Table 13. Overall performance of our framework across different LLM engines, with and without the Validator module, and under different code slicing strategies. The results are grouped into three blocks: baseline models, our framework with Left Flow slicing, and with Full Flow slicing. All metrics are reported as (min, max) across three trials. For baseline models based on T5, min–max ranges are identical due to deterministic decoding using beam search.

Model	**KBI**↑	**FAR**$_1$ ↓	**CPI**$_1$ ↑	**FAR**$_2$ ↓	**CPI**$_2$ ↑
Baseline					
CodeReviewer	(0.00, 0.00)	(97.78. 97.78)	(0.00, 0.00)	–	–
CCT5	(2.22, 2.22)	(97.58, 97.58)	(2.32, 2.32)	(90.91, 90.91)	(3.57, 3.57)
LLaMA-Reviewer	(2.22, 2.22)	(97.62, 97.62)	(2.30, 2.30)	(92.86, 92.86)	(3.39, 3.39)
DISCOREV	(0.00, 0.00)	(97.78, 97.78)	(0.00, 0.00)	–	–
Left Flow					
LLaMa3.1 (70B) [w/o]	(15.56, 24.44)	(80.98, 87.52)	(13.85, 21.39)	(59.22, 75.09)	(19.16, 30.56)
LLaMa3.1 (70B) [w]	(2.22, 2.22)	(37.04, 37.04)	(4.29, 4.29)	(66.67, 66.67)	(4.17, 4.17)
Qwen2 (72B) [w/o]	(35.56, 44.44)	(90.04, 92.16)	(12.84, 16.28)	(81.23, 85.83)	(20.26, 26.39)
Qwen2 (72B) [w]	(22.22, 31.11)	(89.12, 92.03)	(11.73, 16.12)	(78.46, 84.47)	(18.28, 25.46)
Command R+ (103B) [w/o]	(15.56, 26.67)	(90.84, 94.22)	(8.43, 13.63)	(68.00, 80.58)	(17.28, 29.09)
Command R+ (103B) [w]	(4.44, 4.44)	(85.00, 85.00)	(6.86, 6.86)	(62.50, 62.50)	(7.95, 7.95)
Mistral-2407 (123B) [w/o]	(22.22, 31.11)	(89.42, 92.02)	(11.75, 15.79)	(70.39, 79.17)	(21.50, 30.34)
Mistral-2407 (123B) [w]	(24.44, 31.11)	(84.92, 88.22)	(15.90, 20.31)	(63.07, 70.78)	(26.62, 33.77)
LLaMA3.1 (405B) [w/o]	(26.67, 35.56)	(86.09, 89.27)	(15.31, 20.00)	(62.60, 72.95)	(26.86, 36.46)
LLaMA3.1 (405B) [w]	(15.56, 24.44)	(71.07, 79.67)	(17.63, 26.49)	(32.41, 53.02)	(23.38, 35.90)
Full Flow					
LLaMa3.1 (70B) [w/o]	(11.11, 15.56)	(86.96, 89.48)	(10.81, 14.19)	(56.86, 65.99)	(16.75, 22.87)
LLaMa3.1 (70B) [w]	(0.00, 0.00)	(44.44, 48.89)	(0.00, 0.00)	–	–
Qwen2 (72B) [w/o]	(37.78, 46.67)	(90.36, 91.35)	(14.07, 15.98)	(82.07, 84.95)	(21.53, 25.91)
Qwen2 (72B) [w]	(26.67, 33.33)	(91.10, 92.24)	(12.02, 14.05)	(75.47, 80.92)	(22.24, 28.26)
Command R+ (103B) [w/o]	(11.11, 17.78)	(93.58, 95.48)	(6.43, 9.43)	(73.70, 83.89)	(13.15, 21.22)
Command R+ (103B) [w]	(8.89, 8.89)	(76.30, 76.30)	(12.93, 12.93)	(58.33, 58.33)	(14.65, 14.65)
Mistral-2407 (123B) [w/o]	(26.67, 33.33)	(90.05, 91.00)	(13.46, 15.32)	(72.02, 77.29)	(24.53, 30.42)
Mistral-2407 (123B) [w]	(26.67, 33.33)	(84.85, 86.74)	(17.71, 20.83)	(62.82, 69.54)	(28.44, 35.15)
LLaMA3.1 (405B) [w/o]	(31.11, 31.11)	(89.10, 89.57)	(15.63, 16.15)	(71.90, 73.70)	(28.50, 29.53)
LLaMA3.1 (405B) [w]	(20.00, 20.00)	(77.68, 78.10)	(20.91, 21.10)	(66.61, 68.08)	(24.59, 25.02)