KNN Pipeline - Technical Results

Generated: 2025-10-01T21:39:07.943324 UTC

1. Dataset Overview
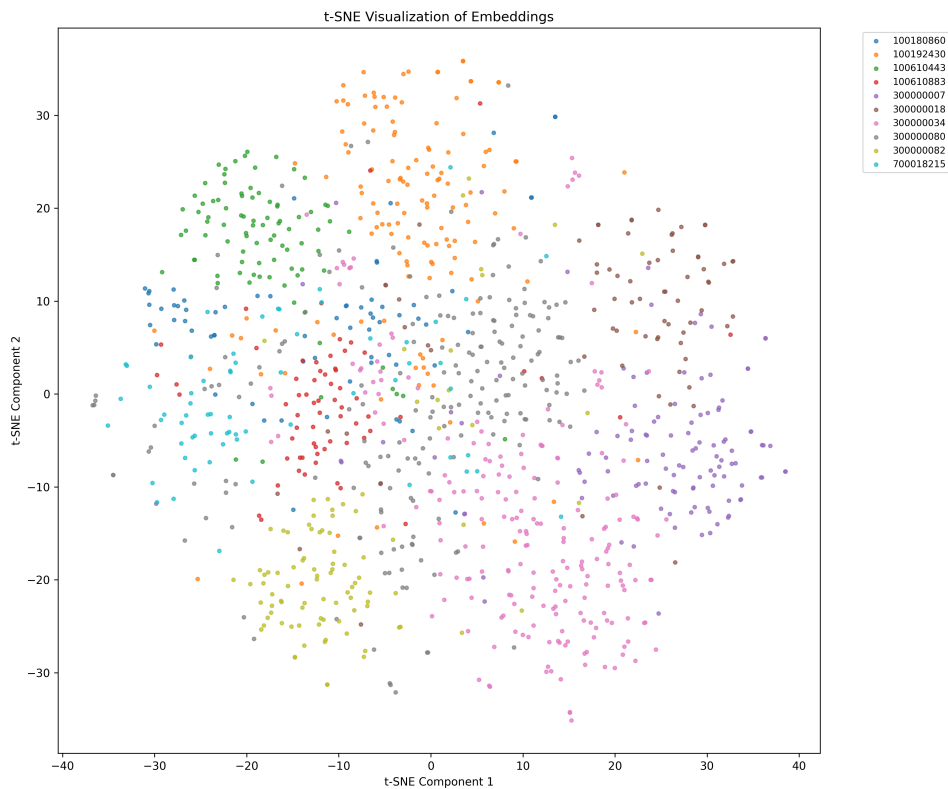
Total samples after cleaning: 1116
Number of syndromes (classes): 10
Embedding dimension: 320
■Top 6 Syndromes by Image Count:
  - 300000034: 210
  - 300000080: 198
  - 100192430: 136
  - 300000007: 115
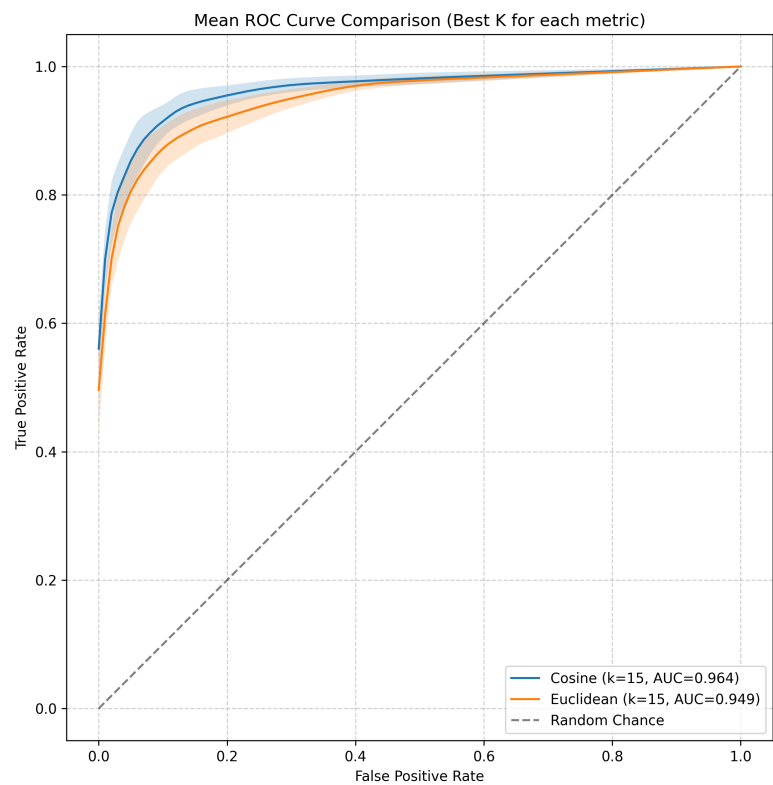 - 300000082: 98
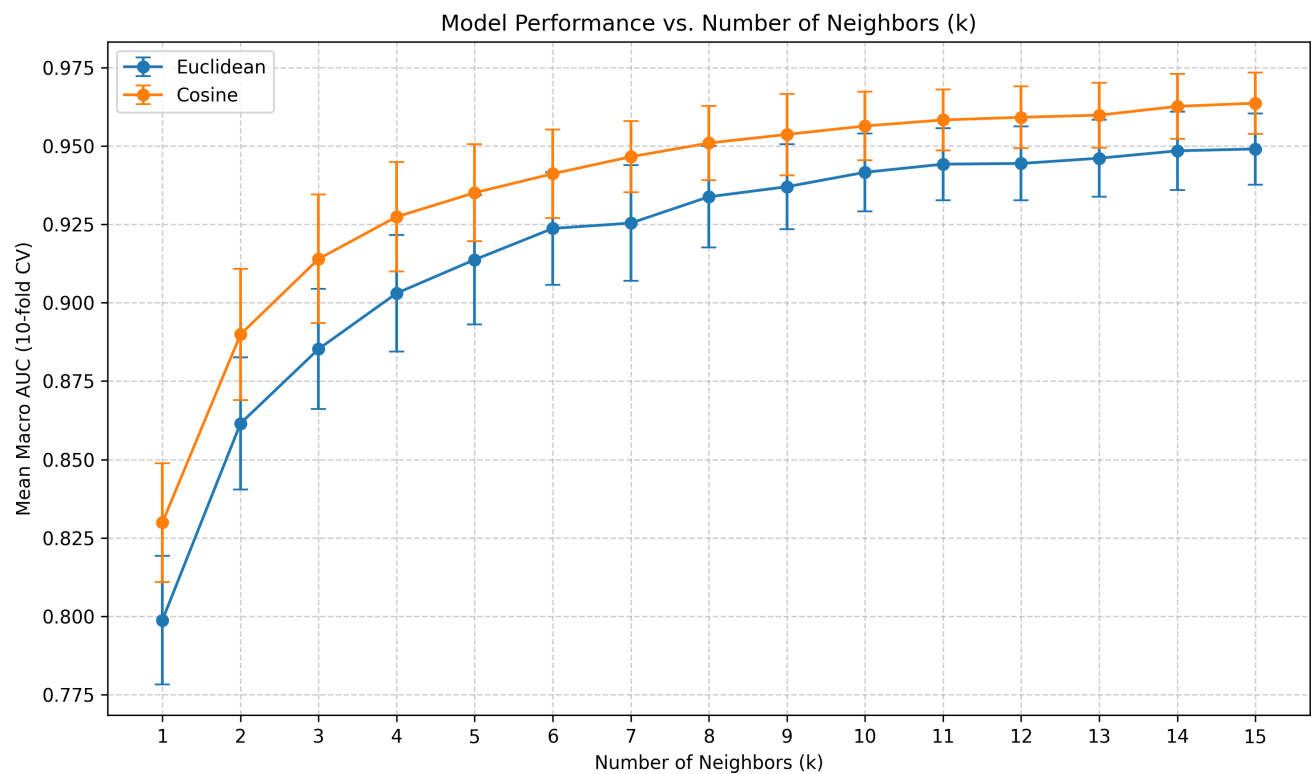 - 100610443: 89

2. t-SNE Visualization of Embeddings

KNN Pipeline - Performance Plots

Generated: 2025-10-01T21:39:08.298648 UTC

3. Mean ROC Curve Comparison (Best K)

Mean ROC Curve Comparison (Best K for each metric)

4. AUC vs. K Neighbors

Model Performance vs. Number of Neighbors (k)

KNN Pipeline - Summary Table

Generated: 2025-10-01T21:39:08.641650 UTC

5. Performance Metrics Across All Runs

Distance	K	AUC Mean	F1 Mean	Top-1 Mean	Top-3 Mean
euclidean	1	0.7988	0.6340	0.6585	0.7293
euclidean	2	0.8615	0.5806	0.6218	0.8279
euclidean	3	0.8852	0.6063	0.6612	0.8664
euclidean	4	0.9030	0.6583	0.6980	0.8772
euclidean	5	0.9137	0.6653	0.7069	0.8763
euclidean	6	0.9237	0.6716	0.7195	0.8870
euclidean	7	0.9254	0.6793	0.7123	0.8907
euclidean	8	0.9338	0.6946	0.7347	0.8969
euclidean	9	0.9370	0.7137	0.7455	0.9086
euclidean	10	0.9416	0.7072	0.7419	0.9113
euclidean	11	0.9442	0.7119	0.7580	0.9122
euclidean	12	0.9444	0.7184	0.7526	0.9131
euclidean	13	0.9461	0.7110	0.7499	0.9139
euclidean	14	0.9484	0.7127	0.7553	0.9157
euclidean	15	0.9490	0.7273	0.7597	0.9184
cosine	1	0.8299	0.6873	0.7141	0.7786
cosine	2	0.8899	0.6613	0.7008	0.8701
cosine	3	0.9140	0.7106	0.7437	0.9032
cosine	4	0.9274	0.7406	0.7715	0.9086
cosine	5	0.9351	0.7524	0.7705	0.9184
cosine	6	0.9411	0.7493	0.7849	0.9220
cosine	7	0.9465	0.7811	0.7920	0.9247
cosine	8	0.9509	0.7663	0.7992	0.9283
cosine	9	0.9537	0.7607	0.8019	0.9337
cosine	10	0.9564	0.7670	0.7921	0.9382
cosine	11	0.9583	0.7642	0.7965	0.9328
cosine	12	0.9591	0.7625	0.7992	0.9337
cosine	13	0.9598	0.7666	0.7939	0.9373
cosine	14	0.9626	0.7604	0.8028	0.9426
cosine	15	0.9636	0.7631	0.8055	0.9435