

Team 4. Analysis of Public Needs through Text Mining on the National Petition Board

1. Project Objective

The National Petition Board of South Korea(국민청원 게시판) is a vital platform where citizens can directly express their opinions and raise social issues.

Through analyzing text data from the National Petition Board using text mining techniques, we could understand key issues and public concerns in Korean society. It also will provide insights to policymakers and highlight sensitive social issues.

2. Background for Topic Selection

- **Increasing Public Participation:** Growing engagement in political and social issues, making the petition board a crucial resource for understanding social trends.
- **Potential for Efficient Data Analysis:** Text mining can efficiently analyze massive data, quantifying public needs and sentiments to inform policy development.

3. Research Methodology

- **Data Collection & Preprocessing:** Use web crawling tools to collect data from “Petition24(<https://www.cheongwon.go.kr/portal/petition/open/view>)” and preprocess with tokenization, stopword removing, stemming, and normalization.
- **Keyword Extraction & Text Representation:** Apply TF-IDF, Bag of Words, and N-gram analysis to extract and analyze keywords.
- **Topic Modeling & Sentiment Analysis:** Utilize LDA for topic classification and sentiment analysis to gauge public sentiment on key issues.
- **KoNLP:** Korean Natural Language Processing package

4. Expected Outcomes

Through this project, we can identify significant issues and trends in Korean society, providing insights to develop policies that align with public demands. Additionally, it enables quick responses to emerging issues and public expectations.