

Topic Modeling for International Patients' Consultations Using Natural Language Processing

Sunmi LEE^a, Sookkyung SUNG^a, Seo Young KANG^a, Sujin HA^a and Eunjoo JEON^{b,1}

^a*International Healthcare Center, Asan medical center, Seoul, Republic of Korea*

^b*Technology Research, Samsung SDS, Seoul, Republic of Korea*

Abstract. We extracted major topic by applying natural language processing and keyword extracting using TF, TF-IDF, TextRank, Yake, KeyBERT. 1452 consultation data were collected from the website and official hospital e-mail. We found six topics categorized into "Medical opinion" related to hospital characteristics and "Non-medical service guidance". Based on this result, it is necessary to establish marketing plan and develop a digital solution for effective consultation.

Keywords. International patients, text consultations, Natural Language Processing, Topic Modeling.

1. Introduction

With the expansion of the global healthcare industry, the number of international patients entering Korea to receive treatment has increased [1]. As such, Korean hospitals were asked various questions related to medical travel thorough the text consultation. The purpose of this study was to organize and elicit meaning from these inquiries and understand the diverse needs of international patients.

2. Methods

From January 1, 2019, to May 31, 2021, 1,452 text consultation in Mongolian, Russian, Chinese, or English received from either website or official hospital e-mail. The text data were translated into English, and then went through the Natural Language Processing as tokenizing, removing stopwords, lemmatization by using Python NLTK library. Next, keyword extraction was performed in five ways (TF, TF-IDF, TextRank, Yake, KeyBERT) and the BERT model was chosen because it reflects accurate meaning with high performance [2]. Lastly, clustering was conducted by the fifty most frequent keywords using Latent Dirichlet Allocation and the final subjects were derived by six health professionals including physician, nurse, and coordinator.

¹ Corresponding Author, Eunjoo Jeon, Technology Research, Samsung SDS, Seoul, Republic of Korea; E-mail: yesredpig@gmail.com.

3. Results

Six topics were initially identified after summarizing the expert’s responses (Table 1).

Table 1. Category, Topics, and Tokens Identified in Text consultation for international patients

Category	Topics	Tokens
Medical opinion	The possibility of liver transplant	liver, suffer, available, treatment, availability, transplant, possible, transplantation, surgical
	Treatment for cancer disease	cancer, chemotherapy, tumor, available, treatment, want, carcinoma, metastasis, treat
	Confirming the exact diagnosis	diagnosis, diagnose, thyroid, breast, lung, blood, disease, pain, ultrasound, biopsy, surgical, diabetes, kidney, fracture, review, bone
Non-medical service guidance	Hospital information	email, clinic, appointment, consultation, surgeon, hospital, professor, doctor
	Length of stay	long, stay, plan, therapy, test, expect, hospitalization
	Estimated cost of treatment	cost, estimate, much, examination, surgical, need, check, therapy, medicine, ultrasound, test

4. Discussion

The main reason for medical opinion category is that the hospital has achieved excellent performance in the field of liver transplantation and cancer treatment [3, 4]. In addition, non-medical inquiry about hospital information, treatment period, and estimated cost was revealed, which can be considered essential at the preliminary text consultation of medical tourism [5].

5. Conclusion

To identify international patients' inquiries to Korean hospitals, this study extracts hidden information from online consultations using NLP. By applying NLP, it is possible to reduce time consuming work and provide objective indicators to establish marketing plan and develop a digital solution (e.g., chatbot) for effective consultation in distant.

References

[1] STATISTA Inc (2021). *Medical tourism in South Korea*. Retrieved January,2022. from <https://www.statista.com/study/72966/medical-tourism-in-south-korea/>

[2] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

[3] Kim S-H, Hwang S, Song G-W, Jung D-H, Ahn C-S, Kim K-H, et al. Institutional standard framework and experience of living donor liver transplantation for overseas non-Korean patients at Asan Medical Center. 2021;35(2):93-9.

[4] 2022 NEWSWEEK DIGITAL LLC (2022). *World's Best Specialized Hospitals 2022*. Retrieved January 2022. from <https://www.newsweek.com/worlds-best-specialized-hospitals-2022/oncology>

[5] Eom T, Yu J, Han H. Medical tourism in Korea – recent phenomena, emerging markets, potential threats, and challenge factors: a review. *Asia Pacific Journal of Tourism Research*. 2019;24(6):563-73. doi: 10.1080/10941665.2019.1610005.