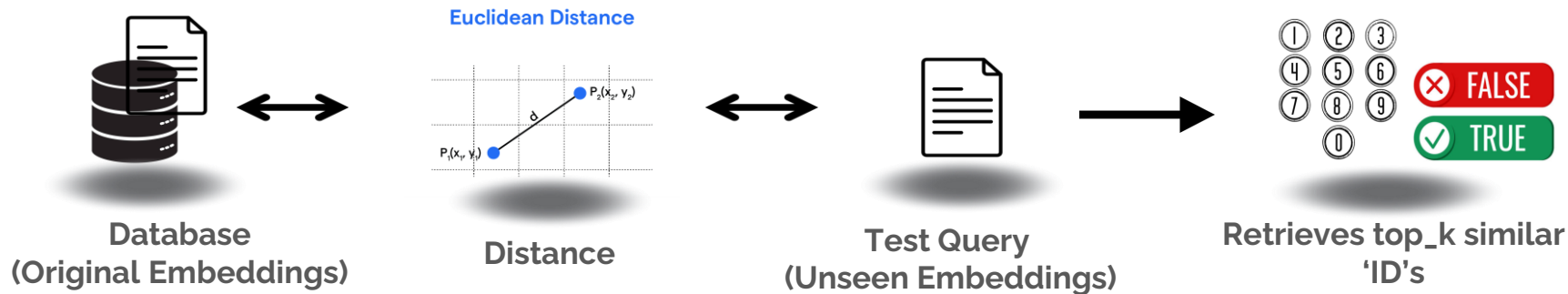Textual description

Sample text

SentenceEncoder (SBERT / SRoBERTa)
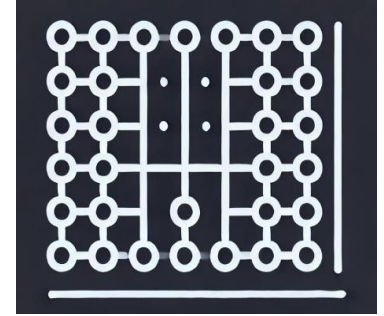
Converting to a fixed-size vector embedding

Similarity Checking-FAISS

Database
(Original Embeddings)

Euclidean Distance

$P_2(x_2, y_2)$

$P_1(x_1, y_1)$

Distance

Test Query
(Unseen Embeddings)

FALSE

TRUE

Retrieves top_k similar
'ID's

Organizing Embeddings in DataFrames

Database

Test Query Set

ID	Embeddings	Successful
ID000064	[-0.001362897455692913, 0.9544798135757446, 0.0...]	TRUE
ID000128	[1.3155205249786377, -0.04855910316109657, 0.0...]	FALSE
ID000192	[-1.1072819232490674, 0.05280644819140434, 0.18...]	FALSE
ID000453	[1.0781232490879674, 0.5228344819140434, 0.23...]	TRUE
ID007895	[-0.023719232490674, 0.01652804819140434, 0.39...]	TRUE

Save as CSV Files

Predictions
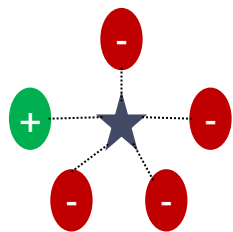
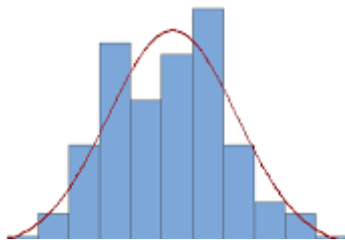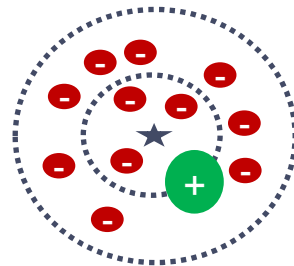Distance-based
(RAC)

LLM Generation
(RAG)

RAC: policies
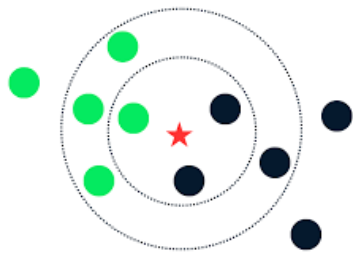
KNN

Sampling

Rarity

RAG: generative models

Llama3.2-3B

Mistral-7B

Llama2-13B-Chat

Do we need RAG?

| ID	Successful	
968	True	The aerial steel pipe exhibits rose-s
712	True	A polyethylene pipe suffered a hol
203	False	A non-sheared linear lesion has deve
064	True	Aerial pipe non-sheared linear lesio
403	False	The pipe sustains rose holes injuries
463	False	Aerial polyethylene pipe shows sign

Zero-Shot

Fixed-Examples

Evaluation
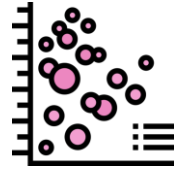
SBERT vs. SRoBERTa

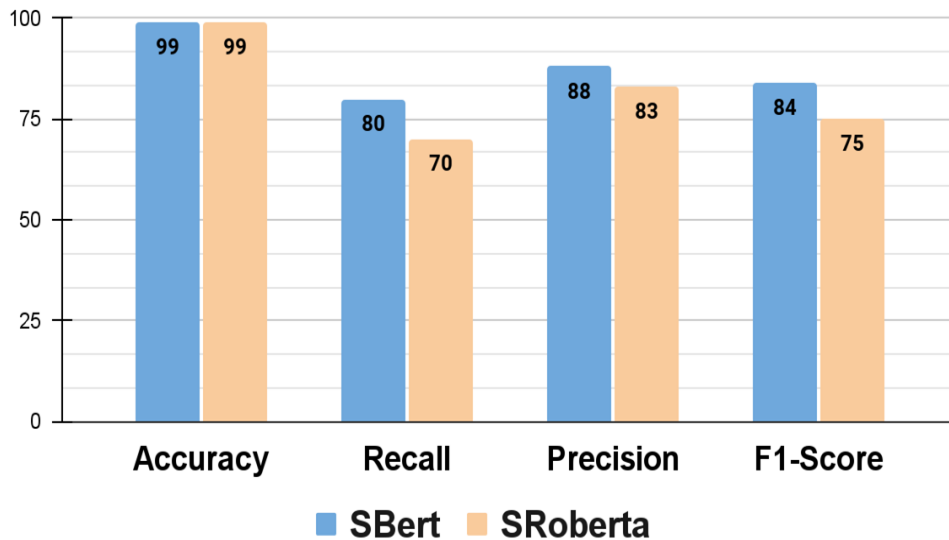**Comparative metrics Graph** for Classification of Labels

Setup 1 : Half-Half Split
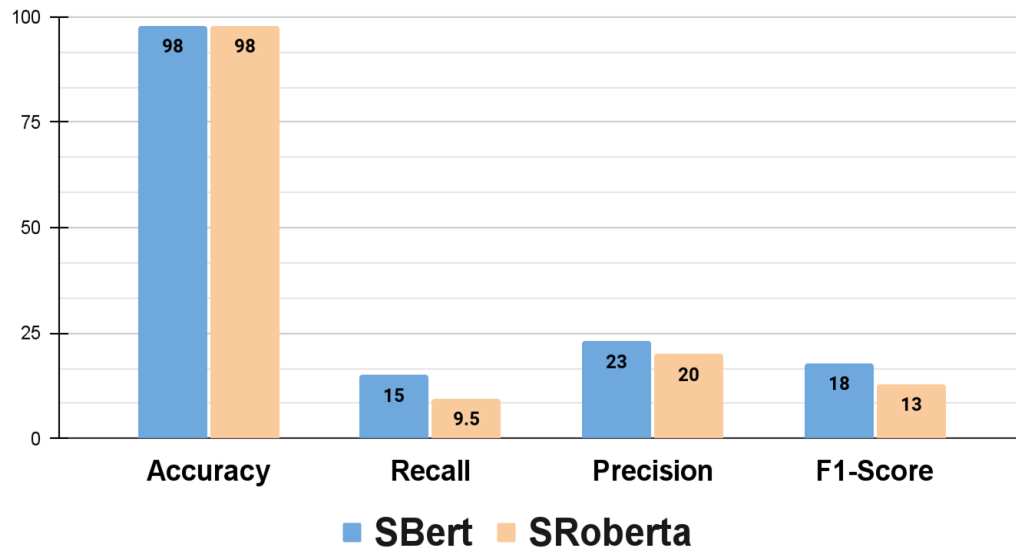
SBERT vs. SRoBERTa

Comparative metrics Graph for Classification of Labels

Setup 2 : Rephrased Split

Test Set Distribution

False (5889)

True (63)

Original Distribution

Test Set Distribution
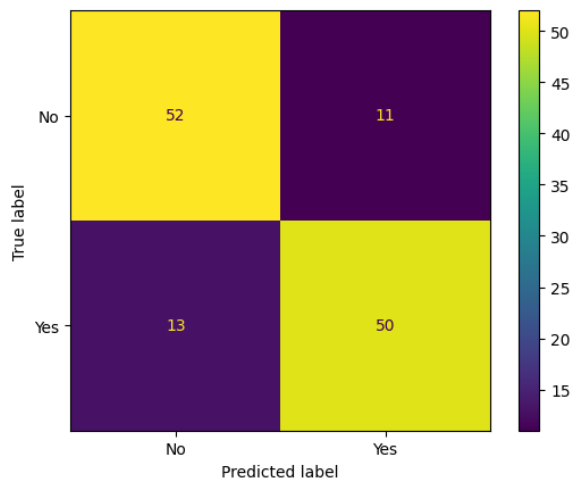
False (63)

True (63)
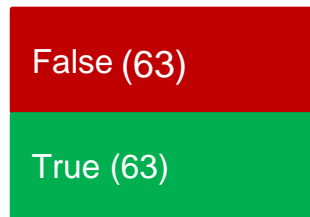
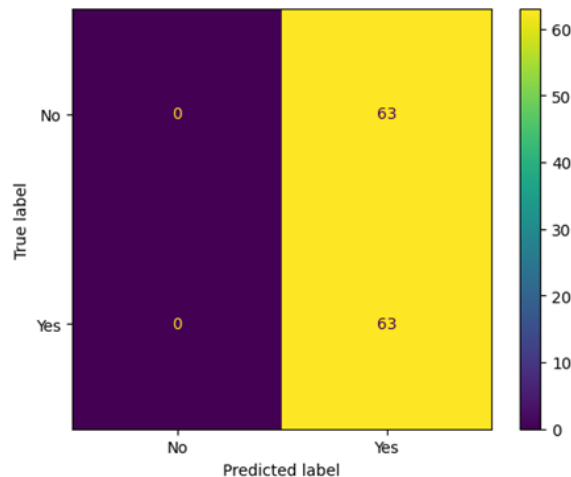Downsampled
Distribution

Zero-Shot

- LLaMa**3.2-3B** seems to understand the concern even without examples.
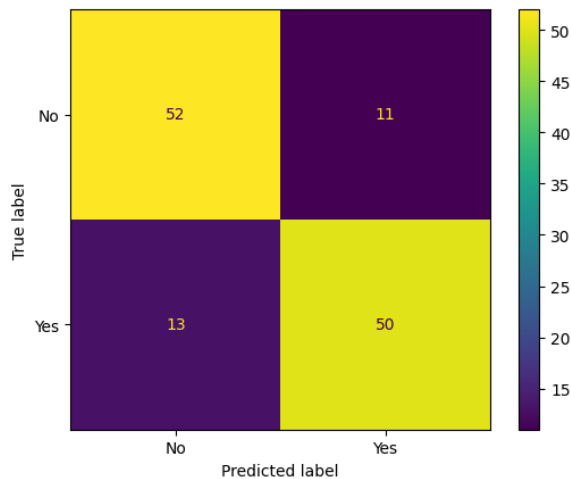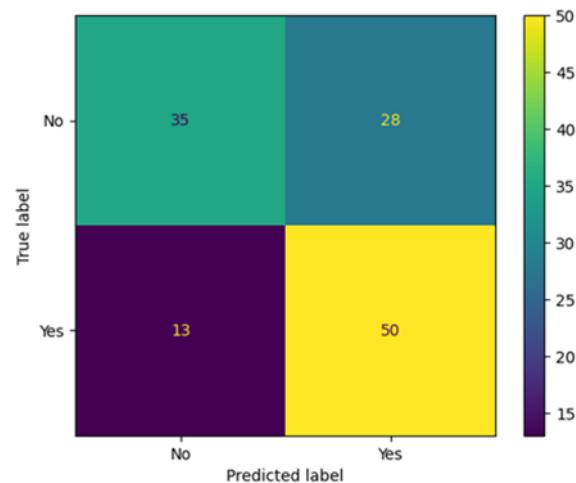

Llama3.2-3B

Llama2-13B-Chat

Zero-Shot vs Few-Shot Fixed-Examples

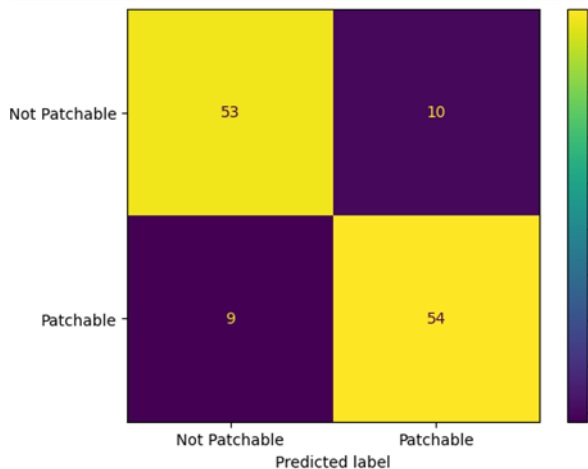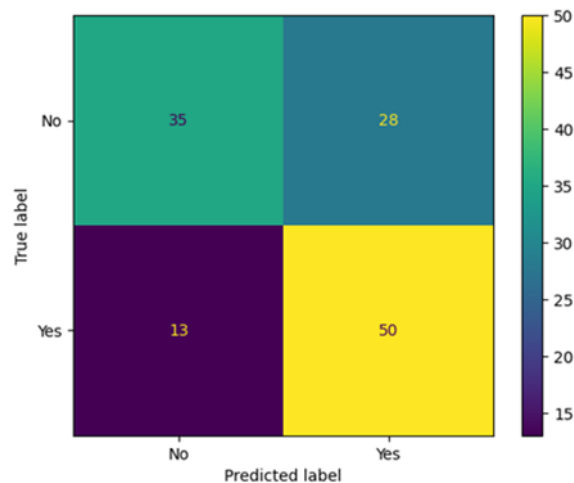- LLaMa3.2-3B seems confused by examples.

Zero-Shot

Few-Shot (FE)

Few-Shot: label names effect

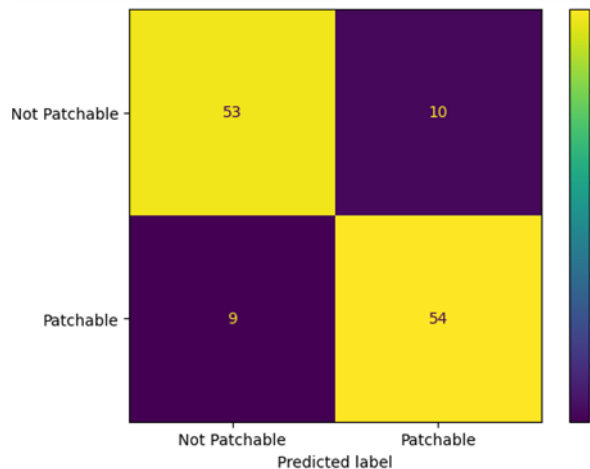- Changing label names **seemed** to help the model to get back on track.
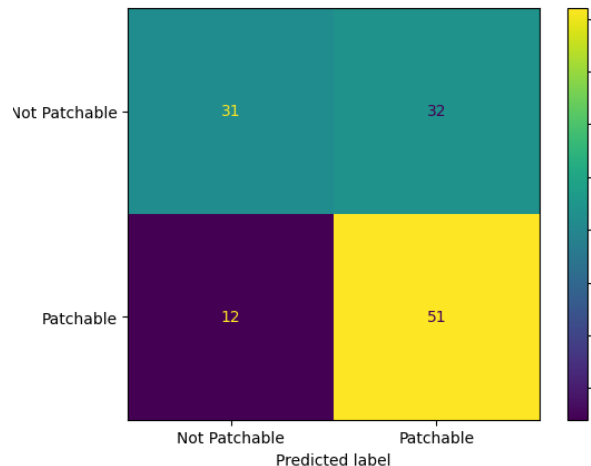

Patchable/Not

Yes/No

Few-Shot: Fixed-Examples vs RAG (LLama3.2-3B)

- Behaviour observed **regardless** label names and k (here, k= 6).
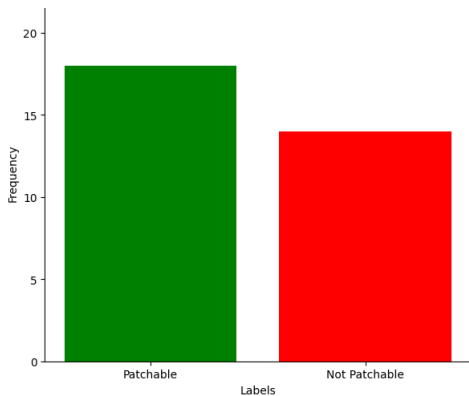

Fixed-Examples

RAG

Prompt Sensitivity

- Llama3.2-3B is asked to label the **same example** 32 times. This behaviour is **not observed** in other models. Results in **zero-shot** setting.

Patchable/Not
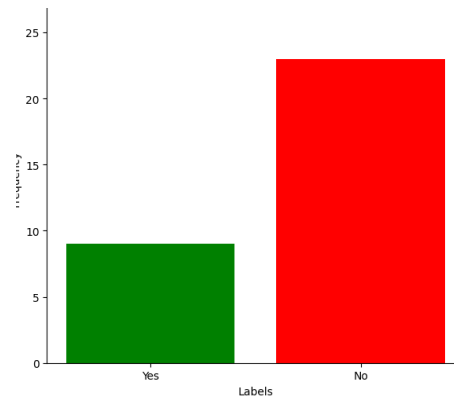
Yes/No

LLaMa3.2-3B Considerations

Llama3.2-3B

- The model seems understand the task without examples;
- Examples confuse the model;
- Models' answers are too chaotic.
- The reason of chaos may lie in its small size.
- The larger Llama3.2 version cannot be tested on Colab (70B)

Few-Shot: Fixed-Examples vs RAG (LLaMa2-13B-Chat)

- The rarity of positive skews models answer towards negative predictions (k=3).
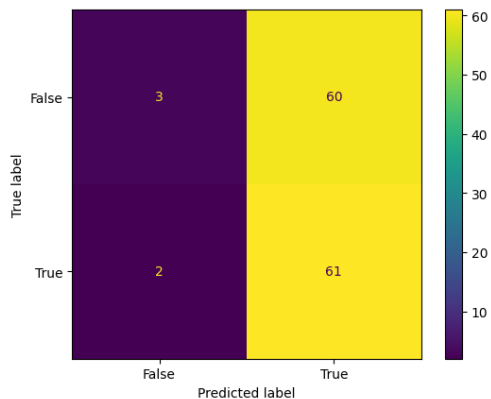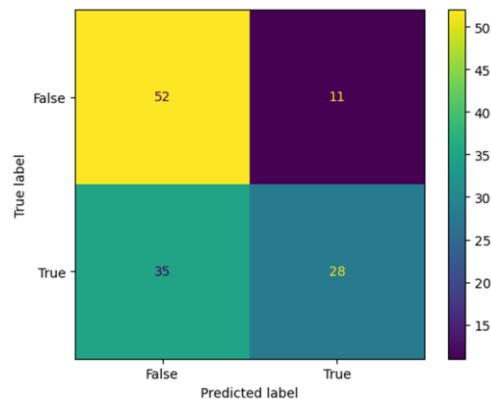
Fixed-Examples

RAG

Few-Shot: Fixed-Examples vs RAG (Mistral-7B)

- The rarity of positive skews models answer towards negative predictions (k=3).
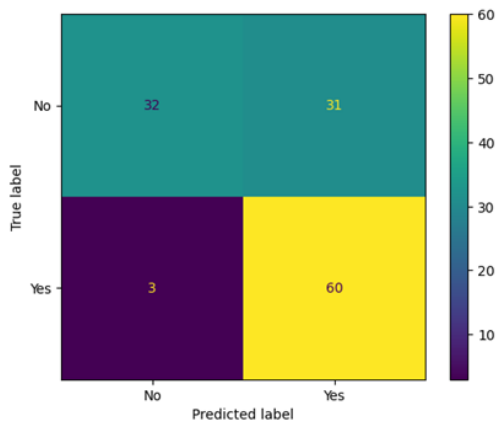
Fixed-Examples
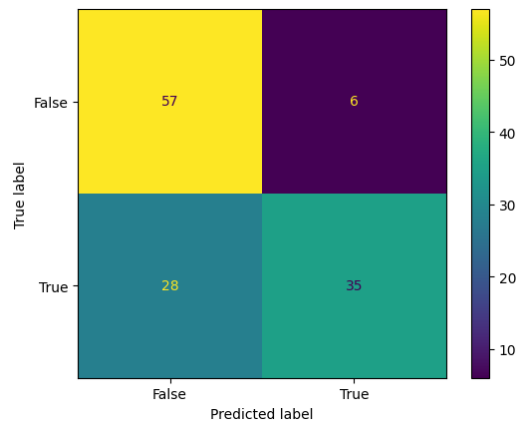
RAG

Few-Shot: Split vs Joint Examples (Mistral 7B)

- Retrieving k-positive closer examples and k-negative closer examples does not help either (joint_k=6, split_k=3).
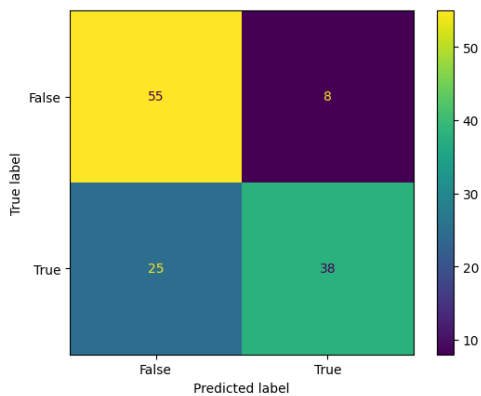
Split
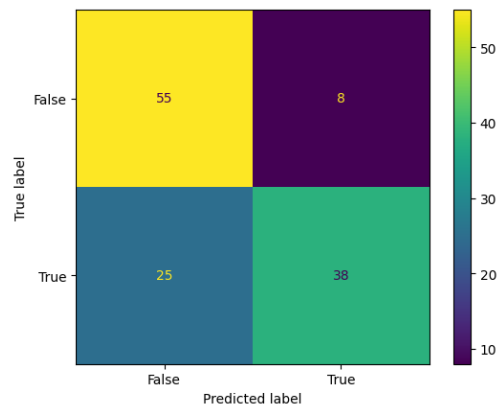
Joint

Larger Models: considerations
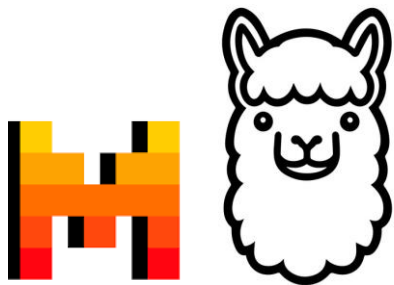
Larger Models

- Answer **stability** goes with model size
- **RAG helps** the model in labelling
- The **scarcity** of positive examples may affect performances (high FN).

RAC: results

- RAC enforces our intuition: the problem is with the scarcity of (+) reparations.
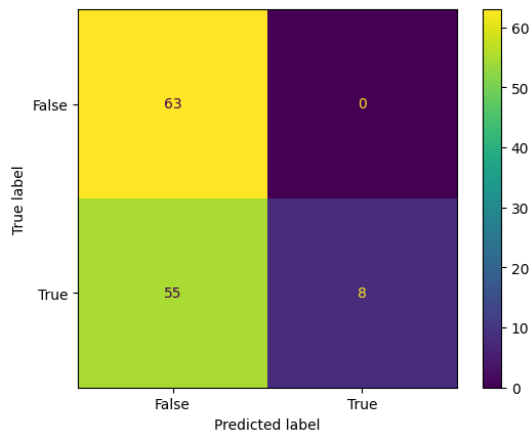
KNN
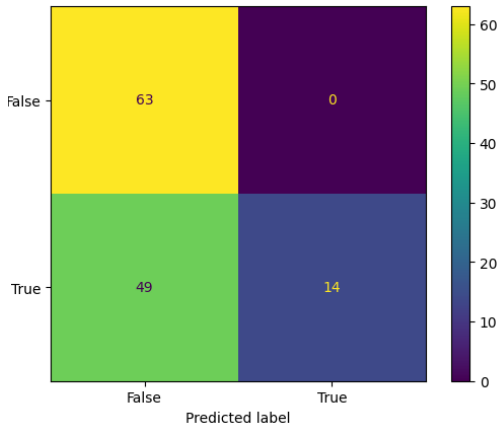
Sampling
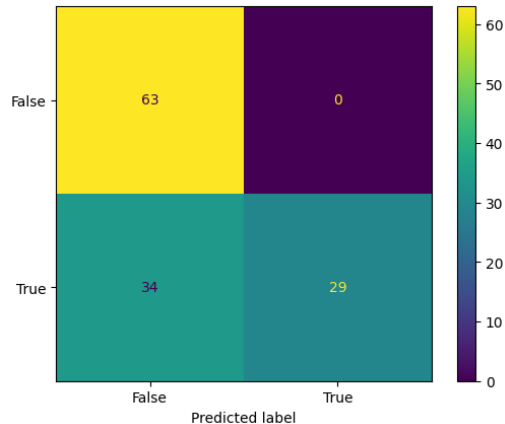
Rarity

What's next?
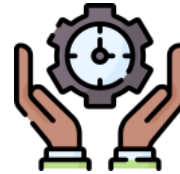
Synthetic Data

- Following our intuition about the *missing-successful* concern, we aim to populate RAG-DB with synthetic-generated positive examples;
- Very-Large-Language-Model such as GPT-4 may be help for generate new *patchable* descriptions;
- GPT-4 should first tested on our test-set in either a zero-shot or few-shot setting. In case of success (99% accuracy), we could consider GPT-4 examples reliable then proceed with synthetic-examples

ChatBot: Explanation

- Further than labelling the example, the chat-bot will also be asked to generate an explanation for the choice;
- We *do not have* examples of reasoning;
- *GPT-4* may judge our chatbot rationale.

ChatBot: Language

- Since CoRe and ItalGas are Italian companies, the main interaction with the chatbot may be in Italian.
- We may think to mount Italian-English translators up/down stream our chatbot

THANK YOU!

Grazie!

ধন্যবাদ

شکریہ

Extra

Zero-Shot Prompt

You are a knowledgeable pipeline maintenance assistant with expertise in gas pipe repairs.
Your task is to determine if a damaged gas pipe is repairable with patches.

Description:
{*description*}   **Input**

Respond with *"{positive_label}"* if the damage can be patched, or *"{negative_label}"* if it cannot. Only provide the label, no explanation.

Few-Shot Prompt

You are a pipeline maintenance assistant specializing in gas pipe repairs. **SYS**
Your task is to decide if a damaged gas pipe can be repaired with patches.

Here some examples:

{examples}

Description: {description}
Label: {label}

RAG/Fixed

Basing your decision on the provided examples, complete by answering either:
- *{positive_label}* if the damage can be repaired by using patches;
- *{negative_label}* if the damage cannot be repaired by using patches.
Answer only and solely the label, do not write other words. **SYS**

Query:
{query}

Description: {description}

Label:

Input

Query and Examples Format

- Examples:
  - Description: {description}
    Label: {label}
- Query:
  - Description: {description}
    Label:

Model Input Example

You are a knowledgable pipeline maintenance assistant specializing in gas pipe repairs.
Your task is to decide if a damaged gas pipe can be repaired with patches.

Here some examples:

Description: The aerial steel pipe exhibits rose-shaped holes in its surface, indicating a faulty section. No high pressure is present within the pipe, and there are no signs of strong corrosion or bad conditions. Additionally, there is no branch near the break, and the pipe is not covered by a wall. There are no ribs present either. This suggests that the damage is localized to this particular area, with no broader systemic issues.
Label: Yes
[…] x (top_k – 1)

Basing your decision on the provided examples, complete by answering either:
- Yes if the damage can be repaired by using patches;
- No if the damage cannot be repaired by using patches.
Answer only and solely the label, do not write other words.

Query:
Description: Aerial pipe non-sheared linear lesion detected at Bad Conditions with Steel material.
Label:

Prompt Sensitivity

'Successful' = False:
The aerial steel pipe has a non-sheared linear lesion caused by strong corrosion, despite high pressure within the pipe. The injury is located near a branch and is covered by a wall, with the presence of a valve close to the break. The pipe also features ribs, indicating that it may be more susceptible to corrosion.
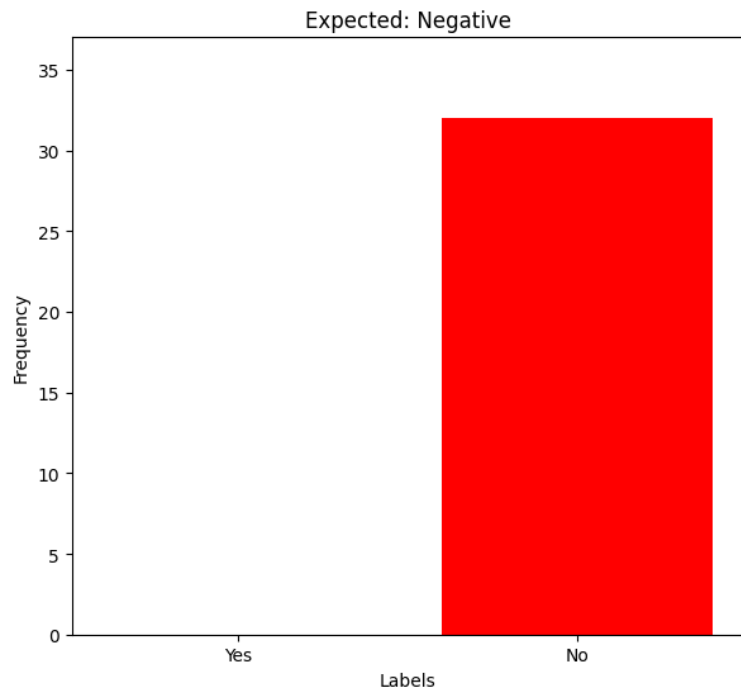
'Successful' = True:
Aerial pipe non-sheared linear lesion detected at Bad Conditions with Steel material.
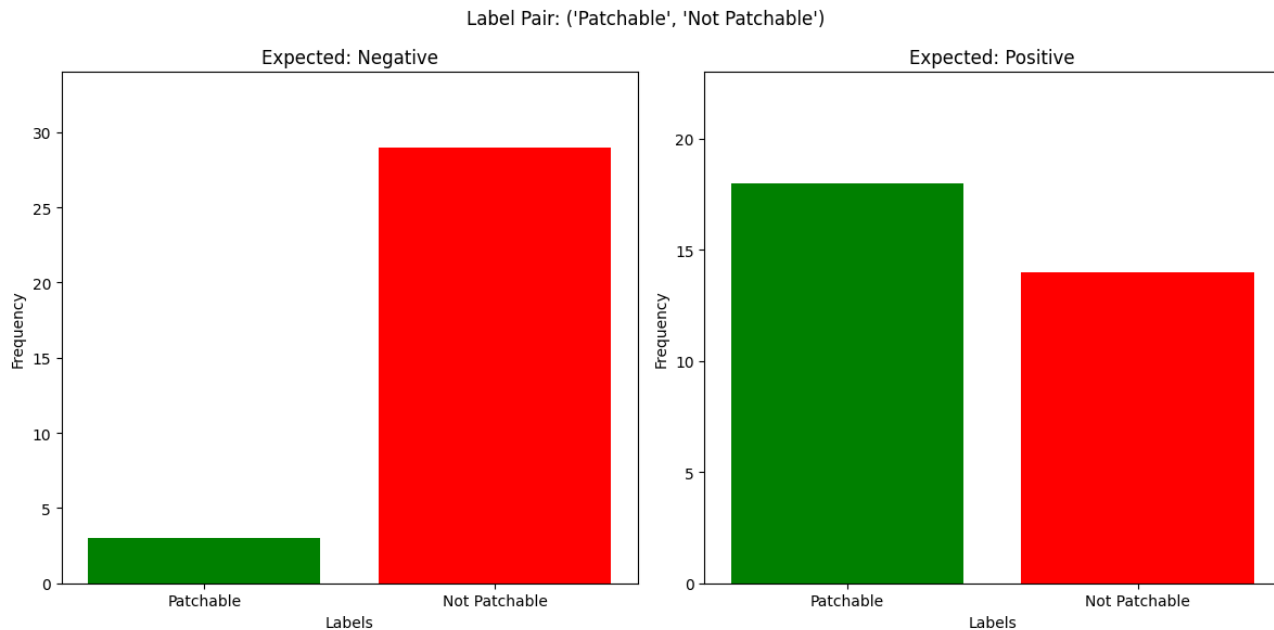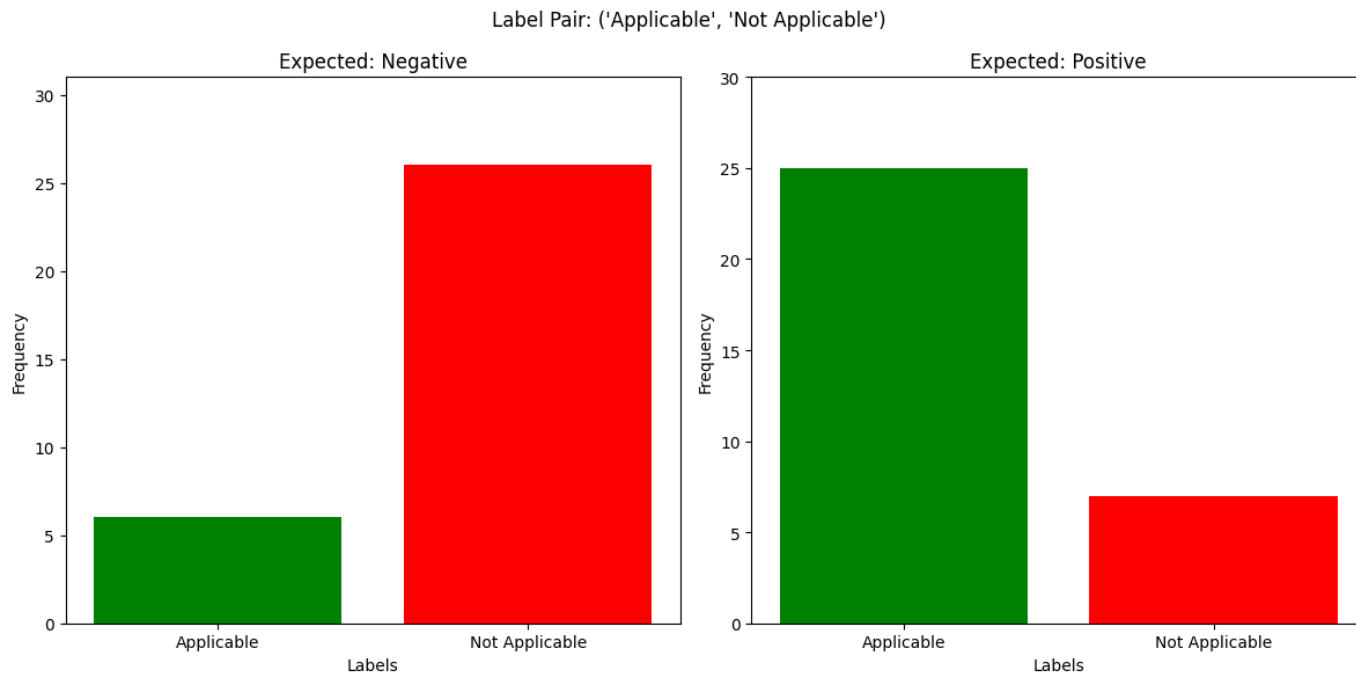
Yes/No

Patchable/Not Patchable

Label Pair: ('Patchable', 'Not Patchable')

Applicable/Not Applicable

Label Pair: ('Applicable', 'Not Applicable')

True/False

Label Pair: ('True', 'False')