ML Training Progress Report

AIAlgoTradeHits.com - 2026-01-10 23:16

Executive Summary

This report summarizes the ML Walk-Forward Validation progress across multiple sectors and stock categories. The validation system uses BigQuery ML with XGBoost classifiers, trained with Essential 8 features and monthly retraining frequency. Overall results demonstrate strong predictive capability with 78% average accuracy across 21,846 predictions.

Metric	Value
Total Predictions	21,846
Correct Predictions	17,039
Overall Accuracy	78.0%
Sectors Validated	8
Unique Stocks	80+
Top Sector Accuracy	98.0% (Space/Robotics)
Model Type	XGBoost Classifier
Features Used	Essential 8
Retraining Frequency	Monthly

Sector Performance Summary

Sector	Accuracy	Predictions	Confidence	Recommendation
Space/Robotics	98.0%	1,090	HIGH	STRONG BUY
Dividend Champions	84.5%	14,447	HIGH	BUY
Semiconductors	81.3%	852	HIGH	BUY
Defense/Aerospace	79.6%	602	MEDIUM	BUY
Healthcare/Biotech	77.0%	1,300	MEDIUM	HOLD
Financials	75.2%	1,438	MEDIUM	HOLD
AI/Cloud	73.6%	1,756	MEDIUM	HOLD
Utilities	66.2%	361	LOW	AVOID

Top Performing Stocks (80%+ Accuracy)

Symbol	Company	Sector	Accuracy
HON	Honeywell	Space/Robotics	93.6%
ROK	Rockwell Automation	Space/Robotics	91.2%
CAT	Caterpillar	Industrials	88.8%
AVGO	Broadcom	Semiconductors	87.4%
RTX	Raytheon	Defense	87.4%
AAPL	Apple	Technology	87.3%
NVDA	NVIDIA	Semiconductors	85.2%
MSFT	Microsoft	AI/Cloud	83.3%
JNJ	Johnson & Johnson	Dividend	84.2%
INTC	Intel	Semiconductors	82.9%

Model Architecture

The ML system uses BigQuery ML with the following configuration:

Model Type: BOOSTED_TREE_CLASSIFIER (XGBoost)
Training Method: Walk-Forward Validation
Retraining: Monthly (optimal for dividend/long-term stocks)

Essential 8 Features:
1. RSI (14-period) - Momentum indicator
2. MACD - Trend direction
3. MACD Histogram - Momentum change
4. Momentum - Price momentum
5. MFI - Money flow index
6. CCI - Commodity channel index
7. RSI Z-Score - Normalized RSI
8. MACD Cross - Crossover signal

Prediction Target: Next-day price direction (UP/DOWN)
Confidence Threshold: 60%+ for signals, 65%+ for trades

Completed ML Tasks

- 5-symbol validation with 97.2% accuracy (AAPL, MSFT, GOOGL, NVDA, TSLA)
- Sector batch training across 8 sectors
- Dividend Champions validation (31 stocks, 84.5% accuracy)
- Cloud Run ML service deployment (60-min timeout)
- Production prediction API endpoints created
- Model caching for 80% cost reduction
- Batch prediction queries (30x fewer API calls)
- Added 26 missing dividend stocks to data pipeline

Recommended Next Steps

Phase 1: Model Improvement (Immediate)

1. Utilities sector feature engineering (add rate sensitivity)
1. Add volume profile features (OBV, VWAP deviation)
1. Implement sector-specific thresholds
1. Test weekly vs monthly retraining for volatiles

Phase 2: Production Deployment

2. Deploy /api/ml/predict endpoint to Cloud Run
2. Implement real-time confidence monitoring
2. Add model versioning for A/B testing
2. Create automated drift detection

Phase 3: Advanced Strategies

3. Multi-timeframe ensemble (hourly + daily signals)
3. Sector rotation based on accuracy rankings
3. Paper trading integration for live validation
3. Quarterly full backtest (2-year historical)

Production API Endpoints

Endpoint	Method	Description
/api/ml/predict	GET/POST	Real-time predictions for symbols
/api/ml/sector-recommendations	GET	Sector recommendations by accuracy
/api/ml/training-status	GET	Current model training status
/api/ml/walk-forward/run	POST	Start new validation run
/api/ml/walk-forward/runs	GET	List all validation runs
/api/ml/walk-forward/runs/<id>	GET	Get run details
/api/ml/walk-forward/runs/<id>/results	GET	Get daily predictions

Service URLs

Service	URL
Trading API	https://trading-api-1075463475276.us-central1.run.app
ML Training Service	https://ml-training-service-1075463475276.us-central1.run.app
Bulletproof Fetcher	https://bulletproof-fetcher-6pmz2y7ouq-uc.a.run.app
Frontend App	https://aialgotradehits.com

Cost Optimization Results

Optimization	Before	After	Savings
Model Caching	$0.50/pred	$0.10/pred	80%
Batch Predictions	30 queries	1 query	97%
Cloud Run (vs Function)	$5/hr	$0.30/hr	94%
Monthly Retraining	Weekly cost	Monthly cost	75%

AIAlgoTradeHits.com - ML Training Progress Report
BigQuery Dataset: aialgotradehits.ml_models
Cloud Run: ml-training-service
Generated: 2026-01-10 23:16