2026 IEEE Chandigarh Subsection International Conference (CHANDICON) | 979-8-3195-4667-8/26/$31.00 ©2026 IEEE | DOI: 10.1109/CHANDICON70877.2026.11767570

When Markets Surprise: Interpretable Machine Learning for Prediction of Indian Small Cap Mutual Fund Return

Rishita Das*, Tapas Si†, *Member, IEEE,* and Rahul Sharma*
*UEMJ Decision Making & Analysis Research Group
School of Management, University of Engineering & Management Jaipur, Rajasthan – 303807, India
Email: rishitadas23@gmail.com ORCID: 0009-0008-5105-5454
rahulsr.sharma@uem.edu.in, ORCID: 0000-0003-3990-2952
†UEMJ Advanced Machine Learning Research Group
AI Innovation Lab, Dept. of Computer Science & Engineering
University of Engineering & Management Jaipur, Rajasthan – 303807, India
Email: tapassi.aiml@gmail.com ORCID: 0000-0001-8267-0304

***Abstract*—Predicting the daily returns of small cap mutual funds in emerging markets like India is difficult because market behaviour is nonlinear, highly volatile and influenced by many macroeconomic factors. This study compares four regression models which are Ordinary Least Squares (OLS), Polynomial Ridge Regression (Poly Ridge), Multi-Layer Perceptron (MLP), and Support Vector Regression (SVR) using SBI Small Cap Fund data from 2016 to 2025 with ≈2,470 observations. Model performance was checked using MAE, MAPE, RMSE and MSE with 10-fold TimeSeriesSplit cross-validation. LIME was used in four scenarios which are best prediction, worst prediction, highest actual return and the most recent observation. Best performance was given by SVR in individual error measures while TOPSIS had shown the Poly Ridge was the most balanced model overall. The results showed that all the models predicted lower values during extreme market movements. LIME had found FII Net Equity and Nifty Small Cap were the major factors influencing the predictions.**

***Index Terms*—Financial Forecasting, Mutual Fund Prediction, Machine Learning, Explainable AI, LIME, TOPSIS**

I. Introduction

The Indian mutual fund industry has grown significantly over the last decade, with Assets Under Management rising from around INR 16 trillion in 2016 to over INR 57 trillion by 2024 [1]. This growth has increased the need for better prediction tools, especially for small cap funds, because they are very volatile in nature and react more to market changes than the large cap funds. Linear regression or other traditional models are still commonly used because they are simple and easy to understand. But these models often fail when markets become unstable [2], [3] because daily financial returns usually do not follow a straight line pattern. On the other hand, machine learning methods can handle better the nonlinear and complex relationships. Recurrent, convolutional networks and other deep learning models have improved financial time-series prediction even more [4], [5]. But, these advanced models are often difficult to explain and trust in practical use.

Our earlier study [6] showed that an optimized Stacked LSTM model had performed very well on the SBI BlueChip Fund. It had achieved a MAPE of 2.80% which was almost 40 times better than the other tree-based models. It had also ranked first in TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution) multi-criteria evaluation. But, that model needed around 165 minutes for training and had limited interpretability. So, it may not be very practical for fund managers to use in everyday decisions. This creates an important question, among faster and easier to explain regression models *which* one works best for predicting Indian small cap fund returns? And also can we understand *why* each model makes the prediction it does under different market conditions? This is where XAI becomes important because it helps us understand why a model gives a particular prediction. LIME (Local Interpretable Model-Agnostic Explanations) [7] is one of the most commonly used XAI techniques. It shows how much each feature contributes to one specific prediction. Study of prediction methods and interpretability methods have improved separately but very few studies have combined both in one complete framework for Indian mutual funds. As far as we know, no previous study has compared multiple regression-based models for Indian small cap fund returns while also giving LIME-based explanations for different market situations. This study fills that gap directly.

A. Objectives

1) To develop and evaluate four regression models for predicting daily returns of the SBI Small Cap Fund.
2) To compare model performance using MAE, MAPE, RMSE, and MSE under a unified cross-validation framework.
3) To analyze model behavior under normal, extreme, and recent market conditions using scenario-specific LIME explanations.

Authorized licensed use limited to: INSTITUTE OF ENGINEERING & MANAGEMENT TRUST. Downloaded on November 18,2026 at 07:06:11 UTC from IEEE Xplore. Restrictions apply.