# Analysis Phase

**📘 Objective of This Phase**

Create a **proof document** for each major stock move (gap up/down), going backwards to:

* Identify **the exact news trigger**
* Check **who published it first**
* Check if it was on **NSE/BSE**, or **news media**
* Grade its **impact strength**
* Analyze **how quickly it was reflected in price**
* Evaluate if our bot could’ve caught it early

**📋 Proof Document Template**

You’ll use this template for each stock event. You can store it as analysis.md.

**🧾 1. Case Title**

Natco Pharma – Q3 FY24 Results Decline – 12th Feb 2024

**🗓 2. Key Dates**

| **Event** | **Date & Time** |
| --- | --- |
| Quarterly results release | 12 Feb 2024, 15:45 PM |
| First media article | 12 Feb 2024, 16:03 PM |
| Market gap down | 13 Feb 2024, 09:15 AM |

**📉 3. Price Movement**

| **Date** | **Open Price** | **Close Price** | **% Change** |
| --- | --- | --- | --- |
| 12 Feb 2024 | ₹ X | ₹ Y | -1.2% |
| 13 Feb 2024 | ₹ Y | ₹ Z | -6.8% |

✅ Save price\_chart.png using TradingView or Zerodha Kite chart screenshot

**📰 4. News Timeline**

**Extract from BSE/NSE:**

* Source: NSE Corporate Announcements
* Filing type: "Outcome of Board Meeting"
* Download and save PDF: bse\_announcement.pdf
* Time: 15:45 IST (after market)

**Media Publication Timing:**

| **Outlet** | **Headline** | **Time** |
| --- | --- | --- |
| Moneycontrol | “Natco Pharma Q3 profit plunges 37% YoY” | 16:03 IST |
| Economic Times | “Natco disappoints with poor earnings” | 16:15 IST |
| Twitter/X (if any) | [screenshot/tweet] | 16:00–16:30 |

**🧠 5. Impact Strength Analysis**

**Event:**

Net profit declined 37% YoY → **Very Bearish**

**Heuristic Rating:**

| **Metric** | **Value** | **Heuristic Score** |
| --- | --- | --- |
| EPS YoY | -37% | -85 |
| Revenue YoY | -12% | -40 |
| Commentary | Negative | -30 |
| Final Sentiment | **-100** | 🔴 (Very Bearish) |

**🚨 6. Could We Have Captured This?**

| **Component** | **Status** | **Could Bot Detect?** | **Notes** |
| --- | --- | --- | --- |
| NSE filing time | ✅ | Yes | Raw text can be parsed immediately |
| PDF structured parsing | ✅ | Yes | Profit drop was mentioned in summary |
| Media coverage | ✅ | Lagged | ~15 mins delay |
| Google News API | ✅ | Yes (with query) | Indexed in <30 mins |
| Price reaction time | ✅ | Next morning | Time to act: ~16 hrs |

**Verdict:**  
✅ YES – Our bot **could have caught** the event **via BSE/NSE scraper**, and **scored -100** before market open next day.

**📌 7. Recommendations**

* Add a rule: "If YoY profit falls > 20%, and revenue also down, score -90 or more"
* Parse .pdf results using regex + PDFTextExtract + keyword matching
* NSE data wins on **timing** vs media

**🔨 Tools for Proof Creation**

| **Task** | **Tool** |
| --- | --- |
| Price charts | TradingView / Kite / Yahoo Charts |
| NSE filings | nseindia.com |
| PDF parsing | pdfplumber, PyMuPDF, camelot |
| News timestamps | Google News Search + timestamps |
| Screenshots | Snipping Tool / ShareX |
| Markdown documentation | Obsidian / Typora / VS Code |

**🧭 Workflow for Each Case Study**

1. Identify a **big move** (gap up/down) in a stock
2. Search NSE/BSE announcements for 24 hrs before
3. Download relevant filings + parse for core data
4. Search Google News + Moneycontrol etc. → Check timing
5. Save screenshots of charts, news
6. Write analysis.md + could\_we\_have\_captured.md
7. Repeat for 10-20 cases

**📌 HIGH-IMPACT CORPORATE EVENTS (India Focused)**

**🧾 A. Earnings & Financial Results**

| **Event** | **Potential Impact** |
| --- | --- |
| Quarterly Earnings (Q1, Q2, Q3, Q4) | 📈📉 Very High |
| Net Profit up/down YoY/QoQ | 📈📉 Very High |
| EPS up/down | 📈📉 High |
| Missed Estimates | 📉 High (if known) |
| Revenue Growth / Drop | 📈📉 High |
| Margins Improvement / Decline | 📈📉 Moderate |
| Auditor Comments or Reservations | 📉 High (negative sentiment) |

**💰 B. Dividends & Capital Returns**

| **Event** | **Potential Impact** |
| --- | --- |
| Interim Dividend Declaration | 📈 Moderate |
| Final Dividend Declaration | 📈 Moderate |
| Special Dividend | 📈 High |
| Dividend Cut | 📉 High |
| Record Date / Ex-Dividend Date | 📉📈 Short-term |
| Bonus Issue | 📈 High |
| Stock Split | 📈 Moderate |
| Buyback Announcement | 📈 High |
| Buyback Completion | 📈📉 Low–Moderate |

**🏛️ C. Board / Corporate Governance Events**

| **Event** | **Potential Impact** |
| --- | --- |
| Board Meeting for Financial Results | 🟡 Neutral (until result) |
| Board Meeting for Fundraising | 📈 Moderate |
| Resignation of CEO/CFO/MD | 📉 Very High |
| Appointment of New CEO | 📈📉 Depends on market perception |
| Change in Promoter Shareholding | 📈📉 High |
| Pledge / Unpledge of Shares | 📉📈 Moderate |
| Board Expansion / Restructure | 🟡 Neutral |

**🔄 D. Mergers, Acquisitions, Demergers**

| **Event** | **Potential Impact** |
| --- | --- |
| M&A Deal Signed | 📈📉 Very High |
| Company Acquired by Another | 📈 High (target company) |
| Acquiring a Loss-Making Business | 📉 High (acquirer) |
| Demerger / Spin-off | 📈 High (if value-unlocking) |
| JV (Joint Venture) Formation | 📈 Moderate |
| Strategic Partnership | 📈 Low–Moderate |

**🧯 E. Regulatory / SEBI / Legal Events**

| **Event** | **Potential Impact** |
| --- | --- |
| SEBI Show Cause Notice | 📉 Very High |
| Income Tax / ED Raids | 📉 Very High |
| Litigation / Court Order | 📉 Moderate–High |
| NCLT / Bankruptcy Filing | 📉 Extreme |
| RBI Restrictions (for NBFCs/Banks) | 📉 High |
| Penalty or Fine | 📉 Moderate |
| Downgrade by Credit Rating Agency | 📉 High (esp. for NBFCs) |

**🧨 F. Operational Disruptions**

| **Event** | **Potential Impact** |
| --- | --- |
| Fire / Factory Accident | 📉 High |
| Production Halt / Plant Shutdown | 📉 Moderate–High |
| Supply Chain Disruption | 📉 Moderate |
| Product Launch Failure | 📉 Low–Moderate |
| Loss of Key Client / Contract | 📉 Moderate–High |

**🚀 G. Growth & Expansion**

| **Event** | **Potential Impact** |
| --- | --- |
| New Plant / Facility Commissioned | 📈 Moderate–High |
| Large Order Win | 📈 High |
| Entry into New Geography | 📈 Moderate |
| Business Expansion Plan | 📈 Moderate |
| Capex Plan Approval | 📈 Low–Moderate |

**🧪 H. Sector-Specific Events**

| **Sector** | **Event Example** | **Impact** |
| --- | --- | --- |
| Pharma | USFDA Approval or Warning Letter | 📈📉 Very High |
| Banking | RBI CRR/SLR Changes, NPA Numbers | 📉📈 High |
| Infra | Order Wins, EPC Contract Awards | 📈 High |
| IT | Attrition Rates, New Client Wins | 📈📉 Moderate |
| NBFC | Liquidity Issue / RBI Restrictions | 📉 Very High |

Coding Phase   
  
✅ Objective

To build a real-time news-based trading system for Indian equities that scores events like earnings, dividends, mergers, CEO resignations, and general news from original sources and ranks them on a -100 to +100 sentiment scale for possible trading decisions.

🔧 System Architecture Overview

┌─────────────┐ ┌──────────────┐ ┌────────────┐

│ Global News │─────▶│ Sentiment │─────▶│ Scoring │

│ + Local │ │ Analysis │ │ Engine │

└─────────────┘ └──────────────┘ └────────────┘

▲ ▲ ▼

│ │ ┌─────────────┐

│ ┌──────┴──────┐ │ Alert & │

│ │ Heuristics │────────▶│ Execution │

│ └─────────────┘ └─────────────┘

│ ▲

┌──────┴─────┐ ┌─────┴──────┐

│ Official │ │ Google │

│ Sources: │ │ News API │

│ NSE/BSE │ └────────────┘

└────────────┘

🧠 Components Breakdown

1. Data Sources

A. Official Sources (First-to-Know Data)

You want the data directly from primary exchanges (NSE/BSE):

* Quarterly Results
* Dividend Announcements
* Board Meetings
* Insider Trades
* Mergers & Acquisitions
* CEO/CXO resignations

🔗 APIs / Scrapers:

* NSE India Announcements
* BSE Corporate Announcements

✅ Action:

* Build a web scraper (or use API if available via NSE/BSE partners)
* Poll every 1-2 minutes to get real-time announcements
* Parse structured data like PDFs with OCR/NLP if needed

B. Google News / Web Scraping

For events that haven't yet hit NSE/BSE:

* Use Google Custom Search API
* Query format: "<Stock Name> news" filtered by past 24 hours
* Example: HDFC Life news past 24 hours

✅ Action:

* Extract top 3-5 headlines per stock
* Filter by keywords: "merger", "resign", "fire", "raid", "upgrade", "downgrade", "SEBI", etc.

2. Sentiment Analysis Layer

A. Custom NLP for Financial Events

General NLP models don’t give enough weight to structured financial events.

✅ Plan:

* Build your own Heuristic Rules Engine:
  + Quarterly earnings growth > 10% → +70
  + Dividend increased YOY → +40
  + CEO Resignation → -60
  + Audit issues / SEBI investigation → -100
  + M&A (younger firm acquiring bigger) → +30 to +80
* Maintain keyword dictionaries and regex extractors for:
  + YoY, QoQ EPS/Profit Growth
  + Management changes
  + Dividend declarations
  + New project launches
  + Regulatory actions

B. Standard NLP for News Headlines

Use pretrained models like:

* FinBERT (fine-tuned for financial sentiment)
* Vader (for quick text scoring)
* Custom fine-tuned models (optional for phase 2)

3. Scoring Engine (-100 to +100)

You’ll build a multi-layered scoring mechanism:

| Source | Data Type | Base Weight | Custom Score |
| --- | --- | --- | --- |
| NSE/BSE | Earnings/Dividend | +80 | + based on growth |
| Google News | Resignation | -60 | via keywords |
| Social Media (Phase 2) | Rumors/Leaks | ±40 | TBD |
| Macros | USD/INR, Nifty, Fed | ±20 | via global data |

Example Rule:

“If net profit is up 20% YoY and revenue up 10% QoQ, then +90 score.”

Aggregated Score =

(NSE Event Weight × Confidence) + (News Sentiment × Relevance)

4. Execution Layer (Alert / Action)

* Once score crosses a threshold, say:
  + +70 or more: Alert for Buy
  + -70 or less: Alert for Sell
* Manual or Automated Order Execution via broker API (Zerodha Kite, Fyers, Angel One)

📡 Global Sentiment Monitor

You want to add global market context:

✅ Sources:

* S&P Futures / Nasdaq Futures
* USD/INR
* Crude Oil Prices
* US Fed Interest Rate Updates

✅ Tools:

* Investing.com API or scraper
* Twitter/X sentiment (optional)

Example:

“If global sentiment is negative (-50), reduce local positive signal by 25%.”

📁 Folder Structure (Dev)

news-trading-bot/

│

|----orchestration/

---|pipeline

├── data\_sources/

│ ├── nse\_scraper.py

│ ├── google\_news.py

│ ├── global\_macro.py

│

├── nlp/

│ ├── sentiment\_analysis.py

│ ├── heuristics.py

│ └── finbert\_model.py

│

├── scoring/

│ └── score\_calculator.py

│

├── alerts/

│ └── telegram\_alert.py

│

├── main.py

└── config.ini

🛠️ Tools & Tech Stack

| Task | Suggested Tool/Library |
| --- | --- |
| Web scraping | requests, BeautifulSoup, Selenium, Playwright |
| PDF parsing | pdfplumber, PyMuPDF, Tesseract OCR |
| NLP / Sentiment | Transformers, FinBERT, spaCy, TextBlob, Vader |
| Real-time news | Google News API, GNews API, SerpAPI |
| Scheduler | APScheduler, Celery, or simple cron |
| Broker Integration | Zerodha Kite Connect, Angel One SmartAPI |
| Alerting | Telegram Bot, Slack, Email |
| Hosting | Local VPS or Cloud (AWS Lambda, EC2, etc.) |

📈 Phase-wise Roadmap

| Phase | Goal | Time |
| --- | --- | --- |
| 1 | Build NSE/BSE scraper + Heuristic scoring | 1 week |
| 2 | Add Google News + NLP headline sentiment | 3-4 days |
| 3 | Scoring + Alert System | 3 days |
| 4 | Add Global Market Factors | 2 days |
| 5 | Integrate with Broker for semi-auto trades | 1 week |
| 6 | Optimize speed (async, caching, proxy) | Ongoing |

✅ Summary of Key Custom Features

* ✅ Fastest reaction to structured financial announcements
* ✅ Custom sentiment scoring tuned for stock-specific weights
* ✅ Real-time Google News scanning per stock
* ✅ Multi-source data fusion (NSE, News, Global)
* ✅ Alert system based on score thresholds

🧩 Optional Add-ons (Future)

* Real-time Twitter/X monitoring
* Order book anomaly detection (sudden volume spikes)
* Earnings Whisper estimates comparison
* Live dashboard for visualizing all events/stocks
* ML model to predict impact of past similar news