

SNS 대화 분석을 통한 주제별 적합 광고 시간대 도출

이지민*, 전예림*, 이지선^o, 우지영*

*순천향대학교 빅데이터공학과,

^o순천향대학교 빅데이터공학과

e-mail: {dfwlals7359, 9261190, 10leegisun}@naver.com, jywoo@sch.ac.kr

When is the best time to run SNS AD per topic?: through conversation data analysis

Jimin Lee*, Yerim Jeon*, Jisun Lee^o, Jiyoung Woo*

*Dept. of Bigdata Engineering, Soonchunhyang University,

^oDept. of Bigdata Engineering, Soonchunhyang University

● 요약 ●

본 논문에서는 시간대와 대화 주제를 활용하여 카테고리별로 적절한 SNS 광고 시간대 예측 방법을 제시한다. 위의 분석으로 광고주들에게 적절한 광고시간을 제안할 수 있다. 연관규칙분석 알고리즘인 apriori를 사용하였다. 주제는 상거래(쇼핑), 미용과 건강, 시사/교육, 식음료, 여가생활로 추려서 분석하였다. 연관분석 결과, 미용과 건강이 18시, 17시, 16시에 가장 활발히 대화를 나누었다. 상거래(쇼핑)이 14시, 16시, 17시 순으로 가장 활발히 대화를 나누었으며, 시사/교육이 15시, 17시, 16시 순으로 많은 대화를 나누었으며, 식음료가 18시, 17시, 19시 순으로 대화를 많이 나누는 것을 확인했다. 마지막으로, 여가생활은 22시, 23시, 21시 순으로 각각의 대화 주제별로 가장 많이 대화를 나누는 시간대가 달라지는 것을 확인할 수 있었다. 이를 통해 소비자 입장에서는 알맞은 광고를 적절한 시간대에 추천받을 수 있다.

키워드: 광고(advertisement), 연관분석(Association analysis), 시간(time), 미용과 건강(beauty and health), 상거래(shopping), 시사/교육(Current affairs/education), 식음료(food and beverage), 여가생활(leisure life)

I. Introduction

현 코로나19 상황이 지속되면서 나날이 경제 피해가 심각하다. 따라 집에서 보내는 시간이 증가하며 코로나 이전 대비 온라인 이용이 증가하는 추세이다. 이로 인해 광고는 지영업자들의 부담을 덜어주기에 적합하여 효과가 극대화되고 있다. 국민의 대부분이 사용하고 있는 SNS(카카오톡)를 활용하여 어떤 카테고리가 어느 시간대에 활발하게 다루어지는지 분석하고자 한다. 분석을 토대로 해당 시간대에 적절한 광고를 노출시키고자 한다.(Fig. 1.)



Fig. 1. 배너 광고

II. Method

한국인의 SNS 데이터(카카오톡 대화 내용)은 AI hub의 음성/자연어 부문에서 가져왔다 [1]. 구축 데이터량은 대략 200만건으로 발화 시간, 본문, 주제, 말차례 수 등으로 이루어진다. 본 데이터는 2020년 10월부터 구축되었으며 4,174,668명이 참여하였다. 대화 내용의 주제는 개인 및 관계, 주거와 생활, 상거래(쇼핑), 식음료 외 12개가 있지만, 그 중 가장 많은 대화가 포함되어 있는 주제 5가지를 선정하였다. 연관규칙분석 알고리즘으로 시간대와 대화 내용 주제 간의 연관분석을 실시했다. 어느 시간대에 어떤 주제의 광고를 넣는 것이 효과적인지 예측가능하다.

III. Results

1. 미용과 건강

dialogueinfo.topic이 미용과 건강일 때 18시, 17시, 16시 순으로 SNS 대화가 활발하다. 연관분석 결과 lift 값이 1이 넘어 유의미함을

확인했다(Fig. 2.) 18시의 대화 본문을 확인하면 치료 및 수술, 운동, 미용 다이어트 등의 키워드를 확인하여 이와 관련된 광고를 노출시켜 마케팅의 효율을 증가시킬 수 있다.

lhs	rhs	support	confidence	coverage	lift	count
[2] {time=18}	=> {topic=미용과 건강}	0.0045363576	0.07098121	0.063909273	1.0472523	68
[3] {topic=미용과 건강}	=> {time=18}	0.0045363576	0.06692913	0.067778519	1.0472523	68
[4] {time=17}	=> {topic=미용과 건강}	0.0045363576	0.08732673	0.087378252	0.9933344	68
[5] {topic=미용과 건강}	=> {time=17}	0.0045363576	0.06692913	0.067778519	0.9933344	68
[6] {time=16}	=> {topic=미용과 건강}	0.0044696464	0.06620553	0.067511674	0.9769723	67
[7] {topic=미용과 건강}	=> {time=16}	0.0044696464	0.06594488	0.067778519	0.9769723	67

Fig. 2. 미용과 건강

2. 상거래(쇼핑)

dialogueinfo.topic이 상거래(쇼핑)일 때 14시, 16시, 17시 순으로 SNS 대화가 활발하다. 연관분석 결과 lift 값이 1이 넘어 유의미함을 확인했다(Fig. 3.) 14시의 대화 본문을 확인하면 중고거래, 의복, 물건 및 가격 등의 키워드를 확인하여 이와 관련된 광고를 노출시켜 마케팅의 효율을 증가시킬 수 있다.

lhs	rhs	support	confidence	coverage	lift	count
[2] {time=14}	=> {topic=상거래(쇼핑)}	6.537692e-03	0.10071942	0.06409940	1.1963425	98
[3] {topic=상거래(쇼핑)}	=> {time=14}	6.537692e-03	0.07765452	0.084189460	1.1963425	98
[4] {time=16}	=> {topic=상거래(쇼핑)}	6.337558e-03	0.09387352	0.067511674	1.1150270	95
[5] {topic=상거래(쇼핑)}	=> {time=16}	6.337558e-03	0.07527734	0.084189460	1.1150270	95
[6] {time=17}	=> {topic=상거래(쇼핑)}	5.937292e-03	0.08814811	0.067378252	1.0466272	89
[7] {topic=상거래(쇼핑)}	=> {time=17}	5.937292e-03	0.07622298	0.06409940	1.0700072	89
[8] {time=15}	=> {topic=상거래(쇼핑)}	5.870580e-03	0.09233998	0.063575717	1.0968116	88
[9] {topic=상거래(쇼핑)}	=> {time=15}	5.870580e-03	0.06973059	0.084189460	1.0968116	88
[10] {time=18}	=> {topic=상거래(쇼핑)}	5.537025e-03	0.08663883	0.063909273	1.0290936	83

Fig. 3. 상거래(쇼핑)

3. 시사/교육

dialogueinfo.topic이 시사/교육일 때 15시, 17시, 16시 순으로 연관분석 결과 lift 값이 1이 넘어 유의미함을 확인했다(Fig. 4.) 15시의 대화 본문을 확인하면 임시, 자기 계발, 외국어 학습 등의 키워드를 확인하여 이와 관련된 광고를 노출시켜 마케팅의 효율을 증가시킬 수 있다.

lhs	rhs	support	confidence	coverage	lift	count
[2] {topic=시사/교육}	=> {time=15}	0.0047364910	0.08142202	0.058172115	1.2807094	71
[3] {time=15}	=> {topic=시사/교육}	0.0047364910	0.07450157	0.063575717	1.2807094	71
[4] {topic=시사/교육}	=> {time=17}	0.0044696464	0.07683486	0.058172115	1.1403511	67
[5] {time=17}	=> {topic=시사/교육}	0.0044696464	0.06633663	0.067378252	1.1403511	67
[6] {topic=시사/교육}	=> {time=16}	0.0040693796	0.06995413	0.058172115	1.0361782	61
[7] {topic=시사/교육}	=> {time=16}	0.0040693796	0.06022869	0.07281553	1.0361782	61
[8] {time=20}	=> {topic=시사/교육}	0.0040026684	0.07281553	0.054969980	1.2517258	60
[9] {topic=시사/교육}	=> {time=20}	0.0040026684	0.06880734	0.058172115	1.2517258	60
[10] {time=21}	=> {topic=시사/교육}	0.0034689793	0.08364749	0.054503002	1.0941237	52

Fig. 4. 시사/교육

4. 식음료

dialogueinfo.topic이 식음료일 때 18시, 17시, 19시 순으로 SNS 대화가 활발하다. 연관분석 결과 lift 값이 1이 넘어 유의미함을 확인했다(Fig. 5.) 18시의 대화 본문을 확인하면 음식, 배달, 식사 메뉴 등의 키워드를 확인하여 이와 관련된 광고를 노출시켜 마케팅의 효율을 증가시킬 수 있다.

lhs	rhs	support	confidence	coverage	lift	count
[2] {time=18}	=> {topic=식음료}	0.0086724483	0.13569937	0.063909273	1.3260323	130
[3] {topic=식음료}	=> {time=18}	0.0086724483	0.08474576	0.102334890	1.3260323	130
[4] {time=17}	=> {topic=식음료}	0.0083388926	0.12376238	0.067378252	1.2093859	125
[5] {topic=식음료}	=> {time=17}	0.0083388926	0.08148631	0.102334890	1.2093859	125
[6] {time=19}	=> {topic=식음료}	0.0080720480	0.15963061	0.050567045	1.5598845	121
[7] {topic=식음료}	=> {time=19}	0.0080720480	0.07887875	0.102334890	1.5598845	121
[8] {time=16}	=> {topic=식음료}	0.0075383589	0.11666008	0.067511674	1.0911242	113
[9] {topic=식음료}	=> {time=16}	0.0075383589	0.07366362	0.102334890	1.0911242	113
[10] {time=11}	=> {topic=식음료}	0.0067378252	0.11716937	0.057505003	1.1449602	101

Fig. 5. 식음료

5. 여가 생활

dialogueinfo.topic이 여가 생활일 때 22시, 23시, 21시 순으로 SNS 대화가 활발하다. 연관분석 결과 lift 값이 1이 넘어 유의미함을 확인했다(Fig. 6.) 18시의 대화 본문을 확인하면 취미, 관심사 등의 키워드를 확인하여 이와 관련된 광고를 노출시켜 마케팅의 효율을 증가시킬 수 있다.

lhs	rhs	support	confidence	coverage	lift	count
[2] {time=22}	=> {topic=여가 생활}	0.0101400934	0.15784008	0.064242829	1.3504697	152
[3] {topic=여가 생활}	=> {time=22}	0.0101400934	0.08675799	0.116877919	1.3504697	152
[4] {time=23}	=> {topic=여가 생활}	0.0094062708	0.17195122	0.054703135	1.4712036	141
[5] {topic=여가 생활}	=> {time=23}	0.0094062708	0.08047945	0.116877919	1.4712036	141
[6] {time=21}	=> {topic=여가 생활}	0.0082721815	0.15137479	0.054503002	1.2985754	124
[7] {topic=여가 생활}	=> {time=21}	0.0082721815	0.07077626	0.116877919	1.2985754	124
[8] {time=15}	=> {topic=여가 생활}	0.0076050700	0.11962225	0.063575717	1.0234803	114
[9] {topic=여가 생활}	=> {time=15}	0.0076050700	0.06506849	0.116877919	1.0234803	114
[10] {time=00}	=> {topic=여가 생활}	0.0075383589	0.15026596	0.050166778	1.2856659	113

Fig. 6. 여가 생활

아래의 그래프를 통해 각 topic별 시간대 support분포를 한눈에 확인할 수 있다. (Fig. 6.)

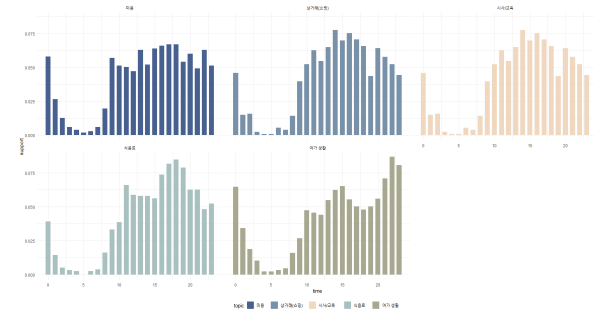


Fig. 7. topic별 시간에 따른 support

IV. Conclusions

본 논문에서는 SNS 대화 시간대와 본문을 토대로 적절한 광고 시간대 예측을 하였다. 소비자는 알맞은 광고를 적절한 시간대에 추천받을 수 있고, 광고주는 높은 효율로 광고가 가능하다. 성별과 연령대에 따른 분석이 추가로 행해진다면 더욱 개인맞춤화 된 광고로 더 좋은 마케팅 효과를 기대해 볼 수 있다.

REFERENCES

[1] AI hub, <https://aihub.or.kr/aidata/30718>.