

정형/비정형 데이터 기반 산업 평가 정보 분석 및 시각화 서비스 구현

*김경원 **정승경 ***조대근 ****윤경로

*,**,***전자부품연구원 *,***건국대학교

*kwkim@keti.re.kr

Development of Structured/Unstructured data-based Industry Evaluation Information Analysis and Visualization Service

*Kyungwon Kim **Seunggyeong Chung ***Daekeun Cho ****Kyoungro Yoon

*,**,***Korea Electronics Technology Institute *,****Konkuk University

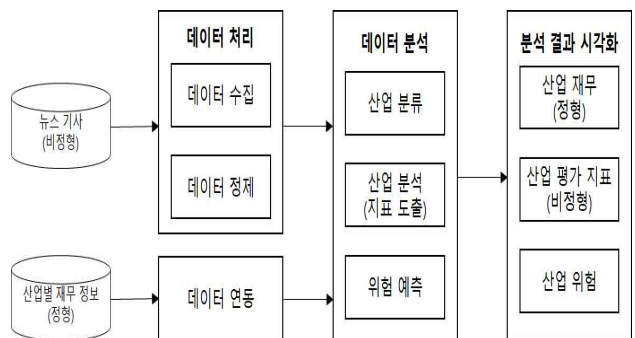
요약

기존 산업평가 방법은 산업별로 분류된 기업의 재무, 비재무 관련 정형 데이터를 기반으로 통계적 기법을 이용하여 각 산업을 평가하고 있다. 이러한 정형 데이터 기반의 산업 평가 방법은 산업별 재무 정보의 집계 및 통계에 오랜 시간이 소요된다. 따라서, 현재 시장 상황을 반영하기 어려운 현실이다. 최근에는 빠르게 변화하는 산업 환경을 반영하기 위해 뉴스 기사와 같은 비정형 데이터를 통해 산업 트렌드를 분석하기 위한 연구가 이루어 지고 있다. 이에, 본 논문에서는 실시간으로 변화하는 산업 트렌드를 반영하여 적시에 산업 분석 정보를 제공하기 위해 정형/비정형 데이터 기반의 산업평가 정보 분석 엔진을 구현하고, 산업별로 분석된 산업평가 정보를 활용하여 사용자가 직관적인 판단을 할 수 있도록 산업평가 정보 시각화 서비스를 제안한다.

1. 서론

기존 산업 평가 서비스는 정형 데이터를 활용한 통계적 분석 방법을 기반으로 제공되고 있다. 기존 정형 데이터 기반 산업 평가는 급변하는 산업 전반의 상황을 제공하지 못하고 있으며, 이를 보완할 수 있는 정보 제공을 통한 긍정적인 평가에 활용할 수 있는 적절한 방안을 제공하지 못하고 있다. 본 논문에서는 기존 정형 데이터 기반의 산업 평가 방법을 보완하기 위해 비정형 데이터 기반 산업 평가 지표 및 이를 이용한 산업 위험 예측 모형의 적용 방안을 제시하고, 기존 산업 평가 서비스를 기반으로 비정형 데이터 분석을 통한 산업 평가 지표와 산업 위험 예측 모형의 융합을 통해 기존 서비스를 보완할 수 있는 정형/비정형 데이터 융합 기반 산업 평가 서비스 제공 방안을 제안한다.

하고, 비정형 데이터 분석 기반으로 도출된 산업 평가 지표인 산업 관심도, 호감도 정보와 이를 기반으로 분석한 산업 위험 예측 정보를 제공할 수 있도록 서비스에 적용하였다. 본 논문에서 제안하는 정형/비정형 데이터 융합 기반 산업 평가 서비스 구성은 <그림 1>과 같다.



<그림 1> 정형/비정형 데이터 융합 기반 산업 평가 서비스 구성

2. 정형/비정형 데이터 기반 산업 평가 서비스

2.1. 산업 평가 서비스 구성

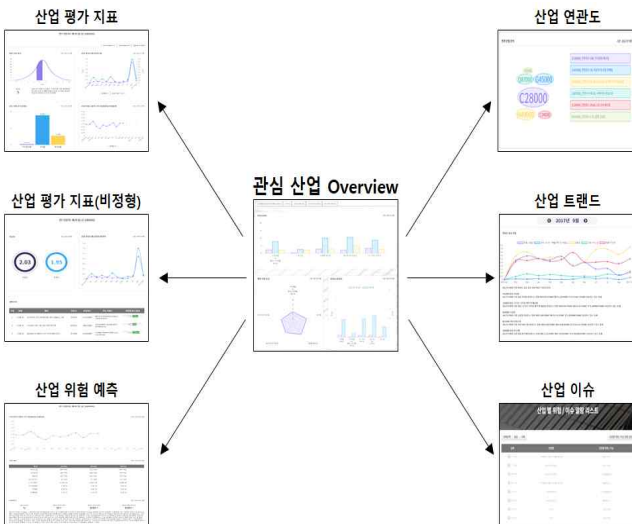
기존 산업 평가 서비스는 각 산업별 통계적 정보를 기반으로 산업별 기본적인 재무 분석 정보를 Overview 형태로 제공하고 있으며, 선택된 특정 산업에 대해 주요 재무 상황, 합산 재무제표, 합산 재무 비율, 시가총액 등의 상세정보를 제공하고 있다. 또한, 산업 평가 서비스에서 가장 중요한 산업별 위험 지표에 대한 정보를 제공하고 있으며, 산업 평가 업체에서 자체적으로 구축한 산업 평가 모형을 기반으로 산업 등급 정보와 산업 관련 다양한 지표와 관련 자료를 제공하고 있다 [1,2]. 본 논문에서는 기존 정형 데이터 기반의 산업 평가와 관련된 기본적인 합산 재무정보, 산업 등급, 산업 위험 지표에 대한 정보를 제공

정형/비정형 데이터 융합 기반 산업 평가 서비스는 정형 데이터 연동 및 비정형 데이터 수집을 통해 데이터를 확보한다. 확보된 비정형 데이터는 정제 과정을 거쳐 분석을 위한 데이터 셋으로 구성된다. 정제된 비정형 데이터 셋은 산업별 다중분류와 분석을 통해 각 산업별 평가 지표인 관심도와 호감도 시계열 데이터 형태로 가공된다[3]. 별도의 인터페이스를 통해 연동된 산업별 재무/위험 지표 등의 정형 데이터와 연계하여 산업 위험 예측을 수행하고 분석 결과를 시각화 모듈을 통해 서비스를 제공할 수 있도록 구성하였다. 분석 결과 시각화 모듈은 기존 정형 데이터 기반 산업 통계 정보와 본 논문에서 제안하는 비정

형 데이터 기반 산업 평가 지표인 관심도, 호감도 시계열 정보, 산업 간 연관 정보, 등의 정형/비정형 데이터를 융합하여 제공함으로써, 기존 정형 데이터 기반 산업 평가 서비스에서 제공하지 못하는 산업 및 시장의 적시적인 정보와 산업 트렌드의 제공을 통해 산업 평가의 효용성과 가치를 높일 수 있다.

2.2. 산업 평가 서비스 시각화 모듈

산업별 분석 결과 시각화 모듈은 정형/비정형 데이터 기반의 산업 분석 결과를 활용하여 의사결정을 할 수 있도록 사용자에게 직관적인 UI를 제공한다.



<그림 2> 산업 평가 서비스 시각화 모듈 구성

<그림 2>는 정형/비정형 데이터 융합 기반 산업 평가 서비스 시각화 모듈 구성을 보여준다. 관심 산업 Overview를 통해 산업 평가 서비스 이용자의 관심 산업에 대한 기본적인 재무 정보, 산업 위험 지표, 산업 등급, 관심도, 호감도 시계열 정보를 비교할 수 있는 화면을 제공한다. 사용자가 선택한 하나의 산업에 대하여 산업 평가 지표에 대한 상세 정보를 제공하며, 본 논문에서 제한한 비정형 데이터 기반 산업 평가 지표인 산업 관심도와 호감도 분석 정보 및 해당 산업에 해당하는 뉴스 기사를 제공한다. 또한, 산업 관심도, 호감도를 기반으로 예측된 산업 위험 지표, 산업간 연관도, 관심 산업에 대한 이슈 리스트 등의 정보를 그래프와 도표를 이용한 다양한 시각화 화면의 제공하여 산업 평가 정보에 대한 사용자의 이해를 높일 수 있도록 구성하였다.

(1) 관심 산업 Overview

<그림 3>의 관심 산업 Overview 화면은 서비스 이용자가 기록한 5개의 관심 산업에 대한 산업별 위험 지표인 부도 발생률, 연체율, 휴폐업 발생률을 그래프로 제공하며, 산업별 등급과 산업별 관심도, 호감도 정보를 비교할 수 있도록 그래프를 제공한다. 정형 데이터 및 비정형 데이터 기반으로 분석된 주요 항목들에 대해 기록된 관심 산업간 비교를 통해서 인사이트를 제공할 수 있도록 구성하였다.



<그림 3> 관심 산업 Overview 화면

(2) 산업 분석 결과 정보

<그림 4>는 서비스 이용자가 선택한 산업에 대한 정형/비정형 데이터 분석 결과를 제공하는 화면이다. <그림 3>의 관심 산업 Overview에서 원하는 산업을 선택하여 활용할 수 있으며, 별도의 산업명 검색을 통해 원하는 산업을 선택 가능하다. 정형 데이터 분석 기반으로 산업 등급과 부도 발생률, 연체율, 휴폐업률 정보를 제공하며, 비정형 데이터 기반으로 도출된 산업 관심도와 호감도 시계열 정보와 이를 기반으로 산업 위험 지표를 예측한 결과를 그래프로 제공한다.



<그림 4> 산업 분석 결과 제공 화면

<그림 4>의 산업 분석 결과를 구성하는 각 항목별로 주요 상세 내용을 별도의 화면으로 제공하며, 비정형 데이터 기반으로 시계열 분석된 산업 관심도와 호감도는 하나의 그래프에 표시하여 비교 분석이 가능하도록 구성하였다. 또한, 정형 데이터 기반의 중요 지표인 산업별 부도율, 연체율 데이터와 비정형 데이터 기반으로 도출된 산업별 관심도, 호감도를 이용하여 미래의 정형지표에 대한 위험을 예측한 결과를 사용자에게 제공한다[4].

(3) 산업간 연관 정보 및 산업 트렌드

<그림 5>는 산업별 키워드를 기반으로 분석된 산업 연관 정보 제공 화면을 보여준다. 해당 산업을 기준으로 연관도가 가장 높은 상위 5개의 산업 정보를 연관도에 따른 타원의 크기로 표현하여 직관성을 높였다[5]. 또한, 연관 산업들의 정형 지표와 비정형 지표에 대한 추이를 비교할 수 있는 그래프를 통해 산업간 관계를 파악할 수 있는 정보를 제공한다.



<그림 5> 산업간 연관도 분석 결과 제공 화면

(4) 산업 이슈 알람

<그림 6>은 서비스 이용자의 관심 산업을 등록하는 화면이다. 이용자는 관심 있는 산업을 등록할 수 있으며, 각 산업별 주요 지표에 대하여 알람 설정이 가능하다. 등록된 산업은 산업 Overview 화면을 통해 관심 산업들을 비교할 수 있는 주요 정보가 제공된다.



<그림 6> 관심 산업 이슈 알람 설정 화면

<그림 6>에서 각 산업별 주요 지표에 대한 모니터링 수치를 등록하면 해당 지표의 등록 수치를 기반으로 해당 지표에 대한 알람을 <그림 7>과 같이 제공한다. 서비스 이용자는 이를 통해 관심 산업에 대한 지표의 변화에 대한 정보를 쉽게 제공 받을 수 있다.



<그림 7> 관심 산업 이슈 알람 제공 화면

3. 결론

기존 산업평가 정보 분석 서비스는 산업별로 해당 산업 분류군에 해당하는 기업들의 재무 정보, 위험 정보 등의 정형 데이터 기반으로 사용자에게 정보를 제공하고 있으나, 뉴스 기사와 같은 현재 상황을 실시간으로 반영하는 정보를 제공하지 못하고 있다. 본 논문에서는 정형 데이터와 뉴스 기사와 같은 비정형 데이터를 융합하여 산업별 분석을 통한 산업 평가 정보를 사용자에게 제공할 수 있는 방안을 제시하고, 산업별 분석 결과를 활용하여 합리적인 의사결정 지원을 위한 시각화 모듈을 구현하였다. 본 논문에서 제안하는 산업 평가 정보 분석 및 분석 결과 시각화 서비스를 통해 사용자에게 산업평가 분석 정보에 대한 인사이트를 제공할 수 있을 것으로 기대된다.

[ACKNOWLEDGEMENT]

이 논문은 2018년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임. (No.2016-0-00414, 산업분석정보 적시 제공을 위한 정형·비정형 빅데이터 융합 기반 산업평가 분석 SW 개발)

[참고 문헌]

[1] Peter J. Sander, Janet Haley, Value Investing For Dummies, 2nded., John Wiley & Sons, Inc., 2008.
 [2] Qinli, Fang and Yizeng, Chen, Stock portfolio analysis based on margin of safety and competitive edge evaluation in the American wine liquor industry. International Conference on Big Data and its Applications, (2016), 1 - 7.
 [3] K. Kim, T. Lim, K. Yoon, "Industry Evaluation Analysis System for Enhanced Industry Analysis Information", Information, vol.20, pp.2537-2542, 2017.
 [4] Haiwei Wu, Dawei Su, Xuesong Huo, Shuang Hu, Zhuodi Wang, Kaiqi Sun, "Using Ranking and Multiple Linear Regression to Explore" the Impact of Social Media Engagement on Student Performance, APPEEC, pp.38-42, 2016.
 [5] Nam, S.-T. et al., A Meta-analysis of the Relationship between Mediator Factors and Purchasing Intention in E-commerce Studies. JICCE, 12 (2014), 257 - 262.