ML Sector Training, Testing & Validation Plan

Executive Summary

This document outlines the comprehensive ML training strategy for walk-forward validation across key market sectors. The plan prioritizes high-growth sectors: Semiconductors, AI/Cloud, Defense/Aerospace, Nuclear Energy, and Quantum Computing.

1. SEMICONDUCTORS (19 Stocks)

Sub-Industry	Symbols	Description
Chip Makers	NVDA, AMD, INTC, AVGO, QCOM, MU, MRVL	GPU, CPU, memory
Equipment	ASML, LRCX, AMAT, KLAC	Lithography, etch
Design/EDA	SNPS, CDNS, ADI, MCHP, TXN, NXPI, ON	Chip design
AI Servers	SMCI	AI infrastructure

2. AI & CLOUD (16 Stocks)

Sub-Industry	Symbols	Description
AI Leaders	NVDA, MSFT, GOOGL, META, AMZN	Major AI platforms
Enterprise AI	CRM, NOW, PANW, CRWD	Business AI apps
AI Infrastructure	PLTR, PATH, DDOG, MDB, TEAM	Data platforms
Quantum	IONQ	Quantum computing

3. DEFENSE & AEROSPACE (6 Stocks)

Sub-Industry	Symbols	Description
Prime Contractors	LMT, RTX, NOC, GD	Defense systems
Space/Drones	RKLB, JOBY	Space launch, eVTOL

4. NUCLEAR & CLEAN ENERGY (4 Stocks)

Sub-Industry	Symbols	Description
Nuclear	CEG	Nuclear power
Solar	ENPH, FSLR	Solar tech
Utilities	NEE	Clean energy

Training Configuration

Parameter	Value
Algorithm	XGBoost Classifier (BigQuery ML)
Features	Essential 8 (rsi, macd, macd_histogram, momentum, mfi, cci, rsi_zscore, macd_cross)
Retraining	Monthly (21 trading days)
Validation Period	60-252 trading days
Model Caching	80% cost reduction
Batch Predictions	30x fewer queries

Estimated Costs

Batch	Stocks	Est. Cost
Semiconductors	19	$3.00
AI & Cloud	16	$2.50
Defense	6	$1.00
Energy	4	$0.75
TOTAL	**45**	**$7.25**

Execution Plan

Phase	Action	Duration
1	Generate PDF Report	Complete
2	Run Semiconductor Batch (19 stocks)	~30 min
3	Run AI/Cloud Batch (16 stocks)	~25 min
4	Run Defense Batch (6 stocks)	~10 min
5	Run Energy Batch (4 stocks)	~8 min
6	Compile Final Results Report	~5 min

Service Information

Cloud Run Service: ml-training-service
Timeout: 60 minutes | Memory: 4GB | Region: us-central1
Project: aialgotradehits

Generated: 2026-01-10