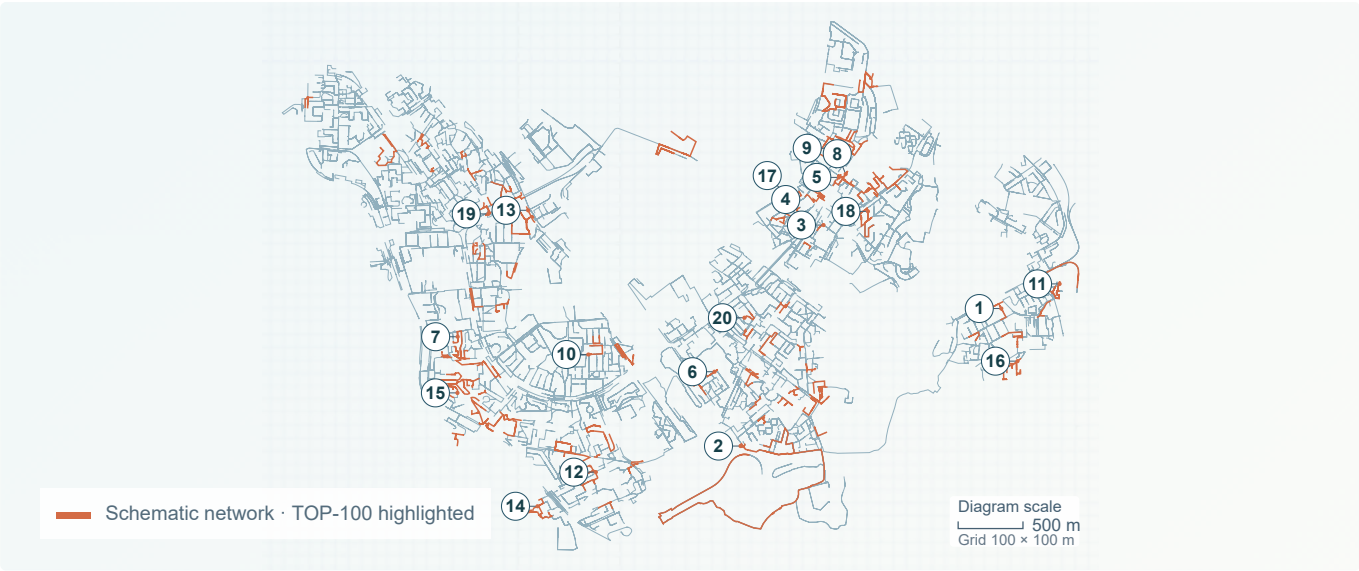
INFRASTRUCTURE · PREDICTIVE ANALYTICS

See the risk.
Set the priority.

How AI turns event history and pipeline passports into a network inspection plan.



84 / 100

LSTM hits in TOP-100

3.9 ×

above random selection

42.3%

of events in 8% of sections

Evaluated against 2025 events · annual and quarterly forecasts for 2026. Heating networks serving an urban agglomeration of around 200,000 residents: supply and return pipelines, underground and above-ground installation.

01 / EXECUTIVE SUMMARY

A result engineers can use

The ranking helps decide which sections to inspect first when crew capacity is limited.

1,250	4,934	165.437	4,977
Section · prediction units	Sys · network segments	km of pipelines	target records over 10 years

What we did	Prepared event history for 2016–2025 and segment passports grouped into Sections. Compared four AI models with the Rivixi Algorithm reference.
How we evaluated	Settings were selected on 2024 data. Frozen predictions were compared with 2025 outcomes: 940 records across 269 Sections.
Main result	Quarterly LSTM identified 84 Sections with events among its top 100. Random selection would average 21.5.
Planning value	The first 100 LSTM Sections — 8% of the sample — contain 42.3% of registered events in the evaluation year.
System output	Inspection priority, segment passports and timelines, model comparisons, annual and quarterly rankings, and a masked network diagram.

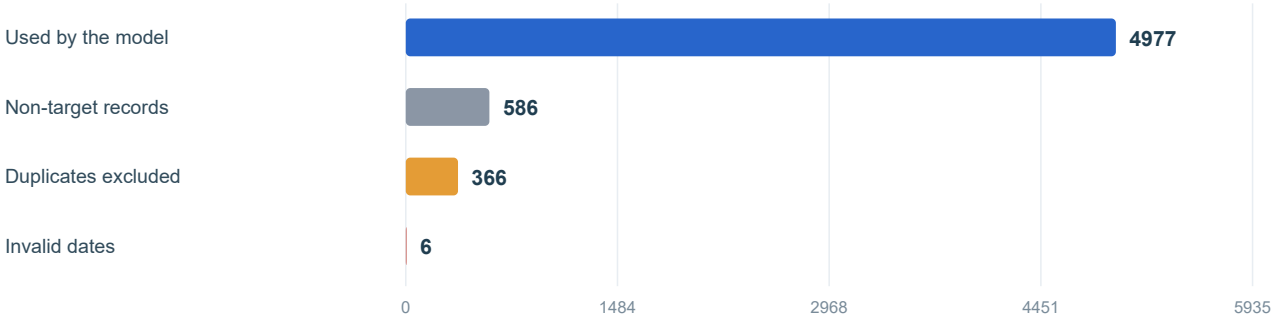
> 84 out of 100 measures hits at the top of the ranking. 42.3% measures the share of event records on those Sections. These are different metrics: one Section can have multiple events.

02 / DATA & PROCESSING

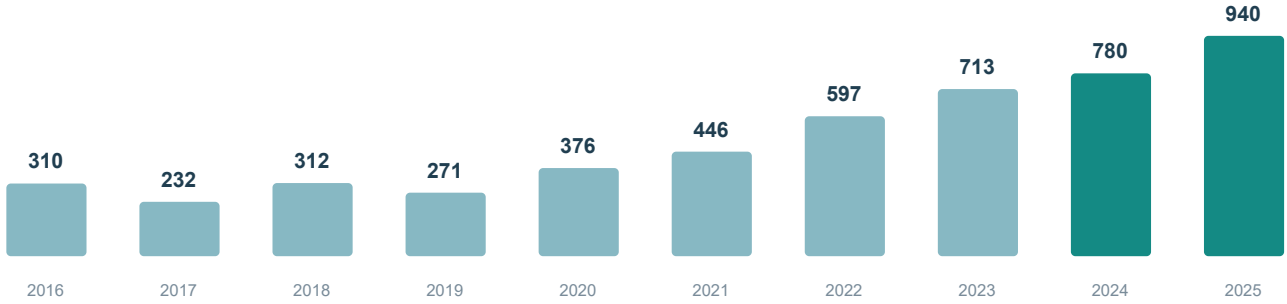
From journal entries to event history

Each source row is screened. Duplicates, excluded records and invalid dates are counted separately.

5,935 source rows → **4,977** target model records



Target event history by year



A target record is a registered event passing the accepted classification. It is a journal record, not automatically a unique physical failure.

03 / PASSPORTS & ANALYSIS UNITS

What the model knows about a pipeline

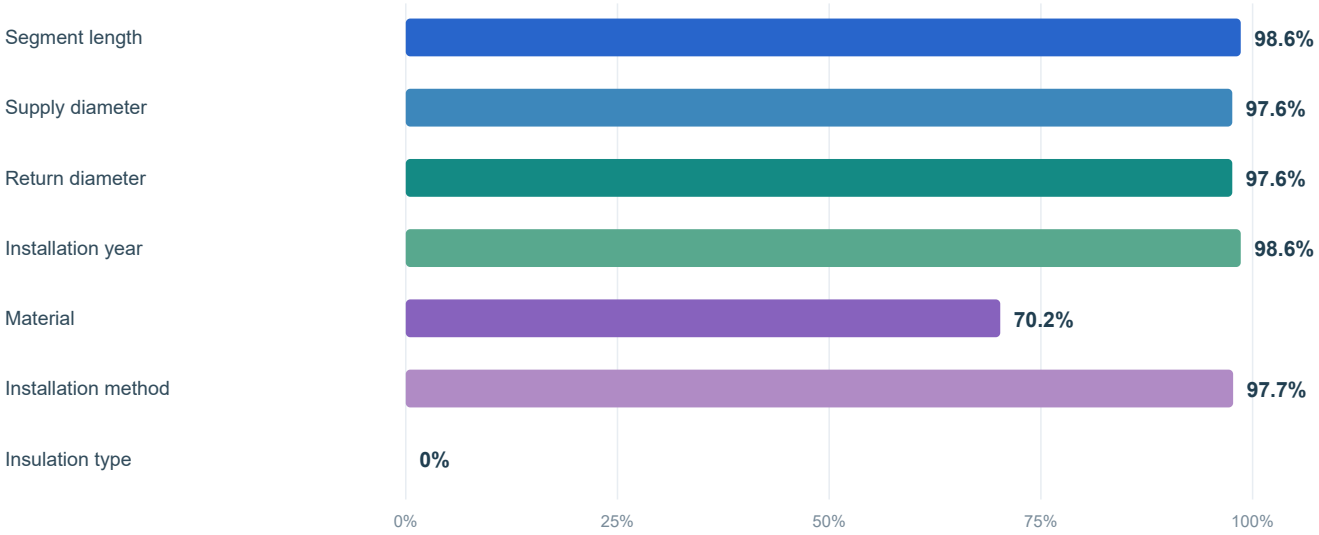
Sys → Section

Sys is an individual segment with its own passport. Multiple segments form a Section, the unit ranked by the model.

History + attributes

The model combines event sequences with passport attributes. Lengths, diameters, installation year and categories describe the physical context of the network.

Source passport completeness



A missing field does not become a confirmed fact. Passport completeness is shown separately from the forecast score and helps identify information that needs clarification.

The chart reflects the source snapshot of the selected pilot area. Completeness is not a measure of forecast accuracy.

04 / HOW THE AI WORKS

Two histories. Four AI models.

The same annual target, with two ways to read history: annually and quarterly.

01	02	03	04
Event history + passports	**LSTM / Gradient Boosting**	**Section score and ranking**	**Inspection plan**

LSTM

Reads event sequences and learns temporal patterns. A passport branch adds section attributes.

Recurrent layer	**64 LSTM**
Passport branch	**64 → 32**
Dropout	**0.42**
Optimizer	**Adam · 0.0008**
Ensemble	**3 runs · averaged scores**
Training	**up to 40 epochs**

Gradient Boosting

Builds decision trees sequentially using history and passport features. Provides an independent comparison with the neural network.

Trees	**200**
Depth	**2**
Learning rate	**0.05**
Subsample	**1.0**
Histories	**annual / quarterly**
Reference	**Rivixi Algorithm**

2017–2023 → 2024
Settings selection

2017–2024 → 2025
Forecast evaluation

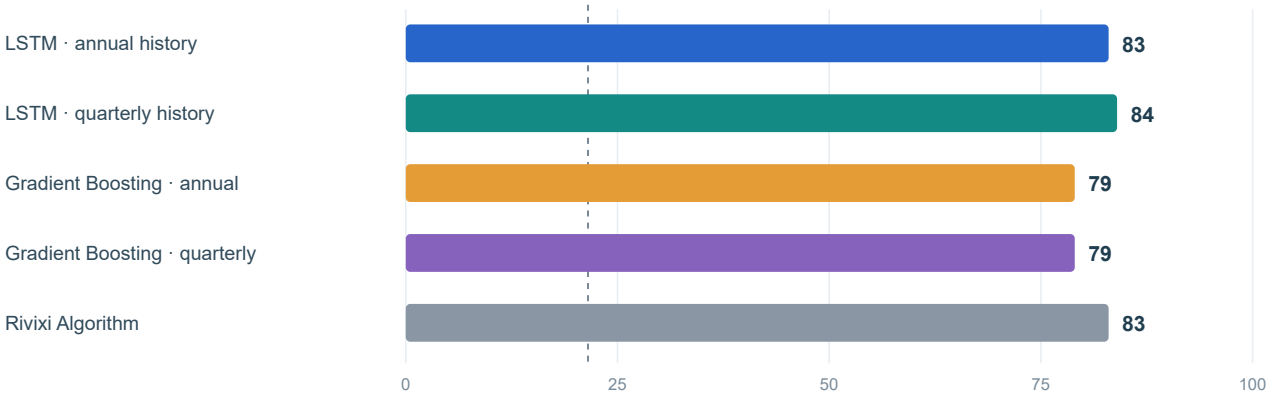
2017–2025 → 2026
Final forecast

The Rivixi Algorithm is a reference used to check AI value. Quarterly history here also predicts an annual target; separate Q1–Q4 outputs are additional quarterly targets.

05 / RESULTS · TOP-100

84 hits among the first 100

How many Sections with 2025 events appeared at the top of each ranking.



Dashed line: random selection averages 21.5 hits per 100, with 269 Sections with events out of 1,250.

Model	ROC-AUC	TOP-20	TOP-100	For inspection[1]	Recall[1]
LSTM · annual history	0.8900	20/20	83/100	437	90%
LSTM · quarterly history	0.8901	20/20	84/100	429	88.8%
Gradient Boosting · annual	0.8879	18/20	79/100	424	89.2%
Gradient Boosting · quarterly	0.8874	19/20	79/100	430	89.2%
Rivixi Algorithm	0.8902	19/20	83/100	454	90.3%

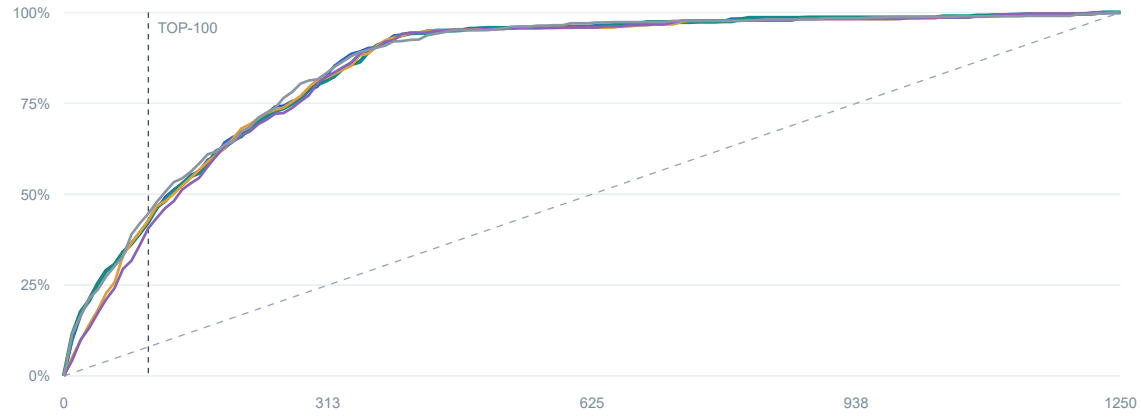
[1] At threshold 0.5. “For inspection” counts all flagged Sections; recall is the share found among the 269 Sections with events. TOP-100 uses rank order and does not depend on this threshold.

3.9× as many hits for quarterly LSTM as random selection. This is a pilot evaluation result, not a guarantee for another network.

06 / EVENT COVERAGE

How many events the ranking captures

Move from highest to lowest scores. The vertical axis is the cumulative share of 2025 event records; the horizontal axis is the number of Sections inspected.



42.3%

TOP-100 · 8% of sections

share of event records on the top 100 quarterly LSTM Sections

What share of events falls on the first N sections

Method	TOP-20	TOP-50	TOP-100	TOP-200
LSTM · annual history	16.6%	29%	42.3%	65.7%
LSTM · quarterly history	17.7%	28.9%	42.3%	64.8%
Gradient Boosting · annual	9.8%	22.8%	42.9%	65.3%
Gradient Boosting · quarterly	10%	21.1%	40.6%	64.4%
Rivixi Algorithm	16.7%	27.3%	44.7%	64.5%
Random order	1.6%	4%	8%	16%

The grey diagonal is random order. The vertical line marks TOP-100. Coverage counts 940 records, not the 269 unique Sections with events.

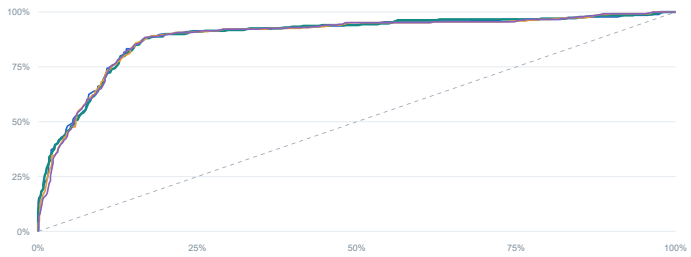
07 / THRESHOLD & EVALUATION

What a threshold of 0.5 means

A threshold turns forecast scores into an inspection list. It does not mean the model is 50% reliable. Below are outcomes at the same threshold.

LSTM · annual history		LSTM · quarterly history		Gradient Boosting · annual		Gradient Boosting · quarterly	
242 correctly identified	27 missed	239 correctly identified	30 missed	240 correctly identified	29 missed	240 correctly identified	29 missed
195 false alarm	786 correctly clear	190 false alarm	791 correctly clear	184 false alarm	797 correctly clear	190 false alarm	791 correctly clear

ROC curves



How to read

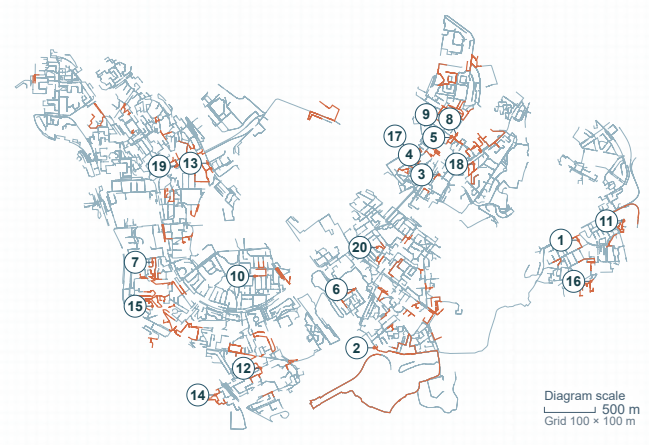
Horizontally: false alarms among Sections without events. Vertically: detections among Sections with events. Closer to the top-left corner means a better balance.

For operational planning, choose TOP-20, TOP-50 or TOP-100 to match crew capacity. The demo provides an interactive ROC curve for threshold analysis.

The same threshold can flag different numbers of Sections across models: score scales are not calibrated to each other. ROC-AUC and TOP ordering do not depend on the chosen threshold.

08 / FORECAST & APPLICATION

From ranking to network operations



Colour: TOP-100 of the quarterly-history LSTM annual forecast for 2026. The virtual diagram is not a GIS pipeline survey.

2026 horizons

Model / target	Threshold 0.5
LSTM · annual history	466 Section
LSTM · quarterly history	440 Section
Gradient Boosting · annual	438 Section
Gradient Boosting · quarterly	441 Section
LSTM · Q1 2026	426 Section
LSTM · Q2 2026	428 Section
LSTM · Q3 2026	438 Section
LSTM · Q4 2026	429 Section

These are forecast selection counts, not actual 2026 event counts. Quarterly lists can overlap and must not be added.

Available in the demo

01 **Map and TOP lists**

02 **Models and evaluation**

03 **Registry and passports**

04 **Data processing**

Temporal comparison

On 2024 evaluation: ROC-AUC 0.9069 · TOP-100 85/100. On 2025 evaluation: 0.8901 · 84/100.

TOP-20 table legend

[1] DL — ductless; D — in a duct; A — above ground; B — basement; T — tunnel. Length is summed across Sys; diameters show ranges of available passport values. Installation years come from passports; differing values are shown as a range.

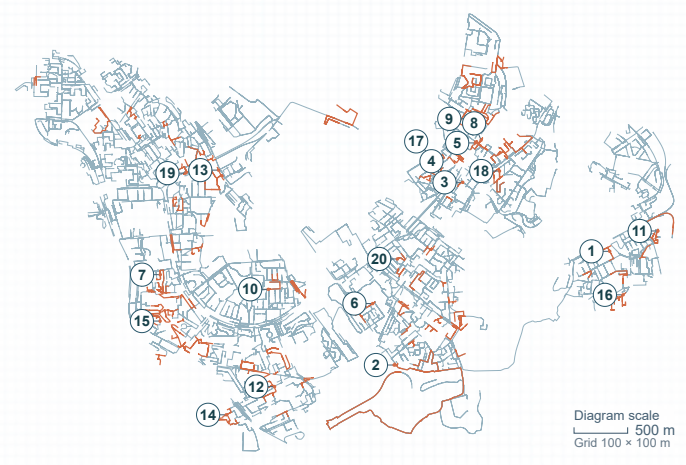
[2] Target journal records for 2016–2025. [3] Records mentioning defect elimination/repair or replacement; these are not unique completed repairs.

[4] F — defect elimination/repair; R — replacement; S — search/excavation; O — other target records. Classes are exclusive: replacement → elimination → search → other. The grid scales the virtual diagram, not a confirmed GIS survey.

09 / TOP-20 · SECTION PASSPORTS

20 priorities on the network diagram

Numbers 1–20 are Section ranks in the quarterly-history LSTM annual forecast for 2026. Each leader line points to a segment in that Section.



No.	Sys	Length, m	Ø supply / return, mm	Laying[1]	Installation year	Events[2]	Works[3]	Types[4]
1	7	153	80–200 / 80–200	B, DL	1970	87	83	F:20 · R:63 · O:4
2	9	97	50–250 / 50–250	B, DL, D	1970	43	40	F:18 · R:22 · S:1 · O:2
3	4	62	80–250 / 70–250	DL, D	1970	60	60	F:22 · R:38
4	6	56	100–200 / 100–200	B, D	1970	79	77	F:17 · R:60 · O:2
5	7	187	150–200 / 100–200	B, DL, D	1970	72	68	F:25 · R:43 · O:4
6	7	360	250–300 / 250–300	A, B, D	1970	59	57	F:18 · R:39 · O:2
7	4	125	50–200 / 50–200	B, DL	1970	82	79	F:24 · R:55 · O:3
8	6	156	100–150 / 100–150	B, D	1970	55	45	F:18 · R:27 · S:8 · O:2
9	8	230	80–150 / 70–150	B, D	1970	32	32	F:9 · R:23
10	6	108	80–150 / 65–125	B, DL, D	1970	22	22	F:3 · R:19
11	10	168	50–150 / 50–150	A, B, D	1970	25	24	F:9 · R:15 · O:1
12	4	92	100–200 / 80–200	D	1970	33	30	F:19 · R:11 · O:3
13	4	246	50–100 / 50–80	DL, D	1970	22	21	F:11 · R:10 · S:1
14	4	258	150–250 / 150–250	B, DL, D	1970	31	30	F:10 · R:20 · O:1
15	3	78	80–250 / 65–250	B, D	1970	39	37	F:9 · R:28 · O:2
16	7	134	50–250 / 50–250	A, B, D	1970	35	32	F:3 · R:29 · S:1 · O:2
17	9	87	50–250 / 50–250	B, DL, D	1970	24	23	F:9 · R:14 · O:1
18	8	137	300–500 / 300–500	B, DL, D	1970	29	28	F:5 · R:23 · S:1
19	5	136	80–100 / 70–80	B, DL	1970	31	30	F:4 · R:26 · S:1
20	7	94	50–150 / 50–150	DL, D	1970	56	54	F:16 · R:38 · O:2

RIVIXI PREDICT

Priorities you can verify.

Open interactive map →