WIKIPEDIA
The Free Encyclopedia

Confusion matrix

In the field of machine learning and specifically the problem of statistical classification, a **confusion matrix**, also known as **error matrix**,[11] is a specific table layout that allows visualization of the performance of an algorithm, typically a supervised learning one; in unsupervised learning it is usually called a **matching matrix**.

Each row of the matrix represents the instances in an actual class while each column represents the instances in a predicted class, or vice versa – both variants are found in the literature.[12] The name stems from the fact that it makes it easy to see whether the system is confusing two classes (i.e. commonly mislabeling one as another).

It is a special kind of contingency table, with two dimensions ("actual" and "predicted"), and identical sets of "classes" in both dimensions (each combination of dimension and class is a variable in the contingency table).

Example

Given a sample of 12 individuals, 8 that have been diagnosed with cancer and 4 that are cancer-free, where individuals with cancer belong to class 1 (positive) and non-cancer individuals belong to class 0 (negative), we can display that data as follows:

Individual Number	1	2	3	4	5	6	7	8	9	10	11	12
Actual Classification	1	1	1	1	1	1	1	1	0	0	0	0

Assume that we have a classifier that distinguishes between individuals with and without cancer in some way, we can take the 12 individuals and run them through the classifier. The classifier then makes 9 accurate predictions and misses 3: 2 individuals with cancer wrongly predicted as being cancer-free (sample 1 and 2), and 1 person without cancer that is wrongly predicted to have cancer (sample 9).

Individual Number	1	2	3	4	5	6	7	8	9	10	11	12
Actual Classification	1	1	1	1	1	1	1	1	0	0	0	0
Predicted Classification	0	0	1	1	1	1	1	1	1	0	0	0

Notice, that if we compare the actual classification set to the predicted classification set, there are 4 different outcomes that could result in any particular column. One, if the actual classification is positive and the predicted classification is positive (1,1), this is called a true positive result because the positive sample was correctly identified by the classifier. Two, if the actual classification is positive and the predicted classification is negative (1,0), this is called a false negative result because the positive sample is incorrectly identified by the classifier as being negative. Third, if the actual classification is negative and the predicted classification is positive (0,1), this is called a false positive result because the negative sample is incorrectly identified by the classifier as being positive. Fourth, if the actual classification is negative and the predicted classification is negative (0,0), this is called a true negative result because the negative sample gets correctly identified by the classifier.

We can then perform the comparison between actual and predicted classifications and add this information to the table, making correct results appear in green so they are more easily identifiable.

Terminology and derivations from a confusion matrix

condition positive (P)
     the number of real positive cases in the data
condition negative (N)
     the number of real negative cases in the data

---

true positive (TP)
     A test result that correctly indicates the presence of a condition or characteristic
true negative (TN)
     A test result that correctly indicates the absence of a condition or characteristic
false positive (FP), Type I error
     A test result which wrongly indicates that a particular condition or attribute is present
false negative (FN), Type II error
     A test result which wrongly indicates that a particular condition or attribute is absent

---

sensitivity, recall, hit rate, or true positive rate (TPR)
$$TPR = \frac{TP}{P} = \frac{TP}{TP + FN} = 1 - FNR$$

specificity, selectivity or true negative rate (TNR)
$$TNR = \frac{TN}{N} = \frac{TN}{TN + FP} = 1 - FPR$$

precision or positive predictive value (PPV)
$$PPV = \frac{TP}{TP + FP} = 1 - FDR$$

negative predictive value (NPV)
$$NPV = \frac{TN}{TN + FN} = 1 - FOR$$

miss rate or false negative rate (FNR)
$$FNR = \frac{FN}{P} = \frac{FN}{FN + TP} = 1 - TPR$$

fall-out or false positive rate (FPR)
$$FPR = \frac{FP}{N} = \frac{FP}{FP + TN} = 1 - TNR$$

false discovery rate (FDR)
$$FDR = \frac{FP}{FP + TP} = 1 - PPV$$

false omission rate (FOR)
$$FOR = \frac{FN}{FN + TN} = 1 - NPV$$

Positive likelihood ratio (LR+)
$$LR+ = \frac{TPR}{FPR}$$

Negative likelihood ratio (LR-)
$$LR- = \frac{FNR}{TNR}$$

prevalence threshold (PT)

$$PT = \frac{\sqrt{FPR}}{\sqrt{TPR} + \sqrt{FPR}}$$

threat score (TS) or critical success index (CSI)

$$TS = \frac{TP}{TP + FN + FP}$$

Prevalence

$$\frac{P}{P + N}$$

accuracy (ACC)

$$ACC = \frac{TP + TN}{P + N} = \frac{TP + TN}{TP + TN + FP + FN}$$

balanced accuracy (BA)

$$BA = \frac{TPR + TNR}{2}$$

F1 score

is the harmonic mean of precision and sensitivity:

$$F_1 = 2 \times \frac{PPV \times TPR}{PPV + TPR} = \frac{2TP}{2TP + FP + FN}$$

phi coefficient (φ or r$_φ$) or Matthews correlation coefficient (MCC)

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

Fowlkes–Mallows index (FM)

$$FM = \sqrt{\frac{TP}{TP + FP} \times \frac{TP}{TP + FN}} = \sqrt{PPV \times TPR}$$

informedness or bookmaker informedness (BM)

$$BM = TPR + TNR - 1$$

markedness (MK) or deltaP (Δp)

$$MK = PPV + NPV - 1$$

Diagnostic odds ratio (DOR)

$$DOR = \frac{LR+}{LR-}$$

Sources: Fawcett (2006),[1] Piryonesi and El-Diraby (2020),[2] Powers (2011),[3] Ting (2011),[4] CAWCR,[5] D. Chicco & G. Jurman (2020, 2021, 2023),[6][7][8] Tharwat (2018).[9] Balayla (2020)[10]

Individual Number	1	2	3	4	5	6	7	8	9	10	11	12
Actual Classification	1	1	1	1	1	1	1	1	0	0	0	0
Predicted Classification	0	0	1	1	1	1	1	1	1	0	0	0
Result	FN	FN	TP	TP	TP	TP	TP	TP	FP	TN	TN	TN

The template for any binary confusion matrix uses the four kinds of results discussed above (true positives, false negatives, false positives, and true negatives) along with the positive and negative classifications. The four outcomes can be formulated in a 2×2 *confusion matrix*, as follows:

		Predicted condition	
Total population = P + N		Positive (PP)	Negative (PN)
Actual condition	Positive (P)	True positive (TP)	False negative (FN)
	Negative (N)	False positive (FP)	True negative (TN)

Sources: [13][14][15][16][17][18][19][20]

The color convention of the three data tables above were picked to match this confusion matrix, in order to easily differentiate the data.

Now, we can simply total up each type of result, substitute into the template, and create a confusion matrix that will concisely summarize the results of testing the classifier:

		Predicted condition	
Total 8 + 4 = 12		Cancer 7	Non-cancer 5
Actual condition	Cancer 8	6	2
	Non-cancer 4	1	3

In this confusion matrix, of the 8 samples with cancer, the system judged that 2 were cancer-free, and of the 4 samples without cancer, it predicted that 1 did have cancer. All correct predictions are located in the diagonal of the table (highlighted in green), so it is easy to visually inspect the table for prediction errors, as values outside the diagonal will represent them. By summing up the 2 rows of the confusion matrix, one can also deduce the total number of positive (P) and negative (N) samples in the original dataset, i.e. $P = TP + FN$ and $N = FP + TN$.

Table of confusion

In predictive analytics, a **table of confusion** (sometimes also called a **confusion matrix**) is a table with two rows and two columns that reports the number of *true positives*, *false negatives*, *false positives*, and *true negatives*. This allows more detailed analysis than simply observing the proportion of correct classifications (accuracy). Accuracy will yield misleading results if the data set is unbalanced; that is, when the numbers of observations in different classes vary greatly.

For example, if there were 95 cancer samples and only 5 non-cancer samples in the data, a particular classifier might classify all the observations as having cancer. The overall accuracy would be 95%, but in more detail the classifier would have a 100% recognition rate (sensitivity) for the cancer class but a 0% recognition rate for the non-cancer class. F1 score is even more unreliable in such cases, and here would yield over 97.4%, whereas informedness removes such bias and yields 0 as the probability of an informed decision for any form of guessing (here always guessing cancer).

According to Davide Chicco and Giuseppe Jurman, the most informative metric to evaluate a confusion matrix is the Matthews correlation coefficient (MCC).[21]

Other metrics can be included in a confusion matrix, each of them having their significance and use.

	Predicted condition		Sources: [22][23][24][25][26][27][28][29][30]	
Total population = P + N	Predicted Positive (PP)	Predicted Negative (PN)	Informedness, bookmaker informedness (BM) = TPR + TNR − 1	Prevalence threshold (PT) $= \frac{\sqrt{TPR \times FPR} - FPR}{TPR - FPR}$
Positive (P)	True positive (TP), hit	False negative (FN), type II error, miss, underestimation	True positive rate (TPR), recall, sensitivity (SEN), probability of detection, hit rate, power $= \frac{TP}{P} = 1 - FNR$	False negative rate (FNR), miss rate $= \frac{FN}{P} = 1 - TPR$
Negative (N)	False positive (FP), type I error, false alarm, overestimation	True negative (TN), correct rejection	False positive rate (FPR), probability of false alarm, fall-out $= \frac{FP}{N} = 1 - TNR$	True negative rate (TNR), specificity (SPC), selectivity $= \frac{TN}{N} = 1 - FPR$
Prevalence $= \frac{P}{P + N}$	Positive predictive value (PPV), precision $= \frac{TP}{PP} = 1 - FDR$	False omission rate (FOR) $= \frac{FN}{PN}$ = 1 − NPV	Positive likelihood ratio (LR+) $= \frac{TPR}{FPR}$	Negative likelihood ratio (LR−) $= \frac{FNR}{TNR}$
Accuracy (ACC) $= \frac{TP + TN}{P + N}$	False discovery rate (FDR) $= \frac{FP}{PP} = 1 - PPV$	Negative predictive value (NPV) $= \frac{TN}{PN}$ = 1 − FOR	Markedness (MK), deltaP (Δp) = PPV + NPV − 1	Diagnostic odds ratio (DOR) $= \frac{LR+}{LR-}$
Balanced accuracy (BA) $= \frac{TPR + TNR}{2}$	F_1 score $= \frac{2PPV \times TPR}{PPV + TPR} = \frac{2TP}{2TP + FP + FN}$	Fowlkes–Mallows index (FM) $= \sqrt{PPV \times TPR}$	Matthews correlation coefficient (MCC) $= \sqrt{TPR \times TNR \times PPV \times NPV}$ $- \sqrt{FNR \times FPR \times FOR \times FDR}$	Threat score (TS), critical success index (CSI), Jaccard index $= \frac{TP}{TP + FN + FP}$

(Left margin label: Actual condition)

Confusion matrices with more than two categories

Confusion matrix is not limited to binary classification and can be used in multi-class classifiers as well.[31] The confusion matrices discussed above have only two conditions: positive and negative. For example, the table below summarizes communication of a whistled language between two speakers, zero values omitted for clarity.[32]

Perceived vowel / Vowel produced	i	e	a	o	u
i	15		1		
e	1		1		
a			79	5	
o			4	15	3
u				2	2

See also

- Positive and negative predictive values

References

1. Fawcett, Tom (2006). "An Introduction to ROC Analysis" (http://people.inf.elte.hu/kiss/11dwhd m/roc.pdf) (PDF). *Pattern Recognition Letters*. **27** (8): 861–874. Bibcode:2006PaReL..27..861F (https://ui.adsabs.harvard.edu/abs/2006PaReL..27..861F). doi:10.1016/j.patrec.2005.10.010 (h ttps://doi.org/10.1016%2Fj.patrec.2005.10.010). S2CID 2027090 (https://api.semanticscholar.o rg/CorpusID:2027090).

2. Piryonesi S. Madeh; El-Diraby Tamer E. (2020-03-01). "Data Analytics in Asset Management: Cost-Effective Prediction of the Pavement Condition Index". *Journal of Infrastructure Systems*. **26** (1): 04019036. doi:10.1061/(ASCE)IS.1943-555X.0000512 (https://doi.org/10.1061%2F%28 ASCE%29IS.1943-555X.0000512). S2CID 213782055 (https://api.semanticscholar.org/CorpusI D:213782055).

3. Powers, David M. W. (2011). "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation" (https://www.researchgate.net/publication/22852930 7). *Journal of Machine Learning Technologies*. **2** (1): 37–63.

4. Ting, Kai Ming (2011). Sammut, Claude; Webb, Geoffrey I. (eds.). *Encyclopedia of machine learning*. Springer. doi:10.1007/978-0-387-30164-8 (https://doi.org/10.1007%2F978-0-387-301 64-8). ISBN 978-0-387-30164-8.

5. Brooks, Harold; Brown, Barb; Ebert, Beth; Ferro, Chris; Jolliffe, Ian; Koh, Tieh-Yong; Roebber, Paul; Stephenson, David (2015-01-26). "WWRP/WGNE Joint Working Group on Forecast Verification Research" (https://www.cawcr.gov.au/projects/verification/). *Collaboration for Australian Weather and Climate Research*. World Meteorological Organisation. Retrieved 2019-07-17.

6. Chicco D.; Jurman G. (January 2020). "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation" (https://www.ncbi.nlm.ni h.gov/pmc/articles/PMC6941312). *BMC Genomics*. **21** (1): 6-1–6-13. doi:10.1186/s12864-019- 6413-7 (https://doi.org/10.1186%2Fs12864-019-6413-7). PMC 6941312 (https://www.ncbi.nlm. nih.gov/pmc/articles/PMC6941312). PMID 31898477 (https://pubmed.ncbi.nlm.nih.gov/318984 77).

7. Chicco D.; Toetsch N.; Jurman G. (February 2021). "The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC786344 9). *BioData Mining*. **14** (13): 13. doi:10.1186/s13040-021-00244-z (https://doi.org/10.1186%2Fs 13040-021-00244-z). PMC 7863449 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7863449). PMID 33541410 (https://pubmed.ncbi.nlm.nih.gov/33541410).

8. Chicco D.; Jurman G. (2023). "The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9938573). *BioData Mining*. **16** (1): 4. doi:10.1186/s13040-023-00322-4 (https://doi.org/10.1186%2Fs13040-023-00322-4). PMC 9938573 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9938573). PMID 36800973 (https://pubmed.ncbi.nlm.nih.gov/36800973).

9. Tharwat A. (August 2018). "Classification assessment methods" (https://doi.org/10.1016%2Fj.aci.2018.08.003). *Applied Computing and Informatics*. **17**: 168–192. doi:10.1016/j.aci.2018.08.003 (https://doi.org/10.1016%2Fj.aci.2018.08.003).

10. Balayla, Jacques (2020). "Prevalence threshold (φe) and the geometry of screening curves" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7540853). *PLOS ONE*. **15** (10): e0240215. arXiv:2006.00398 (https://arxiv.org/abs/2006.00398). Bibcode:2020PLoSO..1540215B (https://ui.adsabs.harvard.edu/abs/2020PLoSO..1540215B). doi:10.1371/journal.pone.0240215 (https://doi.org/10.1371%2Fjournal.pone.0240215). PMC 7540853 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7540853). PMID 33027310 (https://pubmed.ncbi.nlm.nih.gov/33027310).

11. Stehman, Stephen V. (1997). "Selecting and interpreting measures of thematic classification accuracy". *Remote Sensing of Environment*. **62** (1): 77–89. Bibcode:1997RSEnv..62...77S (https://ui.adsabs.harvard.edu/abs/1997RSEnv..62...77S). doi:10.1016/S0034-4257(97)00083-7 (https://doi.org/10.1016%2FS0034-4257%2897%2900083-7).

12. Powers, David M. W. (2011). "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation" (https://www.researchgate.net/publication/228529307). *Journal of Machine Learning Technologies*. **2** (1): 37–63. S2CID 55767944 (https://api.semanticscholar.org/CorpusID:55767944).

13. Fawcett, Tom (2006). "An Introduction to ROC Analysis" (http://people.inf.elte.hu/kiss/11dwhdm/roc.pdf) (PDF). *Pattern Recognition Letters*. **27** (8): 861–874. Bibcode:2006PaReL..27..861F (https://ui.adsabs.harvard.edu/abs/2006PaReL..27..861F). doi:10.1016/j.patrec.2005.10.010 (https://doi.org/10.1016%2Fj.patrec.2005.10.010). S2CID 2027090 (https://api.semanticscholar.org/CorpusID:2027090).

14. Piryonesi S. Madeh; El-Diraby Tamer E. (2020-03-01). "Data Analytics in Asset Management: Cost-Effective Prediction of the Pavement Condition Index". *Journal of Infrastructure Systems*. **26** (1): 04019036. doi:10.1061/(ASCE)IS.1943-555X.0000512 (https://doi.org/10.1061%2F%28ASCE%29IS.1943-555X.0000512). S2CID 213782055 (https://api.semanticscholar.org/CorpusID:213782055).

15. Powers, David M. W. (2011). "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation" (https://www.researchgate.net/publication/228529307). *Journal of Machine Learning Technologies*. **2** (1): 37–63.

16. Ting, Kai Ming (2011). Sammut, Claude; Webb, Geoffrey I. (eds.). *Encyclopedia of machine learning*. Springer. doi:10.1007/978-0-387-30164-8 (https://doi.org/10.1007%2F978-0-387-30164-8). ISBN 978-0-387-30164-8.

17. Brooks, Harold; Brown, Barb; Ebert, Beth; Ferro, Chris; Jolliffe, Ian; Koh, Tieh-Yong; Roebber, Paul; Stephenson, David (2015-01-26). "WWRP/WGNE Joint Working Group on Forecast Verification Research" (https://www.cawcr.gov.au/projects/verification/). *Collaboration for Australian Weather and Climate Research*. World Meteorological Organisation. Retrieved 2019-07-17.

18. Chicco D, Jurman G (January 2020). "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6941312). *BMC Genomics*. **21** (1): 6-1–6-13. doi:10.1186/s12864-019-6413-7 (https://doi.org/10.1186%2Fs12864-019-6413-7). PMC 6941312 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6941312). PMID 31898477 (https://pubmed.ncbi.nlm.nih.gov/31898477).

19. Chicco D, Toetsch N, Jurman G (February 2021). "The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7863449). *BioData Mining*. **14** (13): 13. doi:10.1186/s13040-021-00244-z (https://doi.org/10.1186%2Fs13040-021-00244-z). PMC 7863449 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7863449). PMID 33541410 (https://pubmed.ncbi.nlm.nih.gov/33541410).

20. Tharwat A. (August 2018). "Classification assessment methods" (https://doi.org/10.1016%2Fj.aci.2018.08.003). *Applied Computing and Informatics*. **17**: 168–192. doi:10.1016/j.aci.2018.08.003 (https://doi.org/10.1016%2Fj.aci.2018.08.003).

21. Chicco D, Jurman G (January 2020). "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6941312). *BMC Genomics*. **21** (1): 6-1–6-13. doi:10.1186/s12864-019-6413-7 (https://doi.org/10.1186%2Fs12864-019-6413-7). PMC 6941312 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6941312). PMID 31898477 (https://pubmed.ncbi.nlm.nih.gov/31898477).

22. Balayla, Jacques (2020). "Prevalence threshold (φe) and the geometry of screening curves" (https://doi.org/10.1371%2Fjournal.pone.0240215). *PLOS ONE*. **15** (10): e0240215. doi:10.1371/journal.pone.0240215 (https://doi.org/10.1371%2Fjournal.pone.0240215). PMID 33027310 (https://pubmed.ncbi.nlm.nih.gov/33027310).

23. Fawcett, Tom (2006). "An Introduction to ROC Analysis" (http://people.inf.elte.hu/kiss/11dwhdm/roc.pdf) (PDF). *Pattern Recognition Letters*. **27** (8): 861–874. doi:10.1016/j.patrec.2005.10.010 (https://doi.org/10.1016%2Fj.patrec.2005.10.010). S2CID 2027090 (https://api.semanticscholar.org/CorpusID:2027090).

24. Piryonesi S. Madeh; El-Diraby Tamer E. (2020-03-01). "Data Analytics in Asset Management: Cost-Effective Prediction of the Pavement Condition Index". *Journal of Infrastructure Systems*. **26** (1): 04019036. doi:10.1061/(ASCE)IS.1943-555X.0000512 (https://doi.org/10.1061%2F%28ASCE%29IS.1943-555X.0000512). S2CID 213782055 (https://api.semanticscholar.org/CorpusID:213782055).

25. Powers, David M. W. (2011). "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation" (https://www.researchgate.net/publication/228529307). *Journal of Machine Learning Technologies*. **2** (1): 37–63.

26. Ting, Kai Ming (2011). Sammut, Claude; Webb, Geoffrey I. (eds.). *Encyclopedia of machine learning*. Springer. doi:10.1007/978-0-387-30164-8 (https://doi.org/10.1007%2F978-0-387-30164-8). ISBN 978-0-387-30164-8.

27. Brooks, Harold; Brown, Barb; Ebert, Beth; Ferro, Chris; Jolliffe, Ian; Koh, Tieh-Yong; Roebber, Paul; Stephenson, David (2015-01-26). "WWRP/WGNE Joint Working Group on Forecast Verification Research" (https://www.cawcr.gov.au/projects/verification/). *Collaboration for Australian Weather and Climate Research*. World Meteorological Organisation. Retrieved 2019-07-17.

28. Chicco D, Jurman G (January 2020). "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6941312). *BMC Genomics*. **21** (1): 6-1–6-13. doi:10.1186/s12864-019-6413-7 (https://doi.org/10.1186%2Fs12864-019-6413-7). PMC 6941312 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6941312). PMID 31898477 (https://pubmed.ncbi.nlm.nih.gov/31898477).

29. Chicco D, Toetsch N, Jurman G (February 2021). "The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7863449). *BioData Mining*. **14** (13): 13. doi:10.1186/s13040-021-00244-z (https://doi.org/10.1186%2Fs13040-021-00244-z). PMC 7863449 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7863449). PMID 33541410 (https://pubmed.ncbi.nlm.nih.gov/33541410).

30. Tharwat A. (August 2018). "Classification assessment methods" (https://doi.org/10.1016%2Fj.a ci.2018.08.003). *Applied Computing and Informatics*. **17**: 168–192. doi:10.1016/j.aci.2018.08.003 (https://doi.org/10.1016%2Fj.aci.2018.08.003).

31. Piryonesi S. Madeh; El-Diraby Tamer E. (2020-03-01). "Data Analytics in Asset Management: Cost-Effective Prediction of the Pavement Condition Index". *Journal of Infrastructure Systems*. **26** (1): 04019036. doi:10.1061/(ASCE)IS.1943-555X.0000512 (https://doi.org/10.1061%2F%28 ASCE%29IS.1943-555X.0000512). S2CID 213782055 (https://api.semanticscholar.org/CorpusI D:213782055).

32. Rialland, Annie (August 2005). "Phonological and phonetic aspects of whistled languages". *Phonology*. **22** (2): 237–271. CiteSeerX 10.1.1.484.4384 (https://citeseerx.ist.psu.edu/viewdoc/ summary?doi=10.1.1.484.4384). doi:10.1017/S0952675705000552 (https://doi.org/10.1017%2 FS0952675705000552). S2CID 18615779 (https://api.semanticscholar.org/CorpusID:1861577 9).

■