텍스트 마이닝 기법을 이용한 학습 수요자 요구에 관한 연구 : SNS를 중심으로

이명숙*, 이경미*, 임영규O, 한경임*, 박혜정*
O*계명대학교 교양교육대학
e-mail:hye3255@daum.net*, {woods, mslee, yklim, kyungim}@gw.kmu.ac.krO*

A Study on Learners' Needs Analysis Using Text Mining Techniques : Focusing on SNS

Myung-Suk Lee*, Kyung-Mi Lee*, Youg-Kyu LimO, Kyung-Im Han*, Hye-Jung Park*
O*College Liberal Arts Education, Keimyung University

● 요 약 ●

본 연구는 교양교육에 대한 학습 수요자의 요구와 현재 편성되어 있는 교양교육 교과목들에 대한 차이를 알아본다. 학습 수요자의 다양한 생각들을 SNS를 통해 데이터를 수집하고, 텍스트 마이닝 기법을 이용하여 유용한 정보를 발견하고 시각화 분석을 통해 학습자의 요구를 제시한다. 분석 결과로는 학습자는 교수자와 상호작용 잘되는 수업 방식, 학습자가 참여할 수 있는 수업, 자기주도 학습을 선호하였다. 또한 교양교육 교과목 개설로서는 취업에 필요한 외국어, 자격증 취득이 가능한 과목, 실생활에 적용할 수 있는 실용적인 과목들을 요구하여 실제 균형에 맞게 개설된 교과목과는 차이를 보임을 알 수 있었다.

키워드: 텍스트 마이닝(Text Mining), 소셜네트워크서비스(SNS), 상호작용(Interaction)

I. 서론

과학기술 발달과 맞물려 현재 대학의 교육과정들이 빠르게 변화되고 있으며, 그 속에서 교양교육도 중요한 전환기를 맞고 있다는 데에는 많은 연구자들이 인식하고 있다. 현재 교양교육의 중요성은 그 어느 때보다도 부각되고 있는 것이 사실이다. 이에 따라 각 학교에서 주체적으로 교육과정을 편성하고 교양으로써의 역할을 제대로 수행하려고 노력하고 있다. 현재 기하급수적으로 생겨나고 있는 교양교육 교과목이 수요자인 학생들의 필요성과 요구에 부합되게 개설되고 있는지 이 시점에서 분석해 볼 필요가 있다.

따라서, 본 연구에서는 정형화된 문항 형태인 설문조사 방법 대신 학습자들이 더 자유로운 환경에서 자신들의 의견을 표현할 수 있는 형태인 SNS를 활용하여 답변을 비정형데이터로 수집한다. 수집한 데이터는 텍스트 마이닝 기법을 이용하여 학습자들이 교양과목에서 요구하는 교과목을 분석하고 실제 개설된 교과목과 어떠한 차이가 있는지 분석하고자 한다.

자료 분석과정은 수집한 데이터를 텍스트로 저장하여 R프로그램을 이용하여 전처리과정을 거쳐 시각화 자료를 생성한다. 데이터 마이닝 기법을 이용하여 유용한 정보를 발견하고 시각화된 자료로 표현한다. 분석한 결과를 현재 개설된 교양교과목과 그 차이를 분석한다.

II. 연구 설계

1. 자료생성 및 자료수집

본 연구에서 자료 수집은 K대학교 교양교육 수업을 듣는 재학생을 120명을 대상으로 전공, 학년, 남녀 불문 무작위로 조사하였다. 자료수집 방법은 6개의 개방형 질문에 대하여 학생들이 자유롭게 카카오톡을 통해 응답한 비정형 데이터를 수집하였다. [Table 1]은 본 논문의 목표인 교양교육 교과목에 있어 학습자 요구를 알아보기 위한 방법으로 다음과 같은 질문 문항들을 구성 하였다.

Table 1. Survey Questions

번호	설문 문항
질문1	본인이 선호하는 강의 방식은 무엇입니까?
질문2	그 이유는 무엇입니까?
질문3	여러분들이 교양과목에서 수강하기를 원하는 과목은 무엇입니까?
질문4	위 과목을 원하는 이유를 기술해 주세요.
질문5	기억에 남는 최고의 수업 방식과 최악의 수업 방식은 무엇입니까?
질문6	그 이유를 마음껏 소개해 주세요.

2. 텍스트 마이닝 처리과정

본 연구에서는 텍스트 마이닝의 처리단계를 [Fig. 1]과 같이 4가지 단계로 구성하였다.

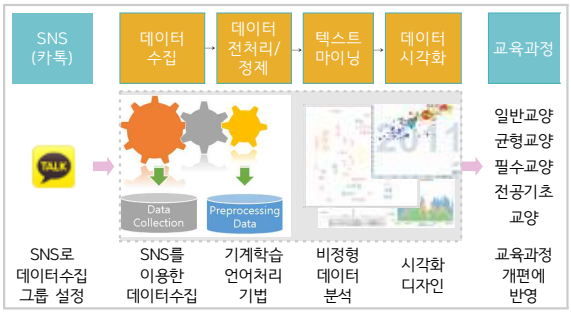

Fig. 1. Text Mining Processing Process

먼저 SNS 중 카카오톡을 이용하여 교수자와 학습자간의 데이터를 수집할 그룹을 설정하면 첫 번째 단계에서 각 질문에 따른 비정형 데이터를 수집하게 된다. 두 번째 단계에서는 수집한 데이터를 전처리와 정제과정을 거쳐 명사들을 제외한 불필요한 데이터들을 제거한다. 세 번째 단계에서는 텍스트 마이닝 단계를 거쳐 빈도별 단어를 분류하여 학습자의 의도를 분석한다. 네 번째 단계에서는 이 값들을 단어별 빈도수로 분류하여 텍스트를 시각화하는 단계를 거치게 된다.

III. 시각화 결과

본 연구의 시각화 결과는 다음과 같이 요약할 수 있다. 첫째, 학생들이 가장 선호하는 강의 방식은 [Fig. 2]와 같이 상호작용이 있는 강의식, 설명식+토론, 자기주도학습, 이론-실습 등의 순으로 나타남을 알 수 있었다. 그 이유로는 학생들이 참여하는 수업, 교수자와 학습자가 상호작용할 수 있기 때문이라고 하였다. 둘째, 학생들이 교양과목에서 수강하기를 원하는 과목의 경우에는 [Fig. 3]과 같이 외국어, 자격증 취득이 가능한 과목, 실생활에 적용할 수 있는 실용적인 과목 등의 순으로 요구하고 있음을 알 수 있었다. 그 이유로는 취업과, 사회에 나가서 바로 활용하는 것이 필요하기 때문이라고 하였다. 셋째, 최고의 수업 방식은 [Fig. 4]와 같이 토론식 수업, 프로젝트형 수업 등이 있었으며, 그 이유로는 상호작용과 소통을 얘기 했다. 넷째, 최악의 수업은 [Fig. 5]와 같이 교수님 혼자 강의, 토론, 종교수업을 들었다. 그 이유로는 상호작용이 없거나 토론 방식이 허술하다는 이유를 들었다.

Fig. 2. A Figure for the Question 1

Fig. 3. A Figure for the Question 2

Fig. 4. A Figure for the Question 5-1

Fig. 5. A Figure for the Question 5-2

VI. 결론

본 연구는 텍스트 마이닝 기법을 이용하여 SNS를 통한 교양교육 교육과정에서 학습자 요구를 분석하여 시각화 결과를 보였다. 학생들이 요구하는 수업방식은 학습자와 교수자가 상호작용을 할 수 있는 수업방식과 학습자가 참여하는 수업방식을 요구함을 알 수 있었다. 또한 교양과목에서 수강하기를 원하는 과목과 실제 개설되어 있는 과목의 차이를 볼 수 있었다. 이는 실제 교양교과목에서 교양으로서 균형되게 수업들이 개설되고 있으나 대부분의 학습자들은 실용성이 떨어진다는 이유를 들었다. 이는 교육과정을 개편할 때 참고하거나 학습자에게 균형 있는 교양교육이 왜 필요한지 설득력 있는 이해가

필요 하다.

향후 과제로서는 더 많은 학습자를 대상으로 자료를 수집하여 학습자들의 성별, 전공별, 학년별로 텍스트 마이닝과 감성분석까지 수행할 필요가 있다.

References

[1] Gyu Ha Kim, Topic analysis in Statistics using text mining, Keimyung Univ. Graduate School, 2015.

[2] TTA Terminology dictionary, http://100.daum.net.

[3] Eun Jeong Yang, Comparative analysis of spatio-temporal unstructured, Kyunghee Univ. Department of Geography, 2015.

[4] InSeong Kang, TaeHo Kim, HongGhul Lee, Data processing techniques applying data mining based on enterprise cloud computing, The Korea Society of Computer and Information, 16(8), 2011.

[5] Hayeon editorial department, "Market prospects for big data and DBMS", 2012.

[6] gsChoi, ysHam, shKim, BigData Visualization, The Korea Society of Computer and Information, 16(8), 2013.