

# Metrics & Methodology

"You get what you measure."

This document defines the ground truth metrics for [PROJECT\_NAME].

## 1. The North Star Metrics

Metric	Unit	Description	Success Threshold
[METRIC_A]	[UNIT]	[Description, e.g., Validation Loss]	< [VALUE]
[METRIC_B]	[UNIT]	[Description, e.g., Recall @ K]	> [VALUE]%
[METRIC_C]	[UNIT]	[Description, e.g., Throughput]	> [VALUE] / sec

## 2. The Audit Levels

We do not trust "training loss". We run explicit audits.

### Level 1: The Smoke Test

- **Goal:** Ensure the code runs without crashing.
- **Dataset:** Random/Synthetic tensor data.
- **Steps:** 10 forward/backward passes.
- **Pass:** Exit Code 0.

### Level 2: The Diagnostic Audit

- **Goal:** Verify the model can learn [SIMPLE\_TASK].
- **Dataset:** [OVERFIT\_DATASET].
- **Pass:** 100% Accuracy / 0.0 Loss.

### Level 3: The SOTA Certification

- **Goal:** Prove superiority over the baseline.
- **Dataset:** Full [TEST\_SET].
- **Pass:** Outperform [PREVIOUS\_SOTA] on [METRIC\_A].

## 3. Terminology Dictionary

- [TERM\_1]: [Definition].
- [TERM\_2]: [Definition].
- [TERM\_3]: [Definition].