Adapting the Elo Rating System for Financial Markets

Siddhartha Srinivas Rentala

Abstract

This document summarizes my investigation into adapting the Elo rating system for ranking securities in financial markets. I have delineated some key assumptions, provided a brief overview of the traditional Elo rating system, adapted its connection to logistic regression as outlined by some online articles, and discussed several extensions necessary to tailor the method for financial applications.

1 Assumptions

- **Pairwise Comparison:** Each security competes in a "game" defined by comparing its monthly average return with that of another security, so each company plays a game against n-1 companies every month.
- **Win Definition:** A security wins a game if its monthly average return exceeds that of its opponent.
- **Continuous vs. Binary Outcomes:** Unlike sports outcomes, financial returns are continuous. This necessitates modifications to a system originally designed for binary win/loss outcomes.
- **Market Dynamics:** The system must account for market volatility, risk, and the time-evolving nature of financial performance. Note that as we introduce new factors to the model, the elo system allows us to effectively accumulate the effect of these over time.

2 The Basic Elo Rating System

The traditional Elo rating system updates a team or player's rating based on the outcome of a game. For teams i and j, the update is given by:

$$\text{elo}_i^{\text{new}} = \text{elo}_i^{\text{old}} + k\,(t_{ij} - \Pr(i \text{ beats } j))$$

where:

- $t_{ij} = 1$ if team i wins, and 0 otherwise.
- The win probability is modeled as:

$$\Pr(i \text{ beats } j) = \frac{1}{1 + 10^{-(\text{elo}_i - \text{elo}_j)/400}}$$

- k is a constant that determines the sensitivity of the rating update.

The number 400 is a scale factor chosen in the original formulation, and the base-10 logarithm is used by historical convention. Given the probabilistic framework given above we can compute expectations for a match-up between two players in the base case of binary outcomes.

3 Connection to Logistic Regression

Steven Morse's blog post reinterprets the Elo rating update as a stochastic gradient descent (SGD) step in a logistic regression framework. Key elements include:

- **Pairwise Data Representation:** For a game between teams i and j, define the input vector:

$$\mathbf{x}_{(ij)} = [x_k] \quad \text{where} \quad x_k = \begin{cases} 1, & k = i, \\ -1, & k = j, \\ 0, & \text{otherwise.} \end{cases}$$

- **Logistic Model:** The probability of team i winning is given by the logistic function:

$$\Pr(y_{(ij)} = 1 \mid \mathbf{x}_{(ij)}; \mathbf{w}) = \sigma(w_i - w_j) = \frac{1}{1 + e^{-(w_i - w_j)}}$$

- **SGD Update:** The SGD update for logistic regression is:

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \alpha \left(\sigma(\mathbf{w}_k^T \mathbf{x}_{(ij)}) - t_{ij} \right) \mathbf{x}_{(ij)}$$

  This mirrors the Elo update, where the rating difference $w_i - w_j$ plays the role of $\text{elo}_i - \text{elo}_j$, and α corresponds to the learning rate (analogous to the k-factor).

Thus, Elo ratings can be interpreted as weights in a logistic regression model, updated through SGD to reflect pairwise results.

4 Extensions for Financial Markets

Adapting the Elo system to financial markets requires several modifications:

- **Continuous Outcomes:** Financial returns are continuous. An extension could involve incorporating the magnitude of the return difference. The traditional system doesn't take into account the magnitude of victory/defeat, by incorporating this we can adapt the model to reward/punish higher degree victories/losses, thus rewarding the players with higher rating increases.
- **Risk and Volatility Adjustment:** Securities with high returns may also exhibit high volatility. Adjustments might be necessary so that high-risk, high-return securities do not distort the rating system; this is something that I have struggled with and right now I am thinking of introducing a volatility drag of sorts.
- **Temporal Dynamics:** Financial performance changes over time. Incorporating mechanisms such as decay of historical data or momentum (autoregressive components) can help capture time-evolving dynamics. This would also help us reduce the bias that may be present in having long-time players playing against new players. A key component of the traditional system in a game like chess is that it is used for 'skill-based matchmaking', meaning that players of around the same rating play each other. In the markets I don't believe we should bias the ratings like that, so it would be more prudent to include an AR component which can help capture time-dependent dynamics.
- **Batching Strategies:** Instead of updating ratings for every individual pairwise comparison, using mini-batches (e.g., grouping all comparisons within a month) may yield more robust updates.
- **Stochastic Components:** To incorporate risk and volatility adjustments, I have considered multi-factor extensions to the model by adding a vector of stochastic factors for each security but how these terms would interact is something that can be discerned after more empirical study.

5 Conclusion

A foundation has been established to rank securities based on monthly returns by reinterpreting the Elo rating system within a logistic regression framework. While the basic Elo model provides a starting point, necessary extensions include handling continuous outcomes, risk adjustments, temporal dynamics, and batching. My future work will focus on empirically validating these modifications and fine-tuning the model for dynamic market environments. Please note that this is a personal project, and I am open and, indeed, looking for advice on how I can proceed.