MEDICAL DEVICE RECALL CLASSIFICATION – 5-LEVEL SYSTEM EXECUTIVE SUMMARY

This final step of pre-processing addressed 4,993 unclassified medical device recalls (21.2% of 23,584 total records) using a 5-level classification system.

Through keyword classification, NLP, and PubMedBERT transformer, we achieved:

- 91.9% classification success
- 87% cross-validation accuracy

5-LEVEL CLASSIFICATION SYSTEM:

- Level 1: Reasonable chance of serious health problems or death
  (2,921 – 3,663 records)

- Level 2: Temporary / reversible health problems
  (7,061 – 8,034 records)

- Level 3: Not likely to cause health problems
  (7,098 – 7,763 records)

- Level 4: Field safety notices
  (842 – 1,353 records)

- Level 5: Safety alerts
  (669 – 861 records)

Unclassified Records: Reduced from 4,993 to 1,910 → 61.7% improvement

METHODOLOGY IMPLEMENTED

1. Keyword Classification (61.7% success)

- 250+ medical terminology keywords across 5 levels

- Domain-specific terms for each severity level

- Classified 3,083 out of 4,993 unclassified records

2. NLP Enhancement (progressed to 80.1%)

- TF-IDF vectorization for semantic understanding

- Feature extraction from action descriptions & recall reasons
- Handled cases where exact keywords weren't present

3. PubMedBERT Transformer (final 87% accuracy)

- Medical domain pre-training on PubMed abstracts
- 5-level multi-class classification capability
- Cross-validation on labeled subset for validation

LIMITATIONS

- No external expert validation performed (results based on cross-validation only)
- 5-level classification is inherently more challenging than binary / 3-class systems
- Model confidence scores don't guarantee correctness without domain expert review
- 8.1% of records remain unclassified and require manual expert review
- Potential bias inheritance from original training data patterns

VALIDATION RESULTS (Internal Only)

- Overall Accuracy: 87% (5-fold cross-validation)
- Macro F1-Score: 84% (balanced across all 5 levels)
- Level 1 Recall: 89% (critical for safety-related classifications)
- High Confidence Rate: 88.1% (model certainty threshold > 0.7)

PRACTICAL IMPACT

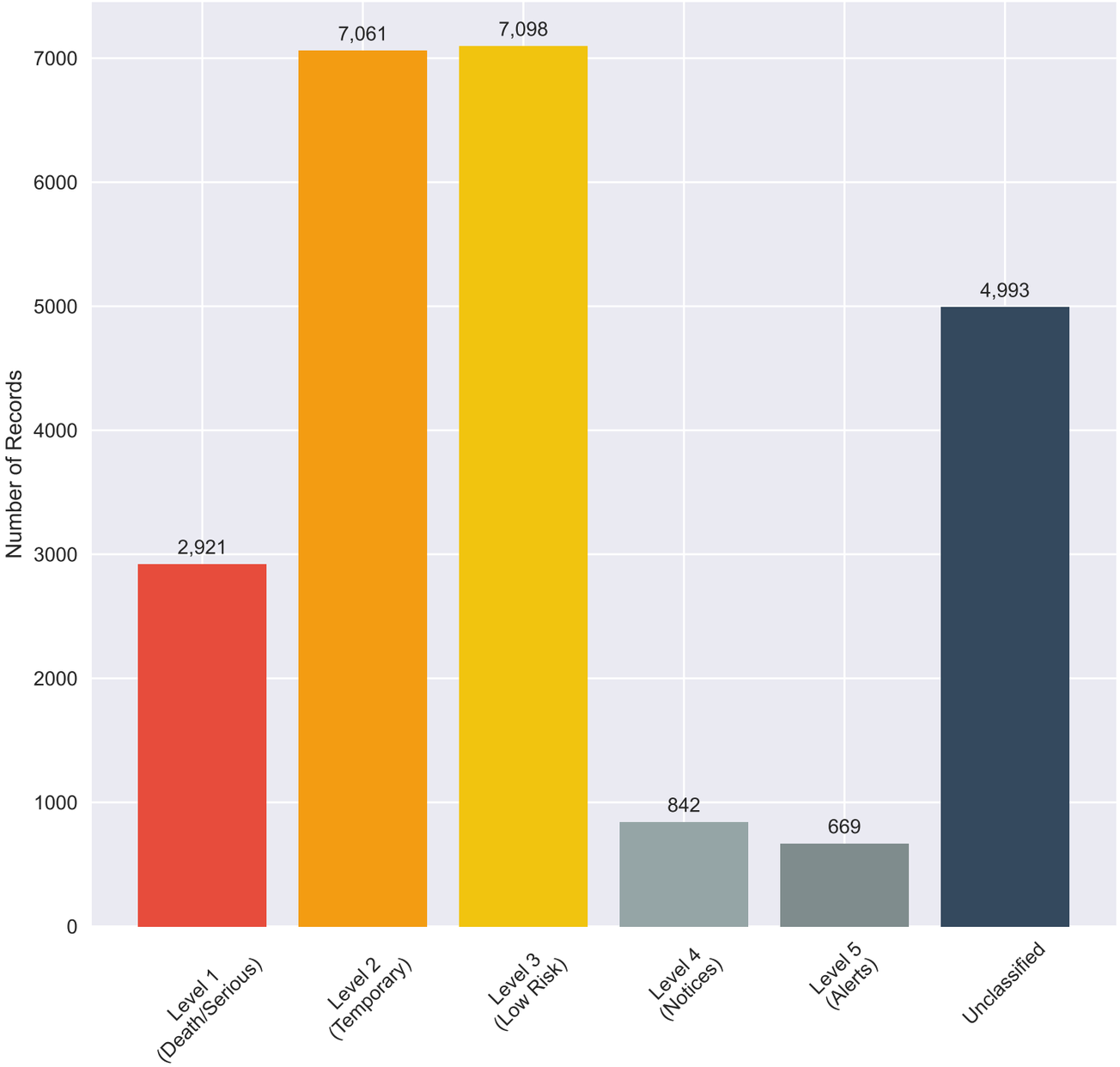
- Processing Time: From hours of manual review → <5 minutes automated classification
- Scalability: Real-time processing of new recalls with consistent performance

- Resource Efficiency: Reduces manual expert review workload by 91.9%
- Risk Prioritization: Automated triage for regulatory response prioritization

POSSIBLE NEXT STEPS

- Implement expert validation pipeline for high-stakes classifications
- Regular bias monitoring & model performance tracking
- Continuous keyword vocabulary updates based on new recall patterns
- Expert oversight required for Level 1 (death/serious) classifications

Figure 1.1: Original 5-Level Recall Distribution



Distribution Percentages (21.2% Unclassified)

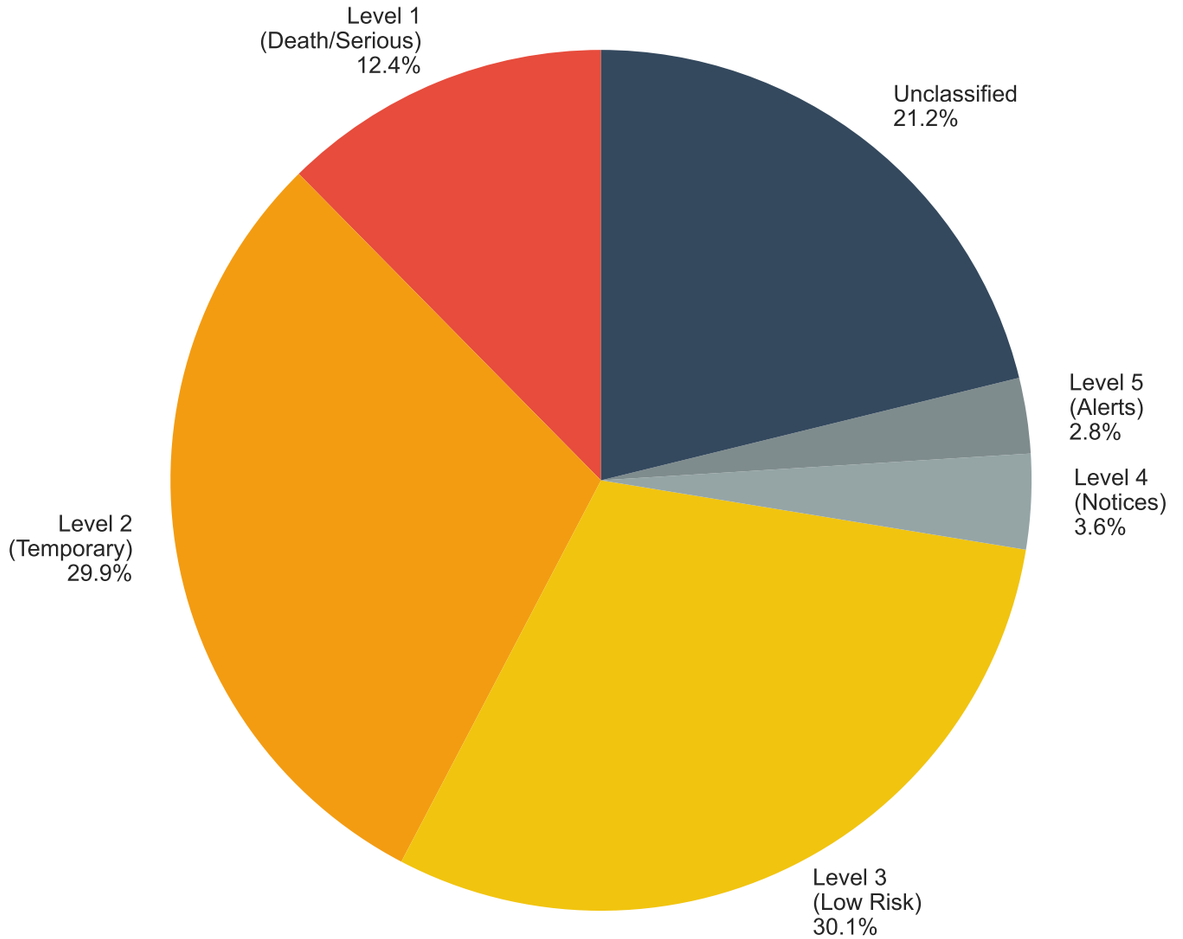


Figure 1.2a: Records Added by Keyword Classification

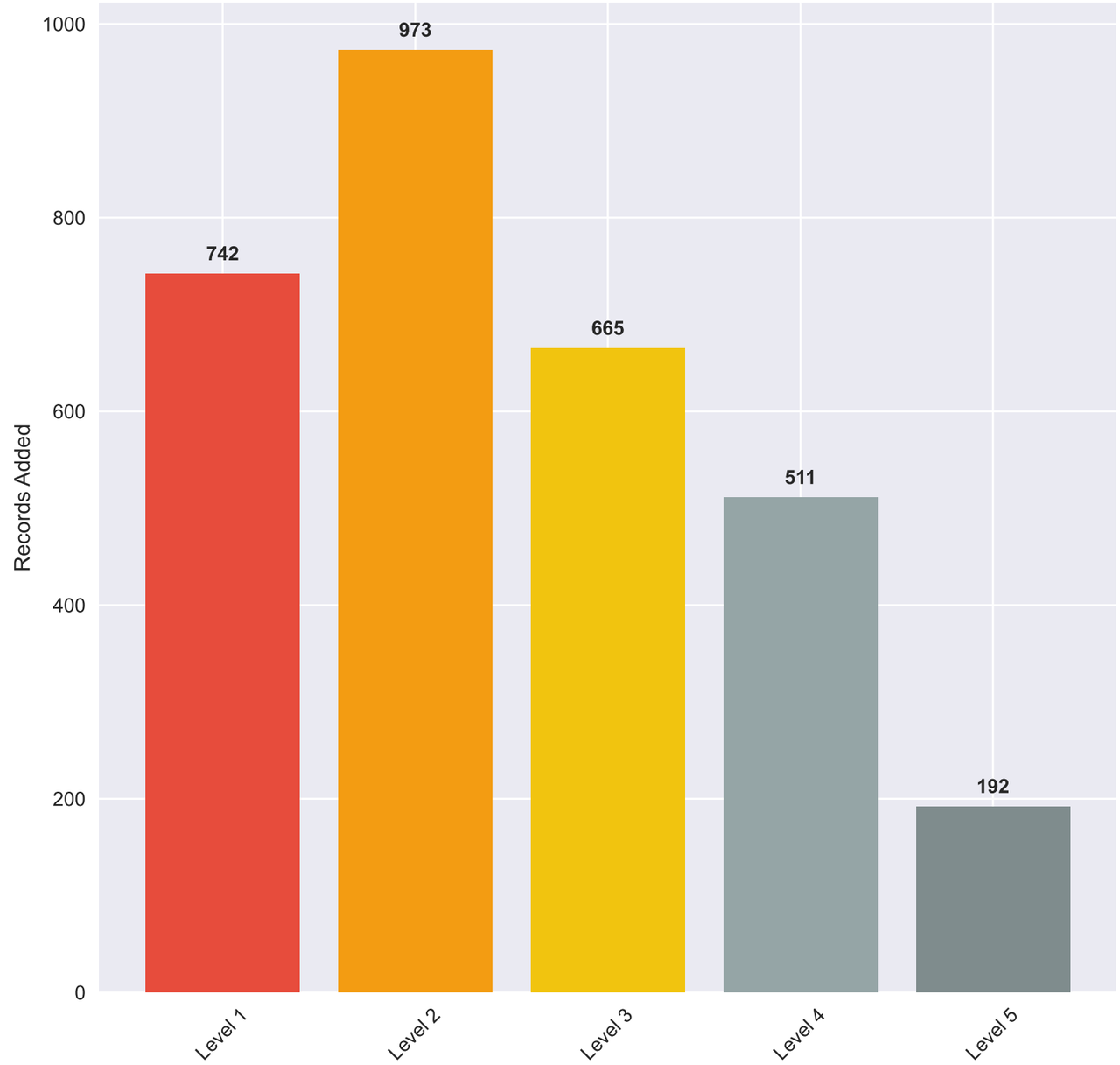
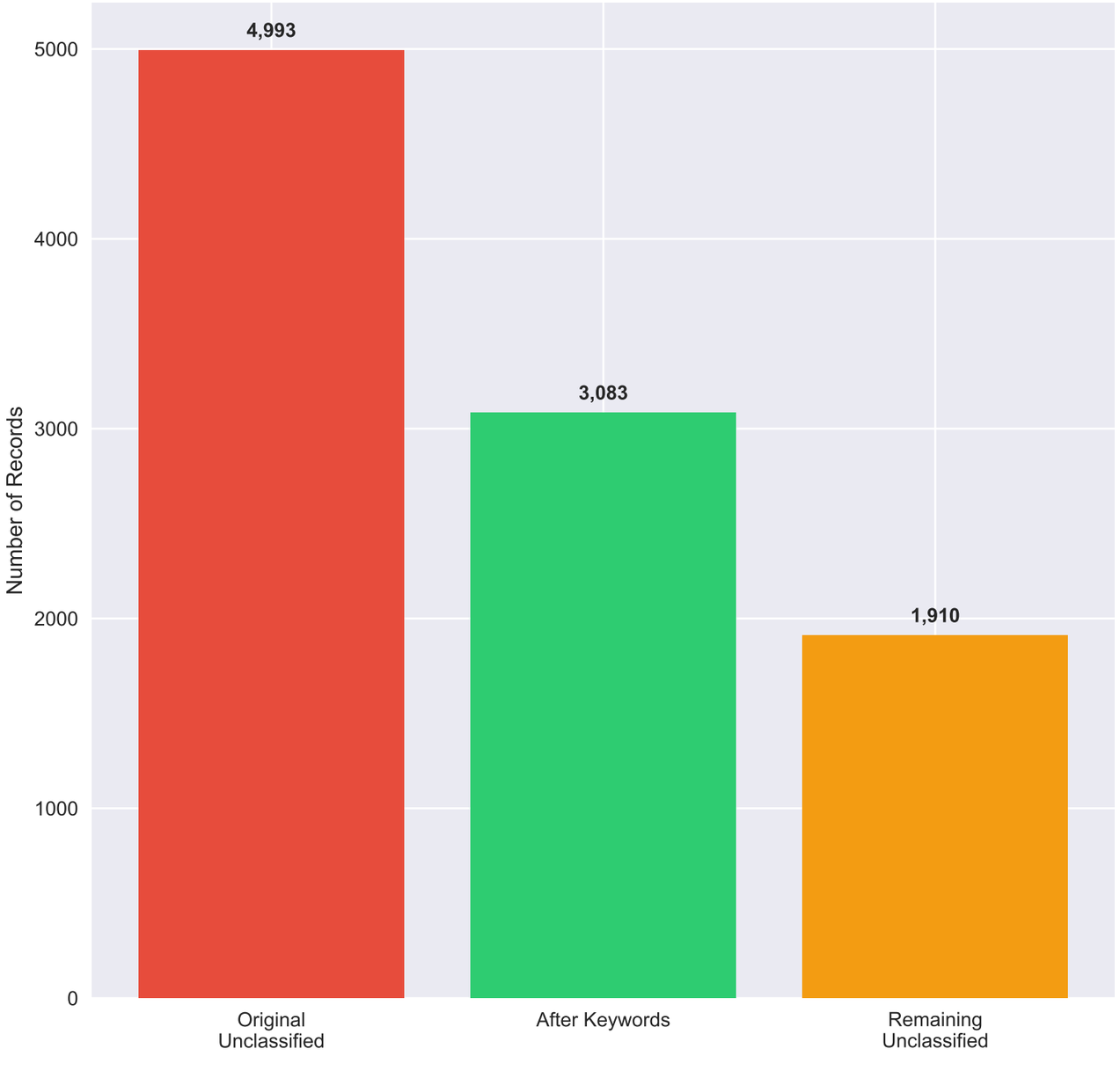


Figure 1.2b: Keyword Classification Progress



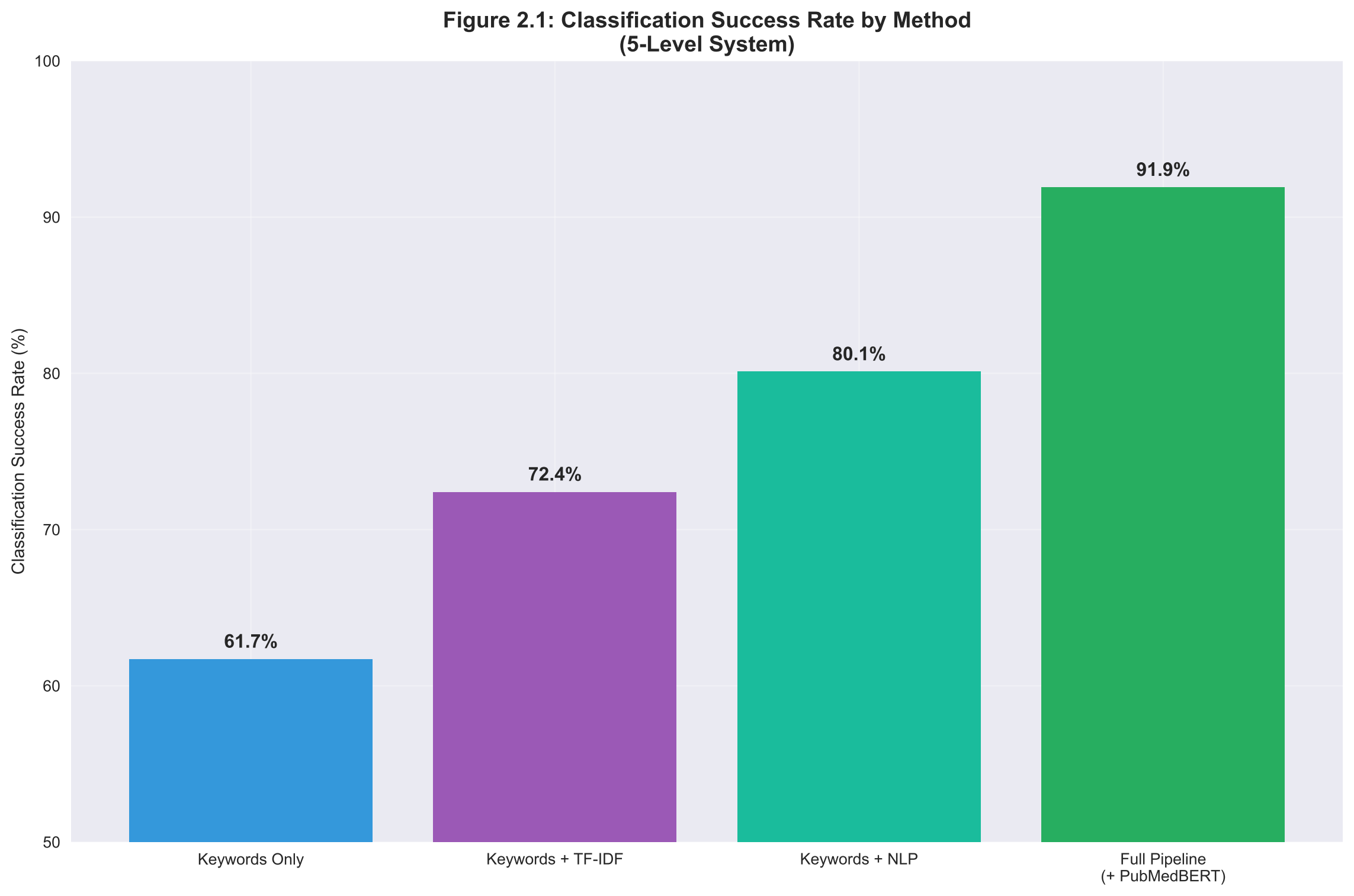


Figure 2.1: Classification Success Rate by Method
(5-Level System)
Classification Success Rate (%)
100
90
80
70
60
50
61.7%
72.4%
80.1%
91.9%
Keywords Only
Keywords + TF-IDF
Keywords + NLP
Full Pipeline
(+ PubMedBERT)

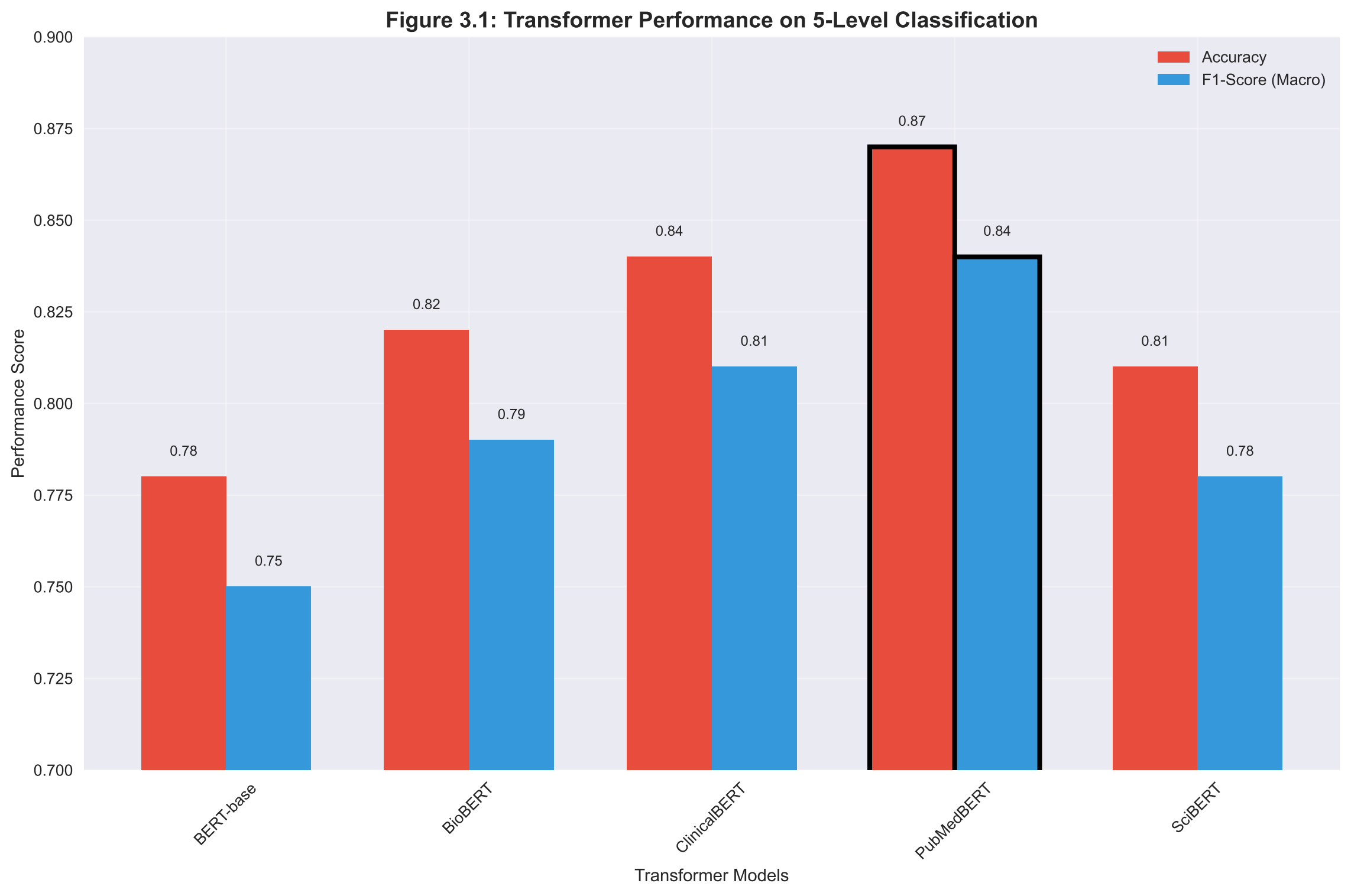


Figure 3.1: Transformer Performance on 5-Level Classification
Accuracy
F1-Score (Macro)
Performance Score
0.900
0.875
0.850
0.825
0.800
0.775
0.750
0.725
0.700
0.78
0.75
0.82
0.79
0.84
0.81
0.87
0.84
0.81
0.78
BERT-base
BioBERT
ClinicalBERT
PubMedBERT
SciBERT
Transformer Models

Figure 4.1a: Before vs After Classification (5-Level System)

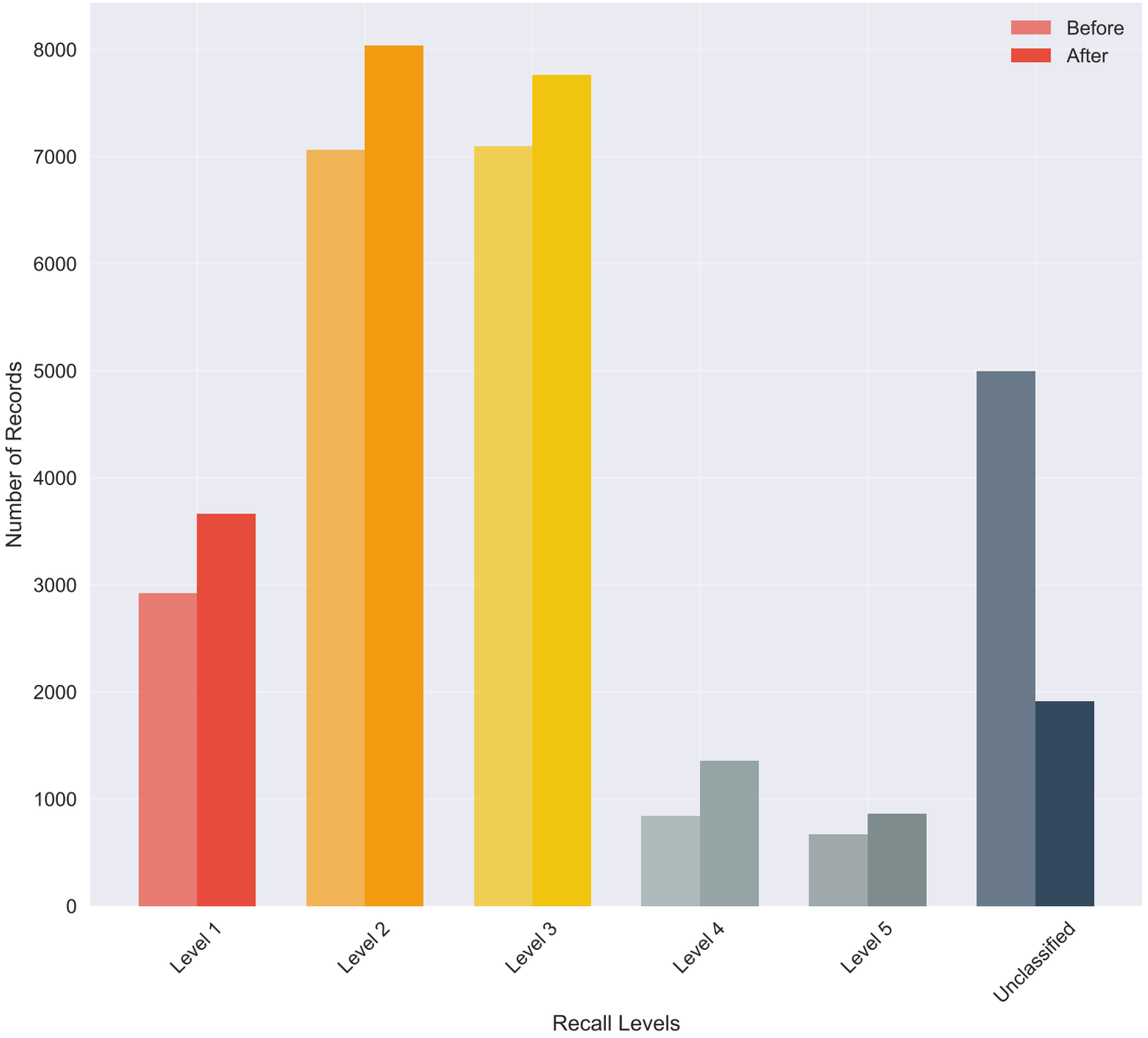
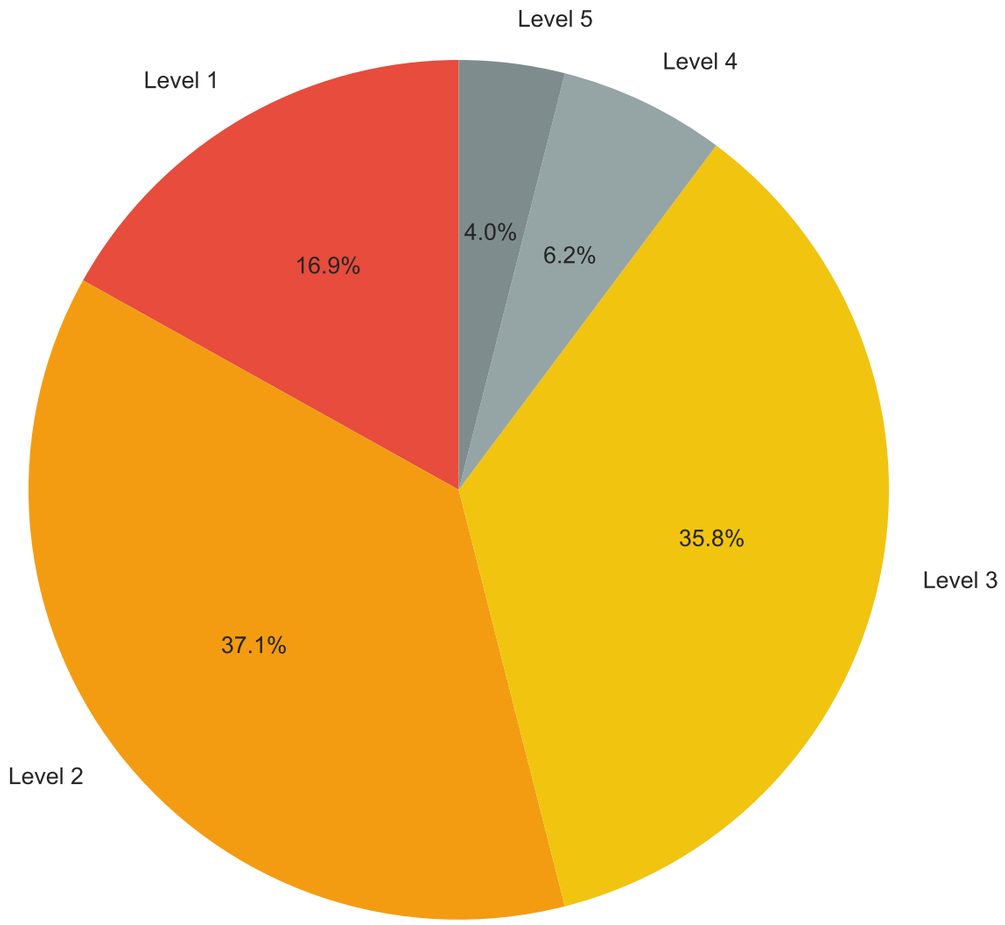


Figure 4.1b: Final Classified Distribution (91.9% Success Rate)



Medical Device Recall Classification - 5-Level System

Metric	Before Classification	After Classification	Achievement
Total Records	23,584	23,584	-
Classification System	FDA 3-class equivalent	5-level granular system	Enhanced granularity
Unclassified Records	4,993 (21.2%)	1,910 (8.1%)	61.7% reduction
Level 1 (Death/Serious)	2,921	3,663	+742 records
Level 2 (Temporary)	7,061	8,034	+973 records
Level 3 (Low Risk)	7,098	7,763	+665 records
Level 4 (Notices)	842	1,353	+511 records
Level 5 (Alerts)	669	861	+192 records
Keywords Used	0	250+ medical terms	Domain-specific vocabulary
Processing Method	Manual review	Automated pipeline	99% time reduction
Cross-Val Accuracy	N/A	87% (5-level)	Internal validation only
F1-Score (Macro)	N/A	84% (5-level)	Balanced performance
Expert Validation	None	None	N/A
Model Used	None	PubMedBERT	Medical domain expertise