Personalized Mutual Fund Recommendation System

Sambhu S S, Sachin Nair, Aarush Kaswan, Souvik Ghosh, and Prof. Raghuram Bharadwaj
IIIT Bangalore
(Dated: May 15, 2024)

Abstract: Mutual funds provide a higher return on investment in spite of having high risk factors as compared to other investment options. Consequently, investors are in a dilemma while choosing the right funds for investing. As a result, an efficient recommendation system along with an explanation for the same is required to aid the investors make the correct choice. In this paper we propose a model that can give recommendations based on user's context. Also, we aim to personalize these recommendations based on the user's interaction with mutual funds.

I. INTRODUCTION

A financial entity that gathers the money of several investors and invests them in different financial securities such as bonds, stocks and short term debt is termed as 'Mutual Fund'. The prediction and estimation of mutual funds' performance is a key ingredient for investors and financial institutions today, but parallelly figuring out the best funds to invest becomes challenging. In this paper, we propose a method for predicting fund prices through a machine learning method which predicts time series data. The fund price is predicted by Prophet model, which is an algorithm for predicting time series data. The daily fund prices for funds are learned by Prophet. The future fund prices are predicted based on given previous prices. The proposed model will consider multiple factors such as expense ratio, risk-level, fund returns and current NAV (Net Asset value) that influence the general performance of mutual funds' markets.

II. PROBLEM

Common mutual fund recommendation systems base their results on market performance rather than personalizing the investment portfolio of the investors. Additionally, the recommendation systems don't have an ability to provide an explicit explanation of why a fund was recommended to the user.

III. CHALLENGES

Since recommender systems are based on high quality data, if we are not able to crunch and analyze it properly, we may not be able to make the most of the recommendation system. Public datasets are scarcely available for the use-case. We have used web scrapping method to get features of different mutual funds and APIs (Application Programming Interface) to collect every day's NAV value of each mutual fund.

IV. PROPOSED SYSTEM

A. Problem Elaboration

Currently, the existing mutual fund recommendation systems present are majorly based on the best performing mutual fund in the market. The personalization of the mutual fund along with predicting the performance of the mutual fund is not considered in depth. Not considering all the factors together reduces the accuracy and doesn't take into account the amount of risk an investor can take. There might be users who are not interested in the best performing mutual fund rather the user shows interest in some stable mutual fund which is having some decent returns. Some users wants to go with low risk level investment. These factors play a decent role in recommending a mutual fund to a user. The proposed model can be developed by combining and preprocessing several data of mutual funds and taking some user demographics followed by finding similarity between the user and mutual funds. The top similar mutual funds are given to Prophet model to predict future returns and the best among them are recommended to the user. Considering the fact that an existing user might change their taste over time, a MAB (Multi-Armed Bandit) approach is implemented to capture user preferences over time and recommending mutual funds similar to user preferences.

B. Algorithm

Mutual Funds are represented using expense ratio, risk-level and fund returns till present day. The expense ratio of a mutual fund is the percentage of the daily investment value that is deducted to cover the fund's management costs. Risk level refers to the level of risk in a mutual fund. User's input includes Investment Amount, Age, Annual Income and type of mutual funds the user prefers. The Investment amount is used to calculate expense ratio the user can give. Age and Annual income calculates user risk level. We have considered a user in younger age can take more risk then an elder user with same annual income. Now using these features, we brought mutual fund and user in the same space. As both are in same space, we have performed cosine similarity

between the user and all mutual funds.

$$similarity(A, B) = \cos(\theta) = \frac{A \cdot B}{||A||\ ||B||}$$

From the similarity matrix, we get the top similar mutual funds for the respective user. These funds are fed to prophet model to predict future returns after one year or three years or five years.

The Prophet model is applied to more accurately predict the future fund price through the relationship between the fund price and the fund's variables. Prophet model is a time series prediction library developed by Facebook and can be used in the form of a Python package. A prediction by Prophet model is made through the following equation.

$$y(t) = g(t) + s(t) + h(t) + ei$$

where y(t), g(t), s(t), and h(t), and ei are a predicted result at time t, a growth function, a seasonality function, a holidays function, and the error assumed to be normally distributed, respectively. The growth function represents a change trend, that is a continuous change of measurement target. Parameters in the function indicate whether the value is continuously increasing or decreasing. The parameters may change when checkpoint is reached because the trend of change can change at a specific point. It is probabilistically determined whether a particular point is the checkpoint. The seasonality function means periodic changes within the overall change trend. The periodic changes are represented by Fourier series function. The holidays function means the effect of an event occurring at a specific point. For example, sales of products increase significantly with each holiday. Prophet model treats the prediction as a curve fitting problem so it can avoid temporal dependence. Therefore, the prediction equation of Prophet model consists only of the sum of several linear or nonlinear functions. This characteristic of Prophet model makes the prediction learning easy and fast. The parameters of component functions in equation are determined by training through the past data. Prophet model predicts not only the future values but also the prediction value ranges. The returns after 't' years is calculated by the given formula:

$$\text{CAGR} = \left(\frac{V_{final}}{V_{begin}}\right)^{1/t} - 1$$

CAGR = Compound Annual Growth Rate
V_{begin} = Beginning Value (Starting NAV value)
V_{final} = Final Value (Predicted NAV value after 't' years)
t = Time in years

The mutual funds with the best predicted returns from the Prophet model are recommended to the user.

MAB Analysis of UCB Algorithm:

As user preferences change over time, a user might be investing in mutual funds not aligned with initial demography. In this approach, mutual funds are represented using expense ratio, risk level as these are constant features of a mutual fund. Mutual funds are clustered using K-Means algorithm. Every day, Prophet model predicts returns of each mutual funds. Returns of a cluster is calculated by taking the average returns of the mutual funds of that respective cluster.

These clusters are represented as the arms in UCB (Upper Confidence Bound) Algorithm. The algorithm is set to explore for 't' number of days. If the user does not invest in any mutual fund, the cluster with highest returns get a weight 'w' as a reward. In the other hand, if the user invest in any mutual fund, the cluster in which the invested mutual fund is in, that cluster gets a reward two times of the weight 'w'. After the exploration completes, the clusters are ranked based on the highest average rewards. The top performing mutual funds of each cluster are recommended to the user. If the user does not invest in any mutual fund during exploration then the mutual funds with high returns will be recommended to the user.

V. CONCLUSION

After reviewing several relevant papers, we have identified the gaps in existing machine learning model approaches for mutual fund recommendation systems. To provide an efficient solution, we have proposed a similarity based time-series model to provide mutual funds aligned to a user's demography. In addition to that, an UCB (Upper Confidence Bound) approach is also proposed to capture user's taste over time and to recommend funds that are similar to user's previous investments.

VI. REFERENCES

[1] Chae-eun Park, Dong-seok Lee, Sung-hyun Nam and Soon-kak Kwon."Implementation of fund recommendation system using machine learning." *Journal of Multimedia Information System*, 8(3):183–190, 2021.

[2] Aayush Shah, Aayushi Joshi, Dhanvi Sheth, Miti Shah and Prof. Pramila M Chawan, "Clustering models for mutual fund recommendation." *International Research Journal of Engineering and Technology*, 10(4):435-441, 2023.