

The Four-Layer Simulation Framework

How the daily stock analysis system works — methodology, theory, and signal definitions

Overview: Every morning after the idio scan surfaces HIGH PAD stocks from SEC 8-K filings, this system runs four analytical layers on each ticker, assembles the signals using a local AI model (Qwen 2.5 7B), then sends the complete context to Claude API to simulate bull, base, and bear scenarios. The result is a structured brief showing what the most likely sequence of events is given the catalyst, who moves first, and what single data point would invalidate the thesis.

Daily Pipeline

9:00 AM	Idio Scan	SEC EDGAR 8-K filings → HIGH PAD tickers surfaced
9:30 AM	Layer 0–3	All four analytical layers run on each HIGH PAD ticker
10:15 AM	Qwen Assembly	Local AI summarizes 16 signals into structured context
10:15 AM	Claude Simulation	Bull / Base / Bear scenarios generated via Claude API
10:30 AM	Report Delivery	Results written to database, website updated

Layer 0 — Market Regime

Before analyzing any individual stock, Layer 0 classifies the macro environment. The same catalyst event plays out completely differently in a bull market vs a bear market. In bull regimes, institutions buy dips and accumulate on good news. In bear regimes, institutions use good news volume as exit liquidity — selling into strength. Layer 0 stamps three labels on every daily analysis.

The Three Regime Dimensions

Macro Trend

SPY and QQQ are measured against their 50-day and 200-day simple moving averages daily.

BULL	Price > 50MA > 200MA for both SPY and QQQ. Institutions buy dips. Positive catalysts get accumulated.
CHOP	Mixed signals — price between MAs or MAs converging. Transitional phase. Signals less reliable.
BEAR	Price < 50MA < 200MA for both. Institutions sell rips. Good news used as exit liquidity.

Sector Regime

For each ticker, its sector ETF (XLK for Tech, XLE for Energy, XLF for Financials, etc.) is compared against SPY performance over the trailing 30 days. This identifies where institutional rotation is occurring.

LEADING	+3% or more vs SPY	Institutional money flowing into this sector. Catalysts get bought.
NEUTRAL	-3% to +3% vs SPY	Sector tracking the broad market. No strong tailwind or headwind.
LAGGING	-3% to -5% vs SPY	Sector underperforming. Institutional interest waning.
DISTRIBUTION	Worse than -5% vs SPY	Active institutional selling. Good news used to reduce exposure.

Liquidity Regime

Three sub-signals each vote +1 (risk-on), 0 (neutral), or -1 (risk-off). A score of +2 or higher = RISK-ON. A score of -2 or lower = RISK-OFF. Otherwise NEUTRAL.

VIX Level Below 15 = risk-on. 15-25 = neutral. Above 25 = risk-off. Fear gauge.

VIX Trend 10-day slope of VIX. Rising = increasing fear. Falling = calming market.

HYG/LQD Ratio High yield vs investment grade bonds. Rising ratio = institutions accepting risk.

Layer 1 — Fundamentals

Layer 1 establishes ground truth about the company — not market perception, but actual financial reality. Data is pulled from yfinance and SEC EDGAR. Four signals are produced.

Quality Score

Three checks, each worth one point. Total score determines the label.

Check	Threshold	Why It Matters
Operating Margins	> 15%	Pricing power. Companies with strong margins can absorb cost shocks.
Revenue Growth	> 5% YoY	Is the business expanding or contracting?
Free Cash Flow	> \$0	Real cash generation vs accounting earnings. Harder to manipulate.

3/3 = STRONG | 2/3 = NEUTRAL | 0-1/3 = WEAK

Valuation

Uses trailing P/E ratio and EV/EBITDA. Either metric being extreme triggers the classification. Expensive stocks have no room for error — any miss is punished severely. Cheap stocks have bad news priced in — positive surprises create asymmetric upside.

CHEAP	PE < 15 or EV/EBITDA < 8
FAIR	PE 15-30 and EV/EBITDA 8-20
EXPENSIVE	PE > 30 or EV/EBITDA > 20

Insider Activity

SEC EDGAR Form 4 filings from the last 90 days. Insider selling is ambiguous — could be divorce, taxes, or portfolio diversification. Insider buying is a much stronger signal: executives rarely buy their own stock unless they believe it is undervalued. Cluster buying (multiple insiders within 30 days) is the most reliable individual stock signal that exists.

Classification: 2+ buys and buys > sells = BUYING. 2+ sells and sells > buys = SELLING. Otherwise = NEUTRAL.

Earnings Track Record

Last 8 quarters of actual EPS vs consensus estimate. Consistent beaters have management teams that under-promise and over-deliver — a key predictor of how the market will react to a Guidance Reset event. Consistent missers have lost credibility; even good news is discounted.

Classification: 6+ beats of 8 = CONSISTENT_BEAT. 3 or fewer beats = MISSER. Otherwise = MIXED.

Layer 2 — Expectations & Sentiment

The core insight behind Layer 2: price moves on changes in expectations, not changes in reality. A company can report excellent earnings and the stock can fall — if expectations were even higher. Layer 2 maps what the market currently believes and identifies where that belief is most likely wrong.

Consensus Gap

Wall Street analyst price targets and recommendation data from yfinance. The upside gap = (mean analyst target - current price) / current price.

UNDERPRICED	<i>Upside > 15%</i>	Market underestimates the stock vs analyst consensus.
FAIRLY_PRICED	<i>Upside -5% to 15%</i>	Current price within normal range of analyst targets.
OVERPRICED	<i>Upside < -5%</i>	Stock trading above analyst targets. Expectations elevated.

Options Flow

Options markets are where institutional money makes directional bets before news becomes public. The signal is the put/call volume ratio on the nearest expiry options chain. Unusually heavy call volume relative to puts indicates institutional accumulation — someone buying cheap leverage in anticipation of an upside move. Heavy put activity indicates hedging or directional downside bets.

ACCUMULATION	<i>Put/Call < 0.7</i>	Call volume nearly double puts. Institutional positioning for upside.
NEUTRAL	<i>Put/Call 0.7-1.3</i>	Balanced flow. No strong directional signal.
DISTRIBUTION	<i>Put/Call > 1.3</i>	Heavy put buying. Institutional hedging or downside positioning.

News Sentiment

Yahoo Finance RSS headlines from the last 7 days, scored using keyword analysis. The trend matters more than any single article. Consistently negative coverage before a positive catalyst = narrative shift potential (powerful price driver). Uniformly positive coverage = expectations may already be elevated.

Positive keywords: beat, upgrade, growth, record, strong, raised, exceeded, bullish, buy. Negative keywords: miss, downgrade, weak, cut, below, disappoints, bearish, sell, loss. Average score > 0.2 = IMPROVING. < -0.2 = DETERIORATING. Otherwise = NEUTRAL.

Retail Sentiment

Reddit post titles scored using the same keyword system. Used primarily as a contrarian signal at extremes. Overwhelming retail bullishness = retail is positioned and waiting, potentially exit liquidity for smarter money. Overwhelming retail bearishness = stock is hated and underowned, often precedes institutional accumulation.

Bull ratio > 0.65 = CROWDED_LONG. Bull ratio < 0.35 = CROWDED_SHORT. Otherwise = NEUTRAL.
Minimum 3 scored posts required to make a directional call.

Layer 3 — Positioning

Layer 3 identifies mechanical forces — buying or selling that happens regardless of fundamentals because actors are forced to act by mandate, margin, or structure. These are the invisible forces that can overwhelm fundamental analysis entirely.

Short Squeeze Potential

When a large percentage of a stock's float is sold short, any positive catalyst creates mechanical forced buying. Short sellers losing money on every tick upward are forced to buy to cover — which pushes the price higher — which forces more covering — which is self-reinforcing. Days to cover (short interest / average daily volume) measures how long the squeeze can run.

CRITICAL	<i>>20% short float AND >10 days to cover</i>	Maximum squeeze fuel. Small move up can trigger massive covering.
HIGH	<i>>15% short float OR >7 days to cover</i>	Significant squeeze potential. Watch for initial covering trigger.
MEDIUM	<i>>8% short float OR >4 days to cover</i>	Moderate short position. Some mechanical upside possible.
LOW	<i>Below all thresholds</i>	Minimal short interest. No mechanical squeeze dynamic.

Institutional Concentration

High institutional ownership by a small number of funds creates coordinated exit risk. When the thesis changes for one large fund, others with the same thesis often follow. Note: passive index funds (Vanguard, BlackRock, State Street) are less concerning than active hedge funds — they don't sell on thesis changes, only on index rebalancing.

Metric: top 10 institutional holders' combined % of shares outstanding. HIGH > 60% | MEDIUM 40-60% | LOW < 40%

ETF Flow Direction

An invisible but powerful force. When money flows into a sector ETF, the ETF must buy every stock in the index proportionally — regardless of individual fundamentals. A stock with great earnings during massive sector ETF outflows can go nowhere as mechanical ETF selling offsets fundamental buyers.

Measured by comparing recent 5-day average ETF volume vs prior 5-day average volume, combined with price direction. Rising volume + rising price = TAILWIND. Rising volume + falling price = HEADWIND. Flat volume = NEUTRAL.

Gamma Exposure

When retail and institutions buy call options, market makers hedge by buying the underlying stock (delta hedging). As price rises toward the strike price, market makers must buy MORE stock to maintain their hedge.

This creates a gamma squeeze — self-reinforcing mechanical buying. High open interest concentrated at strikes near the current price amplifies moves in both directions.

Metric: total open interest (calls + puts) at strikes within 5% of current price on nearest expiry. HIGH > 50,000 contracts | MEDIUM 20,000-50,000 | LOW < 20,000

Simulation Layer — AI Scenario Generation

The simulation layer takes all 16 signals from Layers 0-3 plus the catalyst event from the idio scan and generates structured forward scenarios. It is not a prediction engine — it is a scenario mapping engine that acknowledges uncertainty and models probability-weighted ranges of outcomes.

Two-Step Process

Step 1 — Qwen 2.5 7B Context Assembly (Local)

Qwen 2.5 7B runs locally on the Hetzner server via Ollama. Before sending data to the Claude API, Qwen is used to distill the 16 raw signals into a structured narrative paragraph. Qwen identifies the 3-5 most important signals, flags contradictions (e.g. options showing bullish accumulation but insiders selling), and notes any data gaps. This preprocessing step dramatically reduces the Claude API cost and improves output quality by giving Claude pre-analyzed context rather than raw JSON.

Step 2 — Claude Sonnet Scenario Simulation (API)

Claude Sonnet receives the Qwen-assembled context plus the raw signals and the catalyst event, then generates three structured scenarios.

BULL CASE

The most favorable plausible outcome. Specifies which actor class moves first (short covering, momentum funds, value institutions, retail, market makers, options expiry), the expected price path and timeframe, and what must be true for this scenario to unfold.

BASE CASE

The probability-weighted most likely outcome, often more muted than the bull case due to friction, competing narratives, and sector headwinds. This scenario receives the highest probability weighting.

BEAR CASE

The realistic downside if 1-2 key assumptions prove wrong. Not the worst case — the plausible downside. Specifies what goes wrong and how the price path unfolds.

Fragility Point

The single most important output of the simulation. For every analysis, the system identifies the one assumption the entire thesis depends on most heavily. What single data point, if it changed, would completely invalidate the base case? This tells you exactly what to monitor — not everything, just the specific variable that determines whether the thesis holds.

Watch Variables

2-3 specific, observable events to monitor in the next 48-72 hours that would confirm or deny the base case. These are actionable — pre-market order flow patterns, options activity at specific strikes, competitor reactions, sector ETF volume. Not vague themes, but specific things you can check.

Outcome Tracking & Model Improvement

Every simulation is stored in the database with the date, ticker state, and all signals. When the actual outcome plays out, it is recorded against the prediction. Over time this builds a dataset of where the model is consistently right vs consistently wrong.

What Gets Tracked

Signal	What is recorded
Scenario accuracy	Did the bull, base, or bear case play out?
First mover accuracy	Was the predicted actor class correct?
Price path magnitude	How close was the predicted range to actual?
Fragility point validity	Did the flagged risk actually materialize?
Regime accuracy	Did the regime classification correctly predict behavior?

The goal is to identify regime/event type combinations with consistently high hit rates vs combinations that are essentially guessing. For example: the model may prove highly accurate for M&A; Shock events in BULL regimes but less reliable for Guidance Resets in CHOP regimes. This tells you when to size up vs stay small.

Data Sources & Cost

Source	Used For	Cost
yfinance	Price data, fundamentals, options chains, insider data	Free
SEC EDGAR API	Form 4 insider filings, 8-K event filings	Free
Yahoo Finance RSS	News headlines for sentiment scoring	Free
PullPush / Reddit	Retail sentiment from Reddit posts	Free
Shortvolume.com	Short interest data	Free
Ollama / Qwen 2.5 7B	Local AI context assembly (runs on server)	Free
Claude Sonnet API	Scenario simulation (~\$0.01-0.02 per stock)	~\$15/mo
Hetzner CPX42	8 vCPU / 16GB RAM server in Helsinki	\$22.59/mo

Total new infrastructure cost: ~\$37.59/month on top of existing server.

Idiosyncratic Risk Intelligence System — ericbau.com/risk.html — For informational purposes only. Not financial advice.