



# Coverage Reports Index

**Total Solutions:** 40      **LLM Implementations:** 4  
**Generated:** 2025-11-10 23:12:48

## gemma\_Self\_Planning

Problem	Tests	Line Coverage	Branch Coverage	Interpretation	HTML Report
HumanEval/1	4/4	95.8%	87.5%	Good line coverage but some branches untested	<a href="#">View Report</a> →
HumanEval/17	5/5	95.0%	87.5%	Good line coverage but some branches untested	<a href="#">View Report</a> →
HumanEval/26	3/3	100.0%	100.0%	Excellent coverage - well-tested code	<a href="#">View Report</a> →
HumanEval/77	7/8	100.0%	100.0%	1 test(s) failed - fix failing tests first	<a href="#">View Report</a> →
HumanEval/84	0/5	100.0%	N/A	5 test(s) failed - fix failing tests first	<a href="#">View Report</a> →
HumanEval/94	9/9	94.4%	93.8%	Excellent coverage - well-tested code	<a href="#">View Report</a> →
HumanEval/101	6/6	100.0%	100.0%	Excellent coverage - well-tested code	<a href="#">View Report</a> →
HumanEval/107	8/8	100.0%	100.0%	Excellent coverage - well-tested code	<a href="#">View Report</a> →
HumanEval/114	12/12	85.7%	75.0%	Moderate branch coverage - some	<a href="#">View Report</a> →

Problem	Tests	Line Coverage	Branch Coverage	Interpretation	HTML Report
				conditional branches untested	
HumanEval/156	<b>14/14</b>	<b>100.0%</b>	<b>100.0%</b>	Excellent coverage - well-tested code	<a href="#">View Report</a> →

## gemma\_self\_edit

Problem	Tests	Line Coverage	Branch Coverage	Interpretation	HTML Report
HumanEval/1	<b>4/4</b>	<b>96.2%</b>	<b>87.5%</b>	Good line coverage but some branches untested	<a href="#">View Report</a> →
HumanEval/17	<b>5/5</b>	<b>100.0%</b>	N/A	Good line coverage achieved	<a href="#">View Report</a> →
HumanEval/26	<b>3/3</b>	<b>100.0%</b>	<b>100.0%</b>	Excellent coverage - well-tested code	<a href="#">View Report</a> →
HumanEval/77	<b>0/8</b>	<b>0.0%</b>	N/A	8 test(s) failed - fix failing tests first	<a href="#">View Report</a> →
HumanEval/84	<b>0/5</b>	<b>100.0%</b>	N/A	5 test(s) failed - fix failing tests first	<a href="#">View Report</a> →
HumanEval/94	<b>9/9</b>	<b>95.3%</b>	<b>92.9%</b>	Excellent coverage - well-tested code	<a href="#">View Report</a> →
HumanEval/101	<b>6/6</b>	<b>100.0%</b>	N/A	Good line coverage achieved	<a href="#">View Report</a> →
HumanEval/107	<b>8/8</b>	<b>100.0%</b>	<b>100.0%</b>	Excellent coverage - well-tested code	<a href="#">View Report</a> →

Problem	Tests	Line Coverage	Branch Coverage	Interpretation	HTML Report
HumanEval/114	12/12	100.0%	100.0%	Excellent coverage - well-tested code	<a href="#">View Report →</a>
HumanEval/156	14/14	100.0%	100.0%	Excellent coverage - well-tested code	<a href="#">View Report →</a>

## llama\_self\_edit

Problem	Tests	Line Coverage	Branch Coverage	Interpretation	HTML Report
HumanEval/1	4/4	96.2%	87.5%	Good line coverage but some branches untested	<a href="#">View Report →</a>
HumanEval/17	3/5	88.2%	75.0%	2 test(s) failed - fix failing tests first	<a href="#">View Report →</a>
HumanEval/26	3/3	100.0%	100.0%	Excellent coverage - well-tested code	<a href="#">View Report →</a>
HumanEval/77	0/8	0.0%	N/A	8 test(s) failed - fix failing tests first	<a href="#">View Report →</a>
HumanEval/84	0/5	100.0%	N/A	5 test(s) failed - fix failing tests first	<a href="#">View Report →</a>
HumanEval/94	9/9	100.0%	100.0%	Excellent coverage - well-tested code	<a href="#">View Report →</a>
HumanEval/101	6/6	100.0%	N/A	Good line coverage achieved	<a href="#">View Report →</a>
HumanEval/107	8/8	100.0%	100.0%	Excellent coverage - well-tested code	<a href="#">View Report →</a>

Problem	Tests	Line Coverage	Branch Coverage	Interpretation	HTML Report
HumanEval/114	12/12	100.0%	100.0%	Excellent coverage - well-tested code	<a href="#">View Report →</a>
HumanEval/156	14/14	100.0%	100.0%	Excellent coverage - well-tested code	<a href="#">View Report →</a>

## llama\_self\_planning

Problem	Tests	Line Coverage	Branch Coverage	Interpretation	HTML Report
HumanEval/1	4/4	100.0%	100.0%	Excellent coverage - well-tested code	<a href="#">View Report →</a>
HumanEval/17	2/5	85.7%	75.0%	3 test(s) failed - fix failing tests first	<a href="#">View Report →</a>
HumanEval/26	3/3	95.7%	83.3%	Good line coverage but some branches untested	<a href="#">View Report →</a>
HumanEval/77	0/8	0.0%	N/A	8 test(s) failed - fix failing tests first	<a href="#">View Report →</a>
HumanEval/84	0/5	100.0%	100.0%	5 test(s) failed - fix failing tests first	<a href="#">View Report →</a>
HumanEval/94	9/9	95.3%	92.9%	Excellent coverage - well-tested code	<a href="#">View Report →</a>
HumanEval/101	6/6	100.0%	N/A	Good line coverage achieved	<a href="#">View Report →</a>
HumanEval/107	8/8	100.0%	100.0%	Excellent coverage - well-tested code	<a href="#">View Report →</a>

Problem	Tests	Line Coverage	Branch Coverage	Interpretation	HTML Report
HumanEval/114	12/12	100.0%	100.0%	Excellent coverage - well-tested code	<a href="#">View Report →</a>
HumanEval/156	14/14	100.0%	100.0%	Excellent coverage - well-tested code	<a href="#">View Report →</a>