Complete Professional Guide: TFT Configuration for BTC Day Trading

1. Objective

This document explains how to migrate to Temporal Fusion Transformer (TFT), feed it properly with historical and live data, and configure it for stable BTC/USDT day trading (5m timeframe, 5-step horizon).

2. Historical Data Requirements

Minimum 3–5 years of BTC/USDT 5m candles. Data fields: timestamp, open, high, low, close, volume. Recommended derived features: RSI(14), ATR(14), MACD, rolling volatility (20,50), returns (1,3,5), EMA slopes, volume rolling stats. Store in Parquet format for efficiency.

3. Target Engineering

Future return (5 steps): $(close[t+5] / close[t]) - 1$. Train with Quantile Loss to obtain p10, p50, p90. Expected return = p50. Risk approximation derived from p10 tail.

4. TFT Core Hyperparameters

encoder_length: 60–120 prediction_length: 5 hidden_size: 32–64 attention_head_size: 4 dropout: 0.1–0.3 hidden_continuous_size: 16–32 output_size: 3 (quantiles) Use early stopping (patience 10–20 epochs). Batch size: 64–128 CPU / 256–512 GPU.

5. Training Best Practices

Use walk-forward validation (no random shuffle). Split data chronologically. Monitor validation loss, profit factor, drawdown. Avoid overfitting via dropout and early stopping.

6. Regime & Risk Integration

Combine TFT with Regime Filter (ADX > 18–20). Use ATR-based stop loss and take profit. Limit risk per trade to 1%. Daily max loss around 3%.

7. Retraining Strategy

Retrain daily or weekly. Trigger retrain if performance degrades or distribution drift detected. Version models with timestamp + commit hash.

8. Infrastructure Recommendations

CPU-only minimum: 8 cores, 32GB RAM, 200GB NVMe. Recommended with GPU: 16 cores, 64GB RAM, 12GB VRAM GPU, 500GB NVMe. Inference is lightweight; training benefits from GPU.

9. Monitoring & Dashboard

Track equity curve, rolling profit factor, drawdown, Sharpe ratio, regime frequency, trade distribution. Send Telegram alerts for retrain, drawdown breaches, and drift.