



FidelFolio Investments

Finance + Deep Learning



**MID
PREP**

Introduction

Valuation and investment decisions in equity markets are often influenced by fundamental indicators such as earnings, margins, growth rates, and capital structure. While these indicators are traditionally used by analysts to form qualitative judgments, the opportunity to quantify their relationship with future market performance using machine learning and deep learning opens new frontiers in investment research.

In this challenge, participants will build deep learning models to predict market capitalization growth of listed companies based on a curated dataset of historical fundamental features. The objective is to model complex, non-linear relationships between financial indicators and future market valuation movements.

Problem Statement

Participants must use the provided dataset - consisting of time-series fundamental financial indicators and market capitalization of listed Indian companies - to build a predictive model for market cap growth (targets).

Data Set Explained

Column	Description
Year	Fiscal year of the data snapshot
Company	Name of the listed Indian company
Feature1 to Feature28	A curated list of 28 fundamental financial features
Target1	1-year forward market cap growth (%)
Target2	2-year forward market cap growth (%)
Target3	3-year forward market cap growth (%)

Core Task

1. Build deep learning models to predict all three targets simultaneously or independently and analyse performance across time horizons. Train separate models for each target and compare the results across short-term (Target1) vs long-term (Target3) growth predictions. Comparing predictive performance across 1Y / 2Y / 3Y targets helps understand:
 - a. How fundamental indicators relate to near-term vs long-term value creation.
 - b. Whether different feature sets or model architectures work better for different horizons.
2. Predict the targets using only historical fundamental features up to the prediction date.
3. Build and train a deep learning model (e.g., MLP, LSTM, Transformer, etc.) using this dataset.
4. Evaluate and compare models based on predictive performance.

Challenges involved

1. Feature engineering from raw financial indicators.
2. Avoiding data leakage by ensuring models use only past data to predict future outcomes. Participants must not use any information from future periods ($t+1$, $t+2$, etc.) while building features or imputing data for predictions made at time t .
3. Designing effective model architectures suitable for tabular or time-series data

Key Features Required in the Solution

1. Clean data preprocessing pipeline (handling missing values, outliers (IQR / Winsorisation etc), normalization)
2. Deep learning architecture (can be feed-forward or recurrent or hybrid etc)
3. Clear model training and validation setup (train-test split, cross-validation, etc.)
4. Performance metrics: RMSE

Bonus Features (Optional but Encouraged)

Model Interpretability

1. Which features matter most for predicting growth?
2. Feature attribution techniques (SHAP, LIME, Attention weights, etc.)
3. Visualizations of learned feature patterns or prediction confidence

Final Deliverables

Participants must submit:

1. A GitHub repository containing:
 - a) All source code and scripts
 - b) Trained model files
 - c) ReadMe for setup and running instructions including
 - i. Data preprocessing and assumptions
 - ii. Model design and rationale
 - iii. Evaluation methodology and results
2. A presentation deck explaining:
 - a) Problem approach and model design
 - b) Performance results and interpretation



Evaluation Criteria

Submissions will be evaluated based on:

Criterion	Weightage
Predictive Performance (RMSE)	50%
Model architecture and implementation	30%
Code structure and documentation	10%
Final Presentation	10%

Successful models will be deployed internally at FidelFolio to support investment research!

Dataset

[Access the dataset spreadsheet](#)