



FIELD NOTES / MODEL EVALUATION

High Accuracy Can Hide Missed Leads

Check what the model gets wrong before accepting its headline score.

WHAT YOU WILL DO

Ask for the class totals and all four prediction counts at the cutoff your team intends to use.



SOURCE



INSIGHT



APPLICATION

ARTICLE / FILLABLE WORKSHEET / SOURCE NOTES



Ask what the headline leaves out

A sales leader needs more than a large accuracy percentage before accepting a lead model.

GTMA proposal: ask which qualified leads it caught, which it missed and how many unqualified leads it selected. Use the cutoff intended for the real decision, not an unexplained headline score.

The warning in the research

In a 2025 Frontiers company case, González-Flores and colleagues warn that a model predicting the majority class can achieve high accuracy while missing the minority cases that matter. This is an attributed methodological warning, not a finding that their own model missed every qualified lead.

This warning supplies no universal accuracy target for a SaaS team. The example and checking routine below are GTMA proposals, not methods validated by that paper.

A fictional model gets 95 right and catches none

Fictional example: Maya evaluates 100 historical enquiries. Five qualified under her team's written definition; 95 did not. At the proposed selection cutoff, the model marks every enquiry unqualified. These counts are invented to demonstrate the arithmetic, not research results or a benchmark.

The model gets 95 of 100 answers right: 95% accuracy. But its complete result is zero qualified leads caught, five qualified leads missed, zero unqualified leads incorrectly selected and 95 unqualified leads correctly excluded. It selects nobody. That headline therefore does not demonstrate the job Maya needs: identifying qualified enquiries for attention.

Maya withholds acceptance based on accuracy. She asks the supplier to report all four counts for any alternative it proposes, using the same evaluation group, qualification definition and intended cutoff. She keeps the existing manual review while that comparison is unresolved; she does not treat this example as evidence that manual review is better.



Request a complete result

Record the qualification definition, evaluation group and dates, class totals, cutoff and all four counts. If the tool ranks leads, specify the number the team would work before counting selected and missed cases. Keep enquiries with unknown outcomes separate from known unqualified enquiries.

Agree what missed opportunities and unnecessary work the team can tolerate before choosing a model. These counts make the trade-off inspectable; they do not establish revenue impact, future performance or permission to roll out. If the supplier cannot supply the underlying counts, leave the performance claim unaccepted.



Worksheet: look behind the accuracy score

A GTMA proposed record for deciding whether the supplied model result supports your intended selection task.

Type into the fields, or print. Save a copy to retain your answers.

01 / Define the decision and evaluation group

Record the qualification definition, group, dates and intended threshold or queue size. Keep unknown outcomes separate. Do not enter personal lead records.

02 / Record what actually qualified

Write known qualified and known unqualified totals, plus unknown outcomes separately. Fictional example: five qualified, 95 unqualified, no unknowns.

03 / Record all four prediction counts

Qualified caught; qualified missed; unqualified selected; unqualified excluded. Fictional example: 0, 5, 0, 95. Reconcile these with the group totals.

04 / Finish the acceptance decision

Record acceptable misses and unnecessary work, missing evidence, decision and owner. Fictional example: withhold acceptance; request a comparable complete result; retain manual review pending it.



Read further. Know the limits.

Source findings, GTMA proposals and fictional examples have different roles. Use the original research to examine its methods and context.

[1] The relevance of lead prioritization: a B2B lead scoring model based on machine learning

Frontiers in Artificial Intelligence / Published: 2025-03-07

Single-company case. Supports the attributed warning about accuracy and class imbalance; no universal benchmark or validation of the GTMA worksheet.

<https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2025.1554325/full>

What the source supports

González-Flores, Rubiano-Moreno and Sosa-Gómez (2025) warn in their company case that majority-class predictions can produce high accuracy while missing the minority cases of interest. The selected finding is this methodological caution.

This note does not reproduce model performance scores, claim the paper's own model failed, or approve other statements from the paper.

Scope and uncertainty

The source is a single-company B2B case, not a SaaS benchmark or an independent test of your model. The selected finding supplies no universal pass mark, sales uplift, or forecast of future performance.

Qualification definitions, evaluation periods, populations and cutoff choices must be established for the actual test. Unknown outcomes remain unknown; do not count them as unqualified to fill a table.

GTMA proposal and fictional arithmetic

Maya, the 100 enquiries, the five qualified outcomes and the all-negative predictions are fictional. 95 correct answers out of 100 give 95% accuracy while missing all five qualified enquiries. This illustrates arithmetic, not observed model performance.

The worksheet and decision routine are GTMA proposals, not a method validated by this case. Keep personal lead records in approved systems; use aggregate counts here. Retaining manual review in the example does not prove it performs better.

The worksheet is a GTMA proposal, not a research-validated instrument. Fictional examples are illustrations, not evidence of results.