

Ontology-Based Natural Language Processing of Social Media Data in the Assessment of Health Information Sought During Pregnancy

Joo Yun LEE^{a,1}

^a*College of Nursing, Gachon University, Incheon, Korea*

Abstract. This study analyzed collected social media data from South Korea containing keywords related to “pregnancy” using ontology-based natural language processing. Of the 504,725 documents, those containing concepts related to “maternal emotion” were the most frequent, followed by “family support”. Social media were used as a means of exchanging information and expressing emotions.

Keywords. Pregnant women, social networking, consumer health informatics

1. Introduction

Social media constitute a new platform for health-related content created by users. Social media offer a venue for sharing information and experiences and providing social and emotional support [1]. Pregnant women also use social media for the exchange of information, peer support, and self-empowerment [2]. Women experiencing the life-changing experience of pregnancy often seek assurance as to whether related symptoms or circumstances are normal [2], and the immediate response enabled by social media often satisfies this need. Analyzing social media use by pregnant women is necessary to identify their information needs and to provide appropriate information, mediation, and support. The purpose of this study was to analyze collected social media data of South Korea using an ontology developed to collect and analyze information on the health-related issues of pregnant women discussed in social media [3].

2. Methods

The dataset used for analysis were web documents containing keywords related to “pregnancy” (e.g., pregnant, maternity, mom to be). Data was collected from online communities in two biggest web portal site of South Korea from August 1, 2017, to July 31, 2018 using a crawler. Ontology based NLP was performed for the collected 504,725 documents. The ontology used consists of 13 super-classes and 454 bottom classes (Figure 1). Whether or not a concept appeared in one document was counted by document frequency, and the concept was analyzed by each super-class and bottom class

¹ Corresponding Author, Jooyun Lee, PhD, RN, College of Nursing, Gachon University, 191, Hambangmoe-ro, Yeonsu-gu, Incheon, 21936, South Korea; E-mail: jooyun@gachon.ac.kr.

level. The document can be counted repeatedly for different concepts since multiple concepts can be included in a document.

3. Results

The super-class concepts with high frequency are shown in Table 1. Documents containing the concept related to “maternal emotion” were the most common, with a frequency of 50.2%. Concepts reflecting positive emotions, such as “like” (13.7%) and “excellent” (7.5%), and those indicative of negative emotions, such as “anxiety” (7.4%) and “feel restricted” (6.9%), were included at the class level. Documents containing concepts related to “family support” (25.3%) occurred with the second highest frequency. “husband's parents” (10.1%) and “my parents” (7.2%) were examples at the class level. Concepts related to “maternal health-physiology”, including the class level concepts “pain” (6.5%), “increased heart rate” (4.0%), and “morning sickness” (3.6%) were detected in 24.6% of the documents. Documents containing the concept “infant care preparation” (20.8%) included a high frequency of the concept “breast feeding” (12.3%) at the class level. The examples of concepts with the super-class level “fetal condition” were “monitoring fetal growth” (9.6%), “twin” (4.3%), and “breech (presentation)” (3.0%). The super-class “maternal life-style” included the concepts “food” (6.7%), “exercise” (5.9%), and “exposure to smoking” (4.6%).

Table 1. The super-class concepts with high frequency.

Rank	Super class	Frequency	(%)
1	Maternal emotion	253,336	(50.2)
2	Family support	127,739	(25.3)
3	Maternal physiology	124,252	(24.6)
4	Infant care preparation	104,783	(20.8)
5	Pregnancy stage	96,661	(19.2)
6	Fetal condition	91,051	(18.0)
7	Maternal life-style	87,366	(17.3)
8	Medical check-up	77,529	(15.4)
9	Medical-hospital information	52,604	(10.4)
10	Postpartum care preparation	48,518	(9.6)

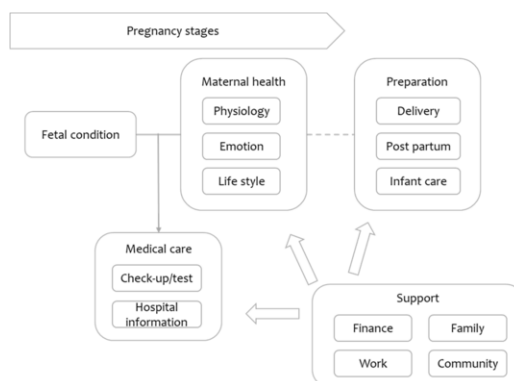


Figure 1. Ontology for analyzing health related issues of pregnant women.

4. Conclusions

In addition to asking questions about pregnancy, pregnant women used social media to express their emotions about pregnancy. These results support the need to provide nursing care related to “maternal emotion” and “family support” during pregnancy. Further studies are necessary to identify the detailed emotional and information needs of pregnant women.

Acknowledgements

This work has supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2018R1C1B5085259).

References

- [1] Benetoli A, Chen TF, Aslani P. Consumer health-related activities on social media: Exploratory study. *J Med Internet Res* 19 (2017), e352.
- [2] Cohen JH, Raymond JM. How the internet is giving birth (to) a new social order. *Information, Communication & Society* 14 (2011), 937-957.
- [3] Lee JY. An ontology for assessing health information needed during pregnancy. *Stud Health Technol Inform* 264 (2019), 1520-1521.