

AI Stock Oracle — DeepCut Improvement Spec

1. Core Differentiator

Not “news aggregation.” Realtime structural signal detector for CAPEX shockwaves across AI + power + water. Output = tradable lead time on illiquid microcaps and private suppliers.

2. Data Layers (Add)

- a) Utility commission dockets (FERC, PUC) – early voltage upgrade plans.
- b) Building permit bulk feeds – parse for data center SIC codes.
- c) Railcar waybill data – spikes in concrete, steel, HVAC loads.
- d) Satellite nighttime light delta (NOAA VIIRS) – hidden rural builds.
- e) Port import manifests – abnormal immersion cooling tank runs.

3. Signal Weighting v2

Score = $0.35 \times \text{Sync}(r) + 0.25 \times \text{Physical footprint} + 0.2 \times \text{Contract} + 0.1 \times \text{Hiring} - 0.1 \times \text{Noise}$.
Physical footprint = permits + rail + night lights. Adds hard asset confirmation.

4. Model Stack Upgrade

- spaCy + Ray for parallel parsing.
- DeepDedup entity resolution.
- STL residual cross correlation.
- LightGBM active learning retrain nightly.

5. Dashboard v2 Minimum

- Heatmap: correlation vs obscurity.
- One click provenance trail.
- Export to TensorBoard projector.

6. Security / Regulatory

- Treat rail/port manifests as CISA sensitive; encrypt at rest.
- FinCEN MSB check if rev > \$1 M.
- Hash personal data in job posting scrape.

7. GoToMarket Stack

Phase 0: free weekly “Orbital Strikes” email.

Phase 1: \$1 K/mo seat.

Phase 2: \$10 K/mo desk licence.

Phase 3: \$250 K/yr quant data licence.

8. Capital / OPEX

Run MVP on single M4 + Lightsail frontend (<\$120/mo).

Satellite/rail feeds ≈ \$800/mo.

Five enterprise pilots cover costs.

9. Execution Criticals (Do■or■Die)

- Archive raw data day 1.
- Build provenance UI first.
- Sell “motion detection,” not sector coverage.
- Fallback scraping for throttled APIs.
- Each alert must cite two evidence types.