📊 Finfluencer Accuracy Leaderboard (India)

Project Documentation

---

1. Project Overview

Project Name

Finfluencer Accuracy Leaderboard (India)

Project Type

AI-powered finance analytics & public accountability platform

Inspiration

Inspired by projects like Alpha Arena (nof1.ai), but focused on **Indian finance YouTubers** and **long-term prediction accuracy** instead of short-term alpha calls.

---

2. Problem Statement

In India, finance YouTubers and influencers (finfluencers) regularly publish market predictions and macroeconomic opinions such as:

- Stock market direction predictions
- Index level expectations (NIFTY, SENSEX)
- Sectoral outlooks
- Asset allocation advice (equity vs gold vs real estate)

These predictions influence **millions of retail investors**, yet:

- There is **no systematic tracking** of prediction accuracy
- Creators highlight successful calls and ignore failed ones
- Viewers lack an objective way to judge long-term credibility

Core Problem

There is **no transparent, data-driven system** to evaluate how accurate these finfluencers actually are over time.

---

3. Project Objective

To build a **public, neutral, evidence-based leaderboard** that:

1. Extracts financial predictions from YouTube videos
2. Compares predictions against real market outcomes
3. Scores prediction accuracy objectively
4. Ranks finfluencers based on historical performance

This platform **does not provide investment advice** and focuses only on **retrospective analysis**.

---

4. Target Audience

Primary Users

- Retail investors
- Finance learners
- Analysts and researchers

Secondary Users

- Finfluencers themselves
- Journalists and media
- Fintech communities

---

5. Initial Creator Scope

Examples of creators to include initially:

- **Akshat Shrivastava**
- **Ankur Warikoo**
- **CA Rachana Ranade**
- Pranjal Kamra
- Asset Yogi

The platform is **creator-agnostic** and expandable.

---

6. Key Insight

Finance videos often contain **implicit or explicit predictions**, such as:

- “Markets will correct soon”
- “Gold will outperform equities”
- “This stock will struggle in the next year”

These statements can be:

- Extracted from transcripts
- Evaluated after sufficient time passes
- Compared against factual market data

This enables **objective scoring**.

---

7. High-Level System Architecture

YouTube API
↓
Video Metadata & Transcripts
↓
Prediction Extraction (LLM)
↓
Prediction Structuring
↓
Market Outcome Retrieval
↓
Prediction Evaluation (LLM)
↓
Scoring Engine
↓
Leaderboard & Creator Pages

---

8. Technology Stack

Frontend

- Next.js (App Router)

- TypeScript
- Tailwind CSS
- shadcn/ui
- Recharts (data visualization)

Backend

- Node.js (via Next.js API routes or separate service)
- Prisma ORM
- PostgreSQL

AI & Data Services

- YouTube Data API
- YouTube Transcript API
- Gemini API (free tier)
- Exa AI or similar search-based retrieval
- Yahoo Finance / NSE data (fallback)

---

9. Functional Requirements

9.1 Creator Management

- Store creator metadata
- Map YouTube channel to creator profile
- Aggregate scores per creator

9.2 Video Ingestion

- Fetch videos from YouTube channels
- Filter finance/market-related videos
- Store metadata (title, URL, publish date)

9.3 Transcript Processing

- Fetch full video transcript
- Segment transcript by timestamps
- Store transcript blocks

9.4 Prediction Extraction

- Identify prediction statements using AI
- Extract:
    - Claim
    - Asset / market

    - Time horizon
    - Confidence level (if mentioned)

9.5 Outcome Evaluation

- Fetch real market data after prediction horizon
- Compare predicted vs actual outcomes
- Generate natural-language explanation

9.6 Scoring System

- Assign numerical accuracy scores
- Aggregate per video and creator

9.7 Frontend Display

- Public leaderboard
- Creator detail pages
- Prediction-level breakdown

---

10. Database Design (Prisma)

Creator Table

```
model Creator {
  id            String   @id @default(uuid())
  name          String
  youtubeHandle String
  channelId     String
  avatarUrl     String
  createdAt     DateTime @default(now())

  videos        Video[]
  score         CreatorScore?
}
```

Video Table

```
model Video {
  id          String   @id
  creatorId   String
  title       String
  url         String
  publishedAt  DateTime
```

```
  predictions  Prediction[]
}
```

Prediction Table

```
model Prediction {
  id            String   @id @default(uuid())
  videoId       String
  timestamp     Int
  transcriptText String

  predictedClaim String
  predictionType String
  horizonMonths  Int

  outcome       Outcome?
  score         Float?
}
```

Outcome Table

```
model Outcome {
  id            String   @id @default(uuid())
  predictionId   String

  actualOutcome  String
  supportingData String
  evaluation     String
  accuracyScore  Float
}
```

CreatorScore Table

```
model CreatorScore {
  creatorId      String @id
  overallScore   Float
  totalEvaluated Int
  lastUpdated    DateTime
}
```

---

11. AI Processing Pipeline

Step 1: Prediction Extraction (Gemini)

Input:

- Transcript segment

Output:

```
{
  "claim": "NIFTY will fall 15% in 2023",
  "asset": "NIFTY50",
  "horizon_months": 12,
  "confidence": "high"
}
```

Step 2: Market Outcome Retrieval

- Search historical market performance
- Store:
    - Price movement
    - Time period
    - Key events

Step 3: Prediction Evaluation (Gemini)

Evaluation Prompt:

- Compare claim vs reality
- Generate explanation
- Score accuracy between **0.0 – 1.0**

---

12. Scoring Methodology

Prediction-Level Score

Factors:

- Direction correctness
- Magnitude closeness
- Time horizon match

- Clarity of prediction

Example:

- Direction correct
- Magnitude partially correct
- Horizon correct

Final Score: **0.70**

Creator-Level Score

Weighted average of:

- Prediction accuracy
- Confidence
- Time horizon
- Recency

---

13. Frontend Pages

`/leaderboard`

- Rank
- Creator name
- Overall accuracy %
- Total predictions evaluated

`/creator/[slug]`

- Creator profile
- Accuracy trend chart
- List of predictions:
    - Video link
    - Timestamp
    - Prediction
    - Outcome
    - Score

`/methodology`

- Explanation of system
- Scoring logic

- AI limitations

---

14. Cron Jobs & Automation

- Daily video ingestion
- Weekly prediction extraction
- Monthly re-evaluation for long horizons
- Score recalculation

---

15. Legal & Ethical Considerations

Mandatory Disclaimers

- Not investment advice
- Retrospective analysis only
- AI-generated evaluations

Ethical Rules

- No mocking or defamatory language
- Evidence-backed explanations only
- Transparent methodology

---

16. MVP Scope

Phase 1

- 3 creators
- 2022–2024 videos
- Manual data verification

Phase 2

- Fully automated pipeline
- Public launch
- Community feedback

---

17. Future Enhancements

- Stock-level accuracy
- Bull vs Bear bias score
- Confidence vs accuracy visualization
- Creator verification & response feature

---

18. Conclusion

This project introduces **accountability and transparency** into Indian finance content using **AI + data + objective scoring**.
It empowers users to judge credibility based on **history, not hype**.
