

## 데이터마이닝 의사결정나무분석을 활용한 세계여자수구선수권대회 승패결정요인 탐색

### Exploring the Determinants of Winner and Defeat of the Women's Water Polo World Championship using Data Mining Decision Tree Analysis

이승훈(한국체육대학교 박사과정) · 박재현(한국체육대학교 교수) · 조은혜\*(한국교원대학교 연구원)

Seunghun Lee · Jae-Hyeon Park *Korea National Sport University* · Eunhye Jo\* *Korea National University of Education*

#### 요약

이 연구의 목적은 세계여자수구선수권대회의 경기기록을 활용하여 승패결정요인을 탐색하는 것이다. 이 연구의 목적을 위해 제18회 FINA 세계여자수구선수권대회의 전체 48경기 기초통계 공식자료를 수집하여 선행연구 및 공식자료를 기반으로 Shooting, Defence, GB(골키퍼 방어), Fouls, Ets 관련 29개 변수를 독립변인으로, 종속변인은 경기승패로 선정하였다. 자료처리분석은 Mann-Whitney U test 및 의사결정나무분석을 수행하였다. 이 연구의 결과는 다음과 같다. 첫째, 승패 집단에 따른 경기변인별 차이를 분석한 결과, 6개의 Shooting 변인, 2개의 Defence 변인, 3개의 GB(골키퍼 방어) 변인, 1개의 Fouls 변인, 2개의 Ets 변인이 통계적으로 유의한 차이를 나타냈다. 둘째, 세계여자수구선수권대회의 승패결정요인으로 GB\_Action goals, Action goals(Shooting), Counterattack goals(Shooting), GB\_Extra player goals로 나타났으며, 4개변인 중 승패를 결정짓는 가장 중요한 변인은 골키퍼의 Action goals 방어율(GB\_Action goals)로 확인되었다. 이 연구의 결과는 세계여자수구경기를 이해하고 분석적 시각에서 접근한 연구로서 스포츠현장에서는 경기력 향상을 위한 유용한 정보로 활용될 수 있을 것이며, 향후 여자수구경기에 대한 연구의 기초정보로서 활용되기를 기대해본다.

#### Abstract

The objectives of this study were to compare water polo game-related statistics by match outcome (winning and losing teams) and to develop a classification tree model explaining the performance of world women's water polo. Forty-eight matches that were played in the 18th FINA World Championships were analyzed. The dependent variable was match outcome and the independent variables were the game-related statistics. To determine the differences between the winning and losing teams, a non-parametric (Mann-Whitney U-test) test was applied. also in order to determine which variables best predict the final outcome, a decision tree was constructed. This was a tree based on the supervised learning method called CHAID. four variables differentiated between winning and losing teams: four were related to the effectiveness of GB\_Action goals, Action goals, GB\_Extra player goals, and Counterattack goals. If the GB\_Action goals are more than 81.81%, the probability of being classified as a winner is 100%. However, GB\_Action goals are less than 46.15%, and the probability of being classified as a defeat is 100%. The decision tree accuracy was classified as 93.8%.

Key words : Women's water polo, Decision tree analysis, Data Mining, Winning and losing determinants

\* eunhyejo3@gmail.com(조은혜)

## I. 서론

오늘날 스포츠 분석학(Sport analytics)분야가 많은 관심을 받는 이유는 스포츠 언론 및 팬들을 위한 흥미정보부터 스포츠 현장에 있는 선수들의 경기력 향상을 위한 기초정보까지 제공하고 있기 때문이다. 예컨대, 스포츠 중계방송을 시청하다보면 선수의 경기전적과 같은 기본정보에서부터 승리확률, 공격성공율, 골키퍼 방어율 등 전문적인 정보에 이르기까지 스포츠판들의 흥미를 위한 정보들이 실시간으로 생성되어 제공되는 것을 확인할 수 있다(김연자, 2011; 나경민, 박재현, 조은혜, 2020; 신진호, 최창환, 2018).

수많은 정보 중에서도 스포츠선수들과 대중이 가장 관심을 갖는 정보는 경기의 승패를 결정짓는 요인일 것이다. 전통적으로 승패결정요인으로 제시되고 있는 요인은 심리, 생체역학, 생리 등 다양한 내외적 요인들이 제시되고 있다. 하지만 대부분의 변인이 매우 역동적인 특성(김혜진, 박재현, 강상조, 2006)을 가지고 있기 때문에 이를 예측하는 것은 쉽지 않은 것이 사실이다. 그럼에도 불구하고 많은 스포츠 전문가들이 승패를 결정짓는 요인을 탐색하기 위한 연구를 지속적으로 수행하는 이유는 첫째, 경기전술, 훈련방식, 선수기용 등 경기력향상을 위한 중요정보로 활용(조은혜, 강용구, 김혜진, 2021)되고 있으며, 둘째, 스포츠판 및 언론의 흥미를 높이기 위한 목적으로 활용되고 있기 때문이다. 마지막으로 연간 평균 약 14조의 매출규모를 나타내는 스포츠배팅 산업시장에서는 금전과 연결되는 구조를 띄고 있어 산업적 가치정보로서도 활용(최창환, 윤지운, 2017; Miller, 2015)되고 있기 때문이다.

한편, 2019년 FINA 세계수영선수권대회 개최국 자격으로 출전권을 얻은 대한민국은 사상 처음으로 여자수구 국가대표팀을 구성하였다. 우리나라 여자수구 국가대표팀은 출전 두 달여를 남기고 대한수영연맹(Korea Swimming federation) 자체선발을 통해 결성되었으나, 헝가리와의 첫 경기에서 0-64로 대패하는 등 세계 강호들을 상대로 5전 전패를 기록, 16개국 중 최하위에 머무르며, 폐막 이후 해체되었다. 이후 대한수영연맹에서는 2023년 열릴 제19회 항저우 아시아경기대회 참가를 목표로 다시 한 번 2022년 연맹 자체 여자수구 국가대표를 선발하여 훈련을 진행하고 있다(대한수영연맹, 2021.9.17. 게시). 우리나라 여자수구 국가대표팀이 국제 대회에 출전해 아시아 강호들과의 경쟁에서 비등한 경기를 이끌어 나가기 위해서는 여자수구의 국제적 흐름을 이해하고, 이를 훈련에 적용해야 한다.

그동안 수구 경기분석과 관련된 선행연구를 살펴보면, 경기기술 및 경기상황을 분석하여 승리 팀과 패자 팀의 차이를 확인한 연구(Escalante et al., 2011; Lupo et al., 2011; Takagi et al., 2005), 최종점수 차이에 따른 경기기술 및 경기상황의 차이를 확인한 연구(Garcia-Marin et al., 2017; Saavedra et al., 2016), 승자 팀과 패자 팀 또는 선수포지션에 따른 슛의 속도 및 정확도의 차이를 확인한 연구(Özkol et al., 2013; Vila et al., 2011), 경기규칙 변경에 따른 경기기술의 차이를 확인한 연구(이승훈, 박재현, 윤효준, 2020; Argudo et al., 2021; Canossa et al., 2022; Saavedra et al., 2020) 등이 보고되었다. 이러한 선행연구 결과는 수구경기의 경기력향상

및 학문적 발전에 기여하였다는 사실은 부정할 수 없다. 그러나 선행연구들의 분석내용을 자세히 살펴보면, 규칙개정 이후 연구들은 모두 남자수구경기를 대상으로 진행한 연구(이승훈 외 2020; Argudo et al., 2021; Canossa et al., 2022; Saavedra et al., 2020)이며, 여자수구경기를 대상으로 한 연구는 대다수가 규칙개정 이전(2019년 FINA 세계수영선수권대회 이전)의 데이터를 기반으로 분석한 정보로서 한계가 있다. 제19회 항저우 아시아경기대회 현장에서 적용할 수 있는 정보를 제공하기 위해서는 경기규칙 개정 이후 수행된 경기기록을 기반으로 승패를 결정짓는 요인을 탐색하는 연구가 수행되어야 할 것이다.

스포츠 경기분석 분야에서 승패를 결정짓는 요인을 탐색하기 위한 목적으로 적용되어진 통계적 기법들을 살펴보면 크게, 판별분석(discriminant), 로지스틱 회귀분석(logistic regression), 인공신경망분석(artificial neural network), 의사결정나무분석 등이 있다. 이중 의사결정나무분석은 의사결정 규칙(decision tree)을 도표화하여 최적의 관련성을 갖는 집단을 소집단으로 예측(prediction)하거나 분류(Classification)를 수행(최종후, 서두성, 1999)하는 통계적 방법이다. 특히 결과가 나무구조 모형으로 제시되기 때문에 해석이 용이하며, 비모수적 모형으로 정규성(normality), 등분산성(equal variance), 선형성(linearity) 등의 가정이 요구되지 않고, 여러 예측변인들을 바탕으로 종속변인에 작용하는 상호작용효과를 탐색할 수 있다는 장점을 가지고 있다. 이러한 장점으로 인해 의사결정나무분석 기법은 다양한 스포츠종목에서 승패를 결정짓는 요인을 탐색하는 방법으로 활용되어지고 있다(김민창 등, 2021; 김혜진 등, 2006; 최창환 등, 2018). 따라서 이 연구는 2019년 FINA 세계여자수구선수권대회의 경기기록을 활용하여 승패결정요인을 탐색하는 것이 목적이다. 구체적으로 세계여자수구선수권대회에서 승패 집단에 따른 경기변인별 차이를 확인하고 데이터 마이닝 기법인 의사결정나무분석을 적용하여 승패결정요인을 탐색하고자 하였다.

## II. 연구방법

### 1. 연구대상자료

이 연구는 세계여자수구선수권대회의 승패를 결정하는 요인이 무엇인지 탐색하는 것이 목적이다. 연구의 목적을 달성하기 위하여 2019년 대한민국 광주에서 열린 제 18회 FINA 세계여자수구선수권대회의 전체 경기기록을 연구대상 자료로 선정하였다. 이 자료는 OMeGa Timing의 공식웹사이트(<http://www.omegatiming.com>)에서 제공하는 기초통계 공식자료이다. 제18회 세계여자수구선수권대회에 참가한 국가는 16개 국가이며, 전체 48경기 결과(승자 48팀, 패자 48팀)의 경기기록을 수집하였다. 연구대상자료의 구체적인 정보는 다음의 <표 1>과 같다.

표 1. 연구대상자료

순위	국가	경기	승	패
1	USA (United States)	6	6	0
2	ESP (Spain)	6	5	1
3	AUS (Australia)	7	5	2
4	HUN (Hungary)	7	4	3
5	RUS (Russia)	6	5	1
6	ITA (Italy)	6	4	2
7	NED (Netherlands)	7	4	3
8	GRE (Greece)	7	3	4
9	CAN (Canada)	6	3	3
10	KAZ (Kazakhstan)	6	2	4
11	CHN (China)	6	2	4
12	NZL (New Zealand)	6	1	5
13	JPN (Japan)	5	2	3
14	RSA (South Africa)	5	1	4
15	CUB (Cuba)	5	1	4
16	KOR (Korea)	5	0	5

## 2. 연구변인

이 연구의 목적을 위하여 선정한 연구변인은 수구경기분석 관련 선행연구와 국제수영연맹의 공식자료를 기반으로 구성하였다. Saavedra 등(2016)의 연구에서는 OMega Timing에서 제공하는 공식 통계자료를 기반으로 슈팅기술별 성공비율, 골키퍼 슈팅기술 상황별 방어비율, 파울, 스틸, 블록샷 등의 변인을 선정하였다. 이에 본 연구에서는 선행연구에서 활용한 경기변인과 공식통계자료에서 제공하는 변인으로 최종변인을 선정하였다. 연구변인을 선정하는 과정에는 10년 이상 수구선수 경력을 보유한 전문가 1인과 체육측정평가 전문가 2인 총 3명의 전문가가 참여하였으며, 의견조율을 통해 총 29개의 최종변인을 구성하였다. 구체적으로 Shooting과 관련된 10개변인, Defence와 관련된 3개변인, GB(골키퍼 방어)와 관련된 10개변인, Fouls와 관련된 2개변인, Ets(기타) 4개 변인으로 구성하였으며, 각 변인에 대한 구체적인 설명은 <표 2>와 같다.

표 2. 연구변인

구분	변인명	변인설명
Shooting	Total	전체 골
	A%	Action goals 득점비율
	C%	Centre goals 득점비율
	D%	Driving goals 득점비율
	X%	Extra player goals 득점비율
	6m%	6m Direct goals 득점비율
	6mF%	6m free throw 득점비율
	PS%	Penalty goals 득점비율
	CA%	Counter Attack goals 득점비율
	CO%	Comer goals 득점비율

Defence	ST	Steals 성공 횟수
	RB	Rebounds 성공 횟수
	BL	Blocked shots 성공 횟수
	Total	전체 골 방어
	A%	Action goals 방어비율
	C%	Centre goals 방어비율
	D%	Driving goals 방어비율
	X%	Extra player goals 방어비율
	6m%	6m goals 방어비율
	6mF%	6m free throw 방어비율
Fouls	PS%	Penalty goals 방어비율
	CA%	Counter Attack goals 방어비율
	CO%	Comer goals 방어비율
	TF	Turnover Fouls
Ets	TPF	Total Personal Fouls
	SP	Sprint (Won/Sprints)
	TO	Time outs
	ES	Exclusions with substitution
	ES4	Exclusions with substitution after 4min

## 3. 자료처리방법

본 연구의 자료처리방법은 다음과 같다. 첫째, 승패 집단에 따른 경기변인별 기술통계와 차이검증을 수행하였다. 차이검증은 비모수(non-parametric)통계기법인 Mann-whitney U test를 실시하였다. Mann-whitney U test를 수행한 이유로는 연구에서 적용된 변인의 속성이 비율형태이며, 일부 변인들이 모수통계의 가정을 충족하지 못하였기 때문임을 밝힌다. 둘째, 승패결정요인을 탐색하기 위하여 의사결정나무분석기법을 적용하였다. 의사결정나무분석 알고리즘은 모형에 포함되어야 할 상호작용 효과를 자동(김세형, 2012)으로 찾아내 분석할 수 있는 CHAID방법을 적용하였다. 의사결정나무분석을 위해 투입한 독립변인은 <표 2>에서 제시한 29개 변인 중 Shooting의 Total(전체 골)과 GB의 Total(골키퍼 전체 골 방어)을 제외한 27개 변인을 투입하였으며, 종속변인은 경기 승패를 투입하였다. 이때 독립변인에서 Shooting의 Total(전체 골)과 GB의 Total(골키퍼 전체 골 방어)은 경기의 승패를 결정짓는 변인으로 알려진 변인이기에 해당 변인을 투입하는 것은 적절하지 않음을 판단하였다. 이에 위 두 변인을 제외한 후 승패 결정요인을 탐색하였음을 밝힌다. 의사결정나무 모형은 다음과 같이 설정하였다. 최대나무깊이(maximum tree depth)는 3단계, 최소사레 수는 부모마디(parent node) 5, 자식마디(child node)는 3의 기준을 적용하였으며, 예측모형의 적절성을 확인하기 위하여 분류정확도(classification accuracy), 오분류 추정치(risk estimate), 표준오차(standard error)를 산출하였다. 모든 통계적 유의수준은 .05로 설정하였고 IBM SPSS(ver 25.0)을 사용하였다.

### III. 연구결과

#### 1. 승패집단에 따른 경기변인별 차이검증

다음 <표 3>은 세계여자수구선수권대회의 승패 집단에 따른 경기변인별 차이를 분석한 결과이다. 결과에 의하면 Shooting 변인 중 6m%(6m goals)를 제외한 Total(Total goals:  $U=198.5$ ,  $z=-6.987$ ,  $p<.001$ ), A%(Action goals:  $U=272.5$ ,  $z=-6.448$ ,  $p<.001$ ), C%(Centre goals:  $U=674$ ,  $z=-3.582$ ,  $p<.001$ ), X%(Extra player goals:  $U=567.5$ ,  $z=-4.291$ ,  $p<.001$ ), PS%(Penalty goals:  $U=884.5$ ,  $z=-2.166$ ,  $p=.030$ ), CA%(Counterattack goals:  $U=649.5$ ,  $z=-3.959$ ,  $p<.001$ )에서 통계적으로 유의한 차이를 나타냈다. 평균을 비교하면 승리 팀의 Shooting 변인이 패배 팀에 비해 높은 것으로 나타났다. Defence 변인은 RB(Rebounds)를 제외한 ST(Steals:  $U=535$ ,  $z=-4.541$ ,  $p<.001$ ), BL(Blocked shots:  $U=859$ ,  $z=-2.175$ ,  $p=.029$ )에서 통계적으로 유의한 차이를 나타냈다. 평균을 비교하면 승리 팀의 Defence 성공 빈도가 패배 팀에 비해 많은 것으로 확인되었다. GB(골키퍼) 변인은 Total(Total goals:  $U=181.5$ ,  $z=-7.112$ ,  $p<.001$ ), A%(Action goals:  $U=179$ ,  $z=-7.137$ ,  $p<.001$ ), X%(Extra player goals:  $U=682.5$ ,  $z=-3.470$ ,  $p<.001$ ) 변인에서 통계적으로 유의한 차이를 나타냈으며, 승리 팀의 GB(골키퍼) 방어 성공률이 패배 팀에 비해 높은 것으로 나타났다. Fouls 변인은 TF(Turnover fouls:  $U=533.5$ ,  $z=-4.545$ ,  $p<.001$ )에서 통계적으로 유의한 차이를 나타냈다. Ets 변인은 SP(Sprints won:  $U=533.5$ ,  $z=-3.034$ ,  $p<.001$ ), TO(Time outs:  $U=712$ ,

$z=-3.473$ ,  $p<.001$ ) 변인에서 통계적으로 유의한 차이를 나타냈으며, SP(Sprints won)은 승리 팀이 패배 팀보다 평균빈도가 높았으며, TO(Time outs)는 승리 팀이 패배 팀에 비해 평균빈도가 적은 것으로 나타났다. 연구변인에서 설정한 Shooting 변인 중 Driving goals, 6mF(6m Free throw), Corner goals, GB 변인 중 Driving goals, 6mF(6m Free throw), Corner goals, Ets 변인 중 Exclusions with substitution after 4min 변인은 빈도수가 적어 p값이 산출되지 않아 결과에 제시하지 않았다.

#### 2. 의사결정나무분석을 적용한 승패결정요인 탐색

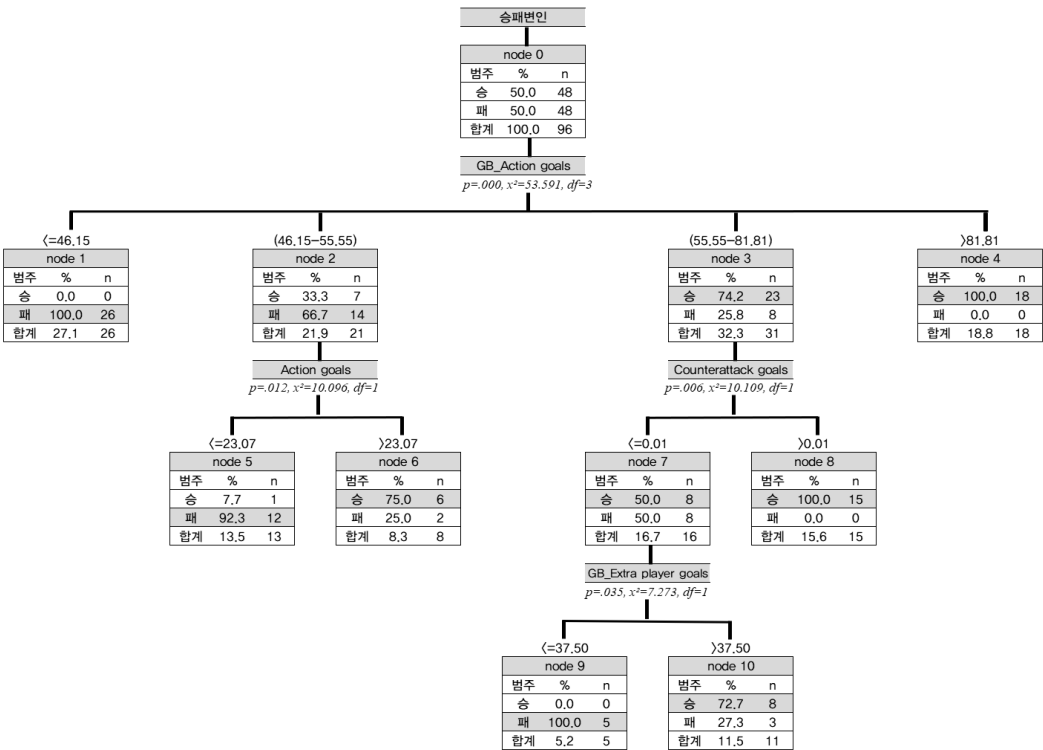
의사결정나무분석 결과는 종속변인(승패)이 가장 위에 위치하며 종속변인과 가장 관련성이 높은 변인들이 가치를 생성하며 계층적으로 위치한다. <그림 1>은 세계여자수구선수권대회의 경기기록을 기반으로 데이터마이닝 의사결정나무분석을 수행한 결과이다. 결과에 따르면 승패를 결정짓는 가장 중요한 변인은 GB\_Action goals( $\chi^2=53.591$ ,  $p<.001$ )로 나타났다. GB\_Action goals(골키퍼 Action goals)방어 성공률이 46.15%이거나 이하일 경우 경기에서 100% 패하는 것으로 나타났으며, 81.81% 이상일 경우에는 100% 승리하는 것으로 분류되었다. GB\_Action goals(골키퍼 Action goals)방어 성공률이 46.15%에서 55.55%일 경우 패할 가능성이 66.7%이지만, Action goals(Shooting) 성공률이 23.07% 이상일 때는 승리할 가능성이 75%로 증가하며, 23.07%이거나 이하일 경우에는 92.3% 패하는 것으로 분류되었다. GB\_Action goals(골키퍼 Action goals)방어 성공률이 55.55%에서 81.81%이고, Counterattack goals(Shooting) 성

표 3. 승패집단에 따른 경기변인별 차이검증 결과

구분	변인	승리집단		패배집단		Mann-Whitney		
		Mean	SD	Mean	SD	U	z	p-value
Shooting	Total	47.15	15.11	22.3	12.19	198.5	-6.987	<.001
	A%	35.92	16.74	14.26	10.28	272.5	-6.448	<.001
	C%	47.31	28.94	26.35	36.33	674.0	-3.582	<.001
	X%	53.80	23.85	32.63	23.5	567.5	-4.291	<.001
	6m%	4.16	17.36	8.33	23.81	1080.0	-1.044	0.296
	PS%	64.06	45.42	42.18	47.86	884.5	-2.166	0.030
	CA%	56.43	40.67	21.31	36.72	649.5	-3.959	<.001
Defence	ST	9.33	7.93	4.89	2.45	535.0	-4.541	<.001
	RB	4.47	2.55	3.91	2.87	1012.5	-1.030	0.303
	BL	3.10	1.71	2.37	2.02	859.0	-2.175	0.029
	Total	65.35	16.16	34.4	15.25	181.5	-7.112	<.001
GB (골키퍼)	A%	75.84	15.50	43.01	18.43	179.0	-7.137	<.001
	C%	46.42	44.24	35.14	29.71	998.0	-1.171	0.241
	X%	42.64	27.32	24.79	22.94	682.5	-3.470	<.001
	6m%	15.62	35.96	7.29	25.24	1054.0	-1.252	0.21
	PS%	7.81	25.34	5.86	18.03	1134.5	-0.232	0.816
	CA%	15.62	32.87	12.53	21.98	1046.5	-0.976	0.329
Fouls	TF	10.68	3.73	17.04	9.16	533.5	-4.545	<.001
	TPF	9.27	4.58	9.95	3.24	1099.5	-0.387	0.698
	SP	2.43	1.36	1.54	1.38	746.5	-3.034	0.002
Ets	TO	0.95	0.77	1.50	0.65	712.0	-3.473	<.001
	ES	0.71	0.84	0.77	0.77	1081.0	-0.564	0.572

공률이 0.01%이거나 이하이며, GB\_Extra player goals(골키퍼 Extra player goals)방어 성공률이 37.50% 이상일 경우 승리할 가능성이 72.7%이며, 37.50%이거나 이하일 경우에는 패할 가능성이 100%로 나타났다. GB\_Action goals(골키퍼 Action goals) 방어 성공률이 55.55%에서 81.81%이고, Counterattack goals(Shooting) 성공률이 0.01%이상일 경우 승리할 가능성이 100%인 것으로 나타났다. <표 4>는 의사결정나무분석 결과를 node의 변인별로 요약한 것이다. 승리 팀을 설명하는 4번 노드(GB\_Action goals 81.81% 이상)는 전체 96경기 중 18경기(18.8%)가 해당되며 18경기 모두(100%)가 승리하는 것으로 나타났다. 8번 노드(GB\_Action goals 55.55%-81.81%이며, Counterattack goals 0.01% 이상)는 전체 96경기 중 15경기(15.6%)가 해당되며 15경기 모두(100%)가 승리하는 것으로 나타났다. 패배를 설명하는 1번 노드(GB\_Action goals 46.15%이거나 이하)는 전체 96경기 중 26경기(27.1%)가 해당되며 26경기 모두(100%)가

패배하는 것으로 나타났다. 9번 노드(GB\_Action goals 55.55%-81.81%이며, Counterattack goals 0.01%이거나 이하, GB\_Extra player goals와 37.50%이거나 이하)는 전체 96경기 중 5경기(5.2%)가 해당되며 5경기 모두(100%)가 패배하는 것으로 나타났다. 5번 노드(GB\_Action goals 46.15%-55.55%이며, Action goals 23.07%이거나 이하)는 전체 96경기 중 13경기(13.5%)가 해당되며 12경기(12.7%)가 패배하는 것으로 나타났다. 이를 통해 세계여자수구선수권대회에서 골키퍼의 GB\_Action goals 방어 성공률이 승패를 결정하는 가장 주요변인이라는 것을 확인할 수 있다. 세계여자수구선수권대회의 승패결정요인에 대한 의사결정나무분석의 분류 정확도는 승리를 예측할 확률은 97.9%, 패배를 예측할 확률은 89.6%, 전체 예측 정확도는 93.8%, 오분류 확률은 정확도의 배반 값인 6.2%로 나타났다.



분류정확도: 93.8%, 오분류율(risk estimate): 6.2%, 표준오차: .025

그림 1. 세계여자수구선수권대회의 승패결정요인에 대한 의사결정나무분석결과

표 4. 세계여자수구선수권대회의 승패결정요인에 대한 의사결정나무분석결과 요약

Nodes	Node		Gain		Resp(%)	GB_Action goals	Action goals	Counterattack goals	GB_Extra player goals
	n	%	n	%					
승리	4	18	18,8	18,8	100,0	>81,81			
	6	8	8,3	6,2	75,0	46,15-55,55	>23,07		
	8	15	15,6	15,6	100,0	55,55-81,81		>0,01	
	10	11	11,5	8,3	72,7	55,55-81,81		<=0,01	>37,50
패배	1	26	27,1	27,1	100,0	<=46,15			
	5	13	13,5	12,7	92,3	46,15-55,55	<=23,07		
	9	5	5,2	5,2	100,0	55,55-81,81		<=0,01	<=37,50

#### IV. 논의 및 결론

이 연구에서는 2019년 FINA 세계여자수구선수권대회의 경기기록을 활용하여 승패결정요인을 탐색하고자 하였다. 구체적으로 2019년 FINA 세계여자수구선수권대회에서 승패 집단에 따른 경기변인별 차이를 확인하고 데이터 마이닝 기법인 의사결정나무분석을 적용하여 승패결정요인을 탐색하고자 하였다. 이 연구의 결과를 바탕으로 세계여자수구경기의 승패를 결정하는 요인에 대해 논하면 다음과 같다.

그동안 수구종목을 대상으로 승패결정요인을 탐색한 선행연구를 살펴보면 (Escalante, Y. et al, 2011; Escalante, Y. et al, 2012; Saavedra, M. et al, 2016) 승리집단과 패배집단을 결정하는 요인으로 전체 슛 득점율(Total goals)과 골키퍼의 전체 골 방어율(GB\_Total goals)을 제시하였다. 반면 수구를 제외한 핸드볼, 농구 등 구기 종목의 공식데이터를 기반으로 승패결정요인을 탐색한 선행연구(김민창 등, 2021; 김세형 등, 2008; 김세형, 2012; 김세형, 2012; 김지용, 이용국, 이승훈, 2022; 신진호, 최창환, 2018; 조은혜 등, 2021)에서는 전체 슛 득점율(Total goals)과 골키퍼의 전체 골 방어율(GB\_Total goals)을 제외한 것을 확인할 수 있었다. 일반적으로 전체 득점율과 전체 방어율을 나타내는 변인은 스포츠 경기에서 승패를 결정하는 변인으로 사실화 되어있기 때문에 보다 면밀하게 승패를 결정하는 요인을 탐색하기 위해서는 전체득점율과 전체 방어율 변인을 투입하는 것은 적절하지 않다고 연구진은 판단하였다. 이에 전체 슛 득점율(Total goals)과 골키퍼의 전체 골 방어율(GB\_Total goals)을 제외한 상태에서 데이터 마이닝 의사결정나무 분석을 활용한 승패결정요인을 탐색하였다.

의사결정나무분석 결과에 따르면 GB\_Action goals(골키퍼 방어)는 세계여자수구선수권대회에서 승패를 결정하는 가장 중요한 변인인 것으로 탐색되었다. 구체적으로 GB\_Action Goals(골키퍼 방어)성공률이 81.81%일 경우 100% 승리하며, 성공률이 46.15%일 경우 100% 패하는 것으로 나타났다. 수구경기에서 Action goals는 Centre goals, Extra player goals, Driving goals, 6m goals, 6mF goals, Penalty goals, Counterattack goals, Corner goals를 제외한 지공상황에서 발생하는 골을 의미한다. 2019년 국제수영연맹(FINA)은 5m 골의 거리를 늘려 6m 골로 변경하였다. 이러한 수구 경기 전술에 중요한 규정변화는 이전과는 다른 공격전술의 다양화를 불러왔다(이승훈 등, 2020). 구체적으로 공격전술에서는 중거리 슛(Middle shot)과 외곽에서의 유기적인 패스 플레이가 중요시되고 있으며, 수비전술에서는 골키퍼의 역할이 중요해졌다. 2020년 올림픽 남자수구경기 데이터를 기반으로 수행된 Enrique, & Carlos의 연구에서는 승패를 결정하는 변인으로 Steals, Action goals, GB\_Action goals가 제시되며 본 연구의 결과를 뒷받침해주었다.

더불어 GB\_Action goals(골키퍼 방어) 다음으로 승패결정요인으로 제시된 변인은 Shooting 변인의 Action goals로 나타났다. GB\_Action goals(골키퍼 방어) 성공률이 46.15%에서 55.55%일 경우

패할 가능성이 66.7%이지만, Action goals 성공률이 23.07% 이상일 때는 승리할 가능성이 75%로 상승하며, 23.07% 이하일 때에는 패할 가능성이 92.3%로 상승하는 것으로 확인되었다. 이는 GB\_Action goals(골키퍼 방어)성공률이 낮더라도 Shooting Action goals의 성공률이 높다면 승리한다는 것을 의미하며, 수구여자경기에서 Action goals 공격, 수비가 경기 승패에 중요변인이라는 것을 보여주고 있다. 따라서 경기력 향상을 위해서는 Action goals의 공격 전술과 방어 전술을 개발해 나가는 것이 중요하다고 판단된다. 2019년 새로운 규칙이 개정되기 전 수행된 Escalante et al.(2011, 2012)연구에서는 GB\_goals(전체 방어율), GB\_5m goals, GB\_Extra player goals가 승패를 결정하는 변인으로 보고되었지만 본 연구에서는 GB\_action goals, action goals, GB\_extra player goal 및 Counterattack goals가 중요한 요인으로 제시되었다. 이는 공격적으로 경기규칙 개정에 따른 전술 및 경기상황의 변화라고도 해석할 수 있다.

이 연구에서 산출된 결과는 제 18회 2019년 세계여자수구선수권대회의 경기기록을 기반으로 연구를 수행하였기 때문에 일반화할 수 없다는 점을 밝힌다. 또한 OMega Timing에서 제공하는 기초통계 공식자료를 기반으로 승패결정요인을 탐색하는 것이 목적이었기 때문에 공식자료 이외의 기록할 수 있는 선수위치변인, 공격전술 등을 고려하지 못하였다. 따라서 후속연구에서는 OMega Timing에서 제공하는 기초통계 공식자료 이외의 보다 구체적인 경기정보변인을 추가하여 승패결정요인을 탐색하는 연구가 이루어져야 할 것이다. 또한, 2019년 경기규칙 변경 이후 수행된 한 개 이상의 세계여자수구대회 결과를 기반으로 승패집단의 스코어 차이에 따른 결정요인 탐색 또는 라운드(예선, 결선 등)에 따른 승패결정요인 탐색 등의 다양한 연구가 수행되길 기대해본다.

종합적으로 이 연구는 데이터마이닝 의사결정나무 분석을 활용하여 세계여자수구선수권대회의 승패결정 패턴을 확인하고자 하였다. 첫째, 세계여자수구선수권대회의 승패 집단에 따른 경기변인별 차이를 분석한 결과, Shooting 변인 중 Total goals, Action goals, Centre goals, Extra player goals, Penalty goals, Counterattack goals, Defence 변인은 Steals, Blocked shots, GB(골키퍼 방어) 변인은 Action goals, Extra player goals, Fouls 변인은 Turnover fouls, Ets 변인은 Sprints won, Time outs 변인에서 통계적으로 유의한 차이를 나타냈다. 둘째, 세계여자수구선수권대회의 승패결정요인으로 GB\_Action goals, Action goals(Shooting), Counterattack goals(Shooting), GB\_Extra player goals로 나타났다. 4개변인 중 승패를 결정짓는 가장 중요한 변인은 GB\_Action goals(골키퍼의 Action goals)방어 성공률로 확인되었다. GB\_Action goals가 81.81% 이상이면 승리집단으로 분류될 확률이 100%, 46.15%이거나 이하이면 패배집단으로 분류될 확률이 100%로 나타났다. 이 연구의 결과는 세계여자수구경기를 분석적 시각에서 접근한 연구로서 스포츠 현장에서는 경기력 향상을 위한 유용한 정보로 사용될 수 있을 것이다.

- 김민창, 최창환, 한민수, 김정호, 김태호(2021). 2021 World Para Ice Hockey Championships의 승패요인 분석: 데이터마이닝 기반 의사결정나무 분석의 적용. **한국특수체육학회지**, 29(3), 105-118.
- 김연자(2011). 배드민턴 경기내용 평가를 위한 기록 분석 시스템 설계. **한국체육측정평가학회지**, 13(3), 103-110.
- 김지웅, 이용국, 이승훈. (2022). 국제경기기록 활용한 남자농구 승패요인 검사: 2019 국제대회를. **한국체육과학회지**, 31 (2), 731-742.
- 김세형 (2012). 한국남자프로농구 경기기록 분석을 통한 승패결정요인 추정방법 비교. **한국체육과학회지**, 21(3), 1347-1360.
- 김세형. (2012). 한국핸드볼 경기기록 분석을 통한 승패결정요인 추정. **체육과학연구**, 23(2), 244-253.
- 김세형, 강상조, 박재현, 김혜진. (2008). 한국프로농구 경기기록 분석에 의한 승패결정요인. **한국체육측정평가학회지**, 10(1), 1-12.
- 김혜진, 박재현, 강상조(2006). 테니스 경기의 득점 및 실점 상황에 의한 승패결정 경기내용분석. **한국체육측정평가학회지**, 8(2), 43-57.
- 나경민, 박재현, 조은혜(2020). 배드민턴 단식경기의 경기분석내용을 활용한 승패결정요인 탐색: 데이터마이닝 기반 의사결정나무분석 적용. **한국체육측정평가학회지**, 22(3), 49-62.
- 대한수영연맹, 2021.9.17. “2022년 연맹 자체 여자 수구 국가대표 선발 대회 관련 공지” <https://www.korswim.co.kr/board/notice/detail/2816> (2022.7.7. 확인)
- 신진호, 최창환(2018). 풋살 경기내용 분석을 통한 승패 결정요인 탐색 : 데이터마이닝 기반 의사결정나무분석의 적용. **한국자료분석학회**, 20(3), 1405-1417.
- 이승훈, 박재현, 윤효준(2020). 수구종목의 경기규칙 개정에 따른 경기 특성 변화분석. **한국체육측정평가학회지**, 22(3), 39-47.
- 조은혜, 강용구, 김혜진(2021). 세계여자핸드볼선수권대회의 라운드별 승패결정요인 탐색 및 비교. **한국체육측정평가학회지**, 23(3), 37-50.
- 최종후, 서두성(1999). 데이터마이닝 의사결정나무 의 응용. **통계분석연구**, 4(1), 61-83.
- 최창환, 윤지운(2017). 경륜 출중정보를 활용한 승자 예측모형 탐색: 데이터마이닝 기반 의사결정나무분석의 적용. **한국체육측정평가학회지**, 19(4), 15-26.
- 최창환, 윤효준, 조은혜, 박재현(2018). 정량적 지표기반 세계 정상급 배드민턴 남자단식 선수들의 경기수행특성. **한국체육측정평가학회지**, 20(1), 1-14.
- Argudo, F. M., García Marín, P., Borges Hernández, P. J., & Ruiz-Lara, E. (2021). Influence of rule changes on shooting performance in balanced matches between two European water polo championship. *International Journal of Performance Analysis in Sport*, 21(1), 61-73.
- Canossa, S., Fernandes, R. J., Estriga, L., Abalde, J. A., Lupo, C., & Garganta, J. M. (2022). Water Polo Offensive Methods after the 2018 FINA Rules Update. *International Journal of Environmental Research and Public Health*, 19(5), 2568.
- Escalante, Y., Saavedra, J. M., Mansilla, M., & Tella, V. (2011). Discriminatory power of water polo game-related statistics at the 2008 Olympic Games. *Journal of sports sciences*, 29(3), 291-298.
- Escalante, Y., Saavedra, J. M., Tella, V., Mansilla, M., García-Hermoso, A., & Dominguez, A. M. (2012). Water polo game-related statistics in Women's International Championships: Differences and discriminatory power. *Journal of sports science & medicine*, 11(3), 475.
- García-Marín, P., & Argudo Iturriaga, F. M. (2017). Water polo: technical and tactical shot indicators between winners and losers according to the final score of the game. *International Journal of Performance Analysis in Sport*, 17(3), 334-349.
- García Ordóñez, E., & Touriño González, C. (2022). Key performance indicators of all-star player and winning teams in elite water polo at the 2020 Olympic Games. *International Journal of Performance Analysis in Sport*, 1-15.
- Lupo, C., Tessitore, A., Minganti, C., King, B., Cortis, C., & Capranica, L. (2011). Notational analysis of American women's collegiate water polo matches. *The Journal of Strength & Conditioning Research*, 25(3), 753-757.
- Miller, T. W. (2015). *Sports analytics and data science: winning the game with methods and models*. FT Press.
- Özkol, M. Z., Turunç, S., & Dopsaj, M. (2013). Water polo shots notational analysis according to player positions. *International Journal of Performance Analysis in Sport*, 13(3), 734-749.
- Saavedra, J. M., Pic, M., Lozano, D., Tella, V., & Madera, J. (2020). The predictive power of game-related statistics for the final result under the rule changes introduced in the men's world water polo championship: a classification-tree approach. *International Journal of Performance Analysis in Sport*, 20(1), 31-41.
- Saavedra, M. J., Escalante, Y., Mansilla, M., Tella, V., Madera, J., & García-Hermoso, A. (2016). Water polo game-related statistics in Women's International Championships as a function of final score differences. *International Journal of Performance Analysis in Sport*, 16(1), 276-289.
- Takagi, H., Nishijima, T., Enomoto, I., & Stewart, A. M. (2005). Determining factors of game performance in the 2001 world Water Polo Championships. *Journal of human movement studies*, 49(2), 333-352.
- Vila, M. H., Abalde, J. A., Alcaraz, E. P., Rodríguez, N., & Ferragut, C. (2011). Tactical and shooting variables that determine win or loss in top-Level in water polo. *International Journal of Performance Analysis in Sport*, 11(3), 486-498.