Adapting Skill Ratings to Luck-Based Hidden-Information Games

Avirup Chakraborty

Indian Statistical Institute

Kolkata, India

Shirsa Maitra

Computer Science Engineering

Heritage Institute of Technology, India

Tathagata Banerjee

Department of Statistics & Data Science

National University of Singapore

Diganta Mukherjee

Sampling and Official Statistics Unit

Indian Statistical Institute, Kolkata

Tridib Mukherjee

Chief Data Scientist and AI Officer

IDfy, India

Abstract

Rating systems play a crucial role in evaluating player skill across competitive environments. The Elo rating system, originally designed for deterministic and information-complete games such as chess, has been widely adopted and modified in various domains. However, the traditional Elo rating system only considers game outcomes for rating calculation and assumes uniform initial states across players. This raises important methodological challenges in skill modelling for popular partially randomized incomplete-information games such as Rummy. In this paper, we examine the limitations of conventional Elo ratings when applied to luck-driven environments and propose a modified Elo framework specifically tailored for Rummy. Our approach incorporates score-based performance metrics and explicitly models the influence of initial hand quality to disentangle skill from luck. Through extensive simulations involving 270,000 games across six strategies of varying sophistication, we demonstrate that our proposed system achieves stable convergence, superior discriminative power, and enhanced predictive accuracy compared to traditional Elo formulations. The framework maintains computational simplicity while effectively capturing the interplay of skill, strategy, and randomness, with broad applicability to other stochastic competitive environments.

Keywords: Elo rating system, skill evaluation, Rummy, stochastic environments, score-based rating

1   Introduction

Player skill evaluation is a fundamental component of competitive gaming ecosystems, guiding matchmaking, player progression, and strategic analysis. Traditional rating systems, such as the Elo [1], Glicko [2], and TrueSkill [3] frameworks, have proven effective in deterministic or near-deterministic games like chess and Go, where outcomes primarily reflect player skill. However, their direct application to luck-influenced, incomplete-information games such as Rummy, Poker, or Bridge introduces notable limitations. In such games, outcomes are not solely determined by player decisions but also by stochastic elements like initial card distributions and variable round dynamics. Consequently, standard Elo-type formulations tend to overestimate short-term fluctuations and underestimate persistent skill differences.

The Rummy environment presents a unique case within this class. It combines probabilistic uncertainty (due to card draws and opponent hands) with strategic depth (melding, discarding, bluffing), leading to outcomes that depend on both tactical skill and random events. This dual dependence makes player-skill inference inherently

noisy, requiring rating mechanisms that can disentangle luck from skill and stabilize ratings over repeated play. Existing systems often fail to achieve this balance. Rapid volatility in ratings discourages consistent player engagement, while slow convergence obscures meaningful skill differentiation. Our study will be focused solely on Rummy.

Recent advances in rating systems, such as TrueSkill Through Time [4] and TrueSkill 2 [5], have focused on improving dynamic updating and parameter learning; however, these models remain primarily suited to deterministic or team-based settings. In contrast, Edelkamp et al (2021) [6] proposed an Elo-based framework for Skat, a game characterized by significant random variation, demonstrating the necessity of correcting standard Elo formulations to account for chance effects. Building on the broader discussion of skill versus luck, Banerjee et al. (2021) [7] introduced a data-driven statistical framework to quantify the relative influence of skill and chance across popular games, revealing systematic differences among Chess, Rummy, Ludo, and Teen Patti. Together, these studies underscore that rating systems for games involving both randomness and hidden information, such as Rummy, must explicitly incorporate chance-adjustment mechanisms to produce fair and stable estimates of player skill.

This paper introduces a modified Elo formulation tailored for Rummy, which can be suitably adjusted to other games played in randomized environments. The method is designed to incorporate luck-adjusted weighting and match-context normalization while maintaining computational simplicity. Further, our method is also adaptable to deterministic games as special cases of parameter choices. Our approach aims to (i) model the probabilistic uncertainty intrinsic to Rummy's deal structure, (ii) reduce the inflation and deflation effects that arise from streaks of fortunate or unfortunate draws, and (iii) preserve interpretability comparable to the classical Elo scale. The proposed method is evaluated through extensive simulations across varying skill distributions, comparing its stability, convergence rate, and predictive reliability against conventional Elo [1], Glicko [2], and TrueSkill [3] frameworks.

Section 2 outlines the methodological framework, reviewing standard rating systems and highlighting their limitations in games influenced by chance before introducing our proposed formulation. Section 3 details the proposed methodology and the corresponding parameter estimation procedures. Section 4 describes the simulation design and presents the empirical findings. Section 5 interprets the results, emphasizing how the proposed rating system more effectively captures underlying skill in stochastic environments compared to existing approaches. Finally, Section 6 concludes with potential extensions and practical applications, particularly for online Rummy platforms.

2 Different Existing Rating Systems

This section provides an overview of several classical rating systems that have been widely used to evaluate player skill.

- **Elo-Based Rating Systems:** Elo-based systems model player skill as a latent variable that is iteratively updated based on game outcomes. In the *classical Elo* formulation for a two-player match between players i and j, the expected outcome for player i is modeled using a logistic link:

$$\alpha_i = \frac{1}{1 + 10^{(R_j - R_i)/400}},$$

and the rating update rule is

$$R_i' = R_i + K(S_i - \alpha_i),$$

where $S_i \in \{0, 0.5, 1\}$ denotes the observed outcome (loss, draw, or win), and K is a fixed sensitivity parameter. Extensions of the classical model address uncertainty and adaptive learning rates. In *Bayesian Elo*, each player's rating is treated as a Gaussian random variable, $R_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$, allowing probabilistic updates. *Elo-MMR* further introduces a player-specific sensitivity factor $K_i = f(\sigma_i, \sigma_j, t)$ that adapts to rating deviation and recency, enabling faster updates for uncertain players and greater stability for consistent ones.

- **Glicko Systems:** While the Elo system assumes a fixed confidence in player skill, it fails to account for uncertainty arising from limited or inconsistent play. The *Glicko* system addresses this by introducing a

rating deviation (RD) that quantifies uncertainty in a player's true strength. A higher RD indicates that a player's current rating is less reliable, typical for newcomers or inactive players, while frequent and consistent play reduces RD, reflecting growing confidence in the estimated skill. From the Elo formula, K is scaled by the rating deviation (larger for uncertain players). *Glicko-2* extends this framework by adding a *volatility parameter* to capture fluctuations in player performance over time, enabling faster correction for unstable or streak-prone players.

- **TrueSkill and TrueSkill-2: TrueSkill and TrueSkill-2:** TrueSkill, developed by Microsoft, extends beyond Elo and Glicko in a Bayesian approach, treating each player's skill as a normal distribution $p_i \sim N(\mu_i, \sigma_i^2)$ rather than a fixed point estimate. In a two-player match:

$$P(i \text{ wins}) = \Phi\left(\frac{\mu_i - \mu_j}{\sqrt{2(\sigma_i^2 + \beta^2)}}\right),$$

where β represents random performance variation within games. After each match, both μ_i and σ_i are updated through approximate Bayesian inference, allowing the model to capture learning and uncertainty dynamically. Unlike Elo or Glicko, which assume scalar and independent ratings, TrueSkill supports multi-player and team-based games by jointly inferring skills from all outcomes. TrueSkill-2 further refines this by enabling multidimensional skill vectors and real-time tracking, though at the cost of higher computational complexity and less transparency compared to classical rating systems.

Despite their success in modeling competitive skill, these systems exhibit limited inferential accuracy in games where outcomes are substantially affected by chance. To overcome this limitation, the next section details our proposed model, designed to incorporate the role of luck in performance evaluation.

3 Methodology

This section outlines the methodological framework underlying our proposed model. Our objective is to develop a rating mechanism that not only captures player skill but also adjusts for randomization inherent in the game of Rummy. For completeness, a detailed description of the two-player Rummy setup is provided in Appendix 7.1.

Challenges

Before presenting the model, we highlight two key challenges that motivate our approach:

- **Performance Differentiation:** Rummy is an incomplete information game, where players independently form melds that may not be directly influenced by their opponents' actions. The final scores of both players provide richer information than just the game outcome: a small score difference suggests comparable play, while a large difference indicates a dominant performance.

- **Influence of Initial Hands:** Rummy involves substantial randomness, and outcomes can be heavily affected by the initial hand. A poor starting hand significantly reduces a player's chance of winning, and ignoring this factor introduces bias into skill estimates.

Formulation

We formulate our model in a stepwise manner as follows:

- We begin with the classical Elo update rule:

$$R_i' \leftarrow R_i + K(S_i - E_i),$$

where $S_i \in \{0, 0.5, 1\}$ denotes the observed match outcome (loss, draw, or win), K is the sensitivity parameter, R_i and R_i' are the pre- and post-match ratings, and E_i is the expected outcome for player i.

In games with high randomness or incomplete information, such as Rummy, the expected outcome α_i (that is, the probability of winning) can vary substantially due to factors such as the initial hand and hidden opponent information.

- To address this limitation, we adopt a **score-based formulation** instead of an outcome-based one. In Rummy, we have already introduced the metrics **MinScore** and **MinDist** (Appendix 7.1), which quantify the *declarability* of a given hand. Modeling these scores enables us to capture the underlying gameplay dynamics of a player more effectively. Accordingly, we reformulate the rating update as

$$R'_i \leftarrow R_i + K(A_i - B_i),$$

where A_i denotes the observed (final) score of player i, and B_i represents a benchmark score modeled for that player. Since these scoring system is inherently bounded, the difference $|A_i - B_i|$ remains within a finite range, allowing K to be scaled appropriately to maintain numerical stability.

It is important to note that in Rummy, a lower **MinScore** or **MinDist** value indicates a better hand, with a score of zero signifying a successful declaration. Therefore, the direction of the rating adjustment must be inverted, effectively rendering the sign of K negative.

- To maintain a *zero-sum* structure analogous to the Elo system, we impose the condition

$$(A_1 - B_1) + (A_2 - B_2) = 0 \quad \implies \quad B_1 + B_2 = A_1 + A_2 = A \text{ (say)}.$$

This establishes a linear dependency between the two benchmark scores that must be jointly determined.

We use the term *benchmark* instead of *expected outcome*, since the benchmark scores B_i depend on the realized total $A_1 + A_2$ observed after the match, unlike the probabilistic expectation in the Elo model, which is computed beforehand.

- To estimate the values of B_i, we consider two key factors. The rating difference between the players, $D_R = R_1 - R_2$, represents the pre-game skill bias, while the difference in their initial hand scores, $D_H = H_1 - H_2$, captures the randomness arising from card distribution. Incorporating both allows us to reduce two types of bias in benchmark estimation.

Using a logistic-type scaling function, we define

$$B_1 = \frac{A}{1 + 10^{-(\alpha \times D_R + \beta \times D_H)}}, \quad B_2 = A - B_1 = \frac{A \cdot 10^{-(\alpha \times D_R + \beta \times D_H)}}{1 + 10^{-(\alpha \times D_R + \beta \times D_H)}},$$

where α and β are model parameters to be estimated empirically. This formulation ensures that $B_i \in [0, A]$ for both players.

- As mentioned earlier, the parameter K changes sign due to the inverse nature of the scoring metrics. Therefore, $|K|$ represents the sensitivity parameter governing the magnitude of rating adjustments. Now, Players who have played fewer games should have ratings that respond more strongly to new outcomes, reflecting greater uncertainty about their true skill level. Therefore, each player i may be assigned an individual sensitivity factor K_i based on their number of games played n_i. A simple formulation is given by:

$$K_i = \begin{cases} k_1, & \text{if } n_i \le p, \\ k_2, & \text{if } p < n_i \le q, \\ k_3, & \text{if } n_i > q, \end{cases} \quad i = 1, 2.$$

Here, $|k_1| > |k_2| > |k_3|$, ensuring that as a player gains more experience (larger n_i), their rating becomes progressively less sensitive to individual match outcomes.

- Finally, we consider the case where player 1 wins (similarly for player 2), implying $A_1 < A_2$. If $A_1 - B_1 > 0$ happens, i.e. the player won but did not play as good as the benchmark, to avoid penalizing the winner we modify the rating updates as follows:

$$R'_1 \leftarrow R_1 + K_1(A_1 - B_1)\mathbf{1}_{\{(A_1 - B_1)(W_1 - 0.5) \le 0\}},$$

$$R'_2 \leftarrow R_2 + K_2(A_2 - B_2)\mathbf{1}_{\{(A_2 - B_2)(W_2 - 0.5) \le 0\}},$$

where W_i is the game outcome indicator (1 for a win and 0 for a loss), and $\mathbf{1}_{\{.\}}$ denotes the indicator function. This ensures that rating adjustments are applied only when the score-based deviation aligns with the actual game result.

Thus, we complete the formulation of our proposed model. We initialize each player with a rating of 1000 and subsequently update these ratings according to the specified rules. Next, we outline the procedure for estimating the model parameters.

Parameter Estimation

We estimate the model parameters—player abilities α and β, and concentration parameters K_i—using the **MinScore** metric. This choice is motivated by MinScore's broader dynamic range (0 to 80, per Appendix 7.1) compared to the more limited **MinDist**. The increased variation in MinScore values provides greater discrimination between player performances, yielding more sensitive parameter estimates.

Choice of K:

In our simulation study (Section 4), we assume that all players participate in approximately the same large number of games. Consequently, the individual K_i values have negligible impact on the long-run results, as they converge after a sufficient number of matches. Hence, we set $K_1 = K_2 = K$.

At the end of a game, one of the players typically achieves a declarable hand, implying $A_i = 0$ for that player, while the maximum possible score is $A_i = 80$. From the model specification, the maximum possible change in rating per game is therefore $80|K|$. To restrict the maximum rating variation to 50 points (i.e., 5% of the initial rating of 1000), we have

$$80|K| = 50 \implies |K| = 0.625.$$

Since $K < 0$ by construction, we set $K = -0.625$. Naturally, choosing a different threshold for the maximum allowable rating change would yield a different value of K, thereby altering the resulting rating distribution.

Choice of Parameter α:

Before estimating the constants α and β, it is essential to understand their interpretation within the model. The core relationship is given by:

$$\frac{B_1}{B_2} = 10^{\alpha D_R + \beta D_H},$$

where a lower benchmark score B_i indicates better performance.

To isolate the effect of the rating difference D_R, we consider the scenario with a neutral hand advantage ($\beta D_H = 0$). The model then simplifies to:

$$\frac{B_1}{B_2} = 10^{\alpha D_R}.$$

If Player 1 has a higher initial rating ($D_R > 0$), we expect better performance, meaning $B_1 < B_2$ and thus $B_1/B_2 < 1$. For the right-hand side, $10^{\alpha D_R}$, to be less than 1, the exponent αD_R must be negative. Since $D_R > 0$, this necessitates that α must be a *negative* constant.

The magnitude of α controls the sensitivity of the model. A smaller absolute value $|\alpha|$ means that a larger rating difference D_R is required to significantly affect the benchmark ratio B_1/B_2. This leads to a system where larger rating differences are needed to distinguish between players, effectively increasing the spread of ratings across the population. However, if $|\alpha|$ is too small, the ratings may diverge excessively, making the system unstable. Therefore, α must be fixed to a carefully chosen value to ensure stable and meaningful ratings.

We calibrate α by considering an extreme, yet plausible, in-game scenario. Let us assume:

- $D_R = 500$, indicating Player 1 is substantially stronger than Player 2.

- $D_H = 0$, indicating both players have hands of identical quality.

- The resulting scores are $B_1 = 2$ (near-optimal) and $B_2 = 80$ (the maximum non-zero score in 13-card Rummy).

Substituting these values into the model yields:

$$\frac{2}{80} = 10^{500\alpha} \quad \implies \quad \alpha = \frac{-\log_{10}(40)}{500} \approx -0.0032.$$

We use this estimate of α in the subsequent simulation study.

Choice of Parameter β:

For parameter β, consider the scenario where $D_R = 0$, indicating players of equal skill levels:

$$\frac{B_1}{B_2} = 10^{\beta D_H}$$

when $D_H > 0$, Player 1 faces a hand disadvantage, and we expect Player 2 to perform better, i.e., $B_1 > B_2$ and consequently $B_1/B_2 > 1$. For the right-hand side $10^{\beta D_H}$ to exceed 1, the exponent βD_H must be positive. Since $D_H > 0$, this requires β to be a *positive* constant.

We estimate β through a systematic tuning procedure with the following steps:

- Fix $\alpha = -0.0032$ as determined from our previous calibration
- Evaluate a range of β values by computing dynamic ratings on simulation data
- Partition the dataset into training and testing subsets
- Train a logistic regression model using only initial rating difference to predict match winners
- Assess model performance on the test set using F1 score as the primary metric
- Select the β value that maximizes predictive performance

To determine the optimal model configuration, we performed a hyperparameter search over a range of values for β. The resulting F1 scores from this tuning process are plotted in Figure 1. The analysis reveals a clear optimum at $\beta = -0.012690$, which yields a maximum F1 score of 0.7927. This value of β optimally calibrates the model's update function, maximizing the harmonic mean between precision and recall.
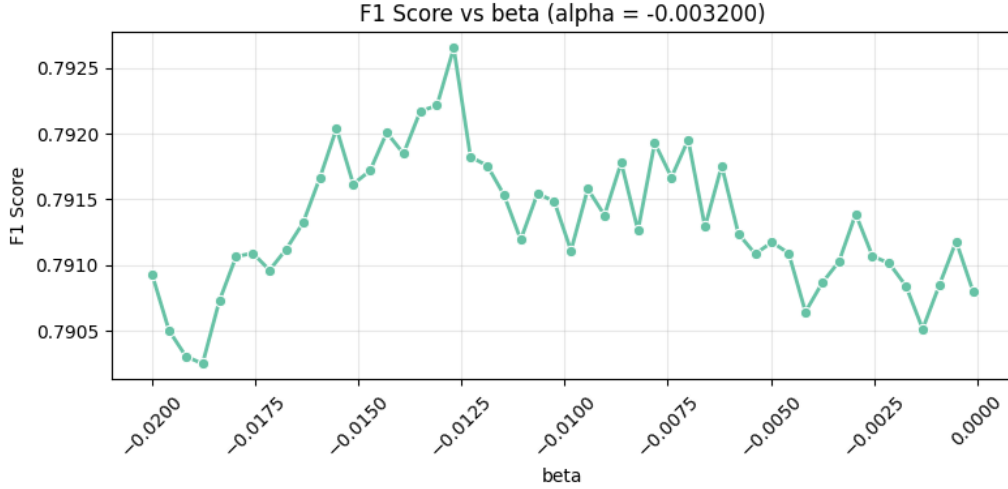

Figure 1

The estimated values for K, α, and β will be used to parameterize the model in the subsequent simulation study.

4   Simulation Study & Results

In this section, we describe the simulation framework employed for the empirical evaluation of the competing strategies. Detailed descriptions of each strategy are provided in Appendix 7.2. These strategies are taken from [9]. Among the considered strategies, the Random and Defeat-Seeking agents serve as naive baselines,

expected to perform poorly due to their lack of strategic sophistication. The MinDist and MinScore agents are based on the core distance and marginal-score metrics and therefore represent the fundamental benchmark strategies. The MinDist+MinScore Hybrid (Mindistscore) and Opponent-Aware MinDist (MindistOpp) extend the MinDist formulation by incorporating hybrid decision rules and opponent modelling, respectively. These enhanced variants are designed to exhibit stronger competitive performance.

The simulation study is structured as follows:

- We simulate 4,500 games for each ordered pair of strategies, resulting in $6 \times 5 = 30$ directed matchups. Consequently, each strategy participates in $4500 \times 5 \times 2 = 45,000$ games in total.

- Each strategy plays an equal number of games as the first and second player to eliminate positional bias and ensure symmetry in competitive conditions.

- The global sequence of games is randomised to avoid ordering effects and to ensure that rating updates evolve in a stochastic and unbiased manner.

- For every match, we record the dynamic rating trajectories and rating increments, which will enable us to analyse rating stability and convergence over time.
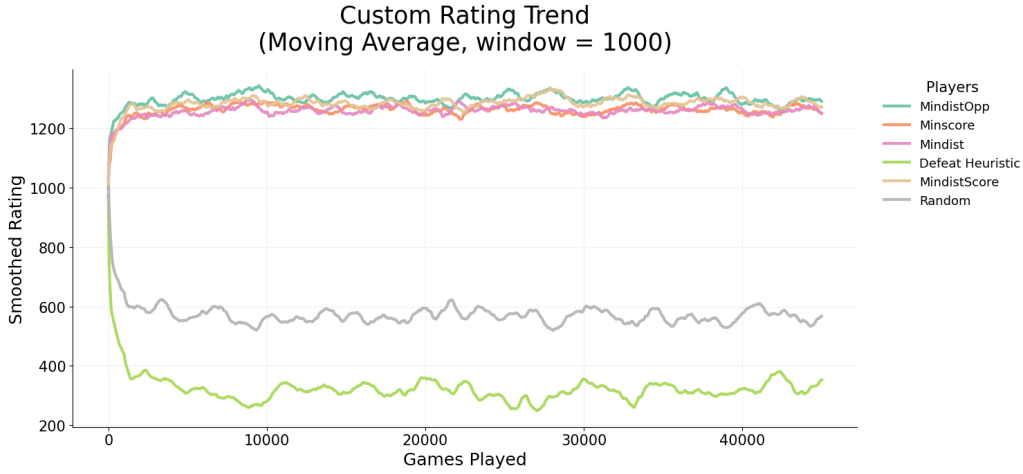
Here are the results of the simulation,

Figure 2: **Rating evolution over simulated games (Custom Elo)**

Figure 2 illustrates how Elo ratings evolve throughout the simulation.

The following observations emerge:

- Ratings begin to stabilise around 5,000 games, indicating convergence toward a consistent skill hierarchy and clear separation among strategies.

- As anticipated, the naive strategies *Defeat Heuristic* and *Random* perform the worst. Their final ratings fall below the neutral Elo baseline of 1000, which confirms that they lose consistently across repeated play.

- Although a moving-average smoother is applied, small fluctuations remain. These can be attributed to heterogeneous matchups, where temporary pairings with stronger or weaker opponents cause short-term rating variation.

- Figure 3 highlights the performance among the stronger strategies. The hybrid strategies *MindistOpp* and *MindistScore* consistently outperform the pure heuristics *Mindist* and *Minscore*, supporting the conclusion that combining distance and score information yields superior decision making capability.

Table 1 reports summary statistics (mean, standard deviation, and coefficient of variation) for the rating paths. The results clearly demonstrate meaningful performance separation, with hybrid strategies achieving both higher and more stable ratings over time.
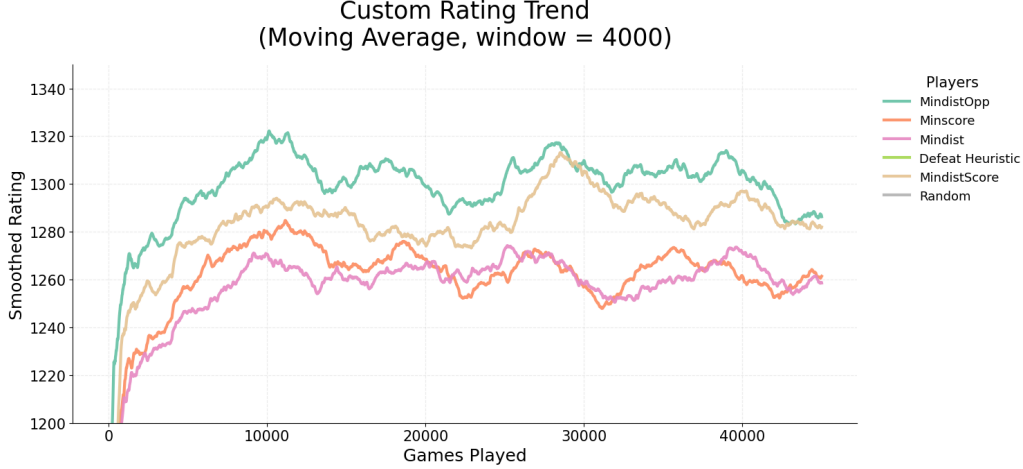
Figure 3: **Zoomed view of rating evolution for top strategies (Custom Elo)**

Strategy	Mean	SD	CV (%)
MindistOpp	1300.788	43.986	3.381
MindistScore	1285.526	43.388	3.375
Minscore	1263.769	39.813	3.150
Mindist	1259.681	41.447	3.290
Random	569.828	42.807	7.512
Defeat Heuristic	320.049	46.737	14.603

Table 1: **Rating Summary of Strategies (Custom Elo)**

The naive strategies show high coefficient of variation which further indicates inconsistent and highly volatile behavior, reinforcing their inferiority as gameplay approaches.

5 Discussion

The results presented in Section 4 demonstrate that our proposed modified Elo rating system effectively addresses the fundamental challenges of skill evaluation in luck-influenced, incomplete-information games such as Rummy. This section interprets these findings, discusses their implications, and contextualizes them within the broader landscape of rating systems research.

Interpretation of Results

The simulation study reveals several key insights into the behavior of our proposed rating mechanism. First, the convergence of ratings around 5,000 games indicates that the system successfully stabilizes despite the inherent stochasticity of Rummy. This convergence behavior is critical for practical deployment, as it ensures that player ratings become increasingly reliable indicators of true skill over time, rather than remaining volatile artifacts of recent luck.

The clear hierarchical separation among strategies, particularly the superior performance of hybrid approaches (MindistOpp and MindistScore) over base cases (Mindist and Minscore), validates a core premise of our model: that a rating system incorporating both score-based performance and initial hand quality can effectively discriminate between different levels of strategic sophistication. The fact that these hybrid and basic strategies consistently achieve ratings above 1350, while naive strategies fall well below 1000, demonstrates that our system rewards genuine strategic depth rather than merely reflecting random variation in outcomes.

Moreover, the coefficient of variation (CV) analysis provides important evidence of rating stability. The top-performing strategies exhibit CVs below 4%, indicating that their ratings, once converged, remain relatively stable despite ongoing matches. In contrast, the Defeat Heuristic strategy shows a CV exceeding 180%, reflecting its fundamentally inconsistent and self-defeating behavior. This contrast underscores the system's ability to distinguish between stable, skill-driven performance and erratic, luck-dependent outcomes.

Advantages Over Traditional Rating Systems Our proposed framework offers several methodological advantages over classical rating systems when applied to stochastic environments. To validate this claim, we implemented the traditional Elo system on the same simulation data, with results presented in Appendix 7.3. The comparison reveals that while traditional Elo successfully establishes a similar hierarchical ordering among strategies, our score-based formulation provides additional benefits. Unlike traditional Elo, which relies solely on binary outcomes and thus treats a narrow victory identically to a dominant one, our approach leverages the richer information contained in final scores, rewarding convincing victories while penalizing losses less severely when a player performs respectably despite an unfavorable initial hand.

The incorporation of initial hand quality through the parameter β addresses a critical limitation in applying standard rating systems to Rummy. As demonstrated by the parameter tuning procedure (Figure 1), optimal calibration of β significantly enhances predictive accuracy, with the F1 score reaching 0.7927. This result indicates that explicitly modeling the influence of randomness, rather than treating it as noise, substantially improves the system's ability to infer true skill from observed outcomes.

Compared to more complex systems such as TrueSkill and Glicko-2, our approach maintains computational simplicity while achieving comparable or superior performance in stochastic settings. The deterministic update rule, parameterized by only three constants (K, α, β), ensures transparency and interpretability, which are essential for player trust and system adoption in competitive gaming platforms. Furthermore, the zero-sum property inherited from classical Elo ensures that the average rating remains constant, preventing rating inflation or deflation over time.

Practical Implications

While our study focuses specifically on Rummy, the proposed framework is readily adaptable to other games characterized by incomplete information and randomness. The key requirement is the availability of a quantifiable performance metric beyond binary outcomes. For instance, in Poker, chip counts or hand strength distributions could serve analogous roles to MinScore in Rummy. In Bridge, trick counts and bidding accuracy might provide similar informational richness.

For online platforms, our proposed rating system offers several practical benefits. By providing more stable and accurate skill estimates, the system can improve matchmaking quality, ensuring that players are paired with opponents of comparable ability. The system's computational efficiency is also advantageous for large-scale deployment. Unlike TrueSkill or Glicko-2, our model computes rating changes through straightforward arithmetic operations, ensuring low latency and scalability for platforms hosting millions of concurrent matches.

Limitations and Considerations

Despite its strengths, our proposed system has several limitations that warrant acknowledgment.

First, the simulation study employs deterministic strategies with fixed behaviors, whereas real human players exhibit learning, adaptation, and psychological variability. While our framework is designed to accommodate such dynamics through the sensitivity parameter K_i, empirical validation with real player data remains essential to assess its robustness in practice.

Second, the choice of K, α, and β involves subjective calibration decisions. Although we have provided principled justifications for our parameter selections, alternative choices could yield different rating distributions. In particular, the threshold for maximum rating change (set at 50 points in our study) directly determines K and consequently affects convergence speed and long-term stability. Practitioners implementing this system must carefully consider these trade-offs in light of their specific application context.

Third, our current formulation assumes two-player matches. Extending the model to multi-player settings introduces additional complexity, as the zero-sum property must be generalized and benchmark scores must be allocated among more than two players. While conceptually feasible, possibly through pairwise comparisons or rank-based allocations, such extensions require further theoretical development and empirical testing.

Finally, the assumption of a static skill distribution throughout the simulation does not reflect the reality of player improvement over time. Incorporating time-dependent skill evolution, perhaps through decay functions or periodic rating adjustments, could enhance the system's long-term accuracy. The Glicko-2 volatility parameter offers one potential approach to modeling such dynamics, though at the cost of increased complexity.

6 Conclusion

This paper has addressed a fundamental challenge in competitive gaming: the accurate evaluation of player skill in environments where outcomes are influenced by both strategic decision-making and random chance. Traditional rating systems, such as Elo, Glicko, and TrueSkill, exhibit significant limitations when applied to incomplete-information, luck-driven games such as Rummy. Our work has developed and empirically validated a modified Elo framework specifically tailored to address these limitations.

The proposed rating system introduces three key innovations: a score-based formulation leveraging richer information from final scores; explicit incorporation of initial hand quality through the parameter β to disentangle skill from luck; and a benchmark-based update rule that maintains the zero-sum property while adapting to stochastic gameplay.

Through an extensive simulation study involving 270,000 total games across six strategies, we demonstrated that the system successfully achieves stable convergence after approximately 5,000 games per strategy, effectively discriminates among strategies (with hybrid approaches achieving ratings above 1350 and naive strategies below 500), and exhibits low volatility once converged (CV below 4% for top strategies). Parameter estimation yielded an F1 score of 0.7927, validating the framework's ability to capture genuine skill differences.

The proposed system offers practical advantages for deployment on online gaming platforms: computational simplicity ensures scalability; interpretability facilitates transparent communication; and stability prevents rating inflation. Beyond Rummy, the framework is readily adaptable to other incomplete-information, stochastic games where quantifiable performance measures exist.

Important limitations include the need for empirical validation with real player data, extension to multi-player settings, and incorporation of time-dependent skill evolution. Future research should address these challenges while exploring adaptive parameters and conducting comparative studies across diverse games.

In conclusion, this work contributes to the growing literature on rating systems for stochastic competitive environments by demonstrating that explicit modeling of randomness and performance quality significantly improves skill inference. The proposed framework offers a principled, computationally efficient, and empirically validated solution with broad potential for application across luck-influenced games. As online gaming continues to expand, rating systems that effectively disentangle skill from chance will become increasingly essential for maintaining competitive integrity, player satisfaction, and engagement.

References

[1] Elo, A. E. (1978). *The Rating of Chessplayers, Past and Present.* Arco Publishing.

[2] Glickman, M. E. (1999). *Parameter estimation in large dynamic paired comparison experiments.* Applied Statistics, 48(3), 377–394.

[3] Herbrich, R., Minka, T., & Graepel, T. (2006). *TrueSkill™: A Bayesian skill rating system.* In *Advances in Neural Information Processing Systems* (pp. 569–576).

[4] Dangauthier, P., Herbrich, R., Minka, T., & Graepel, T. (2007). *TrueSkill Through Time: Revisiting the history of chess.* In *Advances in Neural Information Processing Systems* (Vol. 20).

[5] Minka, T., Cleven, R., & Zaykov, Y. (2018). *TrueSkill 2: An improved Bayesian skill rating system.* Microsoft Research Technical Report.

[6] Edelkamp, S. (2021). *An ELO system for Skat and other games of chance.* arXiv preprint arXiv:2104.05422. https://arxiv.org/abs/2104.05422

[7] Banerjee T., De A., Maitra S., & Mukherjee D. (2021). *Skill vs. Chance Quantification for Popular Card & Board Games.* arXiv preprint arXiv:2410.14363 https://doi.org/10.48550/arXiv.2410.14363

[8] Indian Rummy. Wikipedia https://en.wikipedia.org/wiki/Indian_rummy

[9] Saha P., Chakraborty A., Maitra S., & Mukherjee T. (2025). *Quantitative Rule-Based Strategy Modelling in Classic Indian Rummy: A Metric Optimization Approach.* Manuscript submitted for publication.

7 Appendix

7.1 Indian Rummy

Indian Rummy is a draw-and-discard melding game in which each player is dealt 13 cards and attempts to arrange all cards into valid *sequences* and *sets*; one *pure sequence* (no jokers) is mandatory for a valid declaration. Typical play 2 player version uses one standard 52-card decks together with printed jokers. Below is a brief overview of the rules and scoring system of the rummy game used in our analysis. Details can be found in [8].

Core Rules:

1. *Deal:* each player receives 13 cards; the remaining stock (closed pile) and one face-up discard form the discard mechanism.

2. *Turn play:* on a player's turn they must draw one card (from the closed stock or the open discard) and then discard one card to the open pile.

3. *Valid melds:*

   – *Sequence*: three or more consecutive cards of the same suit (e.g., 7♥, 8♥, 9♥).

   – *Set*: three or four cards of the same rank in different suits (e.g., 8♣, 8♠, 8♦).

4. *Declaration:* to declare (win) a player must show at least two sequences, one of which must be a pure sequence (no jokers). Remaining cards must be arranged into valid sets or sequences.

5. *Jokers & wildcards*: two printed jokers and one hand-determined wildcards (selected from the exposed card under the stock in many Indian variants) may substitute for missing cards to form impure sequences or sets; a pure sequence cannot contain a joker.

Scoring:

– *Face cards:* J, Q, K, A = 10 points each.

– *Numbered cards:* 2-10 = face value in points (e.g., 5 = 5 points).

- *Jokers:* printed jokers and wildcards count as 0 points when used; any un-melded jokers in hand score 0.

- *Round scoring:* after a player declares validly, each opponent totals points from unarranged cards (deadwood). Typical point transfers or losses equal the deadwood counts.

For analysis of the game, we next define a few metrics of a given hand, which will help the simulation strategies to take decisions throughout the game. We mainly consider the $MinScore$ and the $MinDist$ metrics of a hand which will in turn, guide play.

- **MinScore:** The $MinScore(h, wcj)$ represents the minimum achievable score of a hand h when all cards are optimally arranged into valid melds considering the wild-card joker wcj. It quantifies the least possible sum of unmatched card values that can remain after the best possible grouping. A lower $MinScore$ indicates a hand that is closer to a valid declaration, while a higher value reflects a larger number of ungrouped or high-value cards. This metric provides a direct measure of the optimal scoring efficiency of a given hand.

- **MinDist:** The $MinDist(h, wcj)$ metric quantifies the combinatorial distance of a given hand h from the nearest valid configuration considering the wild-card joker wcj. It measures how many single-card substitutions are needed to convert the current hand into a valid declaration, thereby capturing the hand's structural proximity to validity. A smaller $MinDist$ indicates a hand that is nearly complete and strategically strong, while a larger value implies greater deviation from a playable configuration. This measure provides an interpretable proxy for the expected number of turns required to achieve a valid declaration.range

Although both metrics measure proximity to a valid declarable hand in distinct ways, each possesses its own advantages and limitations. The $MinDist$ value indicates the number of cards required to complete a valid meld; however, if the needed cards are already in the discard pile, its practical relevance diminishes. On the other hand, the $MinScore$ metric provides a quantitative measure of card arrangement efficiency but lacks a straightforward strategic interpretation. In the following section, we describe the simulation strategies employed in this study, which are formulated based on these two measures.

7.2   Simulation Strategies

1. **Random**: This agent makes completely random decisions at each step, choosing either the deck or pile card, and discarding a random card from its hand.

2. **Defeat-Seeking**: This agent identifies existing melds in its hand and intentionally breaks them by discarding the lowest-value card. It picks from the pile or deck based on whether the pile card forms a meld; otherwise, it defaults to the deck. The goal is to maximize the $MinScore$ while disrupting potential meld structures.

3. **MinScore-Based**: This agent picks the pile card if doing so improves its $MinScore$ for the best 13 cards by a threshold amount (e.g., 3 points) or completes a valid hand; otherwise, it draws from the deck. It discards the card which reduces $MinScore$ optimally.

4. **MinDist-Based**: This agent selects the pile card if it improves its $MinDist$ for the best 13 cards; otherwise, it draws from the deck. It discards the card which reduces $MinDist$ optimally.

5. **MinDist + MinScore Hybrid**: This agent prioritizes improving $MinDist$, but when both options yield the same $MinDist$, it chooses the action that results in a better $MinScore$. It follows the same discard criteria as the $MinDist$-based agent.

6. **Opponent-Aware MinDist**: This agent extends the $MinDist$-based approach with simple opponent hand modelling. When no improvement in $MinDist$ is possible, it discards cards similar to those recently discarded by the opponent or those likely to form melds with the opponent's dropped cards (e.g., discarding 4 ♣, 6 ♣, or 5 of another suit if the opponent discards 5 ♣). The agent employs the same drop adherence rule as the $MinDist$-based variant.
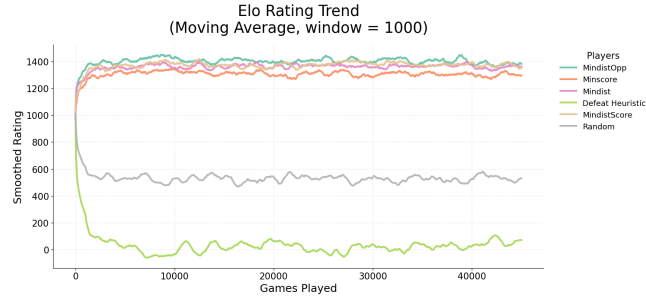
7.3 Simulation Results:

Figure 4: Rating evolution over simulated games (Traditional Elo)

Strategy	Mean	SD	CV (%)
MindistOpp	1402.204	65.425	4.666
MindistScore	1373.527	64.399	4.689
Mindist	1361.234	62.870	4.619
Minscore	1307.487	63.011	4.819
Random	529.055	49.732	9.400
Defeat Heuristic	26.101	67.063	256.939

Table 2: **Rating Summary of Strategies (Traditional Elo)**