Journal of Clinical Trials and Case Studies

Open access

Optimized "Cross-assignment" Program by Leveraging the Optimal Objective Function Method to Improve the Comparability of Scores

Bin Zhao[1,*], Xia Jiang[2]

[1,*]School of Science, Hubei University of Technology, Wuhan, Hubei, China

[2]Hospital, Hubei University of Technology, Wuhan, Hubei, China

RESEARCH ARTICLE

ABSTRACT

Large-scale innovation contests are becoming increasingly popular globally, with many innovators and enterprises actively participating. Work allocation optimization and fair judging are two significant concerns in innovation contests. This academic paper aims to recommend an optimized "cross-assignment" program by leveraging the optimal objective function method to improve the comparability of scores given by various judges. We conducted data-based descriptive statistical analysis and concluded that the two-stage and weighted evaluation schemes are more beneficial than the traditional judging scheme. Nonetheless, there are still some shortcomings that require addressing. To enhance fairness, we propose an improved two-stage evaluation scheme. In the first stage, we normalize the scores with a normal distribution. In the second stage of the process, we implement a system utilizing the Borda sorting technique to categorize submissions into five distinct groups for Judges to evaluate based on their perceptions. We also detail a method for weighting tied scores to determine the final rankings. Testing indicates that this approach yields a Normalized Discounted Cumulative Gain (NDCG) of 0.8667, implying greater fairness and precision in the assessment of submissions

KEYWORDS

Evaluation program, Normalization of a normal distribution, Borda sorting method, Grey correlation analysis, Analysis of variance

Correspondence to: Dr. Bin Zhao, School of Science, Hubei University of Technology, Wuhan, Hubei, China, Email: zhaobin835@nwsuaf.edu.cn

Received: October 24, 2023, **Accepted**: October 31, 2023, **Published**: November 17, 2023

DOI: 10.1042/JCTCS.6.1.014

Citation: Bin Zhao, 2023, Optimized "Cross-assignment" Program by Leveraging the Optimal Objective Function Method to Improve the Comparability of Scores. JCTCS. Vol 6.

INTRODUCTION

Large-scale innovation-based competitions are an effective means of fostering science, technology, innovation, and entrepreneurship development. They attract innovators from different fields to bring innovative solutions to society. The judging program largely determines the success of a competition. Its fairness and transparency are essential to attract more talented participants [1-4]. Therefore, to ensure the sustainability of large-scale innovation competitions, it is necessary to conduct research and improve the effectiveness of the judging program. There is a lack of a standardized judging mechanism in these competitions. A two-stage (online and on-site judging) or three-stage (online judging, on-site judging, and defense judging) process is usually used. The critical aspect of this type of competition is innovation. Innovation refers to the ability to perceive what others do not understand. Evaluation of the same work by different experts can lead to divergent opinions, while innovation leads to novel solutions to problems. It is, therefore, essential to develop an unbiased, impartial, and systematic innovation competition selection scheme to ensure credibility and recognition. The design and improvement of evaluation programs for large-scale innovation competitions have long interested scholars and experts in various disciplines, driven by recent rapid advances in science and technology.

In their 2006 co-authored publication, Henry Chesbrough, Wim Vanhaverbeke, and Joel West introduced the concept of open innovation from its inception [5], discussed its implications for competitions, provided a framework for evaluating large-scale innovation programs, and explored the impact of open innovation for matches. Innovation is the result of discovery, and societal progress results from innovation. The evaluation of innovative ideas has generated much debate, and there are ongoing efforts to establish unbiased evaluation methods, including the framework developed by Poetz and Schreier in 2012. Their model emphasizes the importance of involving regular users to achieve diverse innovations while considering feasibility and social impact. You Qinggen has developed a user-friendly evaluation index system for experts, which provides a theoretical contribution to improving the current evaluation indexes. In addition, Changchao and Minglong [4] investigated the feasibility of evaluation using AHP hierarchical analysis, grey cluster analysis, and other algorithms in the "Challenge Cup" start-up project. In particular, for large innovation competitions, such as joint national and provincial competitions, existing programs are typically based on those developed for smaller competitors, which are not practical to use and tend to be inconsistent in large rounds, leading participants to question the results.

DATA

The data utilized in this paper was provided by a large-scale competition based on innovation. Technical term abbreviations, when used, were explained. The language used was clear, objective, and value-neutral, with a formal register. Biased phrasing was avoided. It was divided into three copies, each undergoing two stages of judging. Five experts assessed the entries in the first stage and generated raw and standard scores. The structure of the paper followed the conventional academic sections and maintained consistent author and institution formatting. The text was precise, free from grammatical, spelling, and punctuation errors, and presented a logical flow of information with causal connections between statements. In the second stage, another panel of three experts reviewed the entries, generating raw and standard scores and a concordance score. In the initial phase, the mean scores of the five experts were calculated, and the pieces positioned within the top 16% of all teams were admitted to the subsequent evaluation stage. After reassessing the standard scores and making appropriate adjustments to the standard scores of a few works with significant differences, the standard scores of the five experts in the first stage and the standard scores of the three experts in the second stage will be averaged into four scores. The resulting scores will then be ranked based on the final total scores to establish the ranking of the works. The dataset comprises 3,000 teams and 125 experts. Each piece of work was randomly assigned to five experts in the first stage, while three experts were given in the second stage. The experts worked independently without interacting with one another throughout the process [6].

DATA CLEANING

- Processing of missing values: For experts who did not rate some of the data points, we remove individual samples with missing values.
- Processing of outliers: Individual samples with outliers are removed from the data. For example, we remove ratings with significant extreme differences in works that did not participate in the second evaluation or that did not win the prize in the second evaluation. We convert the data types to ensure the consistency of our analysis.
- Conversion of data types: Convert different styles in the data to the correct data type. For example, "first prize" in the data should be converted to 1, "second prize" to 2, "third prize" to 3 and "did not win" to 0. Consistency should be maintained throughout the processing.
- Processing for consistency: Only papers that reached the second stage were considered to allow a comparative analysis of the two locations.
- Pre-processing stage: The raw data were technically processed to extract the ten indicators necessary for the model covered by the dataset to establish the evaluation model for this paper. ① the raw score of the first expert in the first evaluation; ② the raw score of the second expert in the first evaluation; ③ the raw score of the third expert in the first evaluation; ④ the raw score of the fourth expert in the first evaluation; ⑤ the raw score of the fifth expert in the first evaluation; ⑤ the raw score of the fourth expert in the first evaluation; ⑤ the raw score of the fifth expert in the first evaluation; ⑥ The raw score of the first expert in the second evaluation; ⑦ The raw score of the second expert in the second evaluation; ⑧ The raw score of the third expert in the second evaluation.

RESULTS AND DISCUSSION

To ensure the fairness, impartiality, and scientific validity of the judging process in the innovation category of the competition, we investigate whether the subjective evaluations of the experts affect the judging results. As a result, we have designed a two-stage scoring system consisting of standard distribution standardization and a Borda ranking-based sorting method. This approach allows us to optimize the judging process and reduce the potential influence of human factors on the final decision [7]. We implemented a two-stage scoring scheme using Matlab based on competition scoring data. In the first stage, we normalized the scores using the normal distribution. To achieve this, we calculated the normal distribution of each judge's score and obtained the normalized result. The figure below shows the normal distribution of some judges' scores and the resulting normalization process.

FUTURE WORK AND IMPROVEMENTS

The judging scheme proposed in this study for large-scale innovation contests can be applied to innovation contests and various other large-scale assessments and evaluations, such as art tests, exams of multiple levels, and elections of public officials, while achieving a more equitable level of accuracy. Its benefits are the enhancement of impartiality, consistency, and objectivity throughout the assessment process. Implementing and designing large-scale innovation competitions is a complex process influenced by several factors that impact each other. This paper focuses on controlling subjective factors that influence the scoring of entries and only provides solutions for controversial entries [8-10]. To ensure the quality and longevity of competitions, further research into additional features for factor analysis is necessary, and a continued focus on improving the judging scheme is imperative. Conflict of interest we have no conflict of interests to disclose and the manuscript has been read and approved by all named authors.

ACKNOWLEDGMENTS

This work was supported by the Philosophical and Social Sciences Research Project of Hubei Education Department (19Y049), and the Staring Research Foundation for the Ph.D. of Hubei University of Technology (BSQD2019054), Hubei Province, China

REFERENCE

1) Annette R, et al. Objective assessment of anesthesiology resident skills using an innovative competition-based simulation approach. A & A Case Rep. 2015, 5:79-87.

2) Chao C, Long M. Research on the evaluation plan of "Challenge Cup" startup competition. Econ Eng. 2011, 13:121-124.

3) Irina V. Multi-attribute decision-making methods as a part of mathematical optimization. Mathe. 2019, 7:915.

4) Suofa Z. A comprehensive evaluation model based on analytic hierarchy process. J Hefei Univ Technol: Natural Science Edition. 2000, 23:582-585.

5) Niu W, Bingru Z, Zhou H. Design of performance evaluation system of financial shared service center in new environment-based on analytic hierarchy method. Frnd Accoun. 2019, 21:135-140.

6) Feng W, Xueqi H. Research on optimization of data storage model of SAAS-based general review system. Comp Appl Res. 2014, 31:2067-2070.

7) Pengcheng X, Bi Q, Zhang H, et al. Construction of data-driven user portrait model of digital library. Libr Infor Ser. 2019, 63:30-37.

8) Segev E. Crowdsourcing contests. Eur J Oper Res. 2020, 281:241-255.

9) Martire KA, Edmond G. Rethinking expert opinion evidence. Melb Univ Law Rev. 2017, 40:967-998.

10) Xiaohui L, Yuhan S, Haoyang L. An experimental study of quantitative evaluations on saliency methods. ACM SIGKDD Conference on Knowledge Discovery and Data Mining. 2021, 3200-3208.