

태권도 뉴스기사의 연도별 주제어 비교분석: 토픽모델링 적용

전민수¹, 임효성^{2*}

¹한국체육대학교 체육과학연구소 학술연구교수, ²상명대학교 산학협력단 특임교수

Comparative Analysis of the Keywords in Taekwondo News Articles by Year: Applying Topic Modeling Method

Jeon, Minsoo¹, Lim, Hyosung^{2*}

¹Academic Research Professor, Institute of Sport Science, Korea National Sport University,

²Special Appointment Professor, Industry-Academy Cooperation Foundation, Sangmyung University

요약 이 연구는 토픽모델링을 적용하여 뉴스기사에 따른 태권도 동향을 연도별로 분석하는 것에 목적이 있다. 언론보도를 통한 태권도 동향을 살펴보기 위해 한국언론재단의 빅카인즈를 통해 뉴스기사와 태권도 전문 언론에 대한 기사를 수집하였다. 검색기간은 2000년 이전, 2001년~2010년, 2011년~2020년 3개의 구간으로 구분하여 검색하여 총 12,124개를 연구자료로 선정하였다. 토픽분석을 위해 전처리 과정을 거쳤으며, LDA 알고리즘을 활용하여 토픽분석을 수행하였다. 이때 모든분석은 python 3을 적용하였다. 그 결과 첫째, 연도별로 다른 언론기사 주제를 분석한 결과 2000년이전 1위는 '세계', 2위는 '남북', 3위는 '올림픽'으로 나타났으며, 2001년~2010년 1위는 '세계', 2위는 '협회', 3위는 '세계태권도연맹'으로 조사되었다. 2011년~2020년 1위는 '세계', 2위는 '시범', 3위는 '국기원'으로 나타났다. 둘째, 2000년이전 뉴스기사를 토픽모델링으로 분석한 결과 토픽은 2가지로 구분되었다. 구체적으로 Topic 1은 '남·북 체육교류', Topic 2는 '올림픽 시범종목 채택'으로 선정되었다. 셋째, 2001년~2010년 뉴스기사를 토픽모델링으로 분석한 결과 토픽은 3가지로 선정되었다. Topic 1은 '태권도 시범공연 및 비리', Topic 2는 '무주태권도공원 조성', Topic 3은 '세계태권도축제'로 선정되었다. 넷째, 2011년~2020년 뉴스기사를 토픽모델링으로 분석한 결과 토픽은 3가지로 선정되었다. Topic 1은 '2018 평창동계올림픽 성공 개최', Topic 2는 '남북 태권도 합동시범공연', Topic 3은 '2017 무주세계태권도선수권대회'로 선정되었다.

주제어 : 토픽모델링, 태권도, 뉴스기사, 연도별, 주제어, 이슈

Abstract This study aims to analyze Taekwondo trends according to news articles by year by applying topic modeling. In order to examine the Taekwondo trend through media reports, articles including news articles and Taekwondo specialized media articles were collected through Big Kinds of the Korea Press Foundation. The search period was divided into three sections: before 2000, 2001~2010, and 2011~2020. A total of 12,124 items were selected as research data. For topic analysis, pre-processing was performed, and topic analysis was performed using the LDA algorithm. In this case, python 3 was applied for all analysis. First, as a result of analyzing the topics of media articles by year, 'World' was the most common keyword before 2000. 'South and North Korea' was next common and 'Olympic' was the third commonest topic. From 2001 to 2010, 'World' was the most common topic, followed by 'Association' and 'World Taekwondo'. From 2011 to 2020, 'World', 'Demonstration', and 'Kukkiwon' was the most common topic in that order. Second, as a result of analyzing news articles before 2000 by topic modeling, topics were divided into two categories. Specifically, Topic 1 was selected as 'South-North Korea sports exchange' and Topic 2 was selected as 'Adoption of Olympic demonstration events'. Third, as a result of analyzing news articles from 2001 to 2010 by topic modeling, three topics were selected. Topic 1 was selected as 'Taekwondo Demonstration Performance and Corruption', Topic 2 was selected as 'Muju Taekwondo Park Creation', and Topic 3 was selected as 'World Taekwondo Festival'. Fourth, as a result of analyzing news articles from 2011 to 2020 by topic modeling, three topics were selected. Topic 1 was selected as 'Successful Hosting of the 2018 Pyeongchang Winter Olympics', Topic 2 was selected as 'North-South Korea Taekwondo Joint Demonstration Performance', and Topic 3 was selected as '2017 Muju World Taekwondo Championships'.

Key Words : Topic modeling, Taekwondo, News article, Yearly, Keywords, Issue

*Corresponding Author : Hyosung Lim(limhyosung@gmail.com)

Received October 6, 2021

Accepted November 20, 2021

Revised October 13, 2021

Published November 28, 2020

1. 서론

1.1 서론

현대사회에서는 다양한 정보를 여러 경로를 통해 전달하고 있다. 정보에 대한 양과 가치가 높아짐에 따라 대중들은 정보를 공유하고 제공받는 일을 반복하여 수행하고 있다[1]. 실제로 사회 및 일상생활에서 생성되는 정보는 일상 및 사회생활을 수행하고 의사소통을 하는데 있어 필수적이라 할 수 있으며, 이러한 정보는 개인 일상의 질을 높이는데 기여하고 있다[2]. 이처럼 현대사회에서 정보의 가치는 지속적으로 높아질 수밖에 없다.

정보를 생성하고 전달하기 위한 수단으로는 다양한 도구들이 활용되고 있다. 일반적으로 대중들에게 정보전달은 언론을 통해 대부분 이루어지고 있으며, 전달매체로는 SNS, TV, 신문 등으로 구분할 수 있다. 특히, 언론에서 보고되는 정보는 사회현상에서 발생하는 주요 내용과 더불어 변화하고 있는 전반적인 내용을 제공함으로써 사전에 대비, 예방할 수 있는 장점이 있다. 또한 언론보도의 경우 현장중심의 많은 정보를 통해 다수 의견을 전달하는 측면에서 특정 분야의 쟁점이 발생할 경우 대표적으로 사용되고 있다[3, 4].

언론에 대한 중요성과 관심이 높아짐에 따라 많은 분야에서는 언론보도 분석에 대한 연구를 수행하고 있다. 교육학분야에서는 교육과정 개정 정책에 대한 언론분석[5]과 코로나19에 따른 온라인 교육에 대한 보도를 분석하였다[6]. 이 외에도 사회학, 정치학[7-9] 그리고 스포츠분야[10-11] 등에서 언론보도 분석에 대한 연구들이 보고되고 있다.

특히, 스포츠분야에서는 각 이슈에 따른 언론보도 분석(코로나, 사건사고 등)을 통해 본질적인 문제를 해결하기 위한 기초자료로 활용되고 있다. 예를 들어 종목에 따른 언론보도 분석에서는 축구, 농구, 야구 종목 등 다양한 종목에서 언론보도 분석을 통해 종목별 동향, 트렌드를 확인하고 있다[10, 12, 13].

한편 스포츠 종목 중 하나인 태권도 종목의 경우 과거부터 현재까지 다양한 이슈들이 발생함에 따라 언론에서는 태권도와 관련된 내용들을 보고하고 있다. 태권도와 관련된 이슈를 살펴보면 남북교류에 대한 언론보도부터 올림픽 정식종목 채택, 전자호구 도입, 경기규칙변화, 아시안게임 품새 종목 채택 등 다양하고 방대한 언론내용들이 보고되고 있다.

이에 따라 태권도분야에서도 언론보도를 바탕으로 전

자호구와 관련된 뉴스기사 분석[14], 태권도시범과 관련된 뉴스기사 분석[15], 2017 무주세계태권도 관련 소셜미디어 내용 분석[16] 등이 보고되고 있다. 현재 태권도 종목과 관련 언론보도 분석은 태권도분야의 발전과 문제점을 확인하는 기초자료로 활용되고 있다. 그러나 대부분 보고된 선행연구들은 하나의 이슈를 통한 언론보도를 분석하고 있어 태권도 분야의 전체적인 동향과 방향을 확인하는데 제한점이 있다. 하나의 특정 분야에서 전체적인 동향을 분석하는 것은 태권도에 대한 정체성을 확립하는 측면과 미래지향적인 발전을 위한 중요한 정보로 활용할 수 있는 장점이 있다[17]. 이미 선행연구에서는 동향분석에 대한 중요성을 강조하고 있어[4], 태권도 분야에서도 시계열적 언론자료에 대한 분석이 필요한 실정이다.

현재 언론보도 및 연구동향 분석에 있어서 다양한 방법론들이 나타나고 있으나, 최근 산업공학과 정보시스템학 등에서는 연구동향을 분석하는 방법으로 토픽모델링(Topic Modeling)을 소개할 수 있다. 토픽모델링은 내용의 주제를 도출하기 위해 텍스트 단어를 분석하는 방법으로 다양한 분야에서 적용되고 있다[18]. 더욱이 [19]가 개발한 토픽모델링의 LDA(latent dirichlet allocation) 알고리즘은 데이터 차원 축소의 유용성과 의미적인 측면에서 일관성을 지닌 주제를 생산하는 많은 장점을 가지고 있어[20], 학계에서 표준으로 적용되고 있는 알고리즘이다[4, 18].

따라서 본 연구에서는 토픽모델링을 적용하여 태권도 동향을 연도별로 구분하여 확인하는 것이 목적이다. 구체적으로 2000년 이전, 2001년~2010년, 2011년~2020년으로 3개의 구간으로 구분하여 태권도 뉴스기사 주제 동향을 확인하였다.

2. 연구방법

2.1 언론보도 자료

이 연구에서는 언론보도를 통한 태권도 동향을 살펴보기 위해 한국언론재단의 빅카인즈를 통해 뉴스기사 자료를 수집하였다. 또한 태권도의 경우 태권도 전문 언론보도에 대한 내용을 포함하기 위해 태권도신문과 무카스의 기사를 함께 수집하였다. 빅카인즈와 태권도 전문 언론에서 검색된 자료는 연구자료로 선정하였으며, 검색에 대한 조건은 다음과 같다. 첫째, 뉴스기사 검색어는 '태권도' 단어로 검색하였다. 둘째, 검색기간은 2000년 이전,

2001년~2010년, 2011년~2020년 3개의 구간으로 구분하여 검색하였다. 이러한 조건을 통해 검색된 기사는 총 13,238개이며, 중복기사를 제외하면 총 12,124개를 연구자료로 선정하였다.

2.2 텍스트 전처리 방법

언론보도 분석에 대한 주제를 텍스트 분석할 시 단어에 대한 정제작업은 반드시 수행되어야 한다[21]. 이에 따라 단어 정제에 대한 전처리 과정을 실시하였으며, 구체적인 내용은 다음과 같다. 첫째, 명사에 대한 단어만 추출하기 위해 Okt 형태소 분석기를 적용하여 단어의 토큰화(tokenization) 과정을 실시하였다. 둘째, 추출된 명사를 바탕으로 불필요한 단어(태권도, 뉴스 등)를 삭제하였다. 셋째, 동일한 의미를 지닌 단어이지만 다르게 표현되는 경우(체육관, 도장)에는 두 개의 단어를 하나의 단어로 통일시켜주는 과정을 거쳤다. 넷째, 단어 발생에 있어 연속적으로 함께 나타나는 단어의 경우 복합명사로 간주하였다. 예를 들어 ‘올림픽’과 ‘종목’에 단어가 연속적으로 빈도가 높을 때 ‘올림픽_종목’으로 변경하였다. 관찰빈도는 5회 이상으로 기준을 설정하였다.

2.3 조사도구의 타당도 및 신뢰도

본 연구에서는 언론보도 주제를 바탕으로 토픽모델링 방법을 적용하였다. 토픽모델링은 수집된 단어를 바탕으로 주제를 추출하는 방법으로 알고리즘은 크게 LSI(latent semantic indexing)와 PLSA(probabilistic latent semantic analysis) 그리고 LDA(latent dirichlet allocation) 등이 있으나, 이 연구에서는 LDA 알고리즘을 활용하였다.

토픽모델링의 경우 토픽 수를 어떻게 설정하는가에 따라 연구결과가 변화할 수 있어 토픽 수를 선정하는 것은 중요한 요소이다[22]. 선행연구 결과에 따르면 연구자 또는 전문성을 지닌 연구자들이 결과 해석 여부, 해석의 적합성을 고려하여 토픽 수를 선정해야 한다고 보고하였다[23-25]. 따라서 이 연구에서는 태권도 교수 2인, 태권도 전문가 3인을 선정하여 토픽 수를 선정하였으며, 이는 토픽 수를 선정하는 과정에 대한 오류를 최소화하기 위해 노력하였다. 최종적으로 선정된 연도별 토픽 수는 2000년 이전 2개, 2001~2010년은 3개, 2011년~2020년 3개의 토픽 수를 선정하여 분석을 실시하였다. 모든 분석은 Python 3을 적용하였다.

3. 연구결과 및 논의

3.1 연도별 언론기사 주제 단어 상위 20

Table 1은 연도별에 따른 언론기사 주제에서 나타나는 단어 상위 10개이다. 2000년 이전 1위는 ‘세계’, 2위는 ‘남북’, 3위는 ‘올림픽’으로 나타났으며, 2001년~2010년 1위는 ‘세계’, 2위는 ‘협회’, 3위는 ‘세계태권도연맹’으로 조사되었다. 2011년~2020년 1위는 ‘세계’, 2위는 ‘시범’, 3위는 ‘국기원’으로 나타났다.

Table 1. Result of Calculating the Top 20 Words in Frequency

Before 2000		From 2001 to 2010		From 2011 to 2020	
Word	n	Word	n	Word	n
World	61	World	628	World	720
North and South(Korea)	50	Association	403	Demonstration	633
Olympics	48	World Taekwondo	383	Kukkiwon	543
Park	32	Competition	377	Association	536
Competition	31	Kukkiwon	320	Competition	386
Culture	31	Gym	226	Gym	323
Host	30	Korea	208	North and South(Korea)	320
North Korea	27	Park	186	Korea	318
Korea	23	Host	144	Host	274
Sydney	21	North Korea	140	Performance	268

3.2 2000년 이전 토픽모델링 분석결과

2000년 이전 뉴스기사를 토픽모델링으로 분석한 결과 토픽주제는 2가지로 선정되었다. Topic 1이 차지하는 비중은 60.5%이며, Topic 2는 39.5%로 나타났다. 2가지 토픽에 대한 주요단어를 요약하면 다음과 같다.

3.2.1 남·북 체육교류

Topic 1에 포함된 키워드를 살펴보면 ‘남북’, ‘세계’, ‘스포츠’, ‘교류’, ‘북한’, ‘한국’, ‘문화’, ‘공원’, ‘유치’, ‘시드니’로 분석되었다. Topic 1에 대한 토픽명칭을 선정하기 위해 토픽 비중도가 높은 원문기사를 살펴본 결과 2000년 이전 국내에서 비중이 높은 스포츠 이벤트들이 개최됨에 따라 남북, 세계, 교류, 북한에 대한 키워드들이 나타났다. 따라서 Topic 1에 대한 명칭은 ‘남·북 체육교류’로 선정하였다.

1988년 서울올림픽이 개최되기 전 남·북 공동 올림픽 개최에 대한 기대감이 나타나면서 체육회담이 본격적으로

Table 2. Topic modeling analysis results (LDA algorithm) through the topics of Taekwondo articles before 2000

Topic	Percentage	Keywords(Probability Distribution)				
Topic 1	60.5%	North and South(Korea)(.156)	World(.145)	Sports(.092)	Exchange(.088)	North Korea(.078)
		Korea(.078)	Culture(.059)	Park(.052)	Host(.050)	Sydney(.029)
Topic 2	39.5%	Olympics(.258)	Event(.097)	Official Event(.063)	Competition(.063)	Permanent Event(.048)
		United States(.048)	Select(.045)	Missionary work(.031)	Association(.029)	Tigers(.026)

Table 3. Topic modeling analysis results (LDA algorithm) through the topics of Taekwondo articles from 2001 to 2010

Topic	Percentage	Keywords(Probability Distribution)				
Topic 1	37.3	Demonstration(.106)	Association(.055)	Performance(.034)	Culture(.028)	Kukkiwon(.019)
		Seoul(.017)	Corruption(.015)	Olympics(.014)	Gym(.013)	Initial(.013)
Topic 2	36.1	World(.123)	North Korea(.058)	World Taekwondo(.032)	Park(.029)	Muju(.029)
		Exchange(.027)	North and South(Korea)(.026)	Festival(.025)	Sports(.025)	Exchange between South and North Korea(.024)
Topic 3	26.6	Competition(.094)	Korea(.059)	United States(.056)	Reunification(.031)	Children(.019)
		Host(.018)	Health(.016)	Athlete(.015)	Gyeokpa(.014)	Host(.014)

로 이루어졌다. 이는 남·북 체육 교류를 이루는데 시발점이 되었으며, 서울올림픽 이후에도 지속적으로 발전되고 있다[26]. 특히 태권도 종목의 경우 1988년 서울올림픽에 시범종목으로 채택되면서 국제올림픽위원회에서는 분리된 태권도기관(세계태권도연맹, 국제태권도연맹)을 통합하기 위해 남·북 태권도기관인 대한태권도협회와 조선태권도위원회에 대한 협력을 강조하였다[27]. 이러한 남·북 체육교류에서는 태권도 종목이 하나의 이슈로 나타난 것으로 판단된다.

3.2.2 올림픽 시범종목 채택

Topic 2에 포함된 키워드를 살펴보면 ‘올림픽’, ‘종목’, ‘정식종목’, ‘대회’, ‘영구종목’, ‘미국’, ‘채택’, ‘선교’, ‘협회’, ‘타이거즈’로 분석되었다. Topic 2에 대한 토픽명칭을 선정하기 위해 토픽 비중도가 높은 원문기사를 살펴본 결과 1988년 서울올림픽 개최와 함께 태권도 시범종목 채택에 따라 올림픽, 종목, 정식종목, 대회에 대한 키워드들이 나타났다. 따라서 Topic 2에 대한 명명은 ‘올림픽 시범종목 채택’으로 선정하였다.

태권도종목은 1988년 서울올림픽대회에서 처음으로 시범종목에 채택되면서 세계적으로 태권도를 알리는 계기가 되었다. 특히, 국내에서 올림픽 개최와 동시에 태권도 종목이 시범종목 채택과 더불어 다음 올림픽에서 정식종목 채택에 대한 논의가 이루어짐에 따라 국내·외적

으로 적지 않은 관심과 가치를 인정받았다[28]. 또한 언론보도에 대한 키워드를 살펴보면 올림픽 잔류에 대한 관심이 높게 나타나는 것으로 해석할 수 있다. 따라서 2000년 이전 태권도 뉴스기사 주제에서는 전반적으로 올림픽과 남·북 관계에 대한 관심을 지닌 것으로 확인할 수 있다.

3.3 2001년~2010년 토픽모델링 분석결과

2001년~2010년 뉴스기사를 토픽모델링으로 분석한 결과 토픽주제는 3가지로 선정되었다. Topic 1에 대한 비중은 37.3%이며, Topic 2는 36.1%, Topic 3은 26.6%로 나타났다. 3가지 토픽에 대한 주요단어를 요약하면 다음과 같다.

3.3.1 태권도 시범공연 및 비리

Topic 1에 포함된 키워드를 살펴보면 ‘시범’, ‘협회’, ‘공연’, ‘문화’, ‘국기원’, ‘서울’, ‘비리’, ‘올림픽’, ‘도장’, ‘처음’으로 분석되었다. Topic 1에 대한 토픽명칭을 선정하기 위해 토픽 비중도가 높은 원문기사를 살펴본 결과 태권도 시범공연과 태권도 시도협회 비리에 대한 내용으로 시범, 협회, 공연에 대한 키워드들이 나타났다. 따라서 Topic 1에 대한 명명은 ‘시범공연 및 비리’로 선정하였다.

태권도 겨루기 종목이 발전함에 따라 태권도 품새, 시범 분야도 경기화를 위한 노력들을 수행하였다. 특히, 시

범 분야에서는 각 해외 시범공연을 통해 태권도 공연문화화를 정착하기 위한 방향과 방법을 모색하고[29], 이에 따라 해외 진출에 대한 영역까지 확장하는 노력을 하였다. 이 외에도 시범종목은 해외 태권도 봉사, 기술전과 등을 통해 태권도를 세계화하는데 이바지한 바 있다. 이러한 노력으로 시범종목은 현재 경기화 되는 과정을 거치고 있으며, 일부 대학에서는 시범대회를 개최하여 진행하고 있는 실정이다. 한편 Topic 1에서 추가로 나타난 '비리' 주제의 경우 시도협회 비리로 인해 승급 심사비, 공금횡령으로 인해 많은 뉴스기사들이 보도된 것으로 사료된다. 태권도 참가 수련생이 높아짐에 따라 각 시도에서는 태권도 자격취득을 위한 심사제도에서 심사비를 높게 산정하여 받는 사건들이 발생하여 사회적으로 적지않은 파장을 불러일으켰다. 이러한 사건을 통해 국기원에서 심사제도 개선에 대한 논의들이 수행된 것을 확인할 수 있었다.

3.3.2 무주태권도공원 조성

Topic 2에 포함된 키워드를 살펴보면 '세계', '북한', '세계태권도연맹', '공원', '무주', '교류', '남북', '축제', '스포츠', '남북교류'로 분석되었다. Topic 2에 대한 토픽명칭을 선정하기 위해 토픽 비중도가 높은 원문기사를 살펴본 결과 무주태권도공원과 관련된 내용으로 나타났다. 따라서 Topic 2에 대한 명칭은 '무주태권도 공원 조성'으로 선정하였다.

국내에서는 태권도공원 조성 유치를 위한 보도와 더불어 무주 지역이 태권도공원 선정됨에 따라 뉴스기사들이 적지 않게 보고되었다. 태권도공원의 경우 태권도 활성화와 함께 전 세계 태권도인들의 성지로 거듭나기 위해 진행하였다. 현재에는 태권도공원을 세계의 관광명소를 자리매김하기 위한 노력들을 수행하고 있으며[30], 이를 위해 태권도와 관련된 다양한 이벤트를(세계태권도대회, 시범공연, 연수 등) 진행하고 있다. 또한 무주태권도 공원과 관련된 '북한'과 '남북'에 대한 키워드는 세계태권도축제가 전주에서 개최되면서 북한 태권도팀이 한국으로 초청되는 내용들이 다루어 졌다. 이때 대회에 참가선수들은 같은 전라북도 지역에 있는 무주태권도공원 조성 예정지에 방문하여 무주태권도공원에 대한 홍보를 수행하면서 남북에 대한 보도가 된 것으로 해석할 수 있다. 따라서 2001년~2010년에는 태권도공원과 관련된 이슈들이 많이 보도된 것으로 판단된다..

3.3.3 세계태권도축제

Topic 3에 포함된 키워드를 살펴보면 '대회', '한국', '미국', '통일', '어린이', '개최', '건강', '선수', '격파', '유치'로 분석되었다. Topic 3에 대한 토픽명칭을 선정하기 위해 토픽 비중도가 높은 원문기사를 살펴본 결과 세계태권도와 관련된 내용으로 나타났다. 따라서 Topic 3에 대한 명칭은 '세계태권도축제'로 선정하였다.

태권도 종목이 세계화됨에 따라 다양한 축제들이 개최되었다. 2001~2010년에는 국내에서 세계태권도축제를 개최하였으며, 이때 평양예술단 공연을 통해 남·북 체육교류를 수행하였다. 또한, 2006년에는 평양 태권도축제 개최 소식을 통해 태권도 종목을 바탕으로 외교발전을 위한 노력 등을 수행한 것으로 나타났다. 따라서 태권도 종목은 단순히 하나의 스포츠 종목이 아닌 '한국 문화브랜드'로 나타나고 있으며 이는 한국의 언어, 생활 그리고 역사를 설명하는 'Korea Brand'로 설명할 수 있다[31]. 또한, 세계태권도의 축제는 남·북을 교류하는 연결점이 되고 있어 앞으로도 남·북 체육교육에 대한 긍정적인 시사점을 제시하고 있다.

3.4 2011년~2020년 토픽모델링 분석결과

2011년~2020년 뉴스기사를 토픽모델링으로 분석한 결과 토픽 주제는 3가지로 선정되었다. Topic 1에 대한 비중은 40.3%이며, Topic 2는 31.4%, Topic 3은 28.3%로 나타났다. 3가지 토픽에 대한 주요단어를 요약하면 다음과 같다.

3.4.1 2018 평창동계올림픽 성공 개최

Topic 1에 포함된 키워드를 살펴보면 '세계', '대회', '한국', '올림픽', '개최', '선수권', '평창', '독도', '한마당', '성공'으로 분석되었다. Topic 1에 대한 토픽명칭을 선정하기 위해 토픽 비중도가 높은 원문기사를 살펴본 결과 평창동계올림픽과 태권도 대회에 대한 내용으로 Topic 1에 대한 명칭은 '2018 평창동계올림픽 성공 개최'로 선정하였다.

2011~2020년에는 2018 평창동계올림픽을 개최하면서 대중들의 스포츠에 대한 관심이 보다 높아지는 시기였다. 태권도 종목에서도 2017 무주세계태권도선수권대회를 성공적으로 개최하면서 2018 평창올림픽에 대한 성공 개최를 기원하는 노력을 함께 하였다. 약 30년 만에 한국에서 개최되는 올림픽대회인 만큼 태권도 종목뿐만 아니라 많은 스포츠 종목에서도 다양한 홍보와 성공기원

Table 4. Topic modeling analysis results (LDA algorithm) through the topics of Taekwondo articles from 2011 to 2020

Topic	Percentage	Keyword(Probability Distribution)				
Topic 1	40.3	World(.093)	Competition(.046)	Korea(.037)	Olympics(.030)	Host(.019)
		Championship(.016)	Pyeongchang(.014)	Dokdo island(.014)	Hanmadang(.012)	Success(.012)
Topic 2	31.4	Demonstration(.177)	Performance(.088)	North Korea(.063)	Hallyu(.027)	Pyongyang(.025)
		United States(.025)	Kukkiwon(.021)	Opening(.019)	Festival(.019)	Festival(.017)
Topic 3	28.3	North and South(Korea)(.082)	Muju(.055)	Culture(.055)	Exchange(.030)	Taekwondowon(.029)
		Seou(.024)	Association(.024)	One(.019)	Joint performance(.018)	Spirit(.016)

을 위한 노력을 수행하였다. 특히 태권도 종목은 평창올림픽 개막식에서 남북 합동 공연을 통해 세계인들에게 남북 화합의 무대를 선보이면서 태권도를 통한 남북 발전된 내용을 전달하였다. 이에 2018 평창올림픽에서는 올림픽을 성공적으로 개최하는 것과 동시에 남북 관계에 대한 내용들이 보도되었다. 실제로 [32]의 연구에 따르면 평창올림픽 키워드를 통한 국내 언론기사를 분석한 결과 중심성이 가장 높은 키워드는 ‘김정은’으로 나타났으며, 2순위는 ‘남북’, 3순위는 ‘대통령’으로 보고하였다. 따라서 2018 평창올림픽에서는 국제 스포츠대회의 성공적인 개최와 더불어 남북관계 개선을 위한 노력들을 수행 하였다. 또한 2018 평창올림픽 경우 스포츠, 올림픽, 남북관계에 대한 연관성이 높은 것으로 해석할 수 있으며, 남북관계에서는 태권도 종목이 관계를 형성하는 매개체로 작용하였음을 확인할 수 있었다.

3.4.2 남북 태권도 합동시범공연

Topic 2에 포함된 키워드를 살펴보면 ‘남북’, ‘무주’, ‘북한’, ‘한류’, ‘평양’, ‘미국’, ‘국기원’, ‘개막’, ‘축제’, ‘페스티벌’로 분석되었다. Topic 2에 대한 토픽명칭을 선정하기 위해 토픽 비중도가 높은 원문기사를 살펴본 결과 남북 합동시범공연에 대한 내용으로 Topic 2에 대한 명칭은 ‘남북 태권도 합동시범공연’으로 나타났다.

2011~2020년에는 남북 태권도 합동시범공연이 평양에서 처음으로 개최되면서 태권도분야 뿐만 아니라 남북 외교적인 측면에서도 의미 있는 교류 행사로 인해 많은 미디어에서 관심을 가져 보도하였다. 또한 평양공연 외에도 서울시청에서 남북 태권도 합동시범공연과 함께 유럽 공연까지 수행하면서 태권도로 인한 남북관계 형성이 간접적으로 전 세계인들에게 긍정적인 메시지를 전달하고 있다. [33]의 연구에서도 남북 태권도 합동시범공연은 태권도 정신과 가치를 세계적으로 알림과 동시에 태권도

올림픽 평화정신을 보다 각인시키는 기회라고 보고하면서 남북 태권도 공연이 세계적인 평화와 관계성이 높음을 강조하였다 더욱이 향후에는 남북 태권도가 하나 되어 전 세계의 평화에 기여할 수 있는 대표적인 스포츠로 자리매김해야 한다고 보고하였다[33].

3.4.3 2017 무주세계태권도선수권대회

Topic 3에 포함된 키워드를 살펴보면 ‘남북’, ‘무주’, ‘문화’, ‘교류’, ‘태권도원’, ‘서울’, ‘협회’, ‘하나’, ‘합동공연’, ‘정신’으로 분석되었다. Topic 3에 대한 토픽명칭을 선정하기 위해 토픽 비중도가 높은 원문기사를 살펴본 결과 무주, 대통령, 북한 대한 내용으로 Topic 3에 대한 명칭은 ‘2017 무주세계태권도선수권대회’로 선정하였다.

2017 세계태권도선수권대회를 국내 무주에서 개최하였다. 세계태권도선수권대회는 대회규모와 더불어 평창동계올림픽을 준비하고 있는 시점이어서 보다 관심이 집중되었다. 개막식에서는 문재인 대통령이 직접 참석하여 축사를 통한 태권도발전, 남북 평화, 평창올림픽 성공 개최에 대한 언급 하였다. 또한 개막식 행사에서는 북한 태권도 시범공연이 진행되어 남북 태권도 발전을 위한 노력들이 수행되었다. 선행연구에서는 2017 무주세계태권도선수권대회와 관련한 기사 분석한 결과 ‘화해’, ‘평화’, ‘최초’에 대한 키워드 빈도가 높은 것으로 보고하고 있어 [16], 언론에서는 태권도를 통한 북한과의 교류에 관심을 두고 있음을 확인할 수 있었다.

4. 결론

이 연구에서는 토픽모델링을 적용하여 뉴스기사에 따른 태권도 동향을 연도별로 확인하였다. 이에 따른 결론은 다음과 같다. 첫째, 연도별에 따른 언론기사 주제를 분

석한 결과 2000년 이전 1위는 '세계', 2위는 '남북', 3위는 '올림픽'으로 나타났으며, 2001년~2010년 1위는 '세계', 2위는 '협회', 3위는 '세계태권도연맹'으로 조사되었다. 2011년~2020년 1위는 '세계', 2위는 '시범', 3위는 '국기원'으로 나타났다. 둘째, 2000년 이전 뉴스 기사를 토픽모델링으로 분석한 결과 '남·북 체육교류'와 '올림픽 시범종목 채택'이 이슈였다. 태권도를 통해 남·북 체육 교류에 대한 논의들이 이루어지면서 남·북 관계에 긍정적인 메시지들이 나타났다. 또한 2000년 태권도 종목이 올림픽 시범종목으로 채택되면서 전 세계인들에게 태권도를 알리는 계기가 되었다. 셋째, 2001년~2010년 뉴스 기사를 토픽모델링으로 분석한 결과 '태권도 시범공연 및 비리', '무주태권도공원 조성', '세계태권도축제'에 대한 내용이 이슈였다. 태권도 종목이 올림픽 정식종목에 채택과 더불어 전 세계인들에게 관심을 받게 되면서 다양한 공연과 축제를 통해 태권도를 알리기 위해 노력하였다. 넷째, 2011년~2020년 뉴스 기사를 토픽모델링으로 분석한 결과 '2018 평창동계올림픽 성공 개최', '남북 태권도 합동시범공연', '2017 무주세계태권도선수권대회'가 이슈였다. 최근에는 국내에서 세계대회와 올림픽과 같은 종합대회를 개최하면서 태권도 국제 행사를 주도적으로 진행하고 있다. 또한 태권도를 통해 남·북에 대한 스포츠 교류를 통해 긴장감을 해소할 수 있는 중요한 역할을 수행하였다.

끝으로 이 연구에서는 시계열적 뉴스 기사를 통해 태권도 동향을 분석하였다는 점에서 적지 않은 의미를 지니고 있다. 태권도 종목의 경우 대중 스포츠로 거듭나기 위해 과거부터 지금까지 지속적인 노력을 수행하였다. 또한 태권도라는 종목을 통해 남·북 관계 개선을 위해 다양한 접근을 시도하였다는 점에서 시사하는 바는 높을 수밖에 없다. 후속연구에서는 뉴스 기사에 대한 내용 외에도 전문가 인터뷰 혹은 학술자료 등에 대한 자료를 수집하여 연구를 수행한다면 보다 의미있는 정보를 생성할 수 있을 것이라 판단한다.

REFERENCES

- [1] S. Sohn, H. Hwang, & Y. Jang. (2011). An Analysis of Korean Media's Education Reporting and News Values. *Media & Education*, 1(1), 115-145.
- [2] H. S. Lee & E. R. Hwang. (2004). *Easy-to-Understand Sports Sociology*. Seoul: Sungshibn Women's University Press.
- [3] J. O. Kim & C. H. Kwon. (2020). Comparative Analysis of News Articles related to Airlines and Staff the Previous Corona19(2019) and After Corona19(2020). *Journal of The Korea Society of Computer and Information*, 25(7), 167-173.
DOI : <https://doi.org/10.978.896020/1576>
- [4] H. J. Yun, J. H. Park, J. Yoon, & M. S. Jeon. (2021). Issue analysis applying topic modeling on sports field due to spread of COVID-19. *The Korean Journal of Measurement and Evaluation in Physical Education and Sports Science*, 23(1), 11-20.
DOI : <https://doi.org/10.21797/ksme.2021.23.1.002>
- [5] Y. Yu. (2017). *Analysis of media coverage on 2015 revised curriculum policy using big data analysis*. Doctoral dissertation, Seoul National University, Seoul.
- [6] S. M. Kim. (2020). Analysis of Press Articles in Korean Media on Online Education related to COVID-19. *Journal of Digital Contents Society*, 21(6), 1091-1100.
DOI : <https://doi.org/10.9728/dcs.2020.21.6.1091>
- [7] J. W. Koo & S. G. Cho. (2017). South Korean Media Coverage of Official Development Assistance: Topic Modeling Analysis of Six Print Media from 1993 to 2016. *Journal of International and Area Studies*, 26(3), 173-210.
- [8] S. Y. Kim & S. H. Oh. (2016). A Study on the News Frame Analysis of Real Estate Policy. *Korean Journal of Contents*, 16(7), 221-230.
DOI : <https://doi.org/10.5392/JKCA.2016.16.07.221>
- [9] K. Kim, S. Lee, & K. Ko. (2020). A Study on Employment Policy Framing using Text Mining: A Comparative Analysis of Newspapers and National Assembly Minutes. *Korean Society and Public Administration*, 30(4), 135-163.
DOI : <https://doi.org/10.53865/KSPA.2020.02.30.4.135>
- [10] S. Yang & J. Kang. (2019). Crisis Communication Strategies of Professional Sport Clubs : A Case Study of Baseball Umpire Money Transaction Scandal by Big Data Analytics Tools. *Korean Society For Sport Management* 24(2), 15-31.
DOI : <https://doi.org/10.31308/KSSM.24.2.2>
- [11] J. H. Lee, J. M. Lee, J. H. Kim, & H. G. Kim (2017). A Study on the Perception Change in Marine Sports by Social Media Big Data Analysis. *Korean Journal of Sport Management*, 22(1), 31-46.
DOI : <https://doi.org/10.31308/KSSM.22.1.3>
- [12] S. K. Yoo, C. H. Lim, & J. Chang. (2018). Media Portrayal of Foreign Coaches in Korea and Vietnam. *Korean Journal of Communication Studies*, 26(4), 27-45.
DOI : <https://doi.org/10.23875/kca.26.4.2>
- [13] S. C. Choi & S. I. Lee. (2019). A Study on Development of Korean Basketball League through Big Data Sentiment Analysis. *Korean Journal of Sports Science*,

- 28(5), 441-462.
DOI : <https://doi.org/10.35159/kjss.2019.10.28.5.441>
- [14] W. J. Lee & C. H. Choi. (2020). Issues in News Articles related to Taekwondo Protector & Scoring System: Topic Modeling and Semantic Network Analysis Approach. *The Korean Journal of Measurement and Evaluation in Physical Education and Sports Science*, 22(4), 41-53.
DOI : <https://doi.org/10.21797/kmse.2020.22.4.004>
- [15] M. H. Kim & T. S. Park. (2017). Analysis of Taekwondo Demonstration Articles as shown on Press Reports. *Korean Journal of Sports Science*, 26(2), 117-130.
DOI : <https://doi.org/10.35159/kjss.2017.04.26.2.117>
- [16] J. H. Park. (2018). A Study on 2017 Muju World Taekwondo Championship Utilizing Big Data-Based on Social Media. *Taekwondo Journal of Kukkiwon*, 9(3), 77-96.
DOI : <https://doi.org/10.24881/tjk.2018.9.3.77>
- [17] K. H. Jung, Y. S. Won, & S. K. Koo. (2015). An Analysis of the Research Trend in Adapted Physical Activity and Exercise : 2011-2014. *Korean Journal of Adapted Physical Activity*, 23(1), 79-90.
DOI : <https://doi.org/10.17006/kjapa.2015.23.1.79>
- [18] S. J. Yang, B. Y. Lee, & H. W. Kim. (2016). A Topic Modeling Approach to the Analysis of Happiness and Unhappiness. *Knowledge Management Research*, 17(2), 165-185.
DOI : <https://doi.org/10.15813/kmr.2016.17.2.007>
- [19] D. M. Blei, A. Y. Ng, & M. I. Jordan. (2003). Latent dirichlet allocation. *The Journal of machine Learning research*, 3, 993-1022.
- [20] Mimno, D., Wallach, H., & McCallum, A. (2008, December). Gibbs sampling for logistic normal topic models with graph-based priors. In *NIPS Workshop on Analyzing Graphs* (Vol. 61).
- [21] H, J Yun, J, H Park, & J, W Yoon. (2019). Introduction of topic modeling for extracting potential information from unstructured text data: Issue analysis on news article of dementia-related physical activity. *Korean Journal of Sport Science*, 30(3), 501-512
DOI : <https://doi.org/10.24985/kjss.2019.30.3.501>
- [22] Y. Lee & J. Park. (2020). Analysis of the Policy Function and Status of Dance Using Topic Modeling. *Asian Journal of Physical Education and Sport Science*, 8(1), 67-81.
DOI : <https://doi.org/10.24007/ajpess.2020.8.1.006>
- [23] J. An, K. Ahn, & M. Song. (2016). Text mining driven content analysis of ebola on news media and scientific publications. *Journal of the Korean Society for Library and Information Science*, 50(2), 289-307.
DOI : <https://doi.org/10.4275/KSLIS.2016.50.2.289>
- [24] Y, S Lee. (2019). *Analysis of Dance Trends in Korea Using Big Data*. Unpublished doctoral dissertation, Doctoral dissertation, Sookmyung Women's University, Seoul.
- [25] H. Choi. (2019). Analysis of Korean Contemporary Dance Research Trends Using Text Mining. *The Korean Journal of arts Education*, 17(4), 103-118.
- [26] H. T. Kim & D. S. Kim. (2006). A Study of the Sports Exchanges and Cooperation between South and North Korea through Seeks an Educational Institution in Roles. *Korean Society Of Sports Science*, 15(4), 73-85.
- [27] S. B. Hong. (2017). A Study on Exchange and Blend of Taekwondo in North and South Korea. *Journal of Martial Art*, 11(2), 83-115.
- [28] J. H. Kim & I. H. Im. (2009). Phenomenological discussion on the aspect of Taekwondo competition. *Philosophy of Movement : Journal of Korean Philosophic Society for Sport and Dance*, 17(4), 203-216.
- [29] D. H. Yoo, J. K. Kim, & S. K. Cho. (2010). Analysis on Important Factors for Presenting the Direction of Taekwondo Demonstration Team Through Delphi/AHP. *Journal of Martial Arts*, 4(1), 63-80.
- [30] B. Yi. (2006). The difference of training motives and training satisfaction of Taekwondo practitioners according to the types of Lifestyles. *Korean journal of physical education*, 45(6), 279-290.
- [31] M. S. Kim, Y. B. Park, & D. K. Kim. (2010). Strategic Plans for The Commercialization of The Seoul General Assembly of The WTF as A Cultural Tourism Product. *Journal of Tourism and Leisure Research*, 22(6), 549-562.
- [32] W. K. Lee, C. K. Heo, & C. K. Jung. (2019). Analysis on Main Key Words of Domestic News Articles Associated Pyeong-Chang Olympics Games using Text Mining. *Korean Society Of Sports Science*, 28(2), 1275-1286.
DOI : <https://doi.org/10.35159/kjss.2019.04.28.2.1275>
- [33] A. Y. Kwak. (2020). A Study on the Meaning of Joint Demonstration and Achieving Process of Taekwondo Demonstration Between South and North Korea. *Journal of Sport and Leisure Studies*, 80, 9-22. DOI : <https://doi.org/10.51979/KSSLS.2020.04.80.9>

전 민 수(Minsoo Jeon)

[정회원]



- 2010년 2월 : 한국체육대학교 체육학과(체육학사)
- 2012년 8월 : 가천대학교 체육교육학과(교육학석사)
- 2018년 8월 : 한국체육대학교 체육학과(이학박사)
- 2020년 9월 ~ 현재 : 한국체육대학교

학술연구교수

- 관심분야 : 빅데이터 분석, 선수평가, 스포츠 경기분석
- E-Mail : minsul144@nate.com

임 효 성(Hyosung Lim)

[정회원]



- 2009년 8월 : 경기대학교 골프지도학과(체육학사)
- 2013년 2월 : 한국체육대학교 체육학과(체육학석사)
- 2017년 2월 : 한국체육대학교 체육학과(이학박사)
- 2021년 6월 ~ 현재 : 상명대학교 산학

협력단 특임교수

- 관심분야 : 스포츠교육, 코칭, 체육정책
- E-Mail : limhyosung@gmail.com