(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2019년07월05일
(11) 등록번호 10-1997111
(24) 등록일자 2019년07월01일

(51) 국제특허분류(Int. Cl.)
G06Q 10/06 (2012.01) *G06F 17/18* (2006.01)
G06Q 30/02 (2012.01)
(52) CPC특허분류
G06Q 10/0637 (2013.01)
G06F 17/18 (2013.01)
(21) 출원번호 10-2018-0002094
(22) 출원일자 2018년01월08일
심사청구일자 2018년01월08일
(56) 선행기술조사문헌
KR1020140133666 A*
KR1020130120029 A*
KR101814005 B1*
최희갑 외 1인, "Zipf 분포를 활용한 지역별 사업체 규모분포 특성", 한국지역학회, 지역연구, 2006년 8월, 제22권, 제2호, pp.275-298.*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
인천대학교 산학협력단
인천광역시 연수구 아카데미로 119 (송도동)
(72) 발명자
김관호
인천광역시 연수구 컨벤시아대로42번길 77 (송도동, 더샵 엑스포)
최영철
경기도 부천시 원미구 소향로 124, 중동신도시포도마을
(뒷면에 계속)
(74) 대리인
김효성

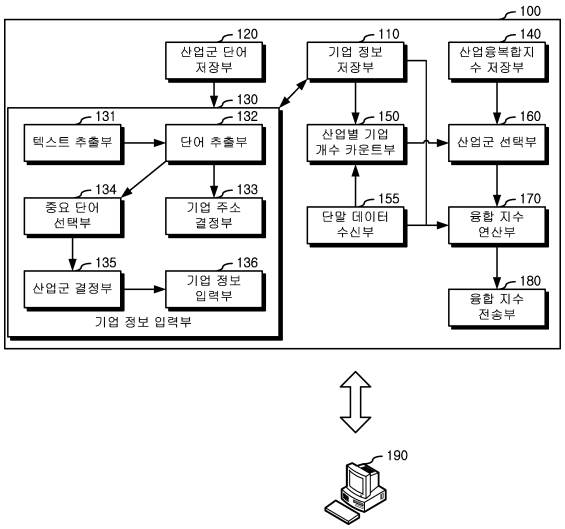
전체 청구항 수 : 총 8 항 심사관 : 태정범

(54) 발명의 명칭 **인공지능 기반의 웹 페이지 분석에 기초한 융합 지수 연산 장치 및 이의 동작 방법**

(57) 요 약

본 발명에 따른 인공지능 기반의 웹 페이지 분석에 기초한 융합 지수 연산 장치 및 이의 동작 방법은 산업융복합지수에 기초하여 제1 지역에 대해 선택된 제1 산업군들과 제2 지역에 대해 선택된 제2 산업군들 간의 산업융복합지수를 결정하고, 결정된 산업융복합지수의 평균 값을 지역 융합 지수로 결정한 후 상기 연산된 지역 융합 지수를 상기 관리자의 단말에 전송함으로써 기업들이 융합 활동을 수행함으로써 기업들이 목표하는 신기술 개발과 경쟁력 제고에 기여할 수 있도록, 여러 산업 군의 기업정보를 추출한 후 별도의 산업연관표를 이용하여 다양한 융합 정보를 분석하고, 분석된 융합 정보를 기업들에게 제공할 수 있다.

대 표 도 - 도1



(52) CPC특허분류
G06Q 30/0201 (2013.01)
(72) 발명자
정민우
인천광역시 서구 청라커낼로 232 LH아파트 419동 502호
곽두희
경기도 광명시 사성로8번길 36, 102호 (광명동, 선문빌라)
고정경
경기도 고양시 일산서구 고양대로 633 동양아파트 103동 302호
박선영
경기도 용인시 기흥구 용구대로2469번길 20, 324호

이 발명을 지원한 국가연구개발사업
과제고유번호 1711039594
부처명 과학기술정보통신부
연구관리전문기관 정보통신산업진흥원
연구사업명 기술확산지원(정보통신)
연구과제명 송도 SW융합 클러스터 (인천 SW융합 BI-Plex)
기 여 율 5/10
주관기관 재단법인인천정보산업진흥원
연구기간 2017.01.01 ~ 2017.12.31
이 발명을 지원한 국가연구개발사업
과제고유번호 WISET-2017-470호
부처명 과학기술정보통신부
연구관리전문기관 한국여성과학기술인지원센터
연구사업명 여성과학기술인 R&D 경력복귀 지원
연구과제명 인공지능을 이용한 광고 및 소셜마케팅 자동화 모델연구
기 여 율 3/10
주관기관 한국여성과학기술인지원센터
연구기간 2017.09.01 ~ 2018.08.31
이 발명을 지원한 국가연구개발사업
과제고유번호 1345263358
부처명 교육부
연구관리전문기관 한국연구재단
연구사업명 개인기초연구(교육부)
연구과제명 인간-공간 상호작용의 전체최적화를 위한 딥러닝 기반의 지능공간 의사결정 모델연구
기 여 율 2/10
주관기관 인천대학교
연구기간 2017.06.01 ~ 2018.02.28

공지예외적용 : 있음

명 세 서

청구범위

청구항 1

복수의 기업들 각각에 대해 기업 주소 및 사전 설정된 복수의 산업군들 중 적어도 하나의 산업군이 매칭되어 있는 기업 정보를 저장하는 기업 정보 저장부;

상기 사전 설정된 복수의 산업군들 간의 산업융복합지수 - 상기 산업융복합지수는 한국은행에서 발행하는 산업연관표 중 국산투입계수 표의 행렬을 전치(Transpose)한 후 행의 합계가 1이 되도록 행렬에 기재된 성분들의 값을 정규화함으로써 산출된 값을 의미함 - 를 저장하는 산업융복합지수 저장부;

관리자의 단말로부터 제1 지역 명칭 및 제2 지역 명칭이 입력되면, 상기 기업 정보 저장부에 저장되어 있는 상기 기업 정보에 매칭되어 있는 기업 주소에 기초하여 상기 복수의 산업군들 각각마다 상기 제1 지역 명칭에 따른 제1 지역 및 제2 지역 명칭에 따른 제2 지역에 위치하는 제1 산업별 기업 개수 및 제2 산업별 기업 개수를 카운트하는 산업별 기업 개수 카운트부;

상기 제1 산업별 기업 개수에 기초하여 상기 복수의 산업군들 중 상기 제1 지역에 위치하는 기업의 개수가 많은 순서대로 N - 상기 N은 1 이상의 정수임 - 개의 제1 산업군들을 선택하고, 상기 제2 산업별 기업 개수에 기초하여 상기 복수의 산업군들 중 상기 제2 지역에 위치하는 기업의 개수가 많은 순서대로 M - 상기 M은 1 이상의 정수임 - 개의 제2 산업군들을 선택하는 산업군 선택부;

상기 산업융복합지수 저장부에 저장되어 있는 상기 복수의 산업군들 간의 산업융복합지수에 기초하여 상기 제1 지역에 대해 선택된 제1 산업군들과 상기 제2 지역에 대해 선택된 제2 산업군들 간의 산업융복합지수를 결정하고, 상기 결정된 산업융복합지수의 평균 값을 상기 제1 지역과 상기 제2 지역 간의 지역 융합 지수로 결정하는 융합 지수 연산부; 및

상기 결정된 지역 융합 지수를 상기 관리자의 단말에 전송하는 융합 지수 전송부

를 포함하고,

상기 융합 지수 연산부는 상기 제1 산업군들과 상기 제2 산업군들 간의 산업융복합지수를 결정하되, 상기 제1 산업군들과 상기 제2 산업군들에 중복되는 제3 산업군 간에 결정된 산업융복합지수를 제외한 산업융복합지수의 평균 값을 상기 제1 지역과 상기 제2 지역 간의 지역 융합 지수로 결정하며,

상기 융합 지수 연산부는 상기 관리자의 단말로부터 제2 기업 명칭 및 제3 기업 명칭이 입력되면, 상기 산업융복합지수 저장부에 저장되어 있는 상기 복수의 산업군들 간의 산업융복합지수에 기초하여 상기 제2 기업 명칭에 따른 제2 기업에 대해 매칭되어 있는 적어도 하나의 산업군과 상기 제3 기업 명칭에 따른 제3 기업에 대해 매칭되어 있는 적어도 하나의 산업군 간의 산업융복합지수를 추가로 결정한 후 상기 추가로 결정된 산업융복합지수의 평균 값을 상기 제2 기업과 상기 제3 기업 간의 기업 융합 지수로 결정하고,

상기 융합 지수 전송부는 상기 결정된 기업 융합 지수를 상기 관리자의 단말에 전송하는

융합 지수 연산 장치.

청구항 2

삭제

청구항 3

제1항에 있어서,

서로 다른 특성 벡터들이 할당되어 있는 사전 설정된 복수의 산업군 단어들을 포함하는 사전 설정된 복수의 산업군 단어 그룹들 - 상기 복수의 산업군 단어 그룹들 각각에 포함된 복수의 산업군 단어들은 서로 간의 특성 벡터를 이용하여 연산되는 유사도가 사전 설정된 기준 유사도 이상인 단어들임 - 을 저장하는 산업군 단어 저장부;

제1 기업의 웹 페이지에 대한 접속 주소가 입력되면, 상기 접속 주소를 기초로 상기 제1 기업의 웹 페이지에 접속하여 상기 제1 기업의 웹 페이지로부터 상기 제1 기업의 웹 페이지 상에 존재하는 복수의 제1 텍스트들을 추출하는 텍스트 추출부;

상기 복수의 제1 텍스트들에 대해 형태소 분석을 수행하여 상기 복수의 제1 텍스트들로부터 복수의 제1 단어들을 추출하는 단어 추출부;

상기 복수의 제1 단어들에 기초하여 상기 제1 기업의 기업 주소를 결정하는 기업 주소 결정부;

상기 복수의 제1 단어들의 상기 제1 기업의 웹 페이지 상에서의 등장 빈도수에 기초하여 상기 복수의 제1 단어들 중 적어도 하나의 중요 단어를 선택하는 중요 단어 선택부;

상기 적어도 하나의 중요 단어가 선택되면, 상기 산업군 단어 저장부를 참조하여, 상기 적어도 하나의 중요 단어에 대한 특성 벡터를 확인하고, 상기 적어도 하나의 중요 단어에 대한 특성 벡터와 상기 복수의 산업군 단어 그룹들 각각에 포함되어 있는 상기 복수의 산업군 단어들 각각에 대한 특성 벡터 간의 유사도를 연산하고, 상기 연산된 유사도에 기초하여 상기 제1 기업에 대해 상기 복수의 산업군들 중 적어도 하나의 산업군을 결정하는 산업군 결정부; 및

상기 제1 기업에 대해 상기 결정된 기업 주소 및 상기 제1 기업에 대해 상기 결정된 적어도 하나의 산업군을 매칭시켜 상기 제1 기업에 대한 기업 정보를 생성한 후 상기 기업 정보 저장부에 저장하는 기업 정보 입력부

를 더 포함하는 융합 지수 연산 장치.

청구항 4

제3항에 있어서,

상기 기업 정보 저장부는 상기 복수의 기업들 각각에 대해 중요 단어가 더 매칭되어 있는 기업 정보를 저장하고,

상기 기업 정보 입력부는 상기 제1 기업에 대해 상기 적어도 하나의 중요 단어가 더 매칭된 상기 제1 기업에 대한 기업 정보를 생성한 후 상기 기업 정보 저장부에 저장하는

융합 지수 연산 장치.

청구항 5

삭제

청구항 6

복수의 기업들 각각에 대해 기업 주소 및 사전 설정된 복수의 산업군들 중 적어도 하나의 산업군이 매칭되어 있는 기업 정보를 저장하는 기업 정보 저장부를 유지하는 단계;

상기 사전 설정된 복수의 산업군들 간의 산업융복합지수 - 상기 산업융복합지수는 한국은행에서 발행하는 산업연관표 중 국산투입계수 표의 행렬을 전치(Transpose)한 후 행의 합계가 1이 되도록 행렬에 기재된 성분들의 값을 정규화함으로써 산출된 값을 의미함 - 를 저장하는 산업융복합지수 저장부를 유지하는 단계;

관리자의 단말로부터 제1 지역 명칭 및 제2 지역 명칭이 입력되면, 상기 기업 정보 저장부에 저장되어 있는 상기 기업 정보에 매칭되어 있는 기업 주소에 기초하여 상기 복수의 산업군들 각각마다 상기 제1 지역 명칭에 따른 제1 지역 및 제2 지역 명칭에 따른 제2 지역에 위치하는 제1 산업별 기업 개수 및 제2 산업별 기업 개수를 카운트하는 단계;

상기 제1 산업별 기업 개수에 기초하여 상기 복수의 산업군들 중 상기 제1 지역에 위치하는 기업의 개수가 많은 순서대로 N - 상기 N은 1 이상의 정수임 - 개의 제1 산업군들을 선택하고, 상기 제2 산업별 기업 개수에 기초하여 상기 복수의 산업군들 중 상기 제2 지역에 위치하는 기업의 개수가 많은 순서대로 M - 상기 M은 1 이상의 정수임 - 개의 제2 산업군들을 선택하는 단계;

상기 산업융복합지수 저장부에 저장되어 있는 상기 복수의 산업군들 간의 산업융복합지수에 기초하여 상기 제1 지역에 대해 선택된 제1 산업군들과 상기 제2 지역에 대해 선택된 제2 산업군들 간의 산업융복합지수를 결정하고, 상기 결정된 산업융복합지수의 평균 값을 상기 제1 지역과 상기 제2 지역 간의 지역 융합 지수로 결정하는

단계;

상기 결정된 지역 융합 지수를 상기 관리자의 단말에 전송하는 단계;

상기 관리자의 단말로부터 제2 기업 명칭 및 제3 기업 명칭이 입력되면, 상기 산업융복합지수 저장부에 저장되어 있는 상기 복수의 산업군들 간의 산업융복합지수에 기초하여 상기 제2 기업 명칭에 따른 제2 기업에 대해 매칭되어 있는 적어도 하나의 산업군과 상기 제3 기업 명칭에 따른 제3 기업에 대해 매칭되어 있는 적어도 하나의 산업군 간의 산업융복합지수를 추가로 결정한 후 상기 추가로 결정된 산업융복합지수의 평균 값을 상기 제2 기업과 상기 제3 기업 간의 기업 융합 지수로 결정하는 단계; 및

상기 결정된 기업 융합 지수를 상기 관리자의 단말에 전송하는 단계

를 포함하고,

상기 지역 융합 지수로 결정하는 단계는 상기 제1 산업군들과 상기 제2 산업군들 간의 산업융복합지수를 결정하되, 상기 제1 산업군들과 상기 제2 산업군들에 중복되는 제3 산업군 간에 결정된 산업융복합지수를 제외한 산업융복합지수의 평균 값을 상기 제1 지역과 상기 제2 지역 간의 지역 융합 지수로 결정하는

융합 지수 연산 장치의 동작 방법.

청구항 7

삭제

청구항 8

제6항에 있어서,

서로 다른 특성 벡터들이 할당되어 있는 사전 설정된 복수의 산업군 단어들을 포함하는 사전 설정된 복수의 산업군 단어 그룹들 - 상기 복수의 산업군 단어 그룹들 각각에 포함된 복수의 산업군 단어들은 서로 간의 특성 벡터를 이용하여 연산되는 유사도가 사전 설정된 기준 유사도 이상인 단어들임 - 을 저장하는 산업군 단어 저장부를 유지하는 단계;

제1 기업의 웹 페이지에 대한 접속 주소가 입력되면, 상기 접속 주소를 기초로 상기 제1 기업의 웹 페이지에 접속하여 상기 제1 기업의 웹 페이지로부터 상기 제1 기업의 웹 페이지 상에 존재하는 복수의 제1 텍스트들을 추출하는 단계;

상기 복수의 제1 텍스트들에 대해 형태소 분석을 수행하여 상기 복수의 제1 텍스트들로부터 복수의 제1 단어들을 추출하는 단계;

상기 복수의 제1 단어들에 기초하여 상기 제1 기업의 기업 주소를 결정하는 단계;

상기 복수의 제1 단어들의 상기 제1 기업의 웹 페이지 상에서의 등장 빈도수에 기초하여 상기 복수의 제1 단어들 중 적어도 하나의 중요 단어를 선택하는 단계;

상기 적어도 하나의 중요 단어가 선택되면, 상기 산업군 단어 저장부를 참조하여, 상기 적어도 하나의 중요 단어에 대한 특성 벡터를 확인하고, 상기 적어도 하나의 중요 단어에 대한 특성 벡터와 상기 복수의 산업군 단어 그룹들 각각에 포함되어 있는 상기 복수의 산업군 단어들 각각에 대한 특성 벡터 간의 유사도를 연산하고, 상기 연산된 유사도에 기초하여 상기 제1 기업에 대해 상기 복수의 산업군들 중 적어도 하나의 산업군을 결정하는 단계; 및

상기 제1 기업에 대해 상기 결정된 기업 주소 및 상기 제1 기업에 대해 상기 결정된 적어도 하나의 산업군을 매칭시켜 상기 제1 기업에 대한 기업 정보를 생성한 후 상기 기업 정보 저장부에 저장하는 단계

를 더 포함하는 융합 지수 연산 장치의 동작 방법.

청구항 9

제8항에 있어서,

상기 기업 정보 저장부를 유지하는 단계는 상기 복수의 기업들 각각에 대해 중요 단어가 더 매칭되어 있는 기업 정보를 저장하는 상기 기업 정보 저장부를 유지하고,

상기 기업 정보 저장부에 저장하는 단계는 상기 제1 기업에 대해 상기 적어도 하나의 중요 단어가 더 매칭된 상기 제1 기업에 대한 기업 정보를 생성한 후 상기 기업 정보 저장부에 저장하는

융합 지수 연산 장치의 동작 방법.

청구항 10

삭제

청구항 11

제6항, 제8항 또는 제9항 중 어느 한 항의 방법을 컴퓨터로 하여금 수행하도록 하는 프로그램을 기록한 컴퓨터 판독 가능 기록 매체.

청구항 12

제6항, 제8항 또는 제9항 중 어느 한 항의 방법을 컴퓨터와의 결합을 통해 실행시키기 위한 저장매체에 저장된 컴퓨터 프로그램.

발명의 설명

기 술 분 야

[0001] 본 발명은 웹 페이지의 분석을 통해 해당 웹 페이지를 운영하는 기업의 키워드 정보를 추출한 후 추출된 기업 키워드 정보 및 산업 융복합 지수에 기초하여 융합 지수를 연산하는 장치 및 이의 동작 방법에 대한 것이다.

배 경 기 술

[0002] 신기술 개발과 신규 시장 창출은 기업의 경쟁력 제고를 위해 반드시 필요한 활동이다. 이러한 신기술 개발과 신규 시장 창출을 효과적으로 수행하기 위해서 기업들은 융합 활동을 통하여 서로 다른 분야의 기술과 아이디어들을 결합해야 할 필요가 있다. 지난 2015년 디지털타임즈와 한국산업기술진흥협회가 공동 진행한 ‘산업현장 융합 활동 현황 설문조사’에 따르면 국내 기업의 80% 이상이 융합 활동이 필요하다고 답하였다. 이를 통하여 기업들 또한 융합에 대한 중요성에 대하여 인식하고 있다는 것을 알 수 있다. 뿐만 아니라 기업들이 융합 활동이 필요하다고 생각한 이유로는 '신기술·신제품 개발'(42.5%)이 가장 많은 것을 확인할 수 있었고 실제 융합 활동을 진행한 기업들 가운데 매출액이 증가한 기업의 비율이 56.1%에 달했다. 이는 감소했다고 답한 5%에 비하여 매우 높은 수치이다. 따라서 융합 활동이 기업들이 목표하는 신기술 개발과 경쟁력 제고에 기여한다는 것을 알 수 있다.

[0003] 현재 정부와 민간단체에서는 기업들의 융합을 위한 정책으로 융합제품 및 서비스의 개발에 대한 금전적 지원이 이루어지고 있고 융합을 통한 새로운 시장을 개척하기 위해 기업 간 교류협력 강화 사업들을 펼치고 있다. 하지만 융합을 저해하는 가장 큰 요인은 융합관련 시장이나 기술 등에 대한 정보의 부족으로 이를 해결하기 위한 노력이 필요한 실정이다.

[0004] 현재 한국은행에서는 정보제공을 위해 국내 산업군 간의 연관성 수치를 알 수 있는 산업연관표를 발행하고 있다. 하지만 기업이나 지역의 융합과 같은 추가적인 정보는 알 수 없다는 단점이 있다. 기존의 국내 연구로는 “우수 중소 제조기업을 위한 스카우팅 시스템과 추천 알고리즘”에 관한 연구가 이루어진 바가 있지만 제조기업에만 국한 된다는 점에서 전체 산업 군에 대한 융합정보 제공이 어렵다는 한계가 있다.

[0005] 따라서, 기업들이 융합 활동을 수행함으로써 기업들이 목표하는 신기술 개발과 경쟁력 제고에 기여할 수 있도록, 여러 산업 군의 기업정보를 추출한 후 별도의 산업연관표를 이용하여 다양한 융합 정보를 분석하고, 분석된 융합 정보를 기업들에게 제공할 수 있는 기법에 관한 연구가 필요하다.

발명의 내용

해결하려는 과제

[0006] 본 발명에 따른 인공지능 기반의 웹 페이지 분석에 기초한 융합 지수 연산 장치 및 이의 동작 방법은 산업융복합지수에 기초하여 제1 지역에 대해 선택된 제1 산업군들과 제2 지역에 대해 선택된 제2 산업군들 간의 산업융

복합지수를 결정하고, 결정된 산업융복합지수의 평균 값을 지역 융합 지수로 결정한 후 상기 연산된 지역 융합 지수를 상기 관리자의 단말에 전송함으로써 기업들이 융합 활동을 수행함으로써 기업들이 목표하는 신기술 개발과 경쟁력 제고에 기여할 수 있도록, 여러 산업 군의 기업정보를 추출한 후 별도의 산업연관표를 이용하여 다양한 융합 정보를 분석하고, 분석된 융합 정보를 기업들에게 제공하고자 한다.

과제의 해결 수단

[0007] 본 발명의 일실시예에 따른 인공지능 기반의 웹 페이지 분석에 기초한 융합 지수 연산 장치는 다수의 기업들 각각에 대해 기업 주소 및 사전 설정된 다수의 산업군들 중 적어도 하나의 산업군이 매칭되어 있는 기업 정보를 저장하는 기업 정보 저장부, 상기 사전 설정된 다수의 산업군들 간의 산업융복합지수를 저장하는 산업융복합지수 저장부, 관리자의 단말로부터 제1 지역 명칭 및 제2 지역 명칭이 입력되면, 상기 기업 정보 저장부에 저장되어 있는 상기 기업 정보에 매칭되어 있는 기업 주소에 기초하여 상기 다수의 산업군들 각각마다 상기 제1 지역 명칭에 따른 제1 지역 및 제2 지역 명칭에 따른 제2 지역에 위치하는 제1 산업별 기업 개수 및 제2 산업별 기업 개수를 카운트하는 산업별 기업 개수 카운트부, 상기 제1 산업별 기업 개수에 기초하여 상기 다수의 산업군들 중 상기 제1 지역에 위치하는 기업의 개수가 많은 순서대로 N - 상기 N은 1 이상의 정수임 - 개의 제1 산업군들을 선택하고, 상기 제2 산업별 기업 개수에 기초하여 상기 다수의 산업군들 중 상기 제2 지역에 위치하는 기업의 개수가 많은 순서대로 M - 상기 M은 1 이상의 정수임 - 개의 제2 산업군들을 선택하는 산업군 선택부, 상기 산업융복합지수에 기초하여 상기 제1 지역에 대해 선택된 제1 산업군들과 상기 제2 지역에 대해 선택된 제2 산업군들 간의 산업융복합지수를 결정하고, 상기 결정된 산업융복합지수의 평균 값을 지역 융합 지수로 결정하는 융합 지수 연산부 및 상기 연산된 지역 융합 지수를 상기 관리자의 단말에 전송하는 융합 지수 전송부를 포함한다.

[0008] 또한, 본 발명의 일실시예에 따른 인공지능 기반의 웹 페이지 분석에 기초한 융합 지수 연산 장치의 동작 방법은 다수의 기업들 각각에 대해 기업 주소 및 사전 설정된 다수의 산업군들 중 적어도 하나의 산업군이 매칭되어 있는 기업 정보를 저장하는 기업 정보 저장부를 유지하는 단계, 상기 사전 설정된 다수의 산업군들 간의 산업융복합지수를 저장하는 산업융복합지수 저장부를 유지하는 단계, 관리자의 단말로부터 제1 지역 명칭 및 제2 지역 명칭이 입력되면, 상기 기업 정보 저장부에 저장되어 있는 상기 기업 정보에 매칭되어 있는 기업 주소에 기초하여 상기 다수의 산업군들 각각마다 상기 제1 지역 명칭에 따른 제1 지역 및 제2 지역 명칭에 따른 제2 지역에 위치하는 제1 산업별 기업 개수 및 제2 산업별 기업 개수를 카운트하는 단계, 상기 제1 산업별 기업 개수에 기초하여 상기 다수의 산업군들 중 상기 제1 지역에 위치하는 기업의 개수가 많은 순서대로 N - 상기 N은 1 이상의 정수임 - 개의 제1 산업군들을 선택하고, 상기 제2 산업별 기업 개수에 기초하여 상기 다수의 산업군들 중 상기 제2 지역에 위치하는 기업의 개수가 많은 순서대로 M - 상기 M은 1 이상의 정수임 - 개의 제2 산업군들을 선택하는 단계, 상기 산업융복합지수에 기초하여 상기 제1 지역에 대해 선택된 제1 산업군들과 상기 제2 지역에 대해 선택된 제2 산업군들 간의 산업융복합지수를 결정하고, 상기 결정된 산업융복합지수의 평균 값을 지역 융합 지수로 결정하는 단계 및 상기 연산된 지역 융합 지수를 상기 관리자의 단말에 전송하는 단계를 포함한다.

발명의 효과

[0009] 본 발명에 따른 인공지능 기반의 웹 페이지 분석에 기초한 융합 지수 연산 장치 및 이의 동작 방법은 산업융복합지수에 기초하여 제1 지역에 대해 선택된 제1 산업군들과 제2 지역에 대해 선택된 제2 산업군들 간의 산업융복합지수를 결정하고, 결정된 산업융복합지수의 평균 값을 지역 융합 지수로 결정한 후 상기 연산된 지역 융합 지수를 상기 관리자의 단말에 전송함으로써 기업들이 융합 활동을 수행함으로써 기업들이 목표하는 신기술 개발과 경쟁력 제고에 기여할 수 있도록, 여러 산업 군의 기업정보를 추출한 후 별도의 산업연관표를 이용하여 다양한 융합 정보를 분석하고, 분석된 융합 정보를 기업들에게 제공할 수 있다.

도면의 간단한 설명

[0010] 도 1은 본 발명의 일실시예에 따른 인공지능 기반의 웹 페이지 분석에 기초한 융합 지수 연산 장치의 구조를 도시한 도면이다.

도 2는 본 발명의 일실시예에 따른 인공지능 기반의 웹 페이지 분석에 기초한 융합 지수 연산 장치의 동작 방법을 도시한 순서도이다.

발명을 실시하기 위한 구체적인 내용

[0011] 이하에서는 본 발명에 따른 실시예들을 첨부된 도면을 참조하여 상세하게 설명하기로 한다. 이러한 설명은 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 각 도면을 설명하면서 유사한 참조부호를 유사한 구성요소에 대해 사용하였으며, 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 본 명세서 상에서 사용되는 모든 용어들은 본 발명이 속하는 기술분야에서 통상의 지식을 가진 사람에 의해 일반적으로 이해되는 것과 동일한 의미를 가지고 있다.

[0012] 도 1은 본 발명의 일실시예에 따른 인공지능 기반의 웹 페이지 분석에 기초한 융합 지수 연산 장치(100)의 구조를 도시한 도면이다.

[0013] 도 1을 참조하면, 본 발명의 일실시예에 따른 인공지능 기반의 웹 페이지 분석에 기초한 융합 지수 연산 장치(100)는 기업 정보 저장부(110), 산업군 단어 저장부(120), 기업 정보 입력부(130), 산업융복합지수 저장부(140), 산업별 기업 개수 카운트부(150), 산업군 선택부(160), 융합 지수 연산부(170) 및 융합 지수 전송부(180)를 포함한다.

[0014] 기업 정보 저장부(110)에는 다수의 기업들 각각에 대해 기업 주소 및 사전 설정된 다수의 산업군들 중 적어도 하나의 산업군이 매칭되어 있는 기업 정보가 저장되어 있다. 또한, 상기 기업 정보 저장부(110)는 상기 다수의 기업들 각각에 대해 중요 단어를 더 매칭시켜 저장할 수 있다.

[0015] 예컨대, 기업 정보 저장부(110)에는 하기의 표 1과 같이 정보가 저장되어 있을 수 있다.

표 1

[0016]

기업 정보 인덱스	기업 주소	산업군 정보	중요 단어
기업 정보 1	주소 1	농림수산품	단어 1, 단어 2, 단어 3
기업 정보 2	주소 2	농림수산품	단어 1, 단어 3
기업 정보 3	주소 3	광산품	단어 4, 단어 5, 단어 6
기업 정보 4	주소 4	음식료품	단어 7, 단어 8
기업 정보 5	주소 5	문화 및 기타 서비스	단어 9, 단어 10, 단어 11
기업 정보 6	주소 6	문화 및 기타 서비스	단어 9, 단어 11, 단어 12
...	...	...	...

[0017] 산업군 단어 저장부(120)는 서로 다른 특성 벡터들이 할당되어 있는 사전 설정된 다수의 산업군 단어들을 포함하는 사전 설정된 다수의 산업군 단어 그룹들을 저장할 수 있다. 여기에서, 상기 다수의 산업군 단어 그룹들 각각에 포함된 다수의 산업군 단어들은 서로 간의 특성 벡터를 이용하여 연산되는 유사도가 사전 설정된 기준 유사도 이상인 단어들일 수 있다.

[0018] 산업군 단어 저장부(120)에는 하기 표 2와 같이 서로 다른 특성 벡터들이 할당되어 있는 사전 설정된 다수의 산업군 단어들을 포함하는 사전 설정된 다수의 산업군 단어 그룹들을 저장할 수 있다.

표 2

[0019]

산업군 단어 그룹 인덱스	산업군 단어	특성 벡터
산업군 단어 그룹 1	컴퓨터	(1, 2, 0, 0, 3)
	휴대폰	(2, 1, -1, -1, 2)
	노트북	(3, 2, 0, -1, 1)
산업군 단어 그룹 2	고등어	(-1, 1, 4, 3, 1)
	꽁치	(0, -1, 3, 6, 2)
	멸치	(-1, -1, 5, 2, -1)
...	...	...

[0020] 여기에서, 산업군 단어 그룹 1은 전기 및 전자기기일 수 있고, 산업군 단어 그룹 2는 농림수산품일 수 있다.

[0021] 여기서, 벡터 사이의 유사도는 하기의 수학식 1에 따라 연산될 수 있다.

수학식 1

[0022]
$$S=\frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n}(A_i)^2} \times \sqrt{\sum_{i=1}^{n}(B_i)^2}}$$

[0023] 여기서, S는 특성 벡터 A와 B 사이의 유사도로 -1에서 1사이의 값을 가지며, 그 값이 클수록 유사한 특성 벡터임을 의미하고, A_i는 특성 벡터 A의 i번째 성분, B_i는 특성 벡터 B의 i번째 성분을 의미한다.

[0024] 예컨대, 상기 표 2에서 "컴퓨터"라는 산업군 단어와 "휴대폰"이라는 산업군 단어에 각각 할당되어 있는 특성 벡터 간의 유사도를 연산하게 되면, 하기의 수학식 2와 같이 연산될 수 있다.

수학식 2

[0025]
$$S=\frac{(1\times 2)+(2\times 1)+(0\times(-1))+(0\times(-1))+(3\times 2)}{\sqrt{1^2+2^2+0^2+0^2+3^2}\times\sqrt{2^2+1^2+(-1)^2+(-1)^2+2^2}}$$
$$=0.8058$$

[0026] 또한, 예컨대, 상기 표 2에서 "컴퓨터"라는 산업군 단어와 "고등어"이라는 산업군 단어에 각각 할당되어 있는 특성 벡터 간의 유사도를 연산하게 되면, 하기의 수학식 3과 같이 연산될 수 있다.

[0027]
$$S=\frac{(1\times(-1))+(2\times 1)+(0\times 4)+(0\times 3)+(3\times 1)}{\sqrt{1^2+2^2+0^2+0^2+3^2}\times\sqrt{(-1)^2+1^2+4^2+3^2+1^2}}$$
$$=0.2020$$

[0028] "컴퓨터"라는 산업군 단어와 "휴대폰"이라는 산업군 단어에 각각 할당되어 있는 특성 벡터 간의 유사도(즉, 0.8058)와 "컴퓨터"라는 산업군 단어와 "고등어"이라는 산업군 단어에 각각 할당되어 있는 특성 벡터 간의 유사도(즉, 0.2020)를 비교하면, "컴퓨터"라는 산업군 단어는 "휴대폰"이라는 산업군 단어와 유사하지만 "고등어"이라는 산업군 단어와는 유사하지 않음을 알 수 있다. 여기에서, 사전 설정된 기준 유사도(예컨대, 0.5000)에 기초하여 유사한 산업군 단어인지를 판별할 수 있다.

[0029] 상기 표 2와 같은 산업군 단어 저장부(120)에 저장되는 복수의 산업군 단어들은 관리자에 의해 임의로 설정된 단어들이며, 각 산업군 단어들에 할당되어 있는 특성 벡터들은 관리자에 의해서 설정된 각 단어들 간의 유사도 기준에 따라 소정의 유사도가 연산되도록 할당된 값일 수 있다. 이때, 각 산업군 단어들 간의 상기 유사도 기준은 웹을 통해 다양한 정보들을 수집하고, 각 정보들에 대한 분석과 학습을 통해서 다양한 산업군 단어들 간의 관계 분석을 수행한 결과에 기반한 기준일 수 있다.

[0030] 기업 정보 입력부(130)는 입력된 제1 기업의 웹 페이지에 대한 접속 주소에 기초하여, 상기 제1 기업의 웹 페이지에 접속한 후 상기 제1 기업의 웹 페이지로부터 상기 제1 기업의 웹 페이지 상에 존재하는 복수의 제1 텍스트들을 이용하여 상기 제1 기업에 대한 기업 정보를 생성한 후 상기 기업 정보 저장부에 저장함으로써 기업 정보 저장부(110)에 저장되어 있는 기업 정보를 갱신할 수 있다. 이를 위하여, 기업 정보 입력부(130)는 텍스트 추출부(131), 단어 추출부(132), 기업 주소 결정부(133), 중요 단어 선택부(134), 산업군 결정부(135) 및 기업 정보 입력부(136)을 포함할 수 있다.

[0031] 텍스트 추출부(131)는 제1 기업의 웹 페이지에 대한 접속 주소가 입력되면, 상기 접속 주소를 기초로 상기 제1 기업의 웹 페이지에 접속하여 상기 제1 기업의 웹 페이지로부터 상기 제1 기업의 웹 페이지 상에 존재하는 복수의 제1 텍스트들을 추출할 수 있다.

[0032] 이때, 본 발명의 일실시예에 따르면, 텍스트 추출부(131)는 상기 제1 기업의 웹 페이지를 구성하는 HTML(Hypertext Markup Language) 코드를 파싱(parsing)하여 상기 HTML 코드 상에서 텍스트 입력과 연관된 태

그(tag)를 통해 삽입되어 있는 텍스트들을 추출함으로써, 상기 제1 기업의 웹 페이지 상에 존재하는 상기 복수의 제1 텍스트들을 추출하되, 상기 HTML 코드 상에 하이퍼링크 태그가 존재하는 경우, 상기 하이퍼링크 태그를 통해 링크되어 있는 서브 페이지에 접속하여 상기 서브 페이지의 HTML 코드로부터 텍스트 입력과 연관된 태그를 통해 삽입되어 있는 텍스트들도 함께 추출함으로써, 상기 제1 기업의 웹 페이지 상에 존재하는 상기 복수의 제1 텍스트들에 대한 추출을 수행할 수 있다.

[0033] 관련해서, 텍스트 추출부(131)는 제1 기업의 웹 페이지를 구성하는 HTML 코드에서 텍스트 입력과 연관된 태그를 통해 삽입되어 있는 텍스트들을 추출하되, "<a href>"와 같은 하이퍼링크 태그가 존재하는 경우, 해당 하이퍼링크 태그를 통해 링크되어 있는 서브 페이지에 접속해서 상기 서브 페이지의 HTML 코드로부터 텍스트 입력과 연관된 태그를 통해 삽입되어 있는 텍스트들을 함께 추출함으로써, 상기 제1 기업의 웹 페이지 상에 존재하는 복수의 제1 텍스트들을 추출할 수 있다.

[0034] 단어 추출부(132)는 상기 복수의 제1 텍스트들에 대해 형태소 분석을 수행하여 상기 복수의 제1 텍스트들로부터 복수의 제1 단어들을 추출할 수 있다.

[0035] 기업 주소 결정부(133)는 상기 복수의 제1 단어들에 기초하여 상기 제1 기업의 기업 주소를 결정할 수 있다. 예를 들어, 상기 복수의 제1 단어들 중 연속된 단어들이 지역 정보를 나타내는 단어들(예 : “서울시”, “영등포구”, “여의도동” 등)인 경우, 기업 주소 결정부(133)는 상기 지역 정보를 나타내는 연속된 단어들을 상기 제1 기업의 기업 주소로 결정할 수 있다.

[0036] 중요 단어 선택부(134)는 상기 복수의 제1 단어들의 상기 제1 기업의 웹 페이지 상에서의 등장 빈도수에 기초하여 상기 복수의 제1 단어들 중 적어도 하나의 중요 단어를 선택할 수 있다.

[0037] 구체적으로, 중요 단어 선택부(134)는 상기 제1 기업의 웹 페이지 상에서의 등장 빈도수에 기초하여 상기 복수의 제1 단어들 중 적어도 하나의 중요 단어를 선택하되, 산업군 단어 저장부(120)에 저장되어 있는 산업군 단어들과 매칭되는 단어를 적어도 하나의 중요 단어로 선택할 수 있다. 이에 따라, “페이”라는 단어가 상기 제1 기업의 웹 페이지 상에서 가장 많이 등장하였다고 하더라도, 산업군 단어 저장부(120)에 “페이”라는 산업군 단어가 저장되어 있지 않은 경우, 중요 단어 선택부(134)는 상기 “페이”를 중요 단어로 선택하지 않을 수 있다.

[0038] 산업군 결정부(135)는 상기 적어도 하나의 중요 단어가 선택되면, 상기 산업군 단어 저장부를 참조하여, 상기 적어도 하나의 중요 단어에 대한 특성 벡터를 확인하고, 상기 적어도 하나의 중요 단어에 대한 특성 벡터와 상기 다수의 산업군 단어 그룹들 각각에 포함되어 있는 상기 다수의 산업군 단어들 각각에 대한 특성 벡터 간의 유사도를 연산하고, 상기 연산된 유사도에 기초하여 상기 제1 기업에 대해 상기 다수의 산업군들 중 적어도 하나의 산업군을 결정할 수 있다. 이때, 산업군 결정부(135)는 상기 적어도 하나의 중요 단어와 상기 다수의 산업군 단어 그룹들 각각에 포함되어 있는 상기 다수의 산업군 단어들 간의 유사도의 평균을 상기 적어도 하나의 중요 단어에 대한 유사도로서 연산할 수 있다.

[0039] 예를 들어, 상기 제1 기업의 웹 페이지 상에서 상기 적어도 하나의 중요 단어가 “휴대폰”, “노트북” 및 “고등어”로 선택된 경우, 산업군 결정부(135)는 “휴대폰”, “노트북” 및 “고등어”에 대한 특성 벡터(즉, (2, 1, -1, -1, 2), (3, 2, 0, -1, 1) 및 (-1, 1, 4, 3, 1))를 확인하고, “휴대폰”, “노트북” 및 “고등어”에 대한 특성 벡터와 상기 다수의 산업군 단어 그룹들 각각에 포함되어 있는 상기 다수의 산업군 단어들 각각에 대한 특성 벡터 간의 유사도를 연산한 후 상기 연산된 유사도에 기초하여 상기 제1 기업에 대해 상기 다수의 산업군들 중 적어도 하나의 산업군을 결정할 수 있다. 이때, 산업군 결정부(135)는 “휴대폰”, “노트북” 및 “고등어”에 대한 특성 벡터와 산업군 단어 그룹 1에 포함되어 있는 산업군 단어들 각각에 대한 특성 벡터 간에 연산된 유사도들의 평균 값과 “휴대폰”, “노트북” 및 “고등어”에 대한 특성 벡터와 산업군 단어 그룹 2에 포함되어 있는 산업군 단어들 각가에 대한 특성 벡터 간에 연산된 유사도들의 평균 값에 기초하여, 상기 제1 기업의 산업군을 결정할 수 있다.

[0040] 기업 정보 입력부(136)는 상기 제1 기업에 대해 상기 결정된 기업 주소 및 상기 제1 기업에 대해 상기 결정된 적어도 하나의 산업군을 매칭시켜 상기 제1 기업에 대한 기업 정보를 생성한 후 상기 기업 정보 저장부에 저장할 수 있다.

[0041] 예를 들어, 상기 제1 기업의 기업 주소가 주소 7이고, 산업군 정보가 전기 및 전자기기로 결정되고, 상기 제1 기업의 웹 페이지 상에서 상기 적어도 하나의 중요 단어가 “휴대폰”, “노트북” 및 “고등어”로 선택된 경우, 기업 정보 입력부(136)는 상기 표 1과 같은 기업 정보를 하기 표 3과 같이 갱신할 수 있다.

표 3

[0042]

기업 정보 인덱스	기업 주소	산업군 정보	중요 단어
기업 정보 1	주소 1	농림수산품	단어 1, 단어 2, 단어 3
기업 정보 2	주소 2	농림수산품	단어 1, 단어 3
기업 정보 3	주소 3	광산품	단어 4, 단어 5, 단어 6
기업 정보 4	주소 4	음식료품	단어 7, 단어 8
기업 정보 5	주소 5	문화 및 기타 서비스	단어 9, 단어 10, 단어 11
기업 정보 6	주소 6	문화 및 기타 서비스	단어 9, 단어 11, 단어 12
기업 정보 7	주소 7	전기 및 전자기기	휴대폰, 노트북, 고등어
...	...	...	...

[0043] 한편, 기업 정보 입력부(136)는 상기 적어도 하나의 중요 단어 중 상기 제1 기업의 산업군 정보에 상응하는 산업군 단어 그룹에 포함되어 있는 상기 다수의 산업군 단어들 각각에 대한 특성 벡터 간의 유사도가 사전 설정된 기준 유사도 이하인 중요 단어의 경우, 기업 정보에 매칭시켜 저장하지 않을 수 있다. 이에 따라, “휴대폰”, “노트북” 및 “고등어” 중 “고등어”는 상기 제1 기업에 대한 기업 정보에 포함되지 않을 수 있다.

[0044] 다음으로, 산업융복합지수 저장부(140)는 상기 사전 설정된 다수의 산업군들 간의 산업융복합지수를 저장할 수 있다. 여기에서, 산업융복합지수는, 예를 들어, 한국은행에서 발행하는 산업연관표 중 국산투입계수 표의 행렬을 전치(Transpose)한 후 행의 합계가 1이 되도록 행렬에 기재된 성분들의 값을 정규화한 값일 수 있다.

[0045] 예컨대, 산업융복합지수 저장부(140)에는 하기의 표 4와 같이 정보가 저장되어 있을 수 있다.

표 4

[0046]

	농림수산품	광산품	음식료품	...	문화 및 기타 서비스	합계
농림수산품	0.130024	0.000059	0.405923	...	0.003776	1
광산품	0.002392	0.000000	0.002877	...	0.010724	1
음식료품	0.408163	0.000272	0.17488	...	0.002731	1
...	...	...	...	...	...	...
문화 및 기타 서비스	0.002662	0.000026	0.059479	...	0.033729	1

[0047] 표 4를 참조하면, 농림수산품의 경우, 농림수산품과 0.130024, 광산품과 0.000059, 음식료품과 0.405923 및 문화 및 기타 서비스와 0.003776의 산업융복합지수를 가지며, 산업들 간의 산업융복합지수의 총 합은 1임을 알 수 있으며, 마찬가지로, 문화 및 기타 서비스의 경우, 농림수산품과 0.002662, 광산품과 0.000026, 음식료품과 0.059479 및 문화 및 기타 서비스와 0.033729의 산업융복합지수를 가지며, 산업들 간의 산업융복합지수의 총 합은 1임을 알 수 있다.

[0048] 산업별 기업 개수 카운트부(150)는 관리자의 단말(190)로부터 제1 지역 명칭 및 제2 지역 명칭이 입력되면, 상기 기업 정보 저장부에 저장되어 있는 상기 기업 정보에 매칭되어 있는 기업 주소에 기초하여 상기 다수의 산업군들 각각마다 상기 제1 지역 명칭에 따른 제1 지역 및 제2 지역 명칭에 따른 제2 지역에 위치하는 제1 산업별 기업 개수 및 제2 산업별 기업 개수를 카운트할 수 있다. 관리자의 단말(190)로부터 입력되는 제1 지역 명칭 및 제2 지역 명칭이 서로 동일한 경우, 산업별 기업 개수 카운트부(150)는 제1 산업별 기업 개수 및 제2 산업별 기업 개수를 동일하게 카운트할 수 있다. 이때, 단말 데이터 수신부(155)가 관리자의 단말(190)로부터 제1 지역 명칭 및 제2 지역 명칭을 수신할 수 있다.

[0049] 예를 들어, 기업 정보 1 내지 3을 각각 가지는 기업 1 내지 3이 상기 제1 지역 명칭에 따른 제1 지역에 위치하고, 기업 정보 4 내지 6을 각각 가지는 기업 4 내지 6이 상기 제2 지역 명칭에 따른 제2 지역에 위치하는 경우, 산업별 기업 개수 카운트부(150)는 제1 지역의 농림수산품에 대해 제1 산업별 기업 개수를 2로 카운트하고, 광산품에 대해 제1 산업별 기업 개수를 1로 카운트하며, 제2 지역의 음식료품에 대해 제2 산업별 기업 개수를 1로 카운트하고, 문화 및 기타 서비스에 대해 제2 산업별 기업 개수를 2로 카운트할 수 있다. 그러나, 예를 들어, 제1 지역 명칭 및 제2 지역 명칭이 서로 동일한 경우, 산업별 기업 개수 카운트부(150)는 제1 및 제2 지역의 농림수산품에 대해 제1 및 2 산업별 기업 개수를 2로 카운트하고, 광산품에 대해 제1 및 2 산업별 기업 개수를 1로 카운트하며, 음식료품에 대해 제1 및 2 산업별 기업 개수를 1로 카운트하고, 문화 및 기타 서비스에 대해 제

1 및 2 산업별 기업 개수를 2로 카운트할 수 있다.

[0050] 산업군 선택부(160)는 상기 제1 산업별 기업 개수에 기초하여 상기 다수의 산업군들 중 상기 제1 지역에 위치하는 기업의 개수가 많은 순서대로 N개의 제1 산업군들을 선택하고, 상기 제2 산업별 기업 개수에 기초하여 상기 다수의 산업군들 중 상기 제2 지역에 위치하는 기업의 개수가 많은 순서대로 M개의 제2 산업군들을 선택할 수 있다. 상기 N 및 M은 1 이상의 정수이다.

[0051] 예를 들어, N 및 M이 1인 경우, 산업군 선택부(160)는 제1 지역에 대해 제1 산업별 기업 개수가 2인 농림수산품을 제1 산업군으로 선택하고, 제2 지역에 대해 제2 산업별 기업 개수가 2인 문화 및 기타 서비스를 제2 산업군으로 선택할 수 있다. 한편, 예를 들어, 제1 지역 명칭 및 제2 지역 명칭이 서로 동일하고, N 및 M이 2인 경우, 산업군 선택부(160)는 제1 및 2 지역에 대해 제1 및 2 산업별 기업 개수가 2인 농림수산품과 문화 및 기타 서비스를 제1 및 2 산업군으로 선택할 수 있다.

[0052] 융합 지수 연산부(170)는 상기 산업융복합지수에 기초하여 상기 제1 지역에 대해 선택된 제1 산업군들과 상기 제2 지역에 대해 선택된 제2 산업군들 간의 산업융복합지수를 결정하고, 상기 결정된 산업융복합지수의 평균 값을 지역 융합 지수로 결정할 수 있다. 구체적으로, 융합 지수 연산부(170)는 제1 산업군들 각각의 제2 산업군들 각각에 대한 산업융복합지수를 결정하고, 제2 산업군들 각각의 제1 산업군들 각각에 대한 산업융복합지수를 결정한 후 상기 결정된 산업융복합지수의 평균 값을 지역 융합 지수로 결정할 수 있다.

[0053] 예를 들어, 제1 지역에 대해 농림수산품이 제1 산업군으로 선택되고, 제2 지역에 대해 문화 및 기타 서비스가 제2 산업군으로 선택된 경우, 융합 지수 연산부(170)는 제1 산업군인 농림수산품의 제2 산업군인 문화 및 기타 서비스에 대한 산업융복합지수를 0.003776으로 결정하고, 제2 산업군인 문화 및 기타 서비스의 제1 산업군인 농림수산품에 대한 산업융복합지수를 0.002662로 결정한 후 상기 결정된 산업융복합지수의 평균 값(즉, 0.003219)을 제1 지역 및 제2 지역 간의 지역 융합 지수로 결정할 수 있다.

[0054] 그러나, 예를 들어, 제1 지역 명칭 및 제2 지역 명칭이 서로 동일하여, 제1 및 2 지역에 대해 농림수산품과 문화 및 기타 서비스가 제1 및 2 산업군으로 선택된 경우, 제1 산업군인 농림수산품의 제2 산업군인 농림수산품 및 문화 및 기타 서비스에 대한 산업융복합지수를 0.130024 및 0.003776으로 결정하고, 제1 산업군인 문화 및 기타 서비스의 제2 산업군인 농림수산품 및 문화 및 기타 서비스에 대한 산업융복합지수를 0.002662 및 0.033729로 결정하고, 제2 산업군인 농림수산품의 제1 산업군인 농림수산품 및 문화 및 기타 서비스에 대한 산업융복합지수를 0.130024 및 0.003776으로 결정하고, 제2 산업군인 문화 및 기타 서비스의 제1 산업군인 농림수산품 및 문화 및 기타 서비스에 대한 산업융복합지수를 0.002662 및 0.033729로 결정한 후 상기 결정된 산업융복합지수의 평균 값(즉, 0.0425477)을 제1 지역(또는 제2 지역)의 지역 융합 지수로 결정할 수 있다.

[0055] 한편, 상기 융합 지수 연산부(170)는 상기 제1 산업군들과 상기 제2 산업군들 간의 상기 산업융복합지수를 결정하되, 상기 제1 산업군들과 상기 제2 산업군들에 중복되는 제3 산업군 간에 결정된 산업융복합지수를 제외한 산업융복합지수의 평균 값을 지역 융합 지수로 결정할 수 있다.

[0056] 예를 들어, 농림수산품과 문화 및 기타 서비스가 상기 제1 산업군들과 상기 제2 산업군들에 동시에 포함되는 경우, 상기 융합 지수 연산부(170)는 제1 또는 제2 산업군인 농림수산품의 제2 또는 제1 산업군인 농림수산품에 대한 산업융복합지수(즉, 0.130024)와 제1 또는 제2 산업군인 문화 및 기타 서비스의 제2 또는 제1 산업군인 문화 및 기타 서비스에 대한 산업융복합지수(즉, 0.033729)를 제외한 산업융복합지수의 평균 값(즉, 0.003219)을 지역 융합 지수로 결정할 수 있다.

[0057] 융합 지수 전송부(180)는 상기 연산된 지역 융합 지수를 상기 관리자의 단말(190)에 전송할 수 있다.

[0058] 상술한 바와 같은 인공지능 기반의 웹 페이지 분석에 기초한 융합 지수 연산 장치(100)는 기업들이 융합 활동을 수행함으로써 기업들이 목표하는 신기술 개발과 경쟁력 제고에 기여할 수 있도록, 여러 산업 군의 기업정보를 추출한 후 별도의 산업연관표를 이용하여 다양한 융합 정보를 분석하고, 분석된 융합 정보를 기업들에게 제공할 수 있다.

[0060] 또한, 상술한 바와 같은 인공지능 기반의 웹 페이지 분석에 기초한 융합 지수 연산 장치(100)는 기업 간의 기업 융합 지수를 연산한 후 상기 연산된 기업 융합 지수를 상기 관리자의 단말(190)에 전송할 수 있다.

[0061] 상기 융합 지수 연산부(170)는 상기 관리자의 단말(190)로부터 제2 기업 명칭 및 제3 기업 명칭이 입력되면, 상기 산업융복합지수에 기초하여 상기 제2 기업 명칭에 따른 제2 기업에 대해 매칭되어 있는 적어도 하나의 산업군과 상기 제3 기업 명칭에 따른 제3 기업에 대해 매칭되어 있는 적어도 하나의 산업군 간의 기업 융합 지수를

연산할 수 있다. 이때, 단말 데이터 수신부(155)가 관리자의 단말(190)로부터 제2 기업 명칭 및 제3 기업 명칭을 수신할 수 있다.

[0062] 예를 들어, 제2 기업 명칭에 대응하는 제2 기업의 산업군이 광산품이고, 제3 기업 명칭에 대응하는 제3 기업의 산업군이 음식료품인 경우, 상기 융합 지수 연산부(170)는 상기 제2 기업 명칭에 따른 제2 기업에 대해 매칭되어 있는 적어도 하나의 산업군과 상기 제3 기업 명칭에 따른 제3 기업에 대해 매칭되어 있는 적어도 하나의 산업군 간의 기업 융합 지수(0.001575)를 연산할 수 있다. 이때, 기업 융합 지수는 제2 기업에 대해 매칭되어 있는 적어도 하나의 산업군의 제3 기업에 대해 매칭되어 있는 적어도 하나의 산업군에 대한 기업 융합 지수(0.002877)와 제3 기업에 대해 매칭되어 있는 적어도 하나의 산업군의 제2 기업에 대해 매칭되어 있는 적어도 하나의 산업군에 대한 기업 융합 지수(0.000272)의 평균 값일 수 있다.

[0063] 상기 융합 지수 전송부(180)는 상기 연산된 기업 융합 지수(0.001575)를 상기 관리자의 단말(190)에 전송할 수 있다.

[0064] 도 2는 웹 페이지 분석에 기초한 키워드 정보 자동추출 방법을 도시한 순서도이다.

[0065] 도 2를 참조하면, 단계(S210)에서는 다수의 기업들 각각에 대해 기업 주소 및 사전 설정된 다수의 산업군들 중 적어도 하나의 산업군이 매칭되어 있는 기업 정보를 저장하는 기업 정보 저장부를 유지할 수 있다.

[0066] 단계(S220)에서는 상기 사전 설정된 다수의 산업군들 간의 산업융복합지수를 저장하는 산업융복합지수 저장부를 유지할 수 있다.

[0067] 단계(S230)에서는 관리자의 단말로부터 제1 지역 명칭 및 제2 지역 명칭이 입력되면, 상기 기업 정보 저장부에 저장되어 있는 상기 기업 정보에 매칭되어 있는 기업 주소에 기초하여 상기 다수의 산업군들 각각마다 상기 제1 지역 명칭에 따른 제1 지역 및 제2 지역 명칭에 따른 제2 지역에 위치하는 제1 산업별 기업 개수 및 제2 산업별 기업 개수를 카운트할 수 있다.

[0068] 단계(S240)에서는 상기 제1 산업별 기업 개수에 기초하여 상기 다수의 산업군들 중 상기 제1 지역에 위치하는 기업의 개수가 많은 순서대로 N개의 제1 산업군들을 선택하고, 상기 제2 산업별 기업 개수에 기초하여 상기 다수의 산업군들 중 상기 제2 지역에 위치하는 기업의 개수가 많은 순서대로 M개의 제2 산업군들을 선택할 수 있다.

[0069] 단계(S250)에서는 상기 산업융복합지수에 기초하여 상기 제1 지역에 대해 선택된 제1 산업군들과 상기 제2 지역에 대해 선택된 제2 산업군들 간의 산업융복합지수를 결정하고, 상기 결정된 산업융복합지수의 평균 값을 지역 융합 지수로 결정할 수 있다.

[0070] 상기 단계(S250)는 상기 제1 산업군들과 상기 제2 산업군들 간의 상기 산업융복합지수를 결정하되, 상기 제1 산업군들과 상기 제2 산업군들에 중복되는 제3 산업군 간에 결정된 산업융복합지수를 제외한 산업융복합지수의 평균 값을 지역 융합 지수로 결정할 수 있다.

[0071] 단계(S260)에서는 상기 연산된 지역 융합 지수를 상기 관리자의 단말에 전송할 수 있다.

[0072] 상술한 바와 같은 본 발명의 일실시예에 따른 웹 페이지 분석에 기초한 키워드 정보 자동추출 방법은 서로 다른 특성 벡터들이 할당되어 있는 사전 설정된 다수의 산업군 단어들을 포함하는 사전 설정된 다수의 산업군 단어 그룹들을 저장하는 산업군 단어 저장부를 유지하는 단계, 제1 기업의 웹 페이지에 대한 접속 주소가 입력되면, 상기 접속 주소를 기초로 상기 제1 기업의 웹 페이지에 접속하여 상기 제1 기업의 웹 페이지로부터 상기 제1 기업의 웹 페이지 상에 존재하는 복수의 제1 텍스트들을 추출하는 단계, 상기 복수의 제1 텍스트들에 대해 형태소 분석을 수행하여 상기 복수의 제1 텍스트들로부터 복수의 제1 단어들을 추출하는 단계, 상기 복수의 제1 단어들에 기초하여 상기 제1 기업의 기업 주소를 결정하는 단계, 상기 복수의 제1 단어들의 상기 제1 기업의 웹 페이지 상에서의 등장 빈도수에 기초하여 상기 복수의 제1 단어들 중 적어도 하나의 중요 단어를 선택하는 단계, 상기 적어도 하나의 중요 단어가 선택되면, 상기 산업군 단어 저장부를 참조하여, 상기 적어도 하나의 중요 단어에 대한 특성 벡터를 확인하고, 상기 적어도 하나의 중요 단어에 대한 특성 벡터와 상기 다수의 산업군 단어 그룹들 각각에 포함되어 있는 상기 다수의 산업군 단어들 각각에 대한 특성 벡터 간의 유사도를 연산하고, 상기 연산된 유사도에 기초하여 상기 제1 기업에 대해 상기 다수의 산업군들 중 적어도 하나의 산업군을 결정하는 단계 및 상기 제1 기업에 대해 상기 결정된 기업 주소 및 상기 제1 기업에 대해 상기 결정된 적어도 하나의 산업군을 매칭시켜 상기 제1 기업에 대한 기업 정보를 생성한 후 상기 기업 정보 저장부에 저장하는 단계를 더 수행할 수 있다.

[0073] 또한, 상술한 바와 같은 본 발명의 일실시예에 따른 웹 페이지 분석에 기초한 키워드 정보 자동추출 방법에 있어서, 상기 기업 정보 저장부를 유지하는 단계는 상기 다수의 기업들 각각에 대해 중요 단어가 더 매칭되어 있는 기업 정보를 저장하는 상기 기업 정보 저장부를 유지하고, 상기 기업 정보 입력 단계는 상기 제1 기업에 대해 상기 적어도 하나의 중요 단어가 더 매칭된 상기 제1 기업에 대한 기업 정보를 생성한 후 상기 기업 정보 저장부에 저장할 수 있다.

[0074] 또한, 상술한 바와 같은 본 발명의 일실시예에 따른 웹 페이지 분석에 기초한 키워드 정보 자동추출 방법은 상기 관리자의 단말로부터 제2 기업 명칭 및 제3 기업 명칭이 입력되면, 상기 산업융복합지수에 기초하여 상기 제2 기업 명칭에 따른 제2 기업에 대해 매칭되어 있는 적어도 하나의 산업군과 상기 제3 기업 명칭에 따른 제3 기업에 대해 매칭되어 있는 적어도 하나의 산업군 간의 기업 융합 지수를 연산하는 단계 및 상기 연산된 기업 융합 지수를 상기 관리자의 단말에 전송하는 단계를 더 수행할 수 있다.

[0075] 이상, 도 2를 참조하여 본 발명의 일실시예에 따른 융합 지수 연산 장치의 동작 방법에 대해 설명하였다. 여기서, 본 발명의 일실시예에 따른 융합 지수 연산 장치의 동작 방법은 도 1을 이용하여 설명한 융합 지수 연산 장치(100)의 동작에 대한 구성과 대응될 수 있으므로, 이에 대한 보다 상세한 설명은 생략하기로 한다.

[0076] 본 발명의 일실시예에 따른 융합 지수 연산 장치의 동작 방법은 컴퓨터와의 결합을 통해 실행시키기 위한 저장매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다.

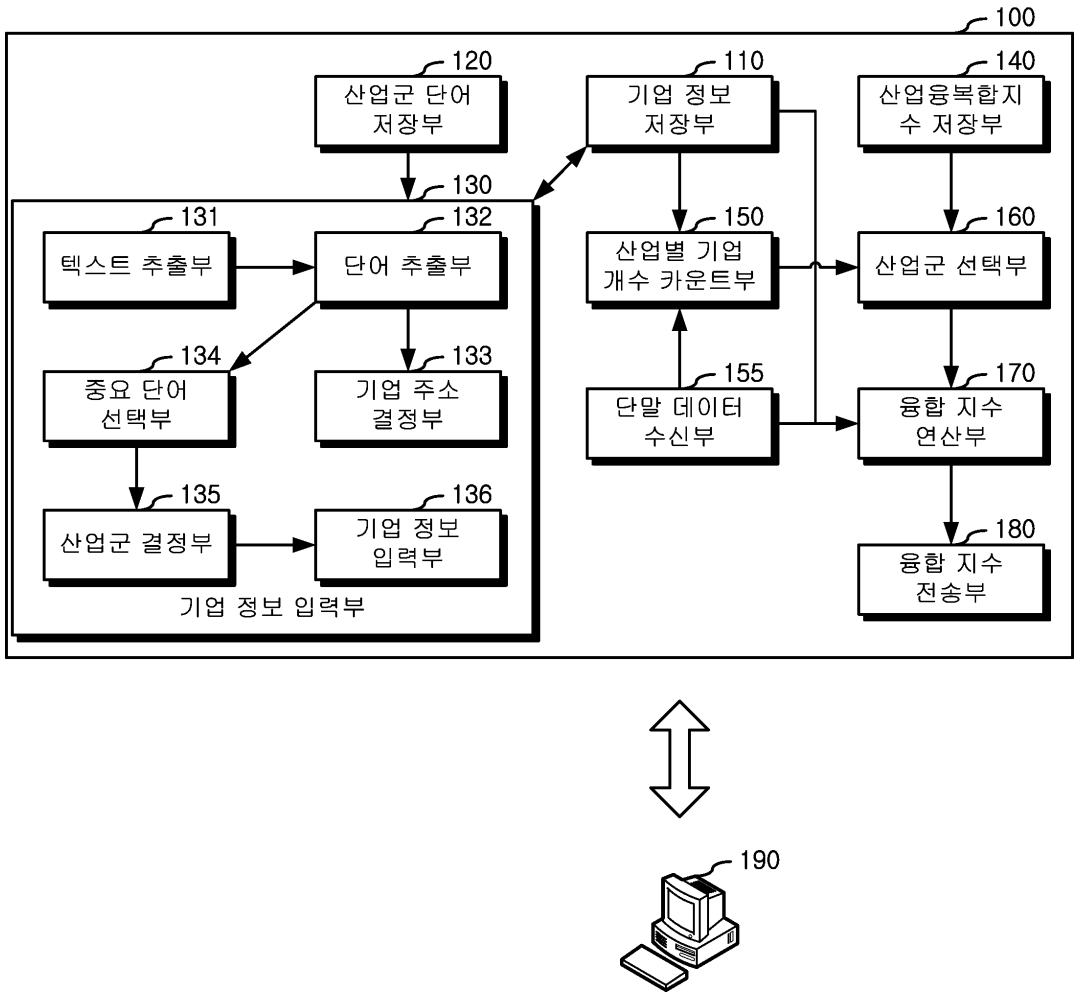
[0077] 또한, 본 발명의 일실시예에 따른 융합 지수 연산 장치의 동작 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 본 발명을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.

[0078] 이상과 같이 본 발명에서는 구체적인 구성 요소 등과 같은 특정 사항들과 한정된 실시예 및 도면에 의해 설명되었으나 이는 본 발명의 보다 전반적인 이해를 돕기 위해서 제공된 것일 뿐, 본 발명은 상기의 실시예에 한정되는 것은 아니며, 본 발명이 속하는 분야에서 통상적인 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및 변형이 가능하다.

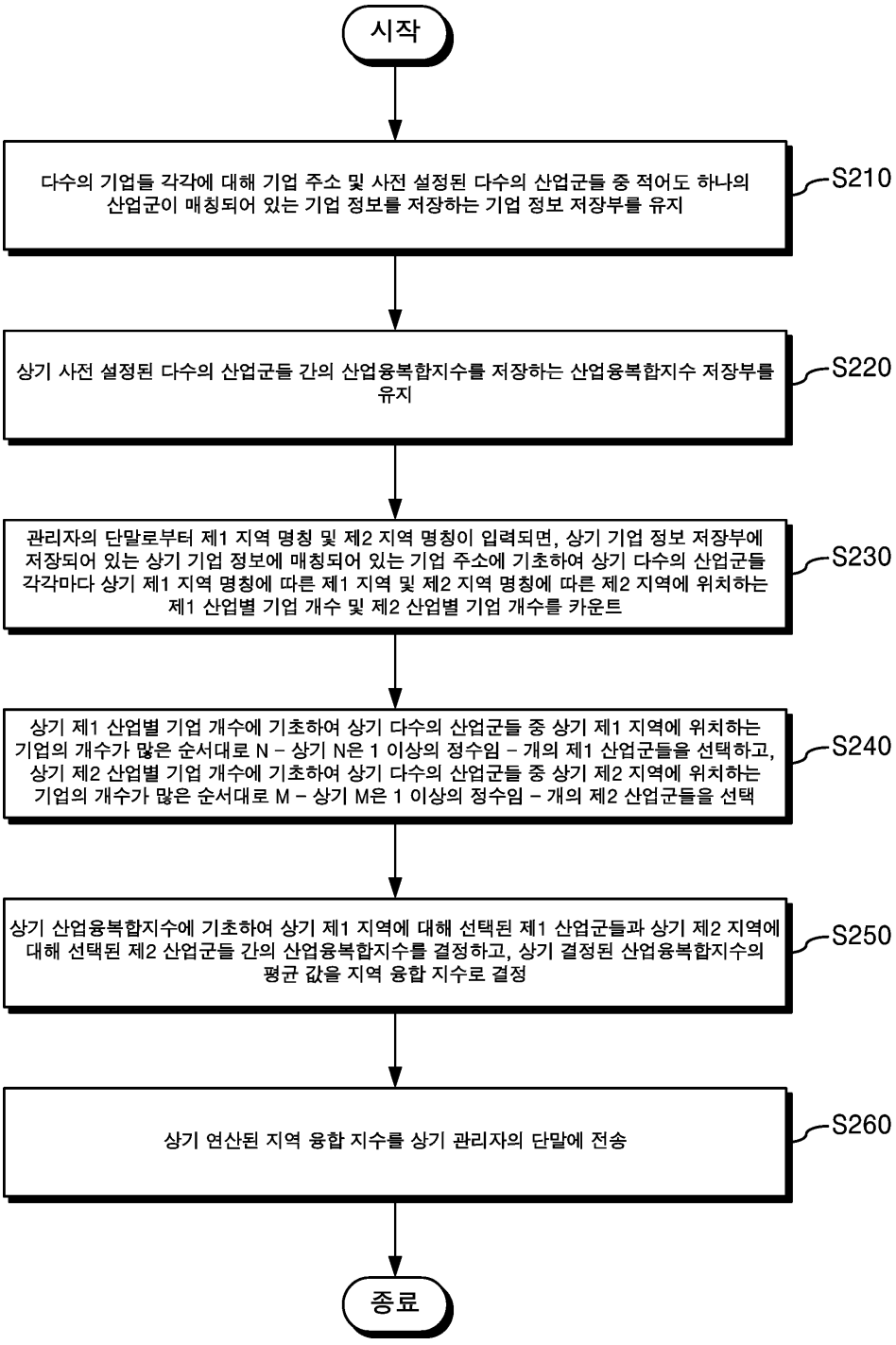
[0079] 따라서, 본 발명의 사상은 설명된 실시예에 국한되어 정해져서는 아니되며, 후술하는 특허청구범위뿐 아니라 이 특허청구범위와 균등하거나 등가적 변형이 있는 모든 것들은 본 발명 사상의 범주에 속한다고 할 것이다.

도면

도면1


100
120
산업군 단어 저장부
110
기업 정보 저장부
140
산업융복합지수 저장부
130
131
텍스트 추출부
132
단어 추출부
150
산업별 기업 개수 카운트부
160
산업군 선택부
134
중요 단어 선택부
133
기업 주소 결정부
155
단말 데이터 수신부
170
융합 지수 연산부
135
산업군 결정부
136
기업 정보 입력부
180
융합 지수 전송부
기업 정보 입력부
190

도면2


시작
다수의 기업들 각각에 대해 기업 주소 및 사전 설정된 다수의 산업군들 중 적어도 하나의 산업군이 매칭되어 있는 기업 정보를 저장하는 기업 정보 저장부를 유지
S210
상기 사전 설정된 다수의 산업군들 간의 산업융복합지수를 저장하는 산업융복합지수 저장부를 유지
S220
관리자의 단말로부터 제1 지역 명칭 및 제2 지역 명칭이 입력되면, 상기 기업 정보 저장부에 저장되어 있는 상기 기업 정보에 매칭되어 있는 기업 주소에 기초하여 상기 다수의 산업군들 각각마다 상기 제1 지역 명칭에 따른 제1 지역 및 제2 지역 명칭에 따른 제2 지역에 위치하는 제1 산업별 기업 개수 및 제2 산업별 기업 개수를 카운트
S230
상기 제1 산업별 기업 개수에 기초하여 상기 다수의 산업군들 중 상기 제1 지역에 위치하는 기업의 개수가 많은 순서대로 N - 상기 N은 1 이상의 정수임 - 개의 제1 산업군들을 선택하고, 상기 제2 산업별 기업 개수에 기초하여 상기 다수의 산업군들 중 상기 제2 지역에 위치하는 기업의 개수가 많은 순서대로 M - 상기 M은 1 이상의 정수임 - 개의 제2 산업군들을 선택
S240
상기 산업융복합지수에 기초하여 상기 제1 지역에 대해 선택된 제1 산업군들과 상기 제2 지역에 대해 선택된 제2 산업군들 간의 산업융복합지수를 결정하고, 상기 결정된 산업융복합지수의 평균 값을 지역 융합 지수로 결정
S250
상기 연산된 지역 융합 지수를 상기 관리자의 단말에 전송
S260
종료