

스포츠 경기 분석 전문가와 AI의 만남

글 | 최형준 chj2812@dankook.ac.kr 체육측정평가를 전공하였으며, 세계스포츠경기분석학회 사무총장직을 수행하고 있습니다. 스포츠 경기 분석과 ICT 기술의 융합에 관심이 많아 스포츠 경기 분석 자료를 시각화하는 데 관심을 가지면서 AI와 접하게 되었습니다. 앞으로도 급속도로 발전하는 AI를 활용하여 스포츠 분야에 적용할 수 있도록 많은 콘텐츠를 개발하고 싶습니다.

필자는 어린 시절부터 운동을 했다. 운동선수로 6년간 활동했고, 대학에서 체육학을 전공했다. 지금도 스포츠를 자주 접하는 곳에서 생활하며 활동한다. 이런 필자에게 인공지능(AI)에 대한 글을 기고하는 기회가 주어졌을 때, 한치의 망설임도 없이 하겠다고 답을 할 수는 없었다. 왜냐하면 스포츠 경기 분석(performance analysis of sport) 분야에서의 인공지능은 여전히 진행형이며, 기술적인 부분에서 많은 전문가들이 존재하기 때문이다. 하지만 필자는 더 많은 사람들이 스포츠 경기 분석이란 분야에 대해 깊게 이해하고, 인공지능과 관련한 전문가들이 스포츠 전공자들과 협업하여 더욱 발전적인 미래를 만들어낼 수 있었으면 하는 바람을 가지고 있다. 이런 여러 가지 이유에서 용기를 내어 이 글을 기고한다.

스포츠 경기 분석

스포츠는 직접 참여함으로써 즐거움을 느낄 수도 있지만, 보는 것만으로도 즐거움을 얻을 수 있는 여가활동 중 하나이다. 스포츠에 대한 전반적인 내용을 모두 설명하기에는 많은 시간이 필요할 것이다. 하지만 스포츠를 업으로 삼고 생활하는 선수들과 지도자들은 그들이 하는 모든 활동 자체에 가치를 부여한다. 선수들 모두에게 스포츠란 그들의 경기력을 향상시키기 위한 활동의 장이자 목표이다.^{*1} 그런 이유에서 스포츠 경기 분석은 선수들의 경기력을 향상시키고, 더 나아가 선수들과 지도자들에게 차기 목표와 달성 방법을 설정하는 데 필요한 기초 자료를 제공해준다.^{*2} 스포츠 경기 분석은 국내로 들어오면서 여러 가지 의미로 해석되어 왔다.^{*3} 스포츠 경기 분석의 개념은 크게 두 가지로 접근할 수 있는데 스포츠 경기의 기록을 분석하는 의미로서의 스포츠 경기 분석(match analysis)과 스포츠 경기 중 관찰될 수 있는 내용을 중심으로 경기력(performance)을 분석하는 의미로서의 스포츠 경기 분석(performance analysis of sport)이다. 최근에는 스포츠에서 수집되는 자료를 바탕으로 역사적인 요인과 통계적인 요인을 병합하여 분석하는 스포츠 분석(sports analytics)^{*4}이란 개념으로 접근하고 있기도 하다. 무엇보다 어떠한 개념으로 스포츠 경기 분석을 접근하더라도 자료를 수집하는 과정에서 필요한 내용과 범위는 유사하다. 스포츠 경기 분석 관점에서는 이러한 자료 수집에 필요한 요인들을 분석인자(performance indicators)^{*5}라 부르며, 관계형 데이터베이스(database) 필드를 설정하는 개념과도 같다. 아래는 스포츠 경기 분석을 수행하기 위해 거쳐야 하는 단계들이다.

스포츠 경기 분석 단계

- (1) 자료 수집
- (2) 전처리
- (3) 자료 분석 및 통계 분석
- (4) 분석 자료의 시각화

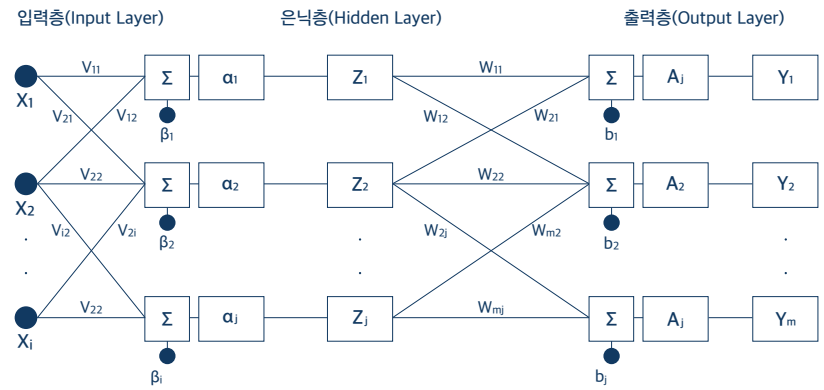
스포츠 경기를 분석할 때 사용하는 여러 기법들은 분석의 목적과 목표에 맞게 적용하며, 경우에 따라서는 간단한 통계 처리부터 복잡한 통계를 접목한 분석기법도 활용되고 있다.

^{*1} 논문 | Fister, I., Fister, D., Deb, S., Mlakar, U., Brest, J., & Fister, I. (2018). Post hoc analysis of sport performance with differential evolution. Neural Computing and Applications, 1-10
^{*2} 논문 | McGarry, T., O'Donoghue, P., & Sampaio, J. (2013). Routledge handbook of sports performance analysis. (2013) Routledge Handbook of Sports Performance Analysis Xvii, New York, NY, US: Routledge/Taylor & Francis Group; US.
^{*3} 논문 | 최형준(2016), <축구 월드컵 경기 공식기록을 통한 데이터의 시각화>, 《한국체육측정평가학회》, pp.83~92
^{*4} 참고 | https://en.wikipedia.org/wiki/Sports_analytics
^{*5} 논문 | Hughes, M., Bartlett, R. (2002). The use of performance indicators in performance analysis. Journal of Sports Sciences. 20, pp.739~754

스포츠 경기 분석과 인공지능

필자는 이 글에서 스포츠 경기 기록과 분석 자료를 토대로 어떠한 방식으로 인공지능을 활용하게 되었는지를 소개하고자 한다. 물론 기술적으로는 기존의 방식을 활용하는 수준이었지만, 스포츠 경기에서 얻을 수 있는 자료의 한계를 고려할 때 결코 쉬운 작업은 아니었다. 먼저 인공지능을 활용하여 2005년 영국 웸블던 테니스 대회의 경기 결과 예측에 관한 연구를 진행하였다. 2004년 127개 경기 중 학습이 불가능한 자료를 제외한 111개 경기 기록을 14개의 분석인자 값으로 입력한 후 학습을 진행하여, 2005년 경기 기록의 결과를 예측하는 방식으로 연구를 진행하였다.

[그림 1] 2005년 영국 웸블던 테니스 대회의 경기 결과 예측에 관한 연구에 사용된 인공신경망 모형*6



[표 1] 경기 결과 예측 학습 과정에서 쓰인 매개변수의 설정값*6

매개변수	설정값
목표 학습 횟수(epochs)	1,000
오차의 목표치(goal)	0
학습률(lr)	0.01
학습률 감소치(lr_dec)	0.7
학습률 증가치(lr_inc)	1.05
최대 학습 실패수(max_fail)	5
최대 오차 증가치(max_perf_inc)	1.04
관성상수(mc)	0.9
최소 경사 감소치(min_gad)	1e - 006

연구결과를 보면 다양한 인공신경망 모델들은 결과를 도출하는 데 있어 매우 광범위한 결과를 가져올 수 있으며, 경기 기록을 토대로 결과를 도출하는 데도 이러한 특성을 고려해야 한다는 사실을 발견하였다. 또한 필자는 경기 기록만으로 예측에 필요한 인공지능을 학습시키는 것은 제한적이며, 이를 개선하기 위해서는 선수들의 특성을 고려한 자료 입력이 필요하다고 판단하였다. 필자는 선수들의 특성을 고려한 자료에 대해서 고민하게 되었다.

따라서 전문적인 기자재를 활용하여 측정한 국가대표 선수들의 체격 및 체력 측정 자료를 인공지능에게 학습시킨다면, ‘인공지능이 선수의 체격 및 체력 데이터를 바탕으로 최적의 스포츠 종목을 알아낼 수 있지 않을까?’라는 고민을 하게 되었다.

스포츠정책과학원에서 수집하고 있는 국가대표 선수들의 체격 및 체력 자료를 이용하여 인공신경망 종류의 하나인 자기구성지도(Self-organizing Map 혹은 Kohen’s feature map)에 적용해보았다. 연구 대상의 체격 및 체력검사 내용은 [표 2]와 같았으며, 연구 대상은 군대 5종, 농구를 비롯한 19개 종목에서 288명의 선수들 자료를 사용하였다.

[표 2] 자기구성지도의 입력층에 입력되었던 연구 대상자의 체격 및 체력검사 내용*7

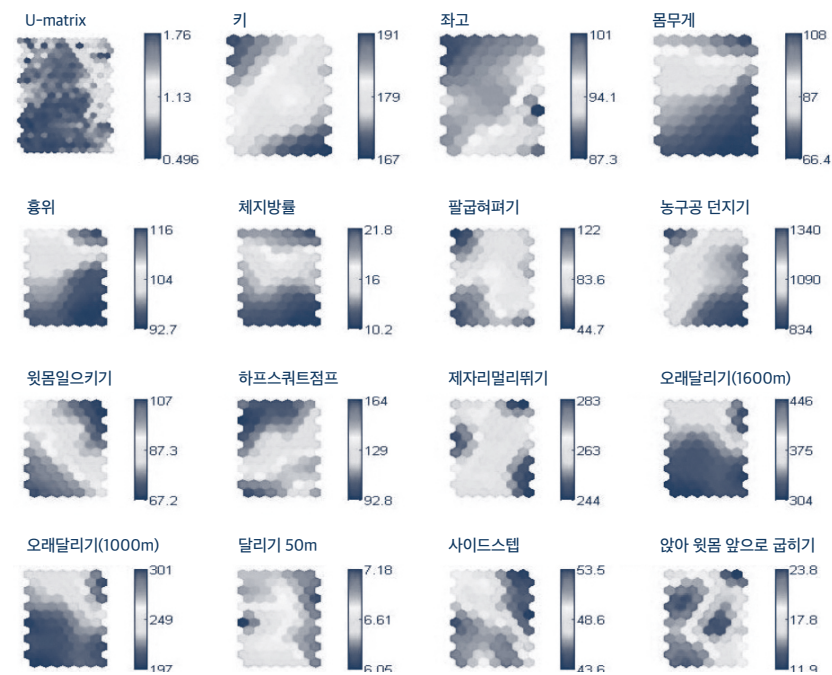
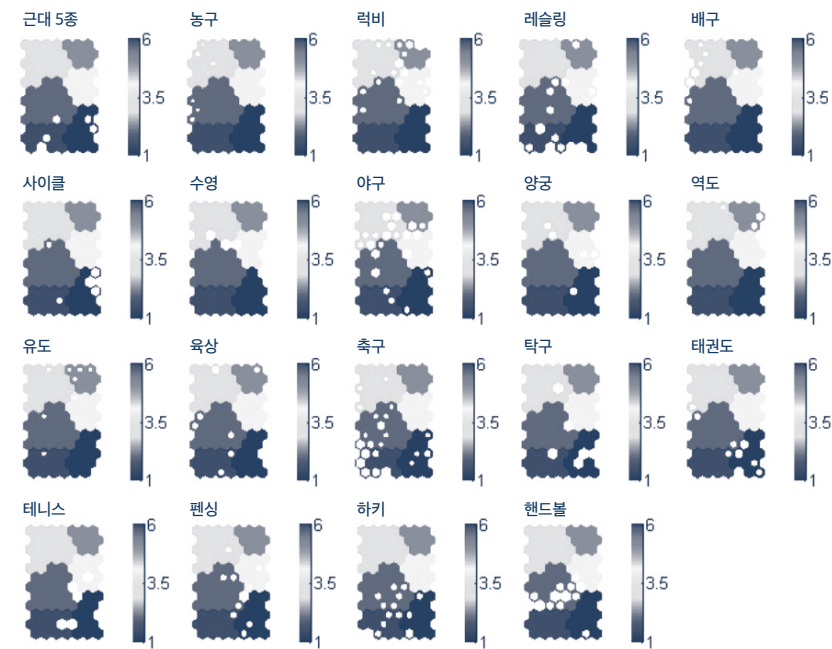
매개변수	요인	검사
체격	길이	신장
	길이	좌고
	무게	체중
	둘레	흉위
	신체 조성	체지방률
체력	상지근지구력	팔굽혀펴기
	상지순발력	농구공 던지기
	복근지구력	윗몸일으키기
	하지근지구력	하프스쿼트점프
	하지순발력	제자리멀리뛰기
	심폐지구력	오래달리기-걸기(1600m)
	심폐지구력	오래달리기(1000m)
	스피드	50m 달리기
	민첩성	사이드스텝 테스트
	유연성	앉아 윗몸 앞으로 굽히기

학습된 자기구성지도에 투영된 각 분석 인자들의 특성은 [그림 2]와 같다. 자기구성지도의 특성상 같은 셀에 있는 분석 인자들의 자료 범위가 모두 포함됨으로써, 학습된 자기구성지도에 새로운 자료가 투영될 때 각 셀에 포함되어 있는 분석 인자들의 특성을 반영하게 된다.

이러한 자기구성지도에 투영된 모든 분석 인자들의 특성을 토대로 군집분석(cluster analysis)을 실시하여 6개 군집으로 축약하였으며, 종목별 선수의 체격 및 체력의 특성을 반영하여 [그림 3]과 같은 결과를 도출하였다.

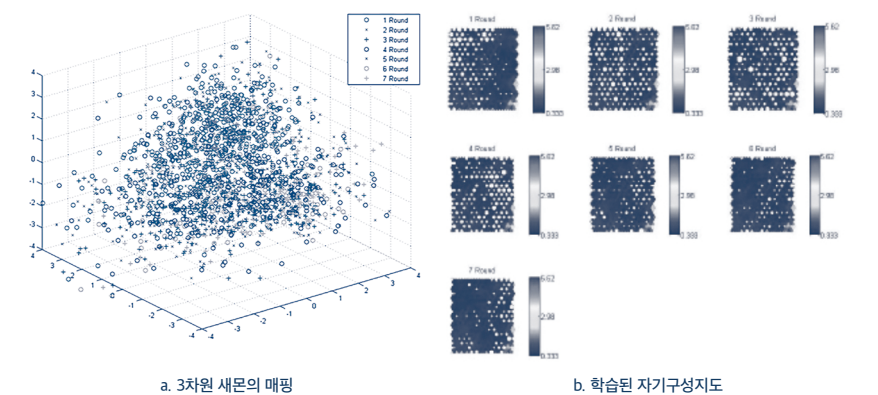
*6 논문 | 최형준·김주학(2006),〈인공 신경망(Artificial Neural Network)을 이용한 2005년도 영국 웸블던 테니스 대회의 경기 결과 예측에 관한 연구〉,《한국체육학회지》45(3), 한국체육학회, pp.459~467

*7 논문 | 최형준·정연성·고병구(2009),〈한국 남자 국가대표 선수의 체격, 체력검사 결과와 자기구성지도(Self-Organized Map)를 이용한 스포츠 종목의 군집분석-사례연구를 중심으로〉,《체육과학연구》20(3), 한국스포츠정책과학원, pp.486~498

[그림 2] 자기구성지도에 투영된 분석인자의 자료 특성 시각화^{*7}[그림 3] 6개 군집으로 나타난 자기구성지도에 투영된 스포츠 종목별 선수들의 원자료^{*7}

결과적으로 자기구성지도에 6개의 군집으로 구분된 종목별 스포츠 선수들의 체격 및 체력 자료는 대부분의 종목에서 가능하였다. 하지만 모든 종목에서 뚜렷한 분류 기준을 밝혀낼 수는 없었다. 그 이유는 포지션별로 선수를 구분하거나 체급이 존재하여 선수들의 체격 및 체력 특성에 차이를 보이는 종목이 존재하기 때문이다. 그러나 보다 괄목할 만한 사실은 기존의 통계분석을 이용해 각 분석 인자에 대한 범위 내에서 분석한 방법에 비해서 인공지능의 포괄적인 분석을 적용할 수 있다는 점과 모든 자료 처리 과정을 시각화하여 살펴볼 수 있다는 점이다.

필자는 스포츠에서 생성되는 자료의 공공성과 인공지능의 자료 처리 및 시각화 기술에 대한 장점을 살려 스포츠 공식기록에서 도출되는 자료의 시각화에 관심을 갖게 되었다. 2005년부터 2010년까지 테니스 4대 그랜드슬램대회(US Open, Austrian Open, Wimbledon Open, French Open)에서 생성되는 자료의 공공성을 토대로 테니스 선수의 경기 기록을 시각화하는 연구를 진행하였다. 필자가 관심 있는 부분은 다차원적 자료(multi-dimensional data)의 시각화인데, 기존의 차원 축소 방법인 요인분석이나 판별분석으로는 자료의 다양한 특성을 시각화하는 데 한계가 있다고 판단하였다. 이에 다차원적 자료를 2차원 혹은 3차원으로 축소시킬 수 있는 새몬의 매핑(Sammon's mapping)과 자기구성지도를 이용하여 2005년부터 2010년까지 테니스 4대 그랜드슬램대회 2060세트를 시각화하였다. 이때 1라운드부터 7라운드까지의 경기를 토너먼트 형식으로 진행하는 테니스 경기의 특성을 반영하여 7라운드까지 진출한 선수의 경기 기록과 1라운드에서 경기를 마무리한 선수의 특성 간에는 분명한 차이가 있을 것이라 가정하였고, 이에 대한 비교를 중점적으로 조명하였다.

[그림 4] 인공지능 기법을 이용한 남자 테니스 선수의 경기 기록 자료^{*8}

연구결과를 살펴보면 새몬의 매핑은 자료의 수가 증가함에 따라서 시각화 진행 후 결과를 해석하는 데 어려움을 보였다. 반면, 학습된 자기구성지도를 통해 나타난 분석 결과는 자기구성지도의 각 셀에서 도출할 수 있는 분석 인자의 범위에 대한 시각화가 가능하다는 점에서 우수성을 보였다. 비록 스포츠 경기 기록 자료를 시각화하는 방법에 대한 간단한 비교였지만, 인공지능을 활용하여 스포츠 경기 기록을 시각화하고 자료를 처리하는 과정에서 도출되는 자료를 토대로 2차 가공이 가능하다는 점에서 의미가 있다고 하겠다.

^{*8} 논문 | 최형준·권민혁(2012), <인공지능(AI: Artificial Intelligence) 기법을 이용한 남자 테니스 선수 경기기록의 시각화>. 《한국체육측정평가학회지》 제14권 제3호, pp.47-56

마치며

인공지능은 산업 분야 전반에 걸쳐 다양하게 적용되고 있고, 그 활용 범위 또한 날이 갈수록 확대되고 있다. 스포츠 경기 분석 분야도 예외는 아니다. 특히 스포츠 경기 기록 분야에서의 활용은 인공지능의 진가를 발휘하기 좋은 주제이다. 최근에는 모든 스포츠 종목에서 경기 기록을 시각화하고 활용하고자 하는 노력이 계속되고 있다. 옵타(Opta)와 SAP 같은 기업들은 이미 스포츠 경기 기록 분야에서 새로운 시장을 창출해내고 있으며, 이를 스포츠 현장에 적용하는 사례도 다양하게 나타나고 있다.

하지만 이와 같은 아웃풋(output)을 내기 위해서는 사람이 판단해야 할 내용도 다수 포함되며, 사람이 판단하는 내용을 보완하기 위해서는 여러 가지 작업이 동시에 필요하다. 스포츠 경기 분석은 목표를 설정하는 단계부터 사람이 많은 작업을 진행한다. 스포츠 전문가와 인공지능 기술의 만남은 앞으로 스포츠 현장을 혁신적으로 뒤바꿔 놓을 것이며, 이를 위해 여러 가지 시도가 계속되어야 한다.



〈스포츠 경기 분석 전문가와
AI의 만남〉브런치로 연결되는
QR 코드입니다.