

학생선수 관련 뉴스기사의 이슈탐색: 토픽모델링 활용

조은혜, 조성인, 김혜진* 한국교원대학교

국문초록

최근 학생선수들의 인권보장이 이슈화되며, 최저 학력기준 및 출석인정 결석허용일수 축소 등 국가정책이 시행되었다. 그러나 스포츠 현장에서는 반발의 목소리가 커지고 있는 실정이다. 이와 더불어 스포츠 학생선수 관련 학교폭력 및 성폭행 등의 이슈는 지속적으로 발생하고 있다. 이에 본 연구에서는 현재 이슈화되고 있는 학생선수 관련 문제들을 객관적으로 파악하고 탐색하기 위해 토픽모델링을 활용하여 학생선수와 관련된 뉴스기사를 기반으로 이슈를 분석하고자 하였다. 본 연구의 목적을 달성하기 위하여 2002년 8월부터 2023년 7월까지 보도된 학생선수와 엘리트선수의 키워드로 검색되는 국내언론기사를 수집하였고, 최종적으로 3,679건의 국내언론기사를 연구대상 자료로 선정하였다. 토픽모델링 분석을 위해 토큰화(tokenization)를 통한 명사단어추출, 불필요단어 제거 등의 데이터 전처리 과정을 진행하였으며, 토픽수 결정을 위한 토픽수별 Coherence Score를 산출하였다. 토픽모델링은 LDA 알고리즘을 적용하였으며, 모든 자료처리는 python 3와 Netminer 4.0프로그램을 사용하였다. 이 연구의 결과는 다음과 같다. 토픽모델링 분석결과 '우수학생선수 대상 훈련개최', '전국대회 참가 및 입상', '학생선수 학습권 및 인권보장', '생활체육 활성화를 통한 선수육성', '우수학생선수 선발 및 육성'이 학생선수 관련 이슈임을 확인할 수 있었다. 이 연구의 결과는 학생선수의 사회적 이슈와 향후 개선 및 발전방안에 관한 다양한 정보를 제공할 수 있을 것이다. 또한 학생선수 관련 정책 및 선수육성의 기초자료로 활용될 수 있을 것이다.

주요어: 학생선수, 이슈, 토픽모델링, 뉴스기사

I. 서론

학교체육진흥법 제2조에 따르면, 학생선수는 “학교운동부에 소속되어 운동하는 학생이나 국민체육진흥법 제33조와 34조에 따른 체육단체에 등록되어 선수로 활동하는 학생을 말한다.”라고 정의하고 있다. 이는 학생의 신분으로 학교에서 학업에 대한 의무를

지키며, 운동선수로서의 활동을 수행해야 한다는 것을 의미한다.

그러나 이해령, 류태호(2021)의 연구에서는 ‘학생선수’를 의미하는 개념이 시대에 따라 변화하였다고 보고하였다. 1970~1980년대에는 ‘학생선수’를 ‘학습하지 않는 학교대표 운동선수’로 정의하였으며, 1990년대에는 ‘학생선수’를 ‘운동해서 대학가는 학교운동부’로 정의하였다고 보고하였다. 2000년대 운동선수도 공부해야 한다는 의견이 제시되며, ‘학생선수’

* 교신저자 김혜진(aries408@knue.ac.kr)

충북 청주시 흥덕구 강내면 태성탑연로 250 한국교원대학교

를 정의하는 개념이 ‘공부하는 학생선수’로 변화하였고, 2010년대에는 ‘정책적 인권 보호 대상 학생선수’로 정의되었다고 보고하였다. 이렇게 ‘학생선수’를 정의하는 개념이 변화한 이유는 학생선수와 관련된 문제들이 끊임없이 대두되고, 이러한 사건·사고를 바라보는 사회의 시선이 변화하였기 때문이다.

그동안 발생한 대표적인 ‘학생선수’ 관련 사건들을 살펴보면, 학생선수(성)폭력사건(이종서, 2022, 장효인, 2021; 정희준, 2010), 부정입학 및 특혜사건(윤진, 2017), 이외에도 학생선수 인권 및 학습권 침해(김처수, 2021; 장달영, 2013) 사건 등이 끊임없이 제기되고 있다.

이에 교육과학기술부와 문화체육관광부는 학생선수를 보호하기 위한 목적으로 ‘공부하는 학생선수 육성’, ‘학생선수 인권보호’라는 국가수준의 정책을 발표하였다(진연경, 유정애, 2015). ‘공부하는 학생선수 육성’ 정책의 세부내용을 살펴보면, 최저학력기준, 정규수업 이수 의무화 및 학력증진 방안 강구 그리고 전국단위 경기 참가 제한(최상협, 이병한, 2020)이 있으며, 이러한 정책들은 2010년 시범운영을 시작으로 2017년부터는 초·중·고 전체로 확대되어 시행되고 있다. ‘학생선수 인권보호’ 정책의 세부내용으로는 학교폭력 및 인권교육 의무화와 주기적 상담(교육부, 2022)이 시행되고 있다.

그러나 학생선수 보호를 목적으로 정책이 만들어지고 시행되었음에도 불구하고 스포츠현장에서는 ‘공부하는 학생선수 육성’ 정책에 반발하고 있는 분위기이다. 특히 최저학력제와 출석인정 결석허용일수에 대한 정책은 스포츠 현장에서 가장 핫한 이슈 중 하나이다. 최저학력제는 학생선수가 국어, 영어, 수학, 사회, 과학 등 주요 교과 성적이 일정수준을 넘지 못하면 전국 대회 출전을 6개월 제한하는 정책으로 실제 최저학력 미 도달 학생선수는 2019년 14.5%, 2020년 13.3%, 2021년 10.9%로 조사되었다(신태현, 2021). 이와 관련하여 스포츠 현장에서는 훈련량을 소화하며, 일괄적

기준의 학습강요는 가혹하다는 의견과 공부는 학생의 본분이며, 은퇴 후 사회진출을 위해 공부는 필수라는 의견(이상서, 2022)이 대립하고 있는 실정이다. 출석 인정 결석허용일수는 학생선수가 대회나 훈련에 참가하기 위해 수업에 빠질 경우 출석으로 인정하는 일수로, 학교급에 따라 허용일수를 다르게 제한하고 있다. 2022년 기준 초등학교 5일, 중학교 12일, 고등학교 25일로 출석인정 결석허용일수를 제한하였으며, 이로 인해 평일보다는 주말을 활용하는 주말리그제가 시행되었다. 그럼에도 불구하고, 탁구선수 신유빈은 고등학교 진학을 포기하였으며, 골프 및 테니스 종목의 학생선수들은 일반 고등학교가 아닌 방송통신고등학교로 진학하는 사례가 나타났다(정이든, 김종호, 2023). 이에 교육부에서는 현장의 목소리를 반영하여 출석인정 결석허용일수를 초등학교는 20일, 중학교는 35일, 고등학교는 50일로 변경하였다(교육부, 2023).

이처럼 학생선수 관련 다양한 이슈들이 발생함에 따라 뉴스기사에서도 관련 내용들을 보고하고 있다. 학생선수 관련 뉴스를 바탕으로 수행된 선행연구로는 학생선수 인권관련 뉴스를 분석한 연구(김양례, 2010)와 성폭력관련 뉴스를 분석한 연구(주형철, 이용건, 2020)로 확인되었다. 그동안 수행된 선행연구들이 학생선수와 관련한 언론의 방향을 확인하는데 기여하였다는 것은 부인할 수 없다. 그러나 선행연구들은 하나의 이슈를 기반으로 한 뉴스를 분석하였기 때문에 학생선수의 전체적인 이슈를 파악하는 것에는 제한점이 있다. 이에 학생선수 관련 전체 이슈들을 객관적으로 파악하고 탐색하는 연구의 필요성이 제기되었다.

현재 비정형 데이터인 뉴스 기사를 분석하는 방법으로는 텍스트 마이닝 기법 중 토픽모델링(Topic Modeling)을 사용할 수 있다(전민수, 임효성, 2021). 토픽모델링은 많은 양의 텍스트로부터 텍스트들간의 관계를 군집화하여 핵심주제를 도출하는 분석기법이다. 특히 Blei, & Jordan(2003)이 개발한 토픽모델링의 LDA(Latent Dirichlet Allocation)알고리즘은

베이지안 기반의 발생확률 추론 모델을 기반으로 토픽을 추출한다는 특징(이원재, 최창환, 2020)을 가지고 있어 학계에서 표준으로 적용되고 있다(윤효준, 박재현, 윤지운, 전민수 2021).

따라서 본 연구에서는 학생선수를 대상으로 보도된 뉴스기사에 대한 이슈를 토픽모델링을 활용하여 탐색하고자 하였다. 학생선수 관련 뉴스기사에 대한 분석 결과는 학생선수 육성 및 정책관련 분야에 기초정보를 제공할 수 있을 것이다.

II. 연구방법

1. 연구자료수집 및 전처리방법

이 연구의 목적을 달성하기 위하여 학생선수를 주제로 보도된 뉴스 기사를 수집하였다. 자료수집 데이터베이스로는 네이버(Naver)와 구글(Google) 포털 사이트를 선정하였다. 자료수집 체계는 첫째, “학생선수”와 “엘리트선수”를 검색 키워드로 선정하였다. 이때 “학생선수”와 “엘리트선수” 키워드가 정확히 일치

하여 제목을 구성하고, 대상이 중·고등학생인 뉴스 기사를 연구자료로 선정하였다. 둘째, 학생선수 키워드로 검색된 뉴스의 시작일인 2002년 8월부터 연구 착수시점인 2023년 7월 15일까지를 검색기간으로 설정하였다. 셋째, 뉴스기사의 출판 언론사에 대한 제한을 두지 않았다. 그러나 2개의 사이트를 통해 수집된 뉴스기사 중 동일한 언론사에서 동일한 제목으로 게재된 뉴스인 경우에는 삭제하였다. 이와 같은 조건으로 총 8,661건 기사가 검색되었으며, 이 중 중복되는 기사 및 학생선수와 관련이 없는 기사 4,987건을 제외하여 최종적으로 3,679건의 뉴스 기사를 연구대상 자료로 선정하였다.

토픽모델링 적용을 위해 3,679개 뉴스기사에 대한 전처리과정을 진행하였다. 첫째, 수집된 뉴스기사의 키워드 명사 추출을 위해 Okt 형태소분석기를 적용하여 단어의 토큰화를 실시하였다. 둘째, 추출된 명사를 바탕으로 불필요한 단어를 삭제하였다. 이 연구에서 검색 키워드로 사용된 학생선수 및 엘리트선수 또한 불필요한 단어로 간주하여 삭제하였다. 이는 연구 목적이 학생선수의 이슈를 탐색하는 것이기 때문에 학생

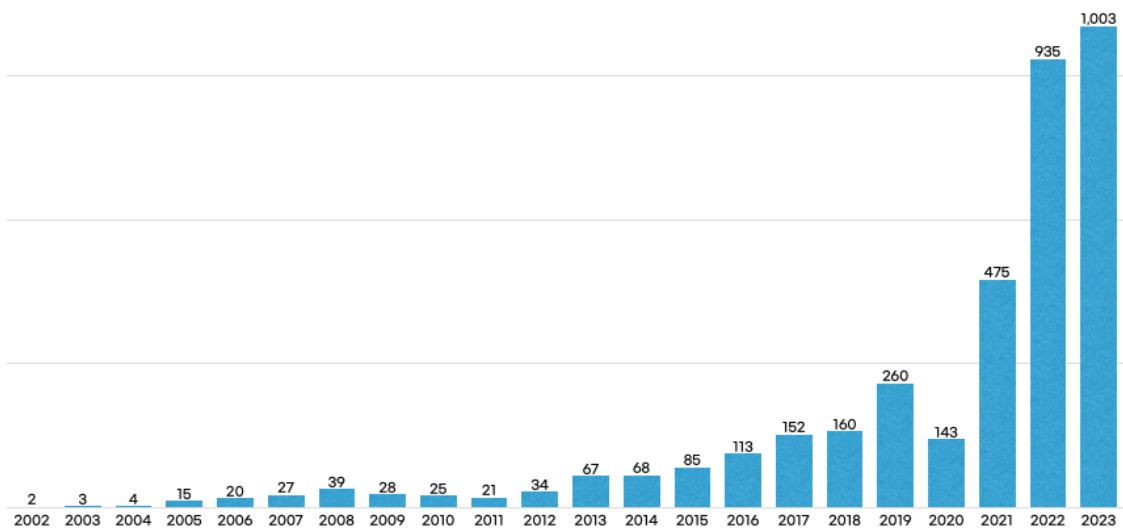


그림 1. 학생선수 관련 뉴스기사의 연도별 빈도

선수 또는 엘리트선수 키워드가 토픽모델을 형성하고 결과를 제시하는데 혼잡을 초래(이원재, 최창환, 2020)할 것으로 판단되어 제거하였다. 셋째, 동일한 의미를 가지고 있지만 다르게 표현되는 유의어는 하나의 단어로 통일시켜주는 과정을 수행하였다. 마지막으로 관찰빈도가 5회 이하인 키워드는 모두 삭제하였다. 연구자료수집 및 전처리를 위해 Python 3을 활용하였다.

2. 자료처리방법

이 연구의 목적은 토픽모델링을 활용하여 학생선수 관련 뉴스기사의 이슈를 탐색하는 것이다. 이 연구의 목적을 달성하기 위하여 수행한 자료처리방법은 다음과 같다. 첫째, 연구자료로 수집된 학생선수 관련 뉴스기사 및 키워드의 빈도를 산출하였다. 빈도산출은 IBM SPSS 25.0을 활용하였다. 둘째 토픽모델링 분석을 위해 키워드의 공출현 빈도와 행렬을 산출하였다. 공출현 빈도와 행렬산출은 Python 3을 활용하였다. 셋째, 토픽모델링은 토픽 수를 사전에 결정해야 한다. 토픽모델링 연구에서 토픽 수를 설정하는 것은 산출되는 결과가 달라질 수 있기 때문에 매우 중요(DiMaggio, Nag, & Blei, 2013)하다. 선행연구에 따르면 토픽 수를 결정하는 방법으로는 연구자 또는 전문성을 지닌 연구자들의 해석여부 및 적합성을 통해 토픽 수를 선정(전민수, 임효성, 2021; Binkley, Heinz, Lawrie, & Overflet, 2014)하는 방법과 Coherence(일관성) Score를 기반으로 토픽 수를 선정(윤효준, 박재현, 윤지운, 2019)하는 방법이 사용되고 있다고 보고하고 있다. 이에 본 연구에서는 Coherence(일관성) Score를 산출하여 토픽 수를 5개(Chherence Score: 0.671)로 설정하였다. 마지막으로 토픽모델링 분석을 위하여 LDA 알고리즘 MCMC(Markov Chain Monte Carlo) 학습방법을 적용하여 토픽과 조건부 확률분포를 산출하였다. 토픽모델링 분석은 Netminer 4.0을 사용하였다.

III. 연구결과 및 논의

1. 학생선수 관련 뉴스기사의 주요 키워드

첫 번째 연구결과는 학생선수 관련 뉴스기사에 대한 기초정보이다. <표 1>은 학생선수 관련 뉴스기사에서 나타난 키워드와 발생빈도를 분석한 결과이다. <표 1>을 살펴보면, ‘참가’ 키워드가 385회로 가장 높은 빈도를 나타냈다. ‘전국소년체전’ 키워드는 345회로 두 번째로 높은 결과를 나타냈으며, ‘개최’ 키워드가 327회로 나타났다. 이외에도 ‘메달획득’, ‘육성’, ‘출석인정결석허용일수’, ‘훈련’ 등의 키워드가 학생선수 관련 뉴스기사에서 공통적으로 발생하는 높은 빈도의 키워드로 나타났다.

표 1. 학생선수 관련 뉴스기사의 빈도상위 20개 키워드

순위	키워드	빈도	순위	키워드	빈도
1	참가	385	11	유소년	137
2	전국소년체전	345	12	전국장애인체육대회	137
3	개최	327	13	출석인정결석허용일수	126
4	메달획득	258	14	꿈나무	115
5	중학교	200	15	격려	115
6	체육	196	16	스포츠클럽	97
7	고등학교	193	17	국가대표	90
8	전국체전	171	18	청소년	88
9	육성	159	19	우승	86
10	초등학교	146	20	훈련	86

2. 우수학생선수 대상 훈련개최

<표 2>는 각 토픽을 구성하고 있는 키워드와 확률분포를 나타낸 것이다. 토픽1(Topic 1)을 구성하는 키워드는 ‘개최’, ‘KBL’, ‘꿈나무’, ‘체육’, ‘훈련’, ‘청소년’, ‘참가’, ‘유소년’, ‘국가대표’ 등으로 나타났다. Topic 1에 대한 토픽명칭을 선정하기 위해 토픽 비중도가 높은 원문기사(원문 중 일부: 박수민, 2012; 최정서, 2018)를 살펴본 결과, 학생선수 중 우수선수로 분류되는 꿈나무, 유소년, 국가대표를 대상으로 훈련이 진행되거나 개최되는 것과 관련된 내용으로 확인되었

표 2. 토픽모델링 분석결과

Topic	Keyword(Probability distribution)				
Topic1	개최(.104)	KBL(.056)	꿈나무(.038)	체육(.034)	훈련(.027)
	청소년(.026)	참가(.025)	유소년(.024)	국가대표(.020)	진행(.019)
Topic2	참가(.149)	메달획득(.074)	격려(.065)	전국소년체전(.052)	개최(.023)
	체육(.023)	전국체전(.022)	성적(.019)	종목(.015)	제주(.015)
Topic3	확대(.035)	실시(.033)	지원(.026)	운영(.026)	교육(.025)
	출석인정결석허용일수(.021)	인권(.020)	진로(.016)	강화(.016)	보호(.016)
Topic4	개최(.078)	메달획득(.064)	우승(.063)	스포츠클럽(.035)	선수권(.034)
	육성(.027)	생활체육(.024)	배드민턴(.023)	경기(.021)	연맹(.015)
Topic5	고등학교(.082)	중학교(.068)	전달(.049)	육성(.043)	장학금(.040)
	초등학교(.028)	우수(.019)	선발(.018)	코치(.014)	활동(.012)

다. 이에 Topic 1에 대한 명명은 ‘우수학생선수 대상 훈련개최’로 선정하였다.

국내에서는 우수학생선수 훈련이 특정지역에서 꾸준히 진행되고 있다. 그리고 우수학생선수 훈련이 진행된다는 뉴스기사 또한 지역별 신문사를 통해 적지 않게 보도되고 있음을 확인할 수 있었다. 국내 특정지역에서 스포츠이벤트가 진행된다는 것은 지역홍보를 넘어 경제적인 활성화에도 영향을 미치기 때문에 관련 내용을 담은 뉴스기사들이 많이 보도되었다고 판단된다. 김민철(2007) 및 유재구, 왕석우, 신태근(2010)은 스포츠이벤트가 지역의 경제효과 및 관광효과에 긍정적인 영향을 미친다고 보고하며, 연구자의 주장을 뒷받침 해주었다. 또한 KBL의 훈련관련 뉴스기사는 선수유치 및 KBL 홍보를 목적으로 보도되었다고 판단된다. 꿈나무, 청소년, 유소년 선수 대상의 훈련을 주관한다는 것은 농구선수를 시작하거나, 선수로 활동하는 학생선수들에게는 동기부여가 될 수 있으며, KBL에 대한 긍정적인 이미지를 심어줄 수 있기 때문이다. 이용식(2008)의 선행연구에서는 스포츠 활성화를 위한 방법으로 종목에 대한 홍보를 강화하고 캠프 등이 활성화 되어야 한다고 보고한 바 있다.

3. 전국대회 참가 및 입상

토픽2(Topic 2)를 구성하는 키워드는 ‘참가’, ‘메

달획득’, ‘격려’, ‘전국소년체전’, ‘개최’ 등으로 확인되었다. Topic 2에 대한 토픽명칭을 선정하기 위해 토픽 비중도가 높은 원문기사(원문 중 일부: 박웅호, 2019; 이상욱, 2012)를 살펴본 결과, 전국대회 참여 및 개최 그리고 대회입상을 통한 메달획득에 관한 내용으로 확인되었다. 이에 Topic 2에 대한 명명은 ‘전국대회 참가 및 입상’으로 선정하였다.

전국소년체육대회는 학생선수에게 국내대회 중 가장 큰 대회로 매년 5월 마지막 주에 개최되며, 50년이 넘는 역사를 가지고 있다. 대회의 중요도가 높은 만큼 시·도에서도 전국소년체전 성적에 많은 관심을 갖고 지원을 하고 있는 실정이다. 이를 뒷받침하듯 지역신문사들이 전국소년체육대회 성적 및 입상선수 관련 뉴스기사들을 보도하고 있다는 것을 확인할 수 있었다. 박준태(2023)의 선행연구에서는 우수선수 발굴이라는 목적으로 준비 및 참여에 따른 시·도간 지나친 경쟁이 발생되고 있다고 보고하며, 본 연구의 결과를 지지 해주었다. 그러나 이러한 지나친 경쟁은 ‘공부하는 학생선수 육성’이라는 정책과 역행하는 사례를 빈번히 보여주고 있어, 전국소년체육대회를 전면 폐지하고 학생체육축전 형식으로 변환하자고 정부에서 요구한 바 있다(문화체육관광부, 2019). 이러한 정책들이 제기되지 않기 위해서는 훈련시스템 변화, 학습권 보장 등 긍정적인 변화가 필요하다고 사료된다.

4. 학생선수 학습권 및 인권보장

토픽3(Topic 3)을 구성하는 키워드는 ‘확대’, ‘실시’, ‘지원’, ‘운영’, ‘교육’, ‘출석인정 결석허용일수’, ‘인권’, ‘진로’ 등으로 확인되었다. Topic 3에 대한 토픽명칭을 선정하기 위해 토픽 비중도가 높은 원문기사(원문 중 일부: 이보람, 2017; 정태화, 2023)를 살펴본 결과, 학생선수 출석인정 결석허용일수 확대 및 인권 교육 지원에 대한 기사로 인한 것임을 확인할 수 있었다. 이에 Topic 3에 대한 명명은 ‘학생선수 학습권 및 인권보장’으로 선정하였다.

이해령, 류태호(2021)의 선행연구에 의하면, 학생선수의 정의가 1970~1980년대 ‘학습하지 않는 학교 대표 운동선수’에서 2010년도부터는 ‘정책적 인권 보호 대상 학생선수’로 변경되었다고 보고하고 있다. 이러한 변화는 (성)폭력, 입시 비리, 학력저하 등 많은 사건·사고(김처수, 2021; 윤진, 2017; 이종서, 2022; 장달영, 2013)들을 이유로 들 수 있다. 정부는 사건·사고를 대처하는 방안으로 ‘공부하는 학생선수 육성’, ‘학생선수 인권보호’라는 정책을 제시하였고 현재 시행 중에 있다. 그러나 대회 출전에 영향을 미치는 출석인정 결석허용일수와 최저 학력제는 스포츠현장에서 반대의견이 많은 정책으로 현재 정책 수정 및 완화가 요구되고 있다. 출석인정 결석허용일수를 이유로 고등학교 진학을 포기한 탁구 국가대표 신유빈 사례(손덕호, 2023), 방송통신고등학교 진학사례(지성배, 2023)를 더 이상 만들지 않기 위해서는 현장의 의견을 반영한 정책이 시행되어야 하며, 스포츠현장에서도 학생선수의 학습에 대한 인식이 전환되어야 한다고 사료된다.

5. 생활체육 활성화를 통한 선수육성

토픽4(Topic 4)를 구성하는 키워드는 ‘개최’, ‘메달 획득’, ‘우승’, ‘스포츠클럽’, ‘선수권’, ‘육성’ 등으로 확인되었다. Topic 4에 대한 토픽명칭을 선정하기 위해

토픽 비중도가 높은 원문기사(원문 중 일부: 전영지, 2023)를 살펴본 결과, 생활체육 한 부분인 스포츠클럽 활성화를 통한 선수육성 제도관련 기사로 인한 것임을 확인할 수 있었다. 이에 Topic 4에 대한 명명은 ‘생활체육 활성화를 통한 선수육성’으로 선정하였다.

스포츠에 대한 대중의 관심이 많아지며, 생활체육인이 증가하고, 관련 대회가 활성화되고 있다. 반면, 엘리트 체육에서는 인기종목을 제외한 대다수의 종목에서 학생선수 부족으로 많은 어려움을 겪고 있는 실정(이종만, 2021)이다. 2019년 문화체육관광부 스포츠혁신위는 ‘스포츠 복지사회 실현을 위한 스포츠클럽 활성화 권고’를 발표하며, 스포츠클럽에서 일반학생과 학생선수 구분 없이 재능, 소질을 발휘하다 특정 시점에서 직업선수로 전환하는 대안적 방식을 제안한 바 있다. 이와 더불어 유승민 전 IOC 선수위원 또한 엘리트체육이 활성화는 생활체육과 유기적으로 연결되어 있기 때문에 생활스포츠 확대정책과 학생선수 지원 정책이 함께 추진되어야 한다고 제시하며, 본 연구의 결과를 뒷받침해주었다.

6. 우수학생선수 선발 및 육성

토픽5(Topic 5)를 구성하는 키워드는 ‘고등학교’, ‘중학교’, ‘전달’, ‘육성’, ‘장학금’ 등으로 확인되었다. Topic 5에 대한 토픽명칭을 선정하기 위해 토픽 비중도가 높은 원문기사(원문 중 일부: 김용우, 2019)를 살펴본 결과, 우수학생선수 선발 및 육성을 위한 장학금 전달과 관련한 기사로 인한 것임을 확인할 수 있었다. 이에 Topic 5에 대한 명명은 ‘우수학생선수 선발 및 육성’으로 선정하였다.

이 연구의 Topic 1의 결과에서 다른 것처럼 국내에서는 우수학생선수, 즉 꿈나무 선수를 발굴하기 위해 다방면으로 노력중이다. 노력의 일환으로 국민체육진흥공단의 한국스포츠정책과학원에서는 매년 각 종목별 ‘꿈나무 선발 측정 평가’를 수행하고 있다. 이외에

도 시도에서는 행정적 차원에서 우수학생선수에게 장학금을 전달하는 경우를 쉽게 찾아볼 수 있으며, 기업에서는 스포츠영재를 선발하고 장학금을 전달하는 과정을 통해 긍정적인 이미지 홍보(정하니, 2016)와 더불어 매출에 영향을 받기도 한다. 김준성, 정호진(2015)의 선행연구에서는 기업의 후원을 받는 스포츠 선수의 성과와 기업 매출의 관계를 분석한 결과, 정적인 영향을 미친다는 것을 검증하였다.

IV. 결론

이 연구에서는 토픽모델링을 활용하여 학생선수를 대상으로 보도된 뉴스기사에 대한 이슈를 탐색하였다. 이에 따른 결론은 다음과 같다.

첫째, 학생선수 관련 뉴스기사의 키워드 빈도가 높았던 키워드로는 ‘참가’, ‘전국소년체전’, ‘개최’, ‘메달획득’ 등 학생선수들의 대회참여 및 입상과 관련된 키워드였으며, 이외에도 ‘육성’, ‘출석인정결석허용일수’, ‘훈련’ 등의 키워드가 빈도가 높은 것으로 확인되었다. 둘째, 학생선수 관련 뉴스기사를 기반으로 토픽모델링을 수행한 결과, 첫 번째 토픽으로 ‘우수학생선수 대상 훈련개최’가 이슈로 탐색되었다. 국내에서는 우수학생선수 훈련이 특정지역에서 꾸준히 진행되고 있으며, 이는 지역홍보를 넘어 경제적인 활성화에도 영향을 미치기 때문이라고 볼 수 있다. 셋째, 두 번째 토픽으로 ‘전국대회 참가 및 입상’이 이슈로 탐색되었다. 우수선수 발굴이라는 목적으로 준비 및 참여에 따른 시·도간 지나친 경쟁이 발생되고 있다. 이러한 지나친 경쟁은 ‘공부하는 학생선수 육성’이라는 정책과 역행하는 사례를 빈번히 보여주고 있어, 전국소년체육대회 폐지론이 제기되고 있다. 이러한 정책들이 제기되지 않기 위해서는 훈련시스템 변화, 학습권 보장 등 긍정적인 변화가 필요하다. 넷째, 세 번째 토픽으로 ‘학생선수 학습권 및 인권보장’이 이슈로 탐색되었다.

정부에서는 ‘공부하는 학생선수 육성’, ‘학생선수 인권보호’라는 정책을 제시하였고 현재 시행 중에 있다. 그러나 대회 출전에 영향을 미치는 출석인정 결석허용일수와 최저학력제는 스포츠현장에서 반대의견이 많은 정책으로 현재 정책 수정 및 완화가 요구되고 있다. 다섯째, 네 번째 토픽으로는 ‘생활체육 활성화를 통한 선수육성’이 이슈로 탐색되었다. 엘리트 체육에서는 인기종목을 제외한 대다수의 종목에서 학생선수 부족으로 많은 어려움을 겪고 있는 실정이다. 엘리트 체육을 활성화시키기 위한 방안으로는 생활스포츠 확대정책과 학생선수 지원정책이 함께 추진되어야 한다고 제시되고 있다. 마지막으로 다섯 번째 토픽으로는 ‘우수학생선수 선발 및 육성’이 이슈로 탐색되었다. 시·도 및 기업에서는 긍정적인 이미지 홍보와 더불어 매출에 영향을 위해 우수학생선수에게 장학금을 전달하는 경우를 쉽게 찾아볼 수 있다. 본 연구의 결과는 학생선수의 사회적 이슈를 확인하는 정보로 학생선수 관련 정책 및 선수육성의 기초자료로 활용될 수 있을 것이다.

참고문헌

- 교육부(2022). 2022년 학교체육 활성화 추진 기본계획(안).
 교육부(2023). 2022년 학교체육 활성화 추진 기본계획(안).
 교육부, 문화체육관광부(2013). 2013년 학교체육 주요업무 계획.
 김민철(2007). 스포츠 이벤트 관여도와 홍보활동이 지역특산물 브랜드에 미치는 영향: 보성녹차를 중심으로. **체육과학연구**, 18(1), 99-110.
 김양례(2010). 학생운동선수 인권 관련 보도의 프레임 분석. **한국스포츠사회학회지**, 23(3), 37-53.
 김용우(2019). (재)유원골프재단, 9년연속 ‘대전지역골프인 재육성’장학금 전달. 충청뉴스, 인터넷기사 출처: <https://www.ccnnews.co.kr/news/articleView.html?idxno=150037>
 김준성, 정호진(2015). 기업 후원을 받는 스포츠 선수의 성과와 기업의 판매 성과 간의 관계 분석. **한국콘텐츠학회논문지**, 15(12), 550-559.

- 김처수(2021). 학생과 운동선수의 경계에 서 있는 학생선수, 인권보호 강화돼야. 사이드뷰, 인터넷기사 출처: <https://www.sideview.co.kr/news/articleView.html?idxno=6367>
- 문화체육관광부(2019). **전국소년체전 폐지 및 체육계비리 근절 대책**.
- 박수민 (2012). KBL 삼성 썬더스, 꿈나무 농구대회 개최. STNTV, 인터넷기사출처: <https://www.stnsports.co.kr/news/articleView.html?idxno=4994>
- 박용호(2019). 영천시, 제48회 전국소년체전 입상선수 격려행사 개최. 대구일보, 인터넷기사 출처: <https://www.idaegu.com/newsView/idg201906090042>
- 박준태. (2023). 텍스트마ining 기법을 활용한 소년체육대회 분석 연구. **한국초등체육학회지**, 29(3), 37-49.
- 손덕호(2023). 탁구 신유빈처럼 고교진학 포기 않도록..학생선수 출석인정 확대. 조선비즈, 인터넷기사출처: https://biz.chosun.com/topics/topics_social/2023/01/19/FIUFBMZ3GVCT7IJ42VOEV2FJCM/
- 신태현(2021). 최저학력 미달 학생선수 6개월 출전 금지 추진, 뉴스토마토, 인터넷기사출처: <https://www.newstomato.com/readNewspaper.aspx?no=1075160>
- 유재구, 왕석우, 신태근(2010). 스포츠이벤트 개최도시의 도시마케팅 효과: 개최도시 주민대상 설문조사를 중심으로. **한국스포츠산업경영학회지**, 15(4), 63-77.
- 윤진(2017). “제2의 정유라 막자”...운동만 잘 하면 안돼. KBS 뉴스. 인터넷기사출처: <https://news.kbs.co.kr/news/pc/view/view.do?ncd=3460970>
- 윤효준, 박재현, & 윤지운. (2019). 비정형 텍스트 자료에서 잠재정보 추출을 위한 토픽모델링 소개: 치매관련 신체활동 뉴스 기사의 이슈 분석. **체육과학연구**, 30(3), 501-512.
- 윤효준, 박재현, 윤지운, 전민수(2021). 토픽모델링을 적용한 코로나19 확산에 따른 스포츠영역 이슈분석. **한국체육측정평가학회지**, 23(1), 11-20.
- 이보람(2017). “제2 정유라, 장시호 막는다” 출석인정결석, 수업일 3분의 1까지. 뉴스핌, 인터넷기사 출처: <https://www.newspim.com/news/view/20170410000011>
- 이상서(2022). [갈곳없는 학생선수들] “운동만도 벼차”vs“공부는 학생본분”. 연합뉴스, 인터넷기사출처: <https://www.yna.co.kr/view/AKR20221028134000501>
- 이상욱(2012). 경주교육지원청 전국소년체전 금메달 유망주 격려. 경북일보, 인터넷기사 출처: <https://www.kyongbuk.co.kr/news/articleView.html?idxno=578276>
- 이용식(2008). 한국 육상 활성화를 위한 정책방안. **체육과학연구**, 19(2), 117-124.
- 이원재, 최창환(2020). 태권도 전자호구 관련 뉴스기사의 이슈: 토픽모델링과 의미연결망 분석의 적용. **한국체육측정평가학회지**, 22(4), 41-53.
- 이종만(2021). <제 102회 전국체전>엘리트 체육 위기에 코로나까지...선수 부족으로 신음중인 학교체육. 인천일보, 인터넷기사출처: <https://www.incheonilbo.com/news/articleViewAmp.html?idxno=1115591>
- 이종서(2022). 안우진 입장문 발표...“할 수 있던 게 침묵 뿐이었다”. 스포츠조선, 인터넷기사출처: <https://www.chosun.com/sports/baseball/2022/11/18/YKGN7EUHEH66CFDYNGNILDW664/>
- 이혜령, 류태호(2021). ‘학생선수’개념의 의미 변천과 교육적 함의 도출. **한국스포츠교육학회지**, 28(2), 129-148.
- 장달영(2013). 제주FC의 편법적 학생선수 학습권 침해, 재발 방지 필요하다. 미디어오늘, 인터넷기사출처: <https://www.mediatoday.co.kr/news/articleView.html?idxno=111798>
- 장효인(2021). “코치가 성폭행” 국대 여중생 폭로에 경찰 수사. 연합뉴스TV, 인터넷기사 출처: <https://m.yonhapnewstv.co.kr/news/MYH20210903019900641>
- 전민수, 임효성(2021). 태권도 뉴스기사의 연도별 주제어 비교 분석: 토픽모델링 적용. **디지털융복합연구**, 19(11), 575-583.
- 전영지(2023). “생활X전문체육 선순환” 문체부, 지정스포츠 클럽 107개소X314개 프로그램지원. 스포츠조선, 인터넷기사 출처: <https://sports.chosun.com/sports-news/2023-03-15/202303150100120780014381?t=n1>
- 정이든, 김종호(2023). 학생선수 출석인정 결석 허용일수 이슈에 관한 고찰: 미디어 프레임 분석을 중심으로. **한국사회체육학회지**, 93, 219-234.

- 정태화(2023). 초등학교5일-20일, 중학교 12일-35일, 고등학교 25일-50일로 학생선수 출석인정결석 허용일수 확대... 대한체육회 등 환영성명. 마니아타임즈, 인터넷기사출처: https://www.maniareport.com/view.php?ud=2023011912143737518e70538d2_19
- 정하니(2016. 11월 25일). ING생명, 국내 스포츠 꿈나무들에게 5억원 지원. **서울경제**, <https://www.sedaily.com/NewsView/1L43VW37TK>
- 정희준(2010). ‘성의 노예’가 된 ‘운동기계’“스포츠, 다 x까라 그래” 그들은 왜 짐승이 되었나?. 프레시안, 인터넷기사출처: <https://www.pressian.com/pages/articles/99984>
- 주형철, 이용건(2020). 빅데이터를 활용한 운동선수 성폭력 인식 분석. **한국스포츠학회**, 18(3), 1105-1114.
- 지성배(2023). [2023 국감] 文정부서 줄인 학생선수 출석인정 일수...방통고 입학 4.5배 폭등으로 이어져. 교육플러스, 인터넷기사출처: <https://www.edpl.co.kr/news/articleView.html?idxno=10507>
- 진연경, 유정애(2015). 국가수준 학교체육정책의 내용과 성과 고찰. **교육과정평가연구**, 18(3), 343-372.
- 최상협, 이병찬(2022). 공부하는 학생선수 만들기. **한국체육 과학회지**, 29(2), 719-730.
- 최정서(2018). KBL-군산시, 2018 KBL 유소년클럽 농구대회 개최 협약. 스포츠한국, 인터넷기사 출처: <https://www.dailysportshankook.co.kr/news/articleView.html?idxno=194229>
- 학교체육진흥법(2023). 학교체육진흥법 법령 제2조. <https://www.law.go.kr/법령/학교체육%20진흥법>.
- Binkley, D., Heinz, D., Lawrie, D., & Overfelt, J. (2014). Understanding LDA in source code analysis. *In Proceedings of the 22Nd Int'l Conf. on Program Comprehension*, 26-36, ACM.
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77-84.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan), 993-1022.
- DiMaggio, P., Nag, M., & Blei, D. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: application to newspaper coverage of U.S. government arts funding. *Poetics*, 41, 570-606.

저자정보

조은혜(Eunhye Jo)

한국교원대학교 학교체육연구소 연구교수
eunhyejo3@gmail.com

조성인(Sung in Jo)

한국교원대학교 체육교육과 석사과정
nanjsn18@gmail.com

김혜진(Hyeoi Jin Kim)

한국교원대학교 체육교육학과 교수
aries408@knue.ac.kr

논문투고일	2023년 11월 30일
심사완료일	2023년 12월 14일
게재확정일	2023년 12월 19일

Abstract

The Korean Journal of Measurement and Evaluation in Physical Education and Sport Science. 2023, 25(4), 1-10

Exploring Issues in News Articles Related to Student Athletes: Utilizing Topic Modeling

Eunhye Jo, Sung in Jo, Hyeoi jin Kim *Korea National University of Education.*

This research aims to objectively assess and explore current issues related to student-athletes by employing topic modeling on news articles. To achieve this objective, a collection of domestic media articles related to student-athletes and elite athletes, searchable using keywords from August 2002 to July 2023, was compiled. Ultimately, 3,679 domestic media articles were selected as the research dataset. Data preprocessing, including tokenization for noun extraction and removal of unnecessary words, was conducted for topic modeling analysis. The determination of the number of topics was achieved using the Coherence Score for each topic, and the analysis employed the LDA algorithm with Python 3 and Netminer 4.0. The results of this study identified the following topics as issues related to student-athletes: 'Training Events for Outstanding Student-Athletes,' 'Participation and Achievements in National Competitions,' 'Learning Rights and Human Rights Protection for Student-Athletes,' 'Fostering Athletes through the Promotion of Recreational Sports,' and 'Selection and Development of Outstanding Student-Athletes.' The findings of this research provide valuable insights into the social issues surrounding student-athletes and can serve as diverse information for future improvements and developments. Furthermore, the results can be utilized as foundational data for policies related to student-athletes and athlete development.

Keywords: student athletes, issue, news articles, topic modeling