

**Project:** Muscle Metrics: Revolutionizing Bodybuilding Competitions with Elo Ratings

**Team:** Jessica De Silva, Andrei Prokhorov

**GitHub:** <https://github.com/jessicadesilva/bodybuilding-ranking>

**Objective:** Create a metric that can be used to predict the outcome of professional bodybuilding competitions. The metric is based on each competitor's historical performance in amateur and professional competitions.

**Motivation:** In 2023, there were 43 occurrences of at least 30 competitors in a given class. Large classes such as this make it difficult for judges to accurately rank competitors. Judges determine rankings after seeing competitors in three distinct comparison situations: (1) in groups of 8 - 10 in alphabetical order, (2) individually in alphabetical order, and then (3) in groups of 8 - 10 based on the head judge's preliminary ranking. An alternative ranking of competitors can be determined using the metric developed in this project. Competitor grouping according to this metric-based ranking can be used as a substitute for either the initial alphabetical grouping, or as a replacement for groupings according to the head judge.

**Data:** Using BeautifulSoup, we scraped the results of amateur and professional bodybuilding competitions for the last 12 years from the website <https://npcnewsonline.com/>. We had 783,708 entries in total, with 57,393 entries in professional competitions.

**Model:** We converted competition rankings into head-to-head matchups whose results are fed into an [Elo](#) rating model. The Elo rating model, originally used for chess matches, assigns a numerical value to each competitor which measures their relative skill level. The model adjusts competitor Elo ratings based on match outcomes, these adjustments are more substantial when the outcome was unlikely based on the competitors' Elo ratings. Each new amateur competitor is assigned the same starting Elo rating, whereas new professional competitors with no amateur data are assigned a starting Elo rating relative to historical averages for new professionals in their division. Predicted rankings are determined by the competitors' relative Elo ratings.

**Results:** The [Kendall Tau Correlation Coefficient](#) measures the similarity between predicted and actual rankings by determining the number of rank inversions required to obtain the actual ranking from the predicted ranking. A perfect predicted ranking gives a correlation coefficient of 1, whereas a perfectly reversed ranking gives a value of -1. The Elo rating rank prediction gave a correlation coefficient of approximately 0.4 averaged across all professional competitions since 2022, whereas alphabetical ranking gave a value of approximately 0. Therefore we can conclude that using an initial grouping based on the Elo rating will give a preliminary ranking much closer to the actual ranking as compared to alphabetical grouping. Further, we used the Precision@K metric with  $K = 5$  to analyze the accuracy of our rankings for professional bikini competition outcomes. The 30-day rolling average of this metric is approximately 65%, implying that, on average, 3 of the top 5 in these competitions can be predicted using Elo ratings alone.

**Observations and Limitations:** Bikini division competitors have notably higher Elo ratings compared to other divisions, likely attributed to the division's popularity and its relatively lower physical strain, enabling them to compete more often. In the 2023 Ms. Bikini Olympia competition, all competitors in the lower Elo rating quartile tied for last place (16th place) and qualified to compete in Ms. Bikini Olympia by winning a professional show outside of the United States. In contrast, 10 of the 12 competitors in the upper Elo rating quartile ranked in the top 11 and all 12 qualified by winning a show in the United States. Elo ratings for international competitors are typically relatively low, however this also seems to align with their performance at Ms. Bikini Olympia.