COMUNICAÇÃO TÉCNICA

Nº 178733

Application of machine learning tecnhiques for identiying emergency conditions in rainwater galleries

Cristina Maria Ferreira da Silva
Adriano Galindo Leal
Luciana Andrea Mori Faria de Moraes
Felipe Silva Silles
Celso Luciano Alves da Silva
Vagner Luiz Gava

Palestra apresentada no IBERO-LATIN CONGRESS ON COMPUTATIONAL METHODS IN ENGINEERING, 44., 2023, Porto
17 slides.

A série "Comunicação Técnica" compreende trabalhos elaborados por técnicos do IPT, palestras apresentadas, apresentados em eventos, publicados em revistas especializadas ou quando seu conteúdo apresentar relevância pública. **PROIBIDO A REPRODUÇÃO, APENAS PARA CONSULTA**.

Instituto de Pesquisas Tecnológicas do Estado de São Paulo S/A - IPT
Av. Prof. Almeida Prado, 532 | Cidade Universitária ou Caixa Postal 0141 | CEP 01064-970
São Paulo | SP | Brasil | CEP 05508-901
Tel 11 3767 4374/4000 | Fax 11 3767-4099

www.ipt.br

XLIV Ibero-Latin American Congress on Computational Methods in Engineering
Porto – Portugal, November 13-16, 2023

APPLICATION OF MACHINE LEARNING TECHNIQUES FOR IDENTIFYING EMERGENCY CONDITIONS IN RAINWATER GALLERIES

Cristina Maria Ferreira da Silva, Adriano Galindo Leal, Luciana Andrea Mori Faria de Moraes, Felipe Silva Silles, Celso Luciano Alves da Silva, **Vagner Luiz Gava**

Artificial Intelligence and Analytics Department – Digital Technologies
IPT – Institute for Technological Research
São Paulo, Brazil

SUMMARY

- 1 Motivation and context
- 2 The methodology applied to the case study
- 3 Machine learning algorithms applied
- 4 Conclusions

1 MOTIVATION AND CONTEXT

- From visual inspection results, identify rainwater drainage galleries in emergency conditions, where observed anomalies with a higher level of criticality and require interventions with urgency.
- In present study, the term "galleries" refers to structures with a span of less than 5 meters, including rainwater galleries, cattle crossings, and underpasses, located in Brazilian highways.

Figure 1. Examples of rainwater gallery inspected (source: IPT)

1 MOTIVATION AND CONTEXT

Examples of works inspected by IPT were named “galleries” in presente study (source IPT)

Cattle crossing

Rainwater gallery with two cells silted (Type 1)

Underpass with a span of less than 5 meters

1 MOTIVATION AND CONTEXT

Examples of rainwater galleries with anomalies (source IPT)

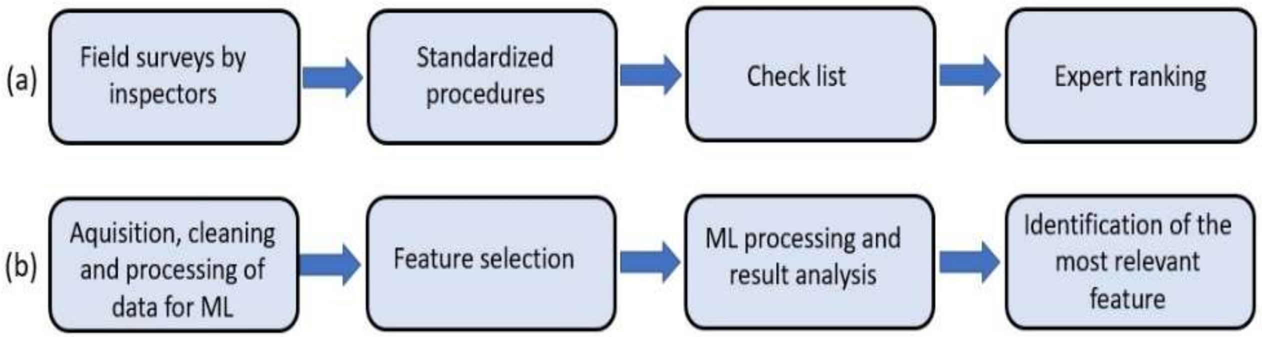
Area with severe erosion on slopes (Type 5)

Arrows pointed erosion under gallery (Type 4)

Flood gallery (Type 1)

2 THE METHODOLOGY APPLIED TO THE CASE STUDY

- Key steps with adding AI methods to analyses of rainwater technical inspection results:



Usual procedures in routine inspection of galleries

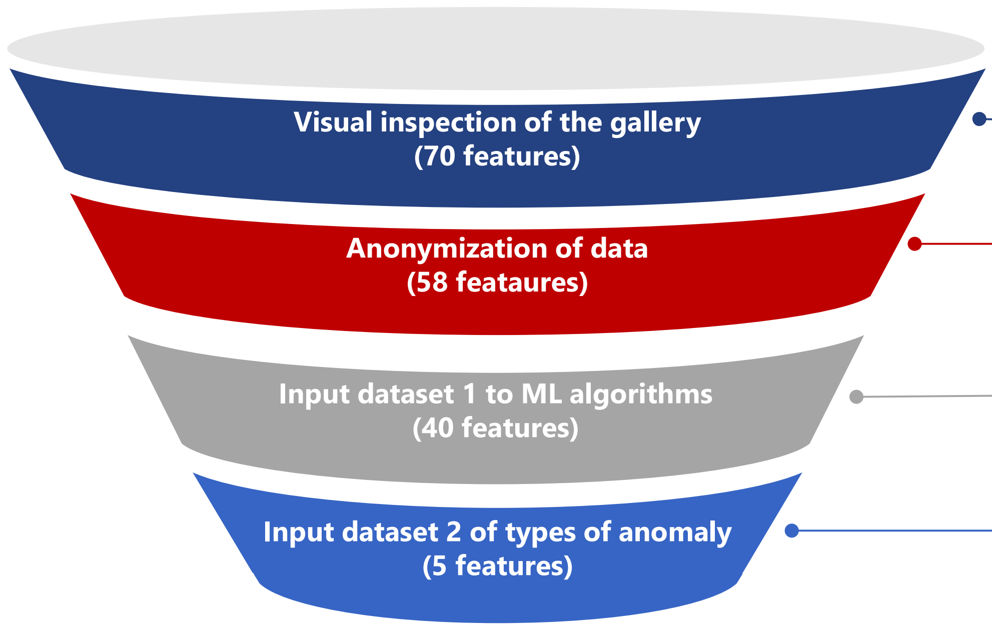
Applying ML algorithms to classify galleries with anomalies in higher level of criticality

Figure 2. Activities performed in the visual inspections (a) and in machine learning processes (b)

- Extract input datasets to case study from results of the visual technical inspection made by IPT comprising about 4000 rainwater galleries in Brazilian highways.

2 THE METHODOLOGY APPLIED TO THE CASE STUDY

- Obtaining data about the rainwater technical inspection:
  How the input datasets were composed?



Checklist of visual inspection with 70 features observed in the field by technical teams

Reduction of the dataset to 58 features, eliminating information about localization, highway, etc. → Shown in Table 1

Reduction of the dataset to 40 features, removing that not decisive in compromising the functioning of the gallery or its structural integrity (sidewalk, guardrail, wheel guard, etc.)

Grouping anomalies in 5 types that occurred more frequently and can compromise structural safety (silting, clogging, flooding, severe erosion, etc.) → Detailed in Table 2

2 THE METHODOLOGY APPLIED TO THE CASE STUDY

- From the checklist of visual inspection of galleries, input datasets were obtained to applied machine learning (ML) algorithms, and the classification of status made by specialists

Results of visual inspection in galleries

Table 1. Features collected in visual inspections of rainwater galleries

Feature index	Feature	Description
1	Gallery identifier	Sequential integer
2	Type of gallery	Rainwater gallery, cattle passage, or underpass
3	Material	Concrete, metallic, or both materials
4, 5	Dimensions	The diameter or transversal dimension (4), length under the highway (5), both in centimeters
6	Cross-section shape	Integer number associated with the format of the gallery, for example, circular, rectangular, among others, considered in this study
7	Depth	The estimated height of the soil layer or pavement above the top surface of the gallery in centimeters
8, 9, 10	The cross-section in concrete and damage observation	The existence or not of concrete elements in cross-section (8), cracking (9), and surface damages (10)
11, 12, 13	Metal piping and damage observation	The existence or not of metallic elements in cross-section (11), deformation (12), corrosion (13)
14 to 26	Anomalies related to the functioning of the gallery	13 features indicating an occurrence of anomalies at extremities and their intensity, such as silting, clogging, and flooding, among others
27 to 31	Concrete retaining walls	The existence or not of retaining walls at the extremities of the gallery and the occurrence of cracks and damage in the surface of the concrete
32 to 36	Runway drainage	The existence or not of drainage components, conditions of cleaning, breaking, or puddles of water
37 to 40	Pavement	The existence or not of pavement on the runway, ripples, damage, and steps on the pavement
41 to 48	Sidewalk	The existence or not of sidewalks, cleaning, and damage
49 to 52	Guardrail	The existence or not of guardrails, alignment, and damage in concrete and metallic components
53, 54	Wheel guard	The existence or not of wheel guards and alignment of elements
55 to 58	Barriers	The existence or not of barriers, alignment, and damage in concrete and metallic components

Table 2. Variables obtained from the features presented in Tab. 1

Variable index	Variable name	Description of anomalies related	Associated features (*)
1	Type 1	Silting, clogging, or flooding with more than 50% of the cross-section obstructed	14, 17 to 26
2	Type 2	Structural damages in piping or ground retaining walls	8, 9, 10, 27 to 31
3	Type 3	Places with corrosion or deformation in metallic gallery	11, 12, 13
4	Type 4	Observation of severe erosion under the gallery	15, 16
5	Type 5	The existence of areas with severe erosion on slopes	32 to 36

(*) The feature index is shown in Tab. 1.

2 THE METHODOLOGY APPLIED TO THE CASE STUDY

14 to 26	Anomalies related to the functioning of the gallery	13 features indicating an occurrence of anomalies at extremities and their intensity, such as silting, clogging, and flooding, among others
27 to 31	Concrete retaining walls	The existence or not of retaining walls at the extremities of the gallery and the occurrence of cracks and damage in the surface of the concrete
32 to 36	Runway drainage	The existence or not of drainage components, conditions of cleaning, breaking, or puddles of water

Table 1. Variables obtained from the features presented in Tab. 1

Variable index	Variable name	Description of anomalies related	Associated features (*)
1	Type 1	Silting, clogging, or flooding with more than 50% of the cross-section obstructed	**14, 17 to 26**
2	Type 2	Structural damages in piping or ground retaining walls	8, 9, 10, 27 to 31
3	Type 3	Places with corrosion or deformation in metallic gallery	11, 12, 13
4	Type 4	Observation of severe erosion under the gallery	15, 16
5	Type 5	The existence of areas with severe erosion on slopes	32 to 36

3 MACHINE LEARNING ALGORITHMS APPLIED

- For over a decade, Big Data and AI algorithms have been used to predict deterioration or damage in civil works from inspection and investigation.
- That activities are intensive, time-consuming, subjective, and costly.

3.1 Conditions considered to apply machine learning algorithms

  - Among the most cited in the literature, eight ML algorithms were selected for application: Random Forest (RF), Gradient Boosting (GB), Logistic Regression (LR), Support Vector Machine (SVM), Naive Bayes (NB), AdaBoost (AB), Decision Tree (DT), and XGBoost (XB).
  - Aiming to classify the state of the galleries into emergency and non-emergency, regarding the intensity of anomalies and the prioritization of interventions.
  - The input datasets were split into 80/20, meaning 80% for training with cross-validation and 20% for testing.

3.2 Phases of study

  - First phase: One input dataset was extracted with 40 features from Table 1, after cleaning and correcting of data
  - Second phase: Five features were defined from the checklist variables and indicated by the inspection specialists (Types 1 to 5 in Table 2), corresponding to anomalies considered the most relevant for evaluating the state of the galleries.

3 MACHINE LEARNING ALGORITHMS APPLIED

3.3 Results of phases and considerations

Table 3. Algorithms results considering the classification of emergency state

Algorithm	Input dataset with 40 features					Input dataset with 5 features				
	Accuracy	Non-Emergency		Emergency		Accuracy	Non-Emergency		Emergency	
		True Positive	False Positive	True Negative	False Negative		True Positive	False Positive	True Negative	False Negative
RF	0.92	710	16	47	54	1.00	726	0	101	0
SVM	0.87	712	14	9	92	1.00	726	0	101	0
LR	0.90	706	20	37	64	1.00	726	0	101	0
GB	0.91	706	20	48	53	1.00	726	0	101	0
NB	0.76	576	150	54	47	1.00	726	0	101	0
DT	0.90	690	36	45	56	1.00	726	0	101	0
AB	0.91	711	15	57	44	1.00	726	0	101	0
XB	0.91	693	33	44	57	1.00	726	0	100	1

First phase

Second phase

3 MACHINE LEARNING ALGORITHMS APPLIED

3.3 Results of phases and considerations

- Focusing on the study and improving operational procedures in visual inspections (Figure 1), the relevance of the input dataset with 5 features was verified.
- As shown in Table 3 for the 8 implemented ML algorithms, the set with 5 features presented better results than 40 features, without classification errors in the respective confusion matrix obtained.
- Achieving 100% accuracy in the classification results based on 5 variables, other aspects should be considered regarding of possibility of overfitting:
  - There is a deterministic link between main inputs and outputs better representing the studied phenomenology, aligning with the problem type and experts' preference for a limited variable set as indicators of criticality level of anomalies.
  - The model validation tests showed no overfitting issues and demonstrated appropriate generalization.
  - It is recommended that further investigation and validation, including testing on independent datasets, would be valuable to corroborate the robustness of the model's performance.

3 MACHINE LEARNING ALGORITHMS APPLIED

3.4 Verifying the essential features using reduction and machine learning

- Using the Mean Decrease in Impurity (MDI) method <u>to determine the relevance among the 5 features.</u>
- Remembering these features corresponding to the types defined by specialists from the 40 features, in predicting the emergency state of galleries.
- Only the Random Forest algorithm was used on the MDI investigation with a training and testing split of 70/30.

3 MACHINE LEARNING ALGORITHMS APPLIED

- The figure below displays the percentage distribution and confusion matrix of the 5 features, which produced 100% accuracy.

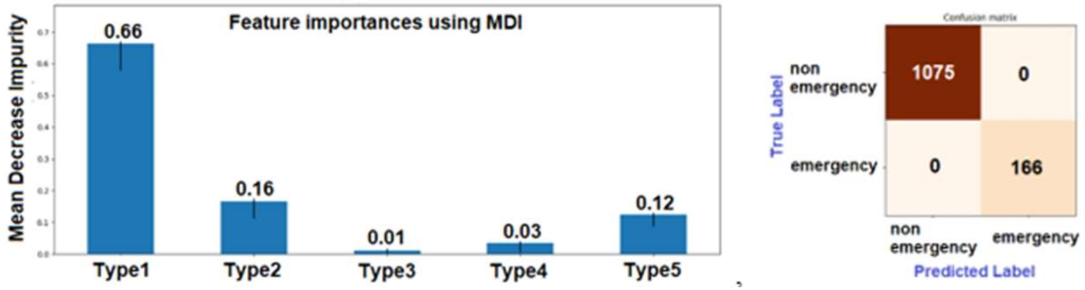


Figure 3. ML with 5 features and its confusion matrix without errors

- This occurs because the aggregation of these variables in the 5 types represents with more accuracy the phenomenology studied.
- The study aims to prioritize the structural safety of the galleries, makes it possible to identify which technical activities can be improved, from field surveys to analyses carried out by specialists, with resource savings and without losing the quality and reliability of results.

4 CONCLUSIONS

The use of ML algorithms to support the treatment of visual inspection results to specifically identify rainwater drainage galleries, where anomalies with a higher level of criticality were observed and requiring more urgent interventions, proved to be interesting in the following aspects:

- **Feature Importance:** the performance of the algorithms varied between the input datasets considered, one with 40 features and the other with 5;
- **Efficiency:** applying ML algorithms enables the rapid and effective identification of galleries requiring intervention with more priority;
- **Safety and Risk Mitigation:** the study approach allows scheduling maintenance and more timely intervention in civil works, increasing road users' safety;
- **Cost-Effectiveness:** based on the verification of the relevance of the features, it is possible to outline training strategies for the technical teams to carry out the visual inspections.

4 CONCLUSIONS

- This research demonstrates the technical practicality of applying ML algorithms and data science techniques, which can be integrated into structural safety assessments of galleries, in alignment with the digital transformation era.
- For the case study examined, it was relevant to identify the primary variables representing the phenomenology and establish the deterministic relationship between them and the output.
- In future works, consider replicating this approach to analyze data from visual inspections in other civil works, like bridges and buildings.

Thank you!

- Vagner Luiz Gava
- vlgava@ipt.br

linkedin.com/school/iptsp/

instagram.com/ipt_oficial/

youtube.com/@IPTbr/

www.ipt.br

INSTITUTO DE PESQUISAS TECNOLÓGICAS

SÃO PAULO GOVERNO DO ESTADO