
REFINING SEARCH SPACES IN HIERARCHICAL TOURNAMENT BRACKETS USING CHALK INDEX

Christopher Toukmaji

University of California, Santa Cruz
ctoukmaj@ucsc.edu

ABSTRACT

In this paper, we propose "Chalk Index", a new metric that quantifies the variation between a "true" bracket and the Chalk Bracket. The Chalk Index is then applied in bracket outcome prediction. We use the previous eleven years of NCAA Division 1 Men's Basketball Tournaments as a case application of Chalk Index. The NCAA Division 1 Men's Basketball Tournament, further referred to as "the tournament", is an annual single elimination bracket-style tournament of 64 college basketball teams, all vying to be crowned the champions of the collegiate basketball world. Each tournament team has each of their game analytics logged by the statistical database "KenPom". However, KenPom updates its analytics as the tournament is played, so historical data will be biased by tournament results. Since KenPom does not store the pre-tournament data, we use the WayBackMachine which allows us to view the web-pages before KenPom has updated its database to reflect the tournament results. We use the combination of KenPom and the WayBackMachine to parse and curate pre-tournament datasets for all the tournament teams from 2011 to 2022 - with the exception of 2020 when the tournament was canceled due to the COVID-19 pandemic. Previous research has shown that there are $2^{63} \approx 9.22 \times 10^{18}$ possible tournament brackets each year, only one of which perfectly predicts each game. We observe that sampling from the distribution of historical true Chalk Indexes improves performance in Stochastic Generation and Logistic Regression prediction methods.

1 Introduction

The NCAA Division 1 Men's Basketball Tournament draws tens of millions of watchers every year [5]. The challenge of predicting a perfect outcome of the NCAA bracket has become a yearly tradition amongst offices, families, and friend groups. The tournament consists of 64 teams placed in a single-elimination bracket. In 2014, Warren Buffett famously hosted a bracket pool and offered one billion dollars to whoever could predict a perfect bracket [1]. However, previous research has shown that there are $2^{63} \approx 9.22 \times 10^{18}$ (9.22 quintillion) distinct bracket combinations [4], so the odds of winning such a prize were slim.

In this paper, we present a new metric along with its derivation and application to automatic bracket generation methods. Section 2 presents a background on the history of the tournament and bracket prediction. In Section 3, we describe our training and testing data, which consists of the results of all NCAA Division 1 Men's Basketball tournament brackets dating back to 1985, since that is when the NCAA moved to a 64-team single elimination tournament bracket.[15]. Section 4 introduces our methodology of deriving a new metric, "Chalk Index". Sections 5 and 6 describe the experiments and results when applying Chalk Index to the three automatic bracket generation methods. Finally, we conclude with Future Extensions and Conclusion in Section 7 and 8.

2 Background

2.1 Tournament Bracket History

The NCAA Division 1 Men's Basketball Tournament, henceforth referred to as the "tournament", is an annual competition amongst 68 college basketball teams all vying to be crowned the champions of the collegiate basketball world. Among the 68 tournament teams, 32 of them qualify for the tournament by winning their local conference championship; these teams are referred to as "automatic qualifiers". The remaining 36 teams are chosen by a panel - or "selection committee" - of twelve Division 1 athletic directors and commissioners [10]; these teams are referred to as "at-large bids".

The selection committee also assigns a seeding - between one and 16 - for each of the 68 teams, where the lower-value seeds are deemed to be better teams. For example, a team ranked as a 1-seed is one of the best teams in the nation, while a 16-seed is one of the worst teams to make the tournament. Each tournament has four regions, and every region has 16 teams with distinct seedings between one and 16. The first round match-ups in each region are determined by the team's seedings. For example, the one seed plays the 16 seed, the two seed plays the 15 seed, and so on and so forth, culminating with the eight seed playing the nine seed. This is a single-elimination tournament, so the winners move onto the next round, and losers are eliminated. The example bracket for a single region is detailed in Figure 1. [2]



Figure 1: The "West" region of the 2021 NCAA Men's Basketball Tournament

There are 64 teams that play in the actual tournament, but there are 68 that qualify. From the original 68 teams, 60 of them are placed into the bracket immediately, while the other eight must participate in play-in games or the "First Four" (circa. 2010) [15], where the winner is placed into the bracket, and the loser is not. Hence, a bracket is comprised of the resulting 64 teams. The play-in games occur between Selection Sunday - the day when the bracket is released - and the first day of the tournament. This is why both teams appear on the bracket (see line 16, where the winner of Norfolk State and Appalachian State gets the right to play Gonzaga). The winner of each of the four regions plays each other in the "Final Four".

2.2 History of Bracket Prediction

The challenge of predicting a perfect outcome of the NCAA bracket has become a yearly tradition amongst offices, families, and friend groups. Predicting the outcome of an NCAA bracket will henceforth be referred to as "bracket". Participants can submit their brackets to several large sports media outlets such as ESPN and Yahoo Sports. These organizations host a worldwide competition where participants can submit their bracket, also known as a "bracket

pool". In 2014, Warren Buffet famously hosted a bracket pool and offered one billion dollars to whoever could predict a perfect bracket [1]. However, previous research has shown that there are $2^{63} \approx 9.22 \times 10^{18}$ (9.22 quintillion) distinct bracket combinations [4], so the odds of winning such a prize were slim.

Bracket pool participants typically make their predictions using empirical data like Win Percentage, Offensive Efficiency, or Defensive Efficiency - or through instinct and partiality towards specific teams, geographical regions, or mascots. There has been previous work to use statistical methods and mathematical models in bracket prediction. In 2014, Google hosted, "March Machine Learning Madness", their first bracket pool on the Data Science community, Kaggle [7]. Participants were given 18 years of historical data including regular season game results, tournament bracket results, and specific team statistics. Participants were tasked to determine probabilities for each possible pair of matchups and were encouraged to use statistical modelling and Machine Learning to do so. The competition continues to run annually now.

Although the brackets submitted via large sports media outlets and competitions such as the "March Machine Learning Madness" constitute a large portion of submitted brackets, there has been novel work when applying algorithms and statistical models for bracket generation and matchup prediction. One popular area within NCAA tournament bracket simulations is 'upset' prediction. An 'upset' is when a higher seed (worse team) surprisingly beats a lower seed (better team). ESPN defines an upset in a similar manner, but the seed of the losing team must be five or lower [6]. In other words, a 11-seed beating a 6-seed, a 10-seed beating a 7-seed, and a 9-seed beating an 8-seed are not considered upsets with ESPN's definition, although these outcomes have higher seeds beating lower seeds. Balance Optimization Subset Selection has been applied to correctly identify upsets with twice as many correct upset predictions as previously used weighted prediction models [3].

3 Description of Data

For our training and testing data, we use the results of all NCAA Division 1 Men's Basketball tournament brackets dating back to 1985, since that is when the NCAA moved to a 64-team single elimination tournament bracket [15]. The NCAA Tournament runs annually which has led to 36 tournaments from 1985 until 2022. There was no tournament in 2020 due to the COVID-19 pandemic. We use 85% of the "true" tournament brackets to derive our proposed metric. We test the effectiveness of our proposed metric on the remaining 15% of brackets which equates to approximately five tournaments. We train on the tournaments from 1985 until 2016, and test on the remaining tournaments from 2017 to 2022.

In some instances, predicting a bracket requires team-specific statistics such as Win Percentage, Offensive Efficiency, and Defensive Efficiency. We obtain the team-specific statistics for all tournament teams from the statistical database 'KenPom' [12]. We use KenPom's data from 2011 to 2022 since any prior data is either biased or non-existent [14]. Therefore, two of our methods use the 36 tournaments from 1985 to 2022 for training, while the remaining method uses 2011 to 2016 for training.

4 Methodology of Chalk Index

We propose a new metric called "Chalk Index" which is derived from two brackets. The first bracket is the "true" bracket of a certain year's tournament, while the second bracket is the hypothesis bracket. When a hypothesis bracket is constructed by choosing the best team of every matchup to win the game, then the hypothesis bracket is referred to as a "Chalk Bracket". While the true bracket changes from year to year, the Chalk Bracket remains the same for every year. The Chalk Index quantifies the variation between a "true" bracket and the Chalk Bracket under two assumptions:

1. Each tournament team has a ranking or seeding.
2. The ranking has an inherent meaning that denotes the strength of the team or some indication of ability to perform well in a tournament. A lower seed is a stronger team than a higher seed (i.e. a 1-seed is better than a 16-seed)

4.1 Calculation of Chalk Index

We define a hypothesis bracket as any possible bracket from the population of all bracket combinations. Given a Chalk Bracket and a hypothesis bracket, we calculate the absolute value of the difference between the seeds of the winners in each matchup, and multiply it by two raised to the power of one less than the round number. The Chalk Index, γ_y , is the sum of these values for all games for a year y . We can represent both the hypothesis bracket and the chalk bracket as matrices as shown below. Brackets can be modeled with matrices of size $(n/2) \times (\log_2 n)$ where n is the number of

teams, since there are a total of $\log_2 n$ rounds until there is one team remaining, and there are $n/2$ teams left after one round. Each column (denoted by j) is a separate round, and each row (denoted by i) is a team in that round. Let α denote the Chalk Bracket and let π denote the hypothesis bracket. The calculation of the Chalk Index is the sum of round-weighted column sums after an element-wise subtraction in the matrix representations of the hypothesis bracket and the Chalk Bracket. More formally, Chalk Index can be defined as:

$$\gamma_y = \sum_{j=1}^{\log_2 n} \left(2^{j-1} * \sum_{i=1}^{n/2} |\alpha_{i,j} - \pi_{i,j}| \right) \iff \exists c \in \mathbb{Z}^+ \text{ s.t. } \log_2 n = c \quad (1)$$

We present a simplified four-team example tournament in Figure 2 and compute the calculation of its Chalk Index below:

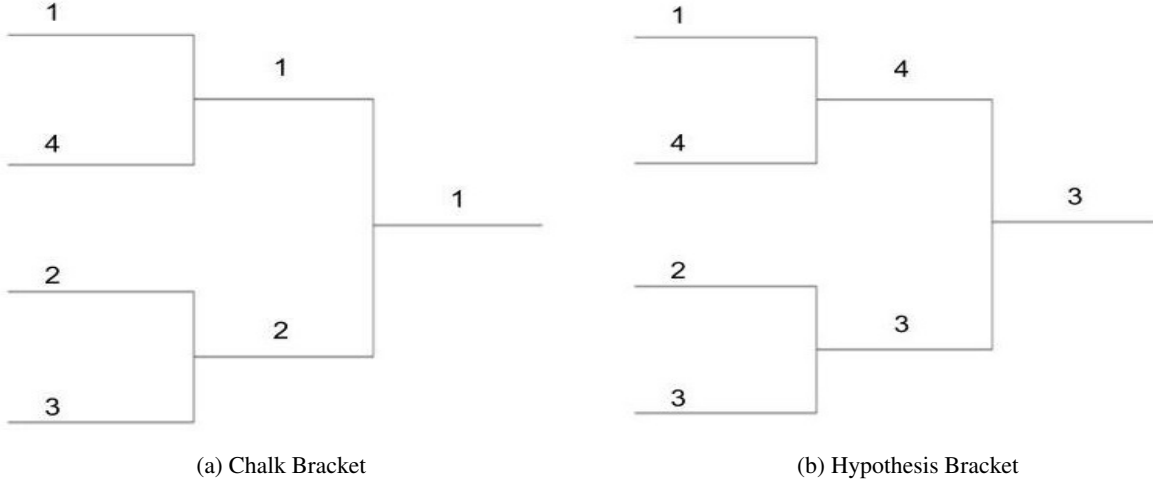


Figure 2: Chalk Bracket and Hypothesis Bracket for a simplified four-team example

$$\alpha = \begin{pmatrix} \text{Chalk Bracket Round 1 Game 1 Winner} & \text{Chalk Bracket Round 2 Game 1 Winner} \\ \text{Chalk Bracket Round 1 Game 2 Winner} & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}$$

$$\pi = \begin{pmatrix} \text{Hypothesis Bracket Round 1 Game 1 Winner} & \text{Hypothesis Bracket Round 2 Game 1 Winner} \\ \text{Hypothesis Bracket Round 1 Game 2 Winner} & 0 \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 3 & 0 \end{pmatrix}$$

$$\begin{aligned} \gamma_y &= \sum_{j=1}^2 \left(2^{j-1} * \sum_{i=1}^2 |\alpha_{i,j} - \pi_{i,j}| \right) \\ &\implies (2^0 * (|1 - 4| + |2 - 3|)) + (2^1 * |1 - 3| + |0 - 0|) \\ &= (1 * (3 + 1) + 2 * (2 + 0)) \\ &= 3 + 1 + 4 + 0 = 8 \end{aligned}$$

Thus, the chalk index for the hypothesis bracket is 8.

4.2 Applying Logistic Regression

We apply Logistic Regression to incorporate team-specific statistics into generating a hypothesis bracket. Logistic Regression is a statistical model that can be used to predict the probability of a certain event to happen. Logistic Regression can be binary such that there is only two possible outcomes of each event, but it can also be expanded to multi-classification cases. Moreover, Logistic Regression is a widely-used statistical model because it can also be expanded to take multiple explanatory variables into account during prediction.

We chose to use Logistic Regression because of its probabilistic nature. In our study, we utilized Binary Logistic Regression where a "1" denotes a win, and a "0" denotes a loss. When fitting our Logistic Regression model, we use game data from only the first round, also known as "Round of 64". For explanatory variables, we use all of the publicly

accessible metrics logged by the statistical database KenPom combined with team-specific information such as the team's conference and whether or not they won their conference. Descriptions of all chosen explanatory variables - along with their trained coefficients - are outlined in Appendix A [13].

Logistic Regression provides coefficients for each of the explanatory variables, then the dot product of the coefficient vector and the given input are passed to the Logistic function outlined in Equation 2.

$$\sigma(t) = \frac{1}{1 + e^{-t}} \quad (2)$$

$$\text{where } t = \beta_0 + \sum_i^m \beta_i x_i.$$

m denotes the number of explanatory variables, and x is the given input vector. β_0 is the bias term which is not multiplied by any explanatory variable.

5 Experiments

5.1 Generation Methods

Sports media outlets like ESPN who host large-scale bracket pools give their users the option to generate a bracket on the user's behalf [8]. ESPN offers users three methods of bracket generation: "Random", "Chalk", and "Smart Bracket". ESPN's "Random" method randomly picks a team to win each matchup with equivalent likelihoods. ESPN's "Chalk" method picks the lower seed in each matchup (the better team) to win. Lastly, ESPN's "Smart Bracket" method uses Bracket Power Index (BPI) simulation when autofilling and generating brackets. BPI is a power ranking system developed by ESPN that uses team-specific statistics to rank the strength of each team.[11] ESPN does not provide an equation, nor the metrics that go are factored in when calculating BPI.

We emulate these generation methods during our experiments. ESPN's "Random" method is trivial to carry over. ESPN's "Smart Bracket" Method is not available to the public; however, we would still like to incorporate team-specific statistics into prediction, so we supply KenPom statistics as explanatory variables in a Logistic Regression model. We train our Logistic Regression model on the 2011 to 2016 tournaments. ESPN's "Chalk" method will not serve much use. It generates the same bracket every iteration because it picks the better team 100% of the time. We introduce a stochastic component such that the lowest seed is not picked 100% of the time, but rather a supplied threshold.

We propose the following generation methods and generate 100 brackets for each method:

1. **Random Generation:** Generate a bracket that picks the winner of each game randomly.
2. **Stochastic Best Team Generation:** Generate a bracket that picks the better team (lower-seed) a certain percentage of the time.
3. **Logistic Regression:** Generate a bracket that weighs each team's statistics to make predictions via Logistic Regression

We observe from Table 1 that the lower seeded teams have a 763-261 record against their higher seeded counterparts in the first round from 1985 to 2016.

Table 1: NCAA First Round Records, 1985 - 2016

Seed Matchup:	Record:
1 vs. 16	128 - 0
2 vs. 15	120 - 8
3 vs. 14	107 - 21
4 vs. 13	102 - 26
5 vs. 12	82 - 46
6 vs. 11	82 - 46
7 vs. 10	78 - 50
8 vs. 9	64 - 64

We use the win percentage of lower seeded teams as the threshold in Stochastic Best Team Generation when testing from 2017 to 2022. Win percentage is calculated as

$$\text{Win Percentage} = \frac{\text{number of wins}}{\text{number of wins} + \text{number of losses}} \quad (3)$$

Via Equation (3), the win percentage of lower seeded teams from 1985 to 2016 is $\frac{763}{1024} \approx 0.745$. So, in our Stochastic Best Team Generation, the lower seed will win approximately 74.5% of the time. In order to test the effectiveness of Chalk Index, we repeat the above methods but with the added constraint that each of the generated brackets must have a Chalk Index that is equal to a prior true bracket's Chalk Index. We contrast the experimental approaches below:

Algorithm 1 Generation Method

```

1:  $k \leftarrow 0$ 
2: for  $k < \text{Number of Brackets to Generate}$  do
3:    $\phi_k = \text{Generate a Bracket using a Generation Method}$ 
4:   yield  $\phi_k$ 
5:    $k \leftarrow k + 1$ 
6: end for

```

Algorithm 2 Generation Method + Chalk Index

```

1:  $\mathbb{S} \leftarrow \text{ChalkIndex}(1985, 1986, \dots, 2016)$ 
2:  $k \leftarrow 0$ 
3: for  $k < \text{Number of Brackets to Generate}$  do
4:    $\rho \leftarrow \text{random}(\mathbb{S})$ 
5:   while True do
6:      $\phi_k = \text{Generate a Bracket using a Generation Method}$ 
7:     if  $\text{ChalkIndex}(\phi_k) == \rho$  then
8:       yield  $\phi_k$ 
9:       break
10:    end if
11:  end while
12:   $k \leftarrow k + 1$ 
13: end for

```

We compare the performance of the brackets that were filled solely with a Generation Method against brackets filled with a Generation Method and sampling from Chalk Index. In all generation methods, we start our selection by picking the winner of final round first, then working backwards to the first round. In hierarchical single-elimination tournaments, the winner of the tournament must have been the winner of all their previous games. So, when the tournament winner is picked, their matchups in previous rounds are propagated backwards as wins. This allows us to streamline the generation of brackets.

5.2 Evaluation Metrics

Simulation results were measured using Bracket Score[9]. As outlined in Table 1, Bracket Score assigns one point for every correct prediction in the first round, two points for correct predictions in the second round, four points for correct predictions in the third round, eight points for correct predictions in the fourth round, 16 points for correct predictions in the fifth round, and 32 points for correct predictions in the final round. This metric takes the relative importance into the game so that predicting the champion correctly is worth more than predicting a less relevant first round game. Additionally, each round has a maximum of 32 possible points. With six total rounds, the maximum Bracket Score is 192.

We report the Mean, Standard Deviation (SD), Maximum (Max), and Minimum (Min) of the Bracket Scores for a given sample of generated brackets $\phi_1, \phi_2, \dots, \phi_n$. Let ω_k be the calculated bracket score for any given bracket ϕ_k .

$$\text{Mean}(\bar{\omega}) = \frac{\sum_{k=1}^n \omega_k}{n} \quad (4)$$

Table 2: "Bracket Score" Metric

Round:	Points for Correct Prediction:	Number of Games in Round:	Total Points Per Round:
1	1	32	32
2	2	16	32
3	4	8	32
4	8	4	32
5	16	2	32
6	32	1	32
		63 total games	192 total points

$$SD = \frac{\sum_{k=1}^n (\bar{\omega} - \omega_k)^2}{n - 1} \quad (5)$$

$$Max(\omega) = \max(\omega_k) \quad (6)$$

$$Min(\omega) = \min(\omega_k) \quad (7)$$

Using the calculated Mean and Standard Deviation, we perform a two sample t-test, also known as an "independent t-test" or an "unpaired t-test", in order to determine if there is statistically significant improvement when using Chalk Index to generate brackets.'

$$t = \frac{\bar{\omega}_1 - \bar{\omega}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (8)$$

where:

- $\bar{\omega}_1$ = Mean Bracket Score of Group 1
- $\bar{\omega}_2$ = Mean Bracket Score of Group 2
- s_1 = Standard Deviation Bracket Score of Group 1
- s_2 = Standard Deviation Bracket Score of Group 2
- n_1 = Size of Group 1
- n_2 = Size of Group 2

We use a two sample t-test because each of our samples are generated from the 2^{63} possible bracket combinations which represents our population.

6 Results

We compare the performance of the randomly generated brackets against the brackets with the Chalk Index constraint. All results are rounded to four significant figures for reporting purposes only.

Table 3: Bracket Scores of (Random Generation) vs. (Random Generation + Chalk Index)

Year	Random Generation: n = 100				Random Generation + Chalk Index: n = 100				
	Mean	SD	Min	Max	Mean	SD	Min	Max	p-value
2017	34.02	15.30	11	104	72.47	29.02	24	149	1.930e-24
2018	32.60	13.24	11	90	57.03	21.61	20	121	2.820e-18
2019	31.42	14.09	11	97	63.03	22.63	28	134	7.028e-26
2021	30.28	11.90	9	93	62.81	25.15	27	136	2.336e-24
2022	31.62	12.49	13	86	61.41	26.88	23	119	1.816e-19

Table 4: Bracket Scores of (Stochastic Best Team Generation) vs. (Stochastic Best Team Generation + Chalk Index)

Year	Stochastic Best Team: $n = 100$				Stochastic Best Team + Chalk Index: $n = 100$				
	Mean	SD	Min	Max	Mean	SD	Min	Max	p-value
2017	68.65	11.69	43	121	64.37	10.08	45	117	0.00609
2018	67.35	14.71	31	109	61.93	14.06	28	98	0.00836
2019	79.01	13.94	47	126	74.83	12.83	41	115	0.0285
2021	87.43	14.93	46	128	86.47	13.11	40	125	0.629
2022	54.29	9.087	36	99	51.13	8.989	29	99	0.0143

Table 5: Bracket Scores of (Logistic Regression) vs. (Logistic Regression + Chalk Index)

Year	Logistic Regression: $n = 100$				Logistic Regression + Chalk Index: $n = 100$				
	Mean	SD	Min	Max	Mean	SD	Min	Max	p-value
2017	50.86	17.93	25	126	66.82	23.22	35	135	1.561e-07
2018	44.14	15.703	19	121	60.91	22.509	21	129	5.185e-09
2019	48.34	15.81	24	105	68.39	26.39	30	138	5.808e-10
2021	44.08	20.22	20	116	64.06	25.38	28	128	4.059e-09
2022	46.47	16.75	20	111	61.46	25.88	25	127	2.357e-06

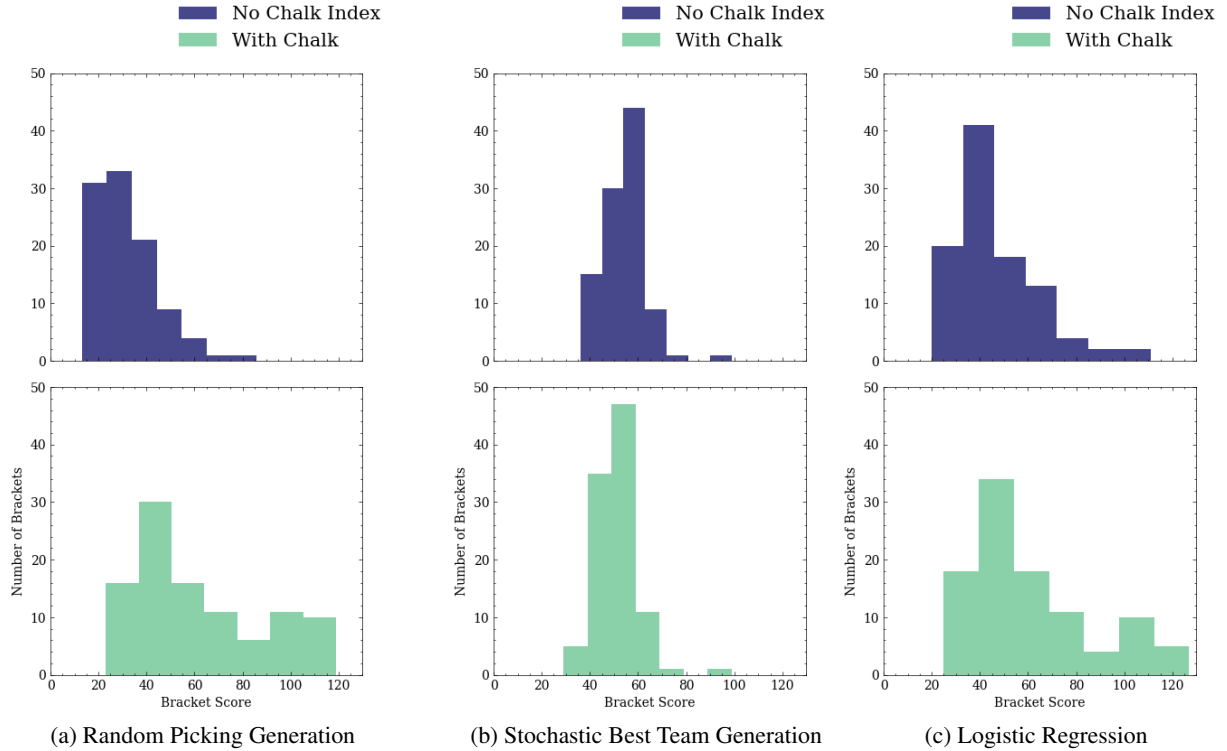


Figure 3: Effect of Chalk Index of 2022 NCAA Bracket Generation

We observe in Table 3 and Table 5 that adding the constraint of sampling from historical Chalk Indexes greatly increases the mean Bracket Scores in both Random Generation and Logistic Regression methods at the small expense of a slight decrease in mean Bracket Scores in Stochastic Best Team Generation. The Bracket Scores in Stochastic Best Team Generation in 2022, as displayed in Figure 3B, seem to follow a similar distribution in both the No Chalk Index and Chalk Index cases.

7 Future Extensions

Chalk Index is a metric for measuring variability between "true" brackets and hypothesis bracket. A lower Chalk Index means lower variability and thus easier predictability when using rankings as the sole indicator of tournament performance. We have started future work that compares the Chalk Indexes of NCAA Tournament Brackets with NBA Playoffs Brackets - all of which are supplied in Appendix B. As opposed to 64-team NCAA Tournament Brackets, NBA Playoff Brackets have 16 teams. Since the range of possible Chalk Indexes is higher for larger tournaments, we can normalize them by dividing by the maximum Chalk Index for each tournament and then compare them to observe which tournament most closely follow its chalk counterpart. As visualized in Figure 4, NBA brackets tend to have more of a chalk nature than NCAA brackets.

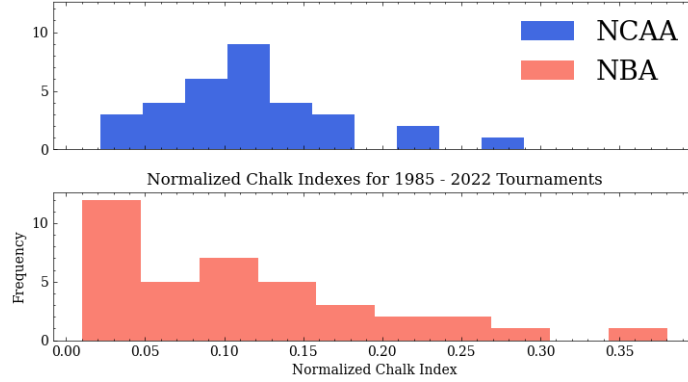


Figure 4: Normalized Chalk Indexes of NCAA Tournament Brackets vs. NBA Playoff Brackets

Future extensions could investigate the variation between Chalk Index distributions in NCAA Tournament Brackets and NBA Playoffs Brackets. Potential explanations could be that NCAA team rankings are decided by a committee [10], while NBA team ranking are determined by their season-long win percentage. This could give insight into the reliability of human rankings against empirical rankings.

In this paper, we combined Chalk Index with the Logistic Regression Model. An additional expansion would be to assess how probabilistic Machine Learning models perform with the addition of Chalk Index.

8 Conclusion

In this paper, we have proposed Chalk Index, a metric that quantifies the variation between a Chalk Bracket and a hypothesis bracket. The Chalk Index is derived for every tournament in the training data and is applied to generate a hypothesis bracket for unseen tournament data. We have shown that combining the Chalk Index with Stochastic Best Team Generation and Logistic Regression significantly improves the mean Bracket Score for our test data.

References

- [1] Dan Carson. *Warren Buffett's 'Billion-Dollar Bracket Challenge' Expanding Entry Number*. Bleacher Report, Mar. 2014. URL: <https://bleacherreport.com/articles/1991405-warren-buffetts-billion-dollar-bracket-challenge-expanding-entree-number>.
- [2] *CBS Sports 2021 NCAA Bracket*. cbssports.com, 2021. URL: <https://htv-prod-media.s3.amazonaws.com/files/cbsw1kyncaa-1616122788.jpg>.
- [3] Shouvik Dutta, Sheldon H. Jacobson, and Jason J. Sauppe. "Identifying NCAA tournament upsets using Balance Optimization Subset Selection". In: *Journal of Quantitative Analysis in Sports* 13 (Jan. 2017). DOI: 10.1515/jqas-2016-0062.
- [4] Jason Gershman. *"The Mathematics of the Billion Dollar Bracket"*. 2014. URL: <https://rusmp.rice.edu/sites/g/files/bxs3761/files/publications/the%20mathematics%20of%20a%20billion%20dollar%20bucket.pdf>.

- [5] Christina Gough. *March Madness NCAA final TV viewership 2021*. Statista, Apr. 2022. URL: <https://www.statista.com/statistics/244249/ncaa-basketball-march-madness-average-audience-per-game/>.
- [6] Peter Keating. *Giant Killing 2.0: An updated guide*. ESPN.com, Mar. 2013. URL: https://www.espn.com/mens-college-basketball/story/_/id/9022008.
- [7] *March Machine Learning Mania*. kaggle.com, 2014. URL: <https://www.kaggle.com/c/march-machine-learning-mania-2014>.
- [8] AJ Mass. *How to fill out a tournament bracket: All the basics so you can join the madness*. ESPN.com, Feb. 2022. URL: https://www.espn.com/fantasy/basketball/story/_/id/26103343/how-fill-tournament-bracket-all-basics-join-madness.
- [9] Nicole R. Matthews et al. “Application of PageRank Algorithm to Division I NCAA men’s basketball as bracket formation and outcome predictive utility”. In: *Journal of Sports Analytics* 7 (Apr. 2021), pp. 1–9. DOI: 10.3233/jsa-200425.
- [10] *NCAA Tournament Selection Committee*. Hoops HD, Jan. 2015. URL: <https://hoopshd.com/ncaa-selection-committee/>.
- [11] Dean Oliver. *Introducing the BPI*. ESPN.com, Feb. 2012. URL: https://www.espn.com/mens-college-basketball/story/_/id/7561413/bpi-college-basketball-power-index-explained.
- [12] Kenneth Pomeroy. *Pomeroy College Basketball Ratings*. kenpom.com. URL: <http://www.kenpom.com>.
- [13] Christopher Toukmaji. *Appendix*. GitHub, Sept. 2022. URL: https://github.com/christoukmaji/CMSAC_Submission/blob/main/OtherDocuments/Appendix.pdf.
- [14] Christopher Toukmaji. *Submission Data*. GitHub, Sept. 2022. URL: https://github.com/christoukmaji/CMSAC_Submission/tree/main/Data.
- [15] Daniel Wilco. *March Madness history - The ultimate guide* | NCAA.com. www.ncaa.com, Jan. 2022. URL: <https://www.ncaa.com/news/basketball-men/article/2021-03-14/march-madness-history-ultimate-guide>.