📦 ADAPT-SQL Batch Processing

Process multiple queries with automatic checkpoints and comprehensive statistics

🎯 Batch Configuration 🔗

Start Index	End Index	Total Queries	Checkpoint Every
0	2146	2147	100

Enabled:

✅ Validation Retry | ⚡ Execution | 📊 Evaluation | 🔄 Full Retry (max 5, score ≥0.8)

🚀 Run Batch Processing

💾 Generating Results CSV...

✅ Batch processing complete! Processed 2147 queries.

📊 CSV saved to: **results/adapt_sql_results_2025-12-14_13-45-06.csv**

💾 Checkpoint saved to: batch_results/final_checkpoint_2025-12-14_13-45-06.pkl

⬇️ Download Results CSV

📊 System Performance Overview

🎯 Key Performance Indicators

Total Queries	Valid SQL	Execution	EX = 1.0	EM = 1.0	Avg Score
2147	100.0%	99.7%	91.8%	34.0%	0.803
	↑ 2147/2147	↑ 2141/2147	↑ 1972/2147	↑ 731/2147	

📈 Complexity Distribution & Performance

✅ EASY	⚠️ NON_NESTED	🔴 NESTED
Count: 585 (27.2%)	**Count:** 1247 (58.1%)	**Count:** 315 (14.7%)
Valid: 585/585 (100.0%)	**Valid:** 1247/1247 (100.0%)	**Valid:** 315/315 (100.0%)
EX=1.0: 565 (96.6%)	**EX=1.0:** 1133 (90.9%)	**EX=1.0:** 274 (87.0%)
Avg Score: 0.868	**Avg Score:** 0.799	**Avg Score:** 0.698

⚡ Execution Performance 🔗

Executed	Success Rate	Avg Time	Failed
2147	99.7%	0.011s	6
	↑ 2141/2147		↓ -6

Common Execution Errors:

• Operational error: 6 occurrences

🎯 Spider Benchmark Results

🥇 Execution Accuracy (EX)

Primary metric - Results match exactly

EX = 1.0

91.8%

↑ 1972/2147
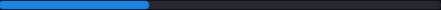
❌ Failed: 175 queries

🥈 Exact-Set-Match (EM)

Secondary metric - SQL structure matches
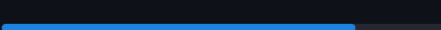
EM = 1.0

34.0%

↑ 731/2147

❌ Failed: 1416 queries

🏆 Composite Score

Weighted: 80% EX + 20% EM

Average

0.803

✅ Perfect (≥0.95): 724

🟢 Good (0.7-0.95): 1248

🟡 Fair (0.5-0.7): 0

🔴 Poor (<0.5): 175

🔄 Full Pipeline Retry Performance

Queries with Retry	Avg Attempts	Retry Success	Max Retries
2147	1.54	91.8%	175
		↑ 1972/2147	

Retry Effectiveness:

Improved 1972 out of 2147 queries that needed retry

🔍 Error Analysis

✅ No validation errors found!

🎲 Generation Strategy Performance

Intermediate Representation

Queries: 1247

Valid: 1247/1247 (100.0%)

EX=1.0: 1133/1247 (90.9%)

Avg Score: 0.799

Simple Few Shot

Queries: 585

Valid: 585/585 (100.0%)

EX=1.0: 565/585 (96.6%)

Avg Score: 0.868

Decomposed Generation

Queries: 315

Valid: 315/315 (100.0%)

EX=1.0: 274/315 (87.0%)

Avg Score: 0.698