Algorithm Pseudocode

GPU-Accelerated Semantic Severity Recalibration

1 Main Recalibration Algorithm

Algorithm 1 Semantic Severity Recalibration (GPU-Accelerated)

Dataset $\mathcal{D} = \{(\text{desc}^{(i)}, d_1^{(i)}, d_2^{(i)}, \hat{y}^{(i)}, c^{(i)})\}_{i=1}^N$ Severity prototypes $\mathcal{P} = \{P_C, P_{Ma}, P_{Mo}, P_{Mi}\}$ Sentence encoder f_θ (all-MiniLM-L6-v2) Drug class table $\mathcal{R}$ Weights $w_s = 0.45$, $w_c = 0.25$, $w_d = 0.30$ Target distribution $\mathcal{T} = [0.05, 0.25, 0.60, 0.10]$ Recalibrated predictions $\{\tilde{y}^{(i)}\}_{i=1}^N$ **Phase 1: Batch encode all descriptions on GPU** $\mathbf{E} \leftarrow GPUBatchEncode f_\theta, \{\text{desc}^{(i)}\}_{i=1}^N$ Shape: $N \times d$ $\mathbf{E} \leftarrow \mathbf{E}/\|\mathbf{E}\|$ Normalize embeddings **Phase 2: Compute prototype centroids** $k \in \{C, Ma, Mo, Mi\}$ $\mathbf{c}_k \leftarrow \text{mean}(f_\theta(P_k))$ Class centroid $\mathbf{c}_k \leftarrow \mathbf{c}_k/\|\mathbf{c}_k\|$ $\mathbf{C} \leftarrow [\mathbf{c}_C; \mathbf{c}_{Ma}; \mathbf{c}_{Mo}; \mathbf{c}_{Mi}]$ Centroid matrix: $4 \times d$ **Phase 3: Vectorized similarity computation** $\mathbf{S}_{sem} \leftarrow \mathbf{E} \cdot \mathbf{C}^T$ Cosine similarity: $N \times 4$ $\mathbf{s}_{sem} \leftarrow ThresholdToScores \mathbf{S}_{sem}$ Map to numeric **Phase 4: Parallel drug class scoring** $\mathbf{s}_{drug} \leftarrow ParallelDrugClassRisk\{(d_1^{(i)}, d_2^{(i)})\}, \mathcal{R}$ **Phase 5: Confidence adjustment** $\mathbf{s}_{conf} \leftarrow AdjustConfidence\{\hat{y}^{(i)}\}, \{c^{(i)}\}$ **Phase 6: Weighted combination** $\mathbf{s}_{final} \leftarrow w_s \cdot \mathbf{s}_{sem} + w_c \cdot \mathbf{s}_{conf} + w_d \cdot \mathbf{s}_{drug}$ **Phase 7: Quantile-based target calibration** $\{\tilde{y}^{(i)}\} \leftarrow QuantileCalibrate \mathbf{s}_{final}, \mathcal{T}$ $\{\tilde{y}^{(i)}\}_{i=1}^N$

2 Semantic Similarity Scoring

Algorithm 2 ThresholdToScores: Convert Similarities to Severity Scores

Similarity matrix $\mathbf{S} \in R^{N\times 4}$ (columns: Contra, Major, Mod, Minor) Thresholds $\tau_C = 0.65$, $\tau_{Ma} = 0.55$, $\tau_{Mo} = 0.45$ Severity scores $\mathbf{s} \in [1.5, 4.0]^N$ $\mathbf{s} \leftarrow \mathbf{1.5}$ Initialize all to Minor score $i \in \{1, \ldots, N\}$ Vectorized in practice $S_{i,0} \geq \tau_C$ $s_i \leftarrow 4.0$ Contraindicated $S_{i,1} \geq \tau_{Ma}$ $s_i \leftarrow 3.2$ Major $S_{i,2} \geq \tau_{Mo}$ $s_i \leftarrow 2.0$ Moderate $\mathbf{s}$

3 Quantile-Based Target Calibration

Algorithm 3 QuantileCalibrate: Exact Target Distribution Matching

Final scores $\mathbf{s}_{final} \in R^N$ Target distribution $\mathcal{T} = [t_C, t_{Ma}, t_{Mo}, t_{Mi}]$ Severity labels $\{\tilde{y}^{(i)}\}_{i=1}^{N}$ $\pi \leftarrow ArgSort\mathbf{s}_{final},$descending Sort indices by score $N \leftarrow |\mathbf{s}_{final}|$ Compute category boundaries from targets $n_C \leftarrow \lfloor N \cdot t_C \rfloor$ Top 5% $n_{Ma} \leftarrow \lfloor N \cdot t_{Ma} \rfloor$ Next 25% $n_{Mi} \leftarrow \lfloor N \cdot t_{Mi} \rfloor$ Bottom 10% $n_{Mo} \leftarrow N - n_C - n_{Ma} - n_{Mi}$ Remaining 60% Assign severities by rank $\tilde{\mathbf{y}} \leftarrow \mathbf{0}^N$ $\tilde{y}_{\pi[1:n_C]} \leftarrow$ "Contraindicated" $\tilde{y}_{\pi[n_C+1:n_C+n_{Ma}]} \leftarrow$ "Major" $\tilde{y}_{\pi[n_C+n_{Ma}+1:N-n_{Mi}]} \leftarrow$ "Moderate" $\tilde{y}_{\pi[N-n_{Mi}+1:N]} \leftarrow$ "Minor" $\tilde{\mathbf{y}}$

4 Confidence Adjustment

Algorithm 4 AdjustConfidence: Penalize Low-Confidence High-Severity

Original predictions $\{\hat{y}^{(i)}\}$ and confidences $\{c^{(i)}\}$ Thresholds $\tau_C = 0.65$, $\tau_{Ma} = 0.50$ Confidence-adjusted scores $\mathbf{s}_{conf}$ $\phi \leftarrow \{C : 4, Ma : 3, Mo : 2, Mi : 1\}$ Label to score map $i \in \{1, \ldots, N\}$ $s_i \leftarrow \phi(\hat{y}^{(i)})$ Base score from prediction $\hat{y}^{(i)} =$ "Contraindicated" **and** $c^{(i)} < \tau_C$ $s_i \leftarrow 3.0$ Downgrade uncertain contraindicated $\hat{y}^{(i)} \in \{$"Contra", "Major"$\}$ **and** $c^{(i)} < \tau_{Ma}$ $s_i \leftarrow 2.5$ Partial downgrade $\mathbf{s}_{conf}$

5 Drug Class Risk Assessment

Algorithm 5 ParallelDrugClassRisk: Pharmacological Risk Scoring

Drug pairs $\{(d_1^{(i)}, d_2^{(i)})\}_{i=1}^{N}$ High-risk classes $\mathcal{R} = \{\text{anticoag}, \text{antiplate}, \text{QT}, \text{MAOI}, \ldots\}$ Drug class scores $\mathbf{s}_{drug}$ Parallel execution on 24 CPU cores **parallel for** $i \in \{1, \ldots, N\}$ **do** $C_1 \leftarrow GetClassesd_1^{(i)}, \mathcal{R}$ $C_2 \leftarrow GetClassesd_2^{(i)}, \mathcal{R}$ $O \leftarrow C_1 \cap C_2$ Overlap "MAOI" $\in C_1 \wedge$ "Serotonergic" $\in C_2$ $s_i \leftarrow 4.0$ Fatal combination "MAOI" $\in O$ $s_i \leftarrow 4.0$ "anticoag" $\in O \vee$ "QT" $\in O$ $s_i \leftarrow 3.5$ $C_1 \neq \emptyset \wedge C_2 \neq \emptyset$ $s_i \leftarrow 3.0$ $C_1 \neq \emptyset \vee C_2 \neq \emptyset$ $s_i \leftarrow 2.5$ $s_i \leftarrow 2.0$ **end parallel for** $\mathbf{s}_{drug}$

6 Complexity Analysis

Phase	**Time Complexity**	**Actual Time**
GPU Batch Encoding	$O(N \cdot d/B)$	45.5s
Centroid Computation	$O(\|\mathcal{P}\| \cdot d)$	<1s
Similarity Matrix	$O(N \cdot 4)$	0.04s
Drug Class Scoring	$O(N/P)$	0.7s
Confidence Adjustment	$O(N)$	<1s
Quantile Calibration	$O(N \log N)$	<1s
Total	$O(N \cdot d/B + N \log N)$	**49.2s**

Table 1: Complexity analysis for $N = 759,774$, $d = 384$, $B = 8192$, $P = 24$ cores

7 Performance Summary

- **GPU**: NVIDIA RTX PRO 5000 (48GB VRAM)
- **CPU**: 24 cores for parallel drug class scoring
- **Memory**: 124GB RAM
- **Throughput**: 15,454 interactions/second
- **Embedding Rate**: 16,696 descriptions/second
- **Total Time**: 49.2 seconds for 759,774 interactions
- **Target Match**: Exact (0% deviation from literature targets)