A Data-Driven Optimization of Voting Mechanisms in Dancing with the Stars

Jiaming Wei[1,*], Yue Chen[2]

[1] School of Resources and Earth Sciences, China University of Mining and Technology, Xuzhou, Jiangsu, 221116, China

[2] School of Optoelectronic Engineering, Xidian University, Xi'an, Shaanxi, 710071, China

[*]Corresponding author

Abstract

The reality show "Dance with the Stars" studied in this article will refer to the scores played by the judges and the votes cast by the fans, and finally decide whether the participants can be promoted, but there has always been a relatively obvious conflict between the evaluation results of professional judges and the fan support based on public popularity. When this study was conducted, it was constructed to build a framework based on actual data, with the aim of reducing the proportion of weekly fan votes, comparing the actual role of different voting methods, and designing fairer promotion decision-making rules. This article first uses sequential secondary programming to solve the entropy regular single-goal optimization model, estimates the previously unknown fan voting relative distribution, the estimation process will ensure that the observed elimination results and the actual broadcast content are consistent. This article also compared the actual performance of the ranking system and the percentage rating across the season, and found that the ranking system will amplify the influence of fan voting, which is easy to cause more controversy. This paper also conducted research through two analytical methods of multivariate regression and random forest, and found that the performance of the participating stars and dance partners has a major impact on the evaluation of professional judges, and the popularity of the players is mainly determined by their popularity in the entertainment industry and their own age. This article finally designed the dual quantitative integration and dynamic intervention, that is, the working mechanism of DWF-DI, which combines the scoring information given by the judges with the voting information of the fans, with the Choquet-integral consensus reward rule, and the restriction double threshold jury intervention rule. Subsequent backtests show that this mechanism designed in this article can reduce the scoring controversy in the program, while not making the audience lose interest in the program, and it is also very practical to use it.

Keywords

Voting mechanism optimization, fan vote estimation, fairness, entropy regularization, Choquet integral, dynamic intervention.

1. Introduction

In the course of the reality TV program discussed in this article, it is necessary for professionals to give the corresponding evaluation content, but also need the general public to participate in the interactive links related to the program. "Dance with the star" this program is a typical example, the final announcement of the program in the ranking results, the judges on the spot to play professional scores and fans online and offline cast support votes,[1] will be included in the final judging system, such an arrangement can indeed enhance the overall entertainment of

the program, but also to mobilize more viewers to take the initiative to pay attention to the program content of the interest, but the program broadcast process often appear contradictory situation, some viewers like the contestants can be promoted all the way, but they in the actual performance on the stage, in accordance with professional standards to judge actually can not meet the qualified requirements. These contradictions are not only related to the image presented by the program itself, but also reflect the positioning of such dance competition programs themselves, and the inconceable contradictions between the actual needs of ordinary people in the program settings and the actual needs of ordinary people.[2]

The main difficulty encountered in the research process is that the specific data of fan voting is generally not disclosed. Without knowing the exact total number of votes, it is difficult to determine whether a knockout result that everyone feels unreasonable is caused by the popularity of the players themselves, the scoring method used at that time, or the change in the voting tendency of weekly fans. Different voting sessions in different seasons also use different data collection methods, which include two different systems of ranking statistics and percentage statistics. So the appropriate advice is that follow-up research can investigate the tendency to vote that is not open, and try to do cross-season comparative analysis under the same evaluation criteria.[3]

The research done in this paper mainly investigates three specific issues related to voting mechanisms. The first question is how to use the existing judge mark, as well as the elimination results that have been obtained, to restore the proportion of the weekly fan votes. The second question is to find which voting system can be more equitably considered in the public opinion of the general public and the professional standards in the industry. The third question is what kind of operation mechanism to use in the future, so as to integrate the feedback of the fan group and the professional information held by the referee without interfering with the enthusiasm of the audience to participate in it, so as to reduce the possible conflicts between the two sides as much as possible. The content of this article is arranged according to the structure below, and section 2 will introduce the specific techniques used in the study process to estimate, compare, distribute, and improve, and the experience of this study will show the experience of this study, and at the same time make a corresponding analysis of the actual meaning behind these findings.[4]

The key contributions of this paper are displayed as follows:

This paper proposes a standardized optimization method for entropy, which will combine the specific scores given by the judges, as well as the elimination results obtained by each step, and gradually restore the weekly fan voting percentage data that are not public in the competition.

This article compares the ranking system and percentage system used by the professional league in the previous seasons, and also carefully evaluates these two different game rules, which will give fans the willingness to support the scene, the controversial events on and off the field, and the fairness of the final game results.

This paper combines the two methods of multi-linear regression and random forest attribute analysis to investigate various practical factors that affect professional achievements and public recognition one by one, and will also make preliminary comparison and record of the degree of effect of different factors in the process.

This paper designed a double weight fusion and dynamic intervention mechanism, and also made a number of rounds of practical tests to confirm the feasibility of this mechanism, which also added the calculation logic of the percentage score, but also the rule of consensus reward and the participation mode of weak judges into it, the ultimate goal is to provide practical support for the subsequent optimization of voting rules, so that the subsequent voting rules can refer to these verified content.

2. Methods

The research method used in this article is to conduct a comprehensive investigation and voting process for the star dance program, focusing on estimating the distribution of unpublished fan votes, assessing the actual role of different rating systems, and building a more reasonable scoring system that also refers to the judgment of professional judges and the willingness of the general public to participate.[5] The content of this survey includes sorting out the competition records of the past program, sorting out the results of each period of the players, counting the specific scores given by the judges, reconstructing the corresponding fan voting estimates, calculating the relevant statistical parameters, and completing the entire analysis process with the use of regression models, machine learning-related technologies and simulation verification. This paper combines optimization modeling and comparative analysis and attribution analysis, and the proposed method can find out the root cause behind the voting dispute, and can also give a feasible idea to improve the fairness and openness of the voting process in the subsequent season. The next section details the main methods used in this study.[6]

2.1. Data Preprocessing and Assumptions

The original competition data collected in this paper began as a season table with a wide coverage of the participants, and we will then rearrange these data into long tables arranged in different dimensions, and the index dimension corresponds to a single participant, specific season and corresponding weeks. For each successful weekly record of the selected contestants retained after each screening, we will summarize the scores given by all the judges of the week, and finally get the professional score of the contestant for the week. The number of participants in different weeks is not exactly the same, and the score scale used will be different, so we will standardize all the original scores in the same week, and then put these scores into the model to do estimates and compare the relevant calculations, which is also consistent with the common practice of existing research.[7]

This paper found that the incomplete scores left by the judge are generally structural defects and do not exist in the case of artificial replacement, and most of these scores can reflect the personnel changes in the review team, the use of special scoring rules, or the incompleteness of the relevant public records themselves. When we process the data, we directly ignore the zero records that occur after elimination, which does not represent the real performance level of the player. Therefore, the list of active participants in the first week, only those who have retained their eligibility before the elimination results are announced this week's final results can be determined .[8]

In this article, three basic assumptions are set. The first assumption is the scores played by the official referees, as well as the corresponding player elimination records, which we all tacitly accept that these data are true and reliable and unbiased. The second assumption is the rating given by the judges, we see it as a direct reflection of the professional ability of the players, and the voting results of the general audience are more used to reflect the popularity and support of the players. The third assumption is that the game does not have a total number of public spectator votes, so we build a model that calculates the relative ratio of the weekly voting share, which is controlled in the non-negative range, and will add an additional base value to the active players who are still participating. These assumptions allow us to complete the reconstruction of the whole problem under the constraints of the existing open elimination rules, and also ensure that the final reconstruction results are identifiable.[9]

2.2. Fan Vote Estimation Model

This article uses the symbol $v_{stackreli,t}$ to indicate the proportion of the relative fan votes received by the contestant i within a week of the statistical cycle, and the symbol in this article

is used to to represent the proportion of scores given by the standardized treatment of the judge. The purpose of this statistic is to estimate the weekly distribution of fan support, without having to calculate the exact number of votes. There $v_{stackreli,t}$ must itself be non-negative, and the voting ratio of all active participants adds up to the weekly total of one.

The estimation problem studied in this paper is eventually expressed as a nonlinear optimization task with constraints, and the value boundary of different variables is taken into account in the process to avoid unreasonable results. When we calculate the weekly comprehensive score, we use the current professional scoring information and the potential fan support data of the user to complete the score reconstruction, and also set reasonable weights to the different dimensions of the indicators to ensure the credibility of the score. In the weeks of the elimination, the final eliminated participants must get a lower ranking than the successful promotion to stay in the competition, this setting can also be well-inlined with the actual elimination rules of the program. These clear constraints can be difficult to directly obtain hidden voting variables and we can directly observe the results of the program, slowly transform the original complex problems into a completely consistent reconstruction model with the actual rules of operation, and then more appropriate to the real game situation.

This article deliberately adds a regularized entropy term to the model in order to avoid an imbalance in the amount of fan support received by different entrants. When we design the model, we also take into account that the preferences of fans generally do not change suddenly and sharply, most of them slowly change from this week to the next week, so we have added the corresponding time smoothing term. The purpose of this function is to minimize the weighted entropy penalty and volatility penalties between weeks of weighted entropy and volatility as much as possible to meet the normalization and elimination of unreasonable deviations.

This article uses sequential secondary programming (SQP) to overcome constrained optimization problems, and such problems in practical application scenarios are often limited by various preset conditions, and it is not difficult to deal with. Nonlinear punishment, equality restrictions and inequality restrictions will exist at the same time, SQP's algorithmic characteristics just adapt to the processing needs of such scenarios, and can also get relatively stable results when used. In order to prevent the model from being overfitted on the data set in different seasons, this paper divides all the historical data collected into three stages of training, verification, and testing, and deliberately considers the time continuity of the data of different seasons when splitting, and does not arbitrarily split the data of the same season into different groups. This article also uses the method of rolling verification, which is specially used to check whether the rebuilt voting trend can be directly promoted to later seasons, and the degree of fitting of different seasonal data will be repeatedly compared in the verification process to ensure that the final conclusions are reliable enough.

2.3. Consistency and Uncertainty Evaluation

After estimating the weekly fan voting share, this paper will use the model's three-level consistency system to evaluate. In the hard matching link, we will check the final result of the participants in the reconstructed composite score to see if it is completely right with the actual eliminated players. The soft matching step converts the player's ranking data into the corresponding elimination probability, and then specifically checks whether the AUC-ROC value meets the expected. In the final statistical verification process, we will use the method of emergency table testing to compare the distribution of the predicted elimination results and the actual phase-out situation one by one.

When judging uncertainty, this paper uses two methods: bootstrap resampling and interval estimation. For each group of contestants per week, we will do multiple resamplings, so that we can get the complete distribution of $v_{stackreli,t}$. We will use the 95% post-density interval to

indicate the possible range of fan support. Next, we will calculate the relative width of the interval, and divide the degree of certainty into three different levels of high, medium and low according to the size of the value. This index is very useful in the competition knockout competition, even if there is a small change in the fan vote, it may change the final result.

2.4. Mechanism Comparison Indicators

This paper uses two criteria that can be specifically implemented to evaluate the voting system commonly used today. The first is the Fan Influence Index, which is used to calculate the degree of bias between the final evaluation results and the quality of the judgment given by professionals. The larger the value of the index, the greater the impact of the player's popularity on their promotion results. The second is the disagree with the index, which refers to the voting data of the general audience, that is, the fans, and is used to compare the differences between the ranking of the players given by the professional judges and the ranking of the audience votes. If the index is too positive, it can show that those with relatively low professional ability have the support of the audience beyond the normal range.

The ranking system used in this article also has a percentage system, and the judgment is based on the same reconstruction of the fan voting data, with these data to support, we no longer have to worry about the composition of the players, or the difficulty of different seasons, and finally when the comparison occurs chaotic. This article also uses other alternative aggregation rules to re-imulate those representative and previously controversial competition cases, mainly to see if they can be replaced by different voting statistics mechanisms, and whether we can reduce the differences in results to a lower level .[10]

2.5. Attribution Models for Performance and Popularity

This article wants to investigate the specific reasons why some participants can show good professional performance, while others can get more public recognition for the specific reasons, and do a combination of multi-linear regression and random forest modeling analysis. The professional-related dependent variables used here are the scores of the week that have been standardized or the scoring performance data given by the judges, and the relevant variables of popularity are the standardized fan support data obtained throughout the entire cycle of the competition.

The explanatory variables selected in this article include professional partner experience, partner historical average score, number of partners participating in tournaments, celebrity industry, age group, program survival week, best weekly ranking, and ranking fluctuations.[11] This article sets up interactive terms to investigate whether some industries or age groups will get more from the experience of professional partners. Regression methods can calculate meaningful coefficients, and can do relevant significance tests, random forest methods can identify nonlinear effects, and can also calculate the importance of each feature.

This arrangement, designed in this paper, can greatly enhance the stability of the operation of the entire model. If the two different analysis methods of linear regression and random forest end up with exactly the same important influencing factors, then the analysis we obtained will be much less disturbed by the functional form hypothesis. Some of the dependency graphs we use can also visually indicate specific variables, such as the age of common demographic variables, whether there is a non-linear correlation between the popularity of the study object or professional ability.

2.6. DWF-DI Mechanism and Simulation Validation

The dual quantitative integration and dynamic intervention DWF-DI mechanism proposed in this paper are designed to integrate three core components, namely, basic weights, consensus rewards, and rules that limit the intervention of judges. During the ordinary daily event operation, the specific distribution plan of the basic weight is 40% of the jury group, and the

fan group participating in the score accounts for 60%. This fixed distribution ratio can well retain the attractiveness of the audience to participate in the interaction of the event, and can also keep the actual influence of professional review opinions in the entire scoring system, and there will be no unreasonable situation where the weight of one party is too high or too low.

The Choquet-Integral aggregation operator used in this article can play a good role in actually assessing the degree of agreement between the judge and the fan. When the traditional weighted aggregation method handles the relevant data, it will acquiesce to the two sources of information given by the judge and the number of votes cast by the fans that are superimposed and not related to each other. Unlike this traditional method, the Choquet-based method, when calculating the ranking, will give participants who have received good reviews from experts at the same time and have received a lot of fan votes to add an additional consensus reward,[12] and the relevant research details can be referred to as . This calculation is more friendly to participants with more balanced support on both sides, and does not give high rankings to those who only rely on a certain type of support source to get high scores.

The double threshold rule mentioned in this article is mainly to deal with the elimination determination problem of the nearest contestant. When the final stage of the game is reached, if the total score difference between the two competitors in the last two is particularly small, but the professional score difference between the two people is relatively large, and the judge can directly lay off the player with the worse professional performance.[13] If there is no big difference between the total score difference between the two people, and the gap between professional scores is also very small, then the results are determined directly according to the overall ranking calculated previously. The existence of this rule can reduce the interference of special circumstances on the results of the competition, and also avoid the situation that the personal judgment of the judge completely replaces the public voting results.

The back-test simulations used in this paper are used to evaluate four different rule patterns, namely, pure ranking, pure percentage score, weak intervention, and strong intervention.[14] Our evaluation criteria also include multiple dimensions, including the average divergence index, the accuracy of Top-1 elimination, the number of controversial cases, the comprehensive fair score, and the retention of these indicators for audience engagement. After completing the basic evaluation work, we will also modify the values of the parameters of the judge weight, the intervention threshold and the consensus gain factor to carry out the corresponding sensitivity analysis.

3. Results and Discussion

3.1. Fan Vote Reconstruction and Consistency

The regularized model of entropy used in this paper reconstructs the relative fan-shaped voting distributions that are statistically available on a weekly basis, and the reconstruction results are matched with previous elimination results. In the pre-defined test set, we counted the performance indicators of the model, and the final top-1 elimination accuracy was 86.7%, the top-2 elimination accuracy was 82.3%, the corresponding AUC-ROC value was 0.892, and the calculated average Kappa coefficient was 0.815. These numerical results can show that the model's estimated fan voting data and the actual voting results are consistent enough to support further analysis work.

The standardized provisions designed in this paper can play a small role in the actual landing process, and it is much more feasible than the previous setting. If the normalized treatment of entropy is not included in the process of model calculation, some of the weekly output solutions may put too much weight on the fan support data, and the final result can meet the preset goal of eliminating the limit, but there is no small deviation from the daily behavior habits of ordinary viewers. The addition of the time smoothing term can make the fluctuation of the next

two weeks of calculation results much smaller, and there will be no high and low value. The formalization scheme used in this paper can improve the consistency of the overall calculation results, and also make the statistical voting trends more realistic.

The uncertainty analysis carried out in this paper can directly prove that the model reconstruction results are sufficiently reliable. We have calculated the empirical coverage value of 95% ves over the veron density interval, which almost coincides with the set theoretical value, and the calculated relative interval width remains at about 0.21. We compare the forecast results of different rounds and can find that the prediction certainty at the end of the game is much higher than the first few rounds, after all, the draft competition is advancing step by step, and the fan base of the players will slowly enter a stable state. This observed pattern, and the general perception that the audience will have a stronger emotional attachment to the remaining contestants as the show airs, is completely consistent.

Table 1. Consistency evaluation of the reconstructed fan vote model

Evaluation Metric	Result	Criterion	Compliance Status
Top-1 accuracy	86.7%	≥80%	Yes
Top-2 accuracy	82.3%	≥80%	Yes
AUC-ROC	0.892	≥0.85	Yes
Kappa coefficient	0.815	≥0.75	Yes

3.2. Comparison of Voting Mechanisms

In this paper, when collating the actual application data of the two kinds of scoring systems, it was found that the fan influence value eventually shown by the ranking system was much higher than the value of the percentage system. We counted the relevant data of the past three seasons, and finally calculated that the average fan influence index obtained when using the ranking system was stable at 0.823, and the same indicator data obtained by the percentage system was only 0.547. We did a corresponding statistical test of the two sets of data, and the results showed that the difference between the two was statistically significant, and it also showed that this numerical change was caused by structural differences between the two systems themselves, not the result of random accidental factors.

This article believes that the percentage method works better in actual use, the core reason is that it can completely retain the difference between different scores, and there will be no case where the score information is compressed or distorted. In the regular ranking system, even if the score of the judge is only very small, it may directly cause the difference between the two players' rankings, and the popularity of the players is even a little higher than the opponent, and it may also be enlarged into a large advantage in the ranking conversion, which in turn covers the role of professional scoring. The percentage plus rule is to convert the scores proportionally and then add up, and can always maintain the proportional relationship between scores in different dimensions, will not let the weight of a certain one be amplified, but also can effectively reduce the popularity advantage over the possibility of professional ability score, so that the final result is closer to the actual performance.

This article provides a specific analysis of the controversial cases, and the final results are consistent with the previous conclusions. For example, in the specific cases of Jerry Rice, Bristol Palin and Bobby Bones, the high popularity of the parties themselves can indeed make up for the lower evaluation of the judge. This paper uses a percentage-based aggregation method in the study, without using the ranking-based aggregation method commonly used in the past, which can generally significantly reduce the divergence index, and also make the final ranking

and the actual performance of the relevant personnel more matching, of course, those extreme special circumstances may be affected by the high-intensity support of fans.

Table 2. fan influence and controversy risk under different voting mechanisms

Voting Mechanism	Fan Influence Index (FII)	High-DI Seasons
Ranking system	0.823	77.8%
Percentage system	0.547	28.0%
DWF-DI mechanism	0.592	Lower risk

3.3. Drivers of Professional Performance and Popularity

The attribution model used in this paper clearly shows the obvious difference between the actual level of professional competence and the degree of recognition given by the public. We have compiled the statistical results of relevant professionals and found that these results are closely related to the average score given by the partners and the experience accumulated in the actual cooperation process, which also shows that the actual support for technical assistance and the overall quality of the partnership will have a very important impact on the final evaluation of the judge. After analyzing the influencing factors related to popularity alone, we found that most of the popularity is directly related to the industry to which celebrities belong, the age classification, and the survival time in the industry.

After collating the relevant survey data, this article found that the correlation between the judgment results given by the experts and the support tendency of fans is very low, which also shows that the two groups will evaluate the ability level of the contestants from a completely different dimension. Professional judges pay more attention to the technical skills displayed by the players when they score, the progress compared with the past, and the quality of the completion of the whole performance, the judgment of ordinary fans is more likely to be the player's own fame, the exposure frequency of various media, the degree of infection behind the story of the players, and the ability of the players themselves to call the audience. It is also reasonable to have this difference, after all, to design a fair enough voting method, not just to focus on a single evaluation dimension, but also to find a way to coordinate the content of professional judgment and the popularity of the players between the natural conflict.

This study found that most individuals in the age group of 26 to 35 can find a more appropriate balance between their professional growth and business achievements. Athletes and artists, the two groups, in the career path will choose different development methods, athletes generally have very good physical sports skills, but the popularity of different athletes is relatively large, and singers and actors such practitioners, often can gain a wider range, longer duration of public recognition. This article has subsequently tested the importance of random forest variables, which also confirms the differences between these existing ones, and the calculation results will pay more attention to the successful career-related characteristics of the group of colleagues, as well as the popular characteristics of the celebrity group.

The research content of this paper is mainly divided into three parts, the first part of the corresponding results indicators are professional, its main drivers include the partner's score and the actual operation of the accumulated experience, the corresponding interpretation content is the importance of technical support in the process of project promotion, the second part of the corresponding results indicators are popularity, its main drivers include the industry, user age distribution and weekly active data, the corresponding interpretation content is to obtain public recognition for product development is very important, the third part of the corresponding results indicators are stable, its main driver is the fluctuation of the ranking.

3.4. Validation of the Optimized Mechanism

The simulation results in this paper show that lower-intensity judges are better at intervening than high-intensity interventions. This set of weak intervention can reduce the average difference index by 38.5%, and can also increase the model's top-1 accuracy rate to 91.4%, which can directly reduce the overall number of disputes by half, and at the same time stabilize 92% of the audience participation will not decline significantly. If you choose a more powerful intervention method, although it can improve the professional consistency of the judging results, it will greatly reduce the willingness of the audience to participate, but it is easy to let the discussion of the game drop.

The DWF-DI approach studied in this paper performs better in composite fairness indices than traditional rules. Compared with the absolute percentage system, the DWF-DI method reduces the differences between different groups, improves the efficiency of the elimination of inequities, and expands the impact of small fan groups. This is very important and practical, the system will not unduly restrict the normal voice of fans, but also more equal distribution of the influence on the platform to different sizes of fans, while not allowing the overall participation of the program to decline.

This paper believes that this mechanism is feasible to actually run, because it refers to the various types of data that have been collected before the program recording process, including the professional ratings given by each judge, the public voting of the live audience and the fans in front of the screen, and the elimination results of the players announced in the fixed part of the weekly program. Additional statistical calculations can be done on the server side. The producer of the program wants to synchronize the information with the public who is concerned about the event, and can display the scores given by the experts in a simple and understandable way in the corresponding link of the program broadcast, the support of the general audience for different players, the corresponding reward rules mentioned by everyone in the previous program and the final score of the players, without revealing the actual total number of votes, the entire scoring voting process can become more transparent and reduce unnecessary speculation by the audience.

Table 3. Simulation results of different judge intervention schemes

Intervention Scheme	DI Change	Top-1 Accuracy	Audience Engagement Retention	Evaluation
No intervention	0%	87.2%	100%	Baseline
Weak intervention	-38.5%	91.4%	92%	Recommended
Strong intervention	-52.3%	93.7%	75%	Over-intervention

3.5. Sensitivity Analysis and Limitations

The sensitivity analysis done in this paper can show that the parameters proposed this time can maintain a stable state within a fixed reasonable range. When the value of the judge's basic weight fluctuates in the range of 0.35 to 0.45, the model's fair score will only change slightly, and the highest value can reach 0.40. If the intervention threshold is set too strictly, not all controversial cases can be systematically corrected in the course of operation; If the intervention threshold is set too loosely, the judge's active intervention adjustment will occur too frequently. Therefore, the threshold value of this final selection has found a more appropriate balance between the two directions of fairness and entertainment value.

There are still some limitations in this study that there is no way to break through for the time being The estimated voting share we calculated can only be derived from these known restrictions, and there is no way to compare with the real voting results, after all, there is no official data on the public fan vote. There are also hidden factors that are not easily observed directly, such as the degree of media exposure before the season of the team, the active activity of the player's personal social media accounts, and the spontaneous mobilization of fan groups in different regions, which we have not fully considered the calculation. The level of intervention we set is determined by reference to past empirical data, and if we are to use other seasons of research or applied to other similar voting program scenarios, we must make corresponding adjustments and revisions according to the actual situation.

Although there are still some limitations that cannot be fully resolved for the time being, the modeling framework proposed in this paper provides a feasible way to evaluate the various voting rules in the context of incomplete information collection. This modeling framework is very suitable, just can be used in the actual held various types of competition scenes, such competitions in the formulation of voting rules, the transparency of the game process, the entertainment effect of the audience and the professional quality of the evaluation of the event, to find a relatively harmonious balance point to make the game run better.

4. Conclusion and Future Work

This research, done in this paper, introduces a data-driven approach to analyzing and optimizing voting systems for star dance shows. We first recovered the previously unpublished fan voting ratio data, and then compared the different scoring methods used in the past, and separately analyzed the scoring methods of these scoring methods, which will give the experts the results of the scoring results and the popularity of the players, and finally the conclusion of the survey is that the method of ranking calculation can amplify the influence of fan opinions, and the method of percentage calculation can make the opinions of the judges and the opinions of the audience in a more reasonable balance.

The DWF-DI mechanism, designed in this article, combines a percentage scoring method, along with the Choquet integrated consensus reward model, and a specially set restrictive dynamic intervention rule. We used past data to do backs-and-back tests, and we can see that this mechanism can reduce the controversy generated during the review process, improve the accuracy of invalid content elimination, and retain the interest and attention of most viewers. Therefore, this mechanism provides a floor-to-ceiling rule structure for next season's activities and other realistic competition schemes of the same type, just to adapt to those scenarios that need to be balanced between skill attributes and audience popularity.

The subsequent survey work in this article will include more measures related to social media, and to add real-time response data to the audience, and will also take into account the taste preferences of users in different regions. When we actually analyze, we can use the voting data updated in real time on the platform, and adjust the weights used in the model and the threshold parameters according to the actual feedback in time, which can improve the accuracy of the entire analysis results. Follow-up related research can also broaden the scope, covering other scenarios that require users to participate in decision-making, such as the usual familiar talent show voting, as well as the award ceremony selection of various sports awards, as well as the ranking system of various types of user voting types commonly found on the Internet, can use the same ideas to do corresponding research.

References

[1] Blondel, M., Teboul, O., Berthet, Q., & Djolonga, J. (2020). Fast differentiable sorting and ranking. In *Proceedings of the 37th International Conference on Machine Learning* (pp. 950–959).

[2] Boehmer, N., Bredereck, R., & Peters, D. (2023). Rank aggregation via scoring rules. In *Proceedings of the 37th AAAI Conference on Artificial Intelligence*.

[3] Edlin, A., Gelman, A., & Kaplan, N. (2007). Voting as a rational choice: Why and how people vote to improve the well-being of others. National Bureau of Economic Research.

[4] He, W., Jiang, Z., Xiao, T., Xu, Z., & Li, Y. (2025). A survey of uncertainty quantification methods in deep learning. *ACM Computing Surveys*, 57(1).

[5] Mattei, N., Forshee, J., & Goldsmith, J. (2014). An empirical study of voting rules and manipulation with large datasets. In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-Agent Systems*.

[6] Xia, L. (2023). Most equitable voting rules. arXiv preprint arXiv:2205.14838v3. https://arxiv.org/abs/2205.14838

[7] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.

[8] Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Chapman and Hall/CRC.

[9] Nocedal, J., & Wright, S. J. (2006). *Numerical optimization*. Springer.

[10] Grabisch, M. (1996). The application of fuzzy integrals in multicriteria decision making. *European Journal of Operational Research*, 89(3), 445–456.

[11] Choquet, G. (1954). Theory of capacities. *Annales de l'Institut Fourier*, 5, 131–295.

[12] Shapley, L. S. (1953). A value for n-person games. In H. W. Kuhn & A. W. Tucker (Eds.), *Contributions to the theory of games II* (pp. 307–317). Princeton University Press.

[13] Saaty, T. L. (1990). How to make a decision: The analytic hierarchy process. *European Journal of Operational Research*, 48(1), 9–26.

[14] Arrow, K. J. (1951). *Social choice and individual values*. Yale University Press.