Data Quality vs Cost: Diminishing Returns

Cost-Effective Zone

Justified for High-Stakes

95-99%
$10,000-20,000
Expert annotation
Extensive QC
Difficult edge cases

90-95%
$2,000-5,000
Validation required
Multiple annotators

80-90%
$500-1,000
Standard annotation

Key Insight
4x cost increase for
last 5% accuracy

Cost per 1000 Examples ($)

Label Accuracy (%)

$20,000

$15,000

$10,000

$5,000

$2,000

$500

80% 85% 90% 95% 97% 99%