Strategy Evaluation Results

Summary

Total Strategy Combinations	2
Total Test Cases	10000
Timestamp	2026-02-13 11:11:13

Strategy Performance

RULE_BASED NLP + DIJKSTRA Path Finder

Classification Metrics

Accuracy	69.93% (6993/10000)
Precision (CORRECT)	71.93%
Recall (CORRECT)	80.80%
F1 Score (CORRECT)	76.11%

Per-Category Accuracy

Category	Accuracy
CORRECT	80.80%
NOT_FRENCH	20.42%
NOT_TRIP	94.24%

Pipeline Timing

Avg Execution Time	17.14 ms

Median Execution Time	10.31 ms
P95 Execution Time	40.56 ms
Min Execution Time	8.45 ms
Max Execution Time	81.08 ms
Std Dev	12.40 ms

NLP Results

Inputs Processed	10000
Successfully Extracted	6658
Extraction Errors	3342
Avg Extraction Time	16.62 ms
Median Extraction Time	9.91 ms
P95 Extraction Time	39.07 ms

Path Finder Results

Paths Computed	6658
Avg Computation Time	0.52 ms
Median Computation Time	0.18 ms
P95 Computation Time	2.75 ms
Avg Path Length	9.35 stations

HF_NER NLP + DIJKSTRA Path Finder

Classification Metrics

Accuracy	52.69% (5269/10000)

Precision (CORRECT)	72.30%
Recall (CORRECT)	32.71%
F1 Score (CORRECT)	45.05%

Per-Category Accuracy

Category	Accuracy
CORRECT	32.71%
NOT_FRENCH	66.44%
NOT_TRIP	100.00%

Pipeline Timing

Avg Execution Time	18.81 ms
Median Execution Time	16.74 ms
P95 Execution Time	30.48 ms
Min Execution Time	12.50 ms
Max Execution Time	1455.58 ms
Std Dev	20.63 ms

NLP Results

Inputs Processed	10000
Successfully Extracted	2682
Extraction Errors	7318
Avg Extraction Time	18.60 ms
Median Extraction Time	16.49 ms

P95 Extraction Time	30.21 ms

Path Finder Results

Paths Computed	2682
Avg Computation Time	0.21 ms
Median Computation Time	0.00 ms
P95 Computation Time	1.16 ms
Avg Path Length	11.14 stations

Results Per Strategy

NLP Results (Global)

The NLP component extracts departure and arrival stations from user input.

Total Inputs Processed	20000
Successfully Extracted	9340
Extraction Errors	10660
Avg Extraction Time	17.61 ms
Median Extraction Time	15.43 ms
P95 Extraction Time	37.47 ms

Path Finder Results (Global)

The path finder computes the shortest route between two stations.

Total Paths Computed	9340
Avg Path Computation Time	0.79 ms
Avg Path Length	9.86 stations

Pipeline Results

End-to-end pipeline performance combining NLP and path finding.

Total Pipeline Executions	20000
Avg Total Time	17.98 ms
Median Total Time	15.62 ms
P95 Total Time	39.08 ms
Min Total Time	8.45 ms
Max Total Time	1455.58 ms

Std Dev	17.04 ms