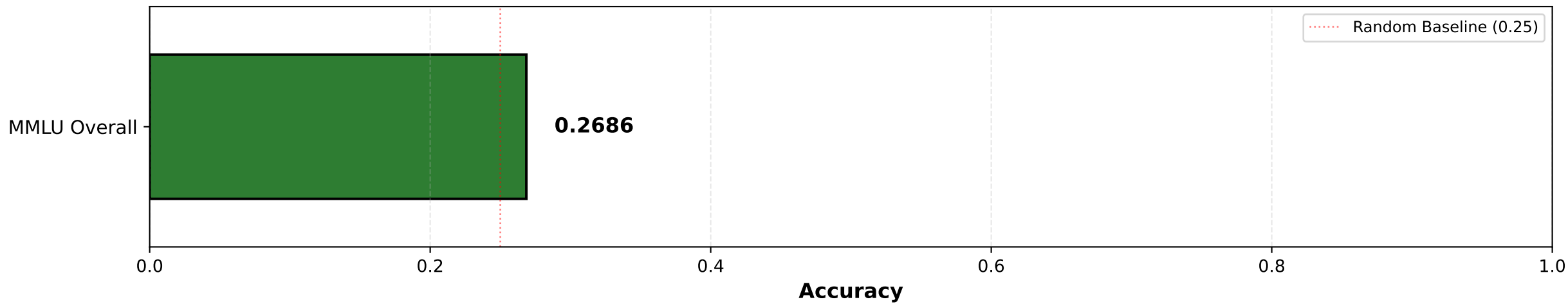
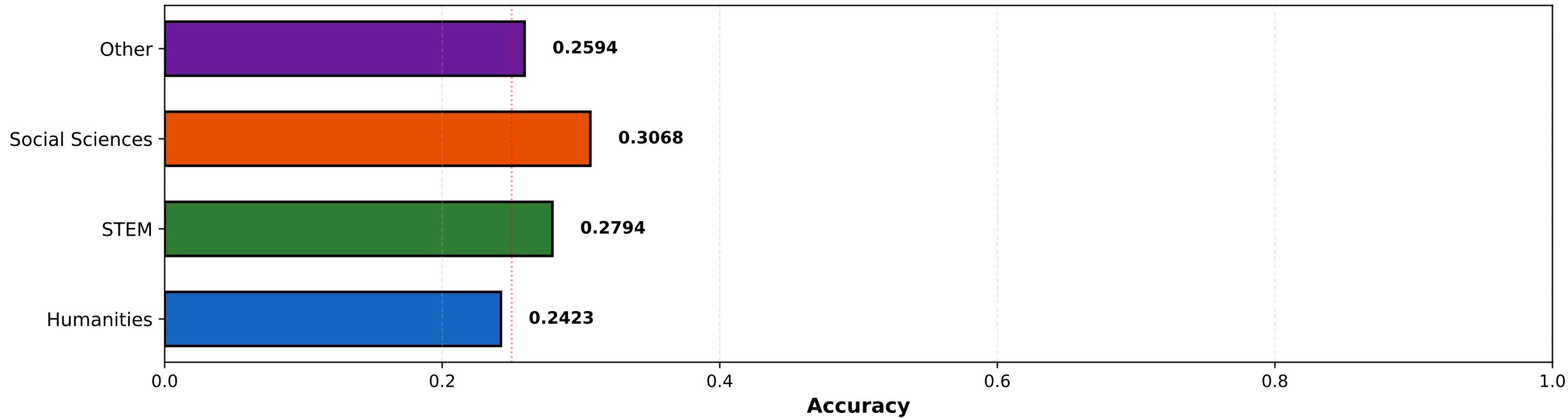


MMLU Performance Analysis: Three-Level Breakdown

Level 1: Overall MMLU Performance



Level 2: MMLU Category Performance



Level 3: Individual MMLU Subtask Performance (Grouped by Category)

