



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2016-0026811

(43) 공개일자 2016년03월09일

(51) 국제특허분류(Int. Cl.)

G06F 17/00 (2006.01) G06F 17/18 (2006.01)

(52) CPC특허분류

G06F 17/00 (2013.01)

G06F 17/18 (2013.01)

(21) 출원번호 10-2015-0123871

(22) 출원일자 2015년09월01일

심사청구일자 2015년09월01일

(30) 우선권주장

1020140115754 2014년09월01일 대한민국(KR)

(71) 출원인

경희대학교 산학협력단

경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)

(72) 발명자

진창호

경기도 성남시 분당구 수내로192번길 25, 402동 1505호 (수내동, 푸른마을쌍용아파트)

아이다 과달루페 메르카도 에르난데스

경기도 용인시 기흥구 덕영대로 1732 (서천동)

(74) 대리인

특허법인 신지

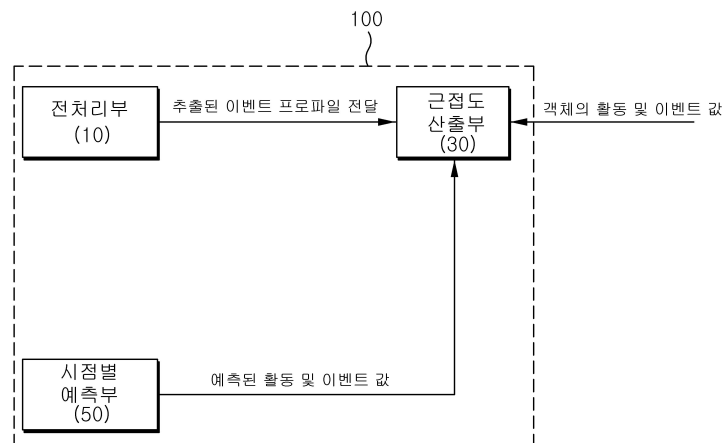
전체 청구항 수 : 총 19 항

(54) 발명의 명칭 목표 달성 진단 장치 및 진단 방법

(57) 요약

본 발명에 따른 목표 달성 진단 장치는 프로세스 목표, 최종성과 유형, 속성, 수행기간 및 평가시점을 정의하고, 정의된 목표, 최종성과 유형, 속성, 수행기간 및 평가시점에 부합되는 하나 이상의 과거 프로세스 인스턴스로부터 속성 값을 수집 또는 확인하며, 프로세스 인스턴스의 속성 값으로부터 이벤트 프로파일을 추출하는 전처리부 및 추출된 이벤트 프로파일에 기초하여 사전확률 및 사후확률을 산출하고, 산출된 사전확률 및 사후확률을 이용하여 정의된 평가시점에서의 현재 진행 중인 프로세스 인스턴스의 시점별 최종 성과 유형별 근접도를 산출하는 근접도 산출부를 포함한다.

대표도 - 도1



명세서

청구범위

청구항 1

프로세스 목표, 최종성과 유형, 속성, 수행기간 및 평가시점을 정의하고, 상기 정의된 목표, 최종성과 유형, 속성, 수행기간 및 평가시점에 부합되는 하나 이상의 과거 프로세스 인스턴스로부터 속성 값을 수집 또는 확인하며, 상기 프로세스 인스턴스의 속성 값으로부터 이벤트 프로파일을 추출하는 전처리부; 및

상기 추출된 이벤트 프로파일에 기초하여 현재 진행 중인 프로세스 인스턴스의 사전확률 및 사후확률을 산출하고, 상기 산출된 사전확률 및 사후확률을 이용하여 상기 정의된 평가시점에서의 현재 진행 중인 프로세스 인스턴스의 시점별 최종 성과 유형별 근접도를 산출하는 근접도 산출부;

를 포함하는 것을 특징으로 하는 목표 달성 진단 장치.

청구항 2

제1항에 있어서,

상기 프로세스 인스턴스의 속성은 환경 속성, 객체 속성, 활동 속성 및 이벤트 속성 중에서 적어도 하나 이상을 포함하는 것을 특징으로 하는 목표 달성 진단 장치.

청구항 3

제1항에 있어서,

상기 전처리부는 상기 과거 프로세스 인스턴스에 대해 하나 이상의 활동 속성 및 하나 이상의 이벤트 속성 중에서 적어도 하나의 속성에 기초하여 군집분석을 실시하여 군집(cluster)을 형성하고, 각 군집으로부터 군집(cluster) 내 과거 프로세스 인스턴스의 속성값이 형성하는 하나 이상의 이벤트 프로파일을 추출하는 것을 특징으로 하는 목표 달성 진단 장치.

청구항 4

제1항에 있어서,

상기 근접도 산출부는 상기 추출된 하나 이상의 이벤트 프로파일에 대해 상기 현재 진행 중인 프로세스 인스턴스가 각 이벤트 프로파일에 속할 사전확률을 산출하고, 상기 산출된 사전확률에 우도(Likelihood) 함수를 적용하여 사후확률을 산출하며, 상기 산출된 사전확률 및 사후확률에 베이저언 정리(Bayesian Theorem)를 적용하여 상기 현재 진행 중인 프로세스 인스턴스의 시점별 최종 성과 유형별 근접도를 산출하는 것을 특징으로 하는 목표 달성 진단 장치.

청구항 5

제4항에 있어서,

상기 시점별 최종 성과 유형별 근접도는

$$P(O_i | D_i) = \sum_{j=1}^n P(I \in EP_j | D_i) P(O_i | I \in EP_j \cap D_i)$$

에 의해 산출되며,

O_i 는 i 번째 최종성과 유형, 상기 D_i 는 t 번째 평가시점까지 프로세스 인스턴스 I 에 의해 발생한 활동에 따른 하나 이상의 이벤트의 값으로 구성된 집합, 상기 I 는 현재 진행 중인 프로세스 인스턴스, 상기 EP_j 는 j 번째 이벤트 프로파일을 나타내는 것을 특징으로 하는 목표 달성 진단 장치.

청구항 6

제1항에 있어서,

상기 근접도 산출부는 목표를 달성하기 위해 필요한 프로세스 수행기간이 다양하게 발생할 수 있는 경우, 모든 발생 가능한 수행기간 각각에 대해 각 수행기간을 가진 프로세스에 속하는 과거 프로세스 인스턴스의 속성값을 수집 또는 확인하여 상기 수행기간에 대응하는 최종 성과 유형별 근접도 및 상기 수행기간이 발생할 확률을 산출하고, 수행기간에 대응하는 최종 성과 유형별 근접도 및 상기 수행기간이 발생할 확률에 기초하여 최종성과 유형의 발생 기대확률을 산출하는 것을 특징으로 하는 목표 달성 진단 장치.

청구항 7

제6항에 있어서,

상기 최종성과 유형별 기대확률은

$$\sum_{L=T}^{\infty} P(\text{project_length} = L | L \geq T) * P(O_{i,L} | D_{i,L})$$

에 의해 산출되며,

상기 L은 상기 프로세스 수행기간, 상기 T는 최소 프로젝트 수행기간, $O_{i,L}$ 은 수행기간 L에 대한 i번째 최종성과 유형, 상기 $D_{i,L}$ 은 수행기간 L에 대한 t번째 평가시점까지 프로세스 인스턴스 I에 의해 발생한 활동에 따른 하나 이상의 이벤트의 값으로 구성된 집합을 나타내는 것을 특징으로 하는 목표 달성 진단 장치.

청구항 8

제1항에 있어서,

현재 진행 중인 프로세스 인스턴스의 활동 속성값 및 활동에 따른 이벤트 속성값에 기초해서 미래의 후속 평가시점에서의 활동 속성값 및 상기 활동에 따른 이벤트 속성값을 예측하는 시점별 예측부;

를 더 포함하는 것을 특징으로 하는 목표 달성 진단 장치.

청구항 9

제8항에 있어서,

상기 현재 진행 중인 프로세스 인스턴스의 미래 후속 평가시점에서의 활동 속성값 및 활동에 따른 이벤트 속성값 예측에 사용될 이벤트 프로파일 유형은,

$$s = \arg \max_j P(I \in EP_j | D_t)$$

에 의해 도출되며,

상기 EP_j 는 j번째 이벤트 프로파일, I는 현재 진행 중인 프로세스 인스턴스, D_t 는 t번째 평가시점까지 프로세스 인스턴스 I에 의해 발생한 활동에 따른 하나 이상의 이벤트의 값으로 구성된 집합, s는 최근 평가시점 t까지 획득된 이벤트 값들이 벡터 D_t 를 형성하는 경우에 프로세스 인스턴스 I가 속할 확률이 가장 높은 이벤트 프로파일 유형을 나타내는 것을 특징으로 하는 목표 달성 진단 장치.

청구항 10

프로세스 목표, 최종성과 유형, 속성, 수행기간 및 평가시점을 정의하는 단계;

상기 정의된 목표, 최종성과 유형, 속성, 수행기간 및 평가시점에 부합되는 하나 이상의 과거 프로세스 인스턴스로부터 속성 값을 수집 또는 확인하는 단계;

상기 프로세스 인스턴스의 속성 값으로부터 이벤트 프로파일을 추출하는 단계;

상기 추출된 이벤트 프로파일에 기초하여 사전확률 및 사후확률을 산출하는 단계;

상기 산출된 사전확률 및 사후확률을 이용하여 상기 정의된 평가시점에서의 현재 진행 중인 프로세스 인스턴스의 시점별 최종 성과 유형별 근접도를 산출하는 단계;

를 포함하는 것을 특징으로 하는 목표 달성 진단 방법.

청구항 11

제10항에 있어서,

상기 프로세스 인스턴스의 속성은 환경 속성, 객체 속성, 활동 속성 및 이벤트 속성 중에서 적어도 하나 이상을 포함하는 것을 특징으로 하는 목표 달성 진단 방법.

청구항 12

제10항에 있어서,

상기 과거 프로세스 인스턴스의 속성 값으로부터 이벤트 프로파일을 추출하는 단계는,

상기 과거 프로세스 인스턴스에 대해 하나 이상의 활동 속성 및 하나 이상의 이벤트 속성 중 적어도 하나의 속성에 기초하여 군집분석을 실시하여 군집(cluster)을 형성하고, 각 군집에 대해 해당 군집에 속한 과거 프로세스 인스턴스들의 속성 값이 형성하는 하나 이상의 이벤트 프로파일을 추출하는 것을 특징으로 하는 목표 달성 진단 방법.

청구항 13

제10항에 있어서,

상기 추출된 이벤트 프로파일에 기초하여 사전확률 및 사후확률을 산출하는 단계는,

상기 추출된 하나 이상의 이벤트 프로파일에 대해 과거 프로세스 인스턴스가 각 이벤트 프로파일에 속할 사전확률을 산출하는 단계; 및

상기 산출된 사전확률에 우도(Likelihood) 함수를 적용하여 사후확률을 산출하는 단계;

를 포함하는 것을 특징으로 하는 목표 달성 진단 방법.

청구항 14

제13항에 있어서,

상기 시점별 최종 성과 유형별 근접도를 산출하는 단계는,

상기 산출된 사전확률 및 사후확률에 베이지언 정리(Bayesian Theorem)를 적용하여 현재 진행 중인 프로세스 인스턴스의 시점별 최종 성과 유형별 근접도를 산출하는 것을 특징으로 하는 목표 달성 진단 방법.

청구항 15

제14항에 있어서,

상기 시점별 최종 성과 유형별 근접도는

$$P(O_i | D_i) = \sum_{j=1}^n P(I \in EP_j | D_i) P(O_i | I \in EP_j \cap D_i) \text{에 의해 산출되며,}$$

O_i 는 i 번째 최종성과 유형, 상기 D_i 는 t 번째 평가시점까지 프로세스 인스턴스 I 에 의해 발생한 활동에 따른 하나 이상의 이벤트의 값으로 구성된 집합, 상기 I 는 현재 진행 중인 프로세스 인스턴스, 상기 EP_j 는 j 번째 이벤트 프로파일을 나타내는 것을 특징으로 하는 목표 달성 진단 방법.

청구항 16

제10항에 있어서,

목표를 달성하기 위해 필요한 프로세스 수행기간이 다양하게 발생할 수 있는 경우, 모든 발생 가능한 수행기간 각각에 대해 각 수행기간을 가진 프로세스에 속하는 과거 프로세스 인스턴스의 속성값을 수집 또는 확인하여 상기 수행기간에 대응하는 최종 성과 유형별 근접도 및 상기 수행기간이 발생할 확률을 산출하는 단계; 및

수행기간에 대응하는 최종 성과 유형별 근접도 및 상기 수행기간이 발생할 확률에 기초하여 최종성과 유형의 발생 기대확률을 산출하는 단계;

를 더 포함하는 것을 특징으로 하는 목표 달성 진단 방법.

청구항 17

제16항에 있어서,

상기 최종성과 유형별 기대확률은

$$\sum_{L=T}^{\infty} P(\text{project_length} = L | L \geq T) * P(O_{i,L} | D_{i,L})$$

에 의해 산출되며,

상기 L은 상기 프로세스 수행기간, 상기 T는 최소 프로젝트 수행기간, $O_{i,L}$ 은 수행기간 L에 대한 i번째 최종성과 유형, 상기 $D_{i,L}$ 은 수행기간 L에 대한 t번째 평가시점까지 프로세스 인스턴스 I에 의해 발생한 활동에 따른 하나 이상의 이벤트의 값으로 구성된 집합을 나타내는 것을 특징으로 하는 목표 달성 진단 방법.

청구항 18

제10항에 있어서,

현재 진행 중인 프로세스 인스턴스의 활동 및 활동에 따른 이벤트 속성값에 기초해서 후속 평가시점의 활동 및 상기 활동에 따른 이벤트 속성값을 예측하는 단계;

를 더 포함하는 것을 특징으로 하는 목표 달성 진단 방법.

청구항 19

제18항에 있어서,

상기 현재 진행 중인 프로세스 인스턴스의 미래 후속 평가시점에서의 활동 속성값 및 활동에 따른 이벤트 속성값 예측에 사용될 이벤트 프로파일 유형은,

$$s = \arg \max_j P(I \in EP_j | D_t)$$

에 의해 도출되며,

상기 EP_j 는 j번째 이벤트 프로파일, I는 현재 진행 중인 프로세스 인스턴스, D_t 는 t번째 평가시점까지 프로세스 인스턴스 I에 의해 발생한 활동에 따른 하나 이상의 이벤트의 값으로 구성된 집합, s는 최근 평가시점 t까지 획득된 이벤트 값들이 벡터 D_t 를 형성하는 경우에 프로세스 인스턴스 I가 속할 확률이 가장 높은 이벤트 프로파일 유형을 나타내는 것을 특징으로 하는 목표 달성 진단 방법.

발명의 설명

기술 분야

[0001] 본 발명은 상황 예측 방법에 관한 기술로서, 보다 상세하게는 과업 수행 진행 상황을 예측 및 진단하는 방법과 관련한 기술이다.

배경 기술

[0002] 다양한 공정에서의 결과를 예측하고, 다가올 미래 시점에 어떤 활동을 취할지 예측하기 위한 기술들이 개발되고 있다. 이와 같은 예측 기술의 하나로 베이저안 정리(Bayesian Theorem)가 많이 활용되고 있다. 베이저안 정리는 제조업 분야에서 제품의 품질관리에 활용되어 생산물의 불량률을 예측하는데 활용되고 있으며, 기상예보 및 임상시험 등과 같이 다양한 분야에서 예측을 위한 기술로서 활용되고 있다. 베이저안 정리를 이용한 예측 기술은 상술한 예측 분야의 활용에 국한되는 것이 아니라, 사람의 행동이나 목표를 예측하거나, 미래 시점의 행동을 예측하는 분야에도 활용이 되고 있다. 하지만, 종래의 예측 기술은 정확도가 높지 않으며, 이러한 예측을 위한 구체적인 알고리즘을 포함하고 있지 않다.

[0003] 대한민국 공개특허 제10-2011-0043277호는 발전설비 각 부품의 고장을 예측하기 위한 기술이 기재되어 있다. 상기 특허는 발전설비 각 부품의 고장 이력에 대한 데이터베이스를 구축하고, 이를 토대로 베이저안 통계기법을 사용하여 발전설비 각 부품들의 고장확률을 예측한다. 하지만, 상기 특허는 베이저안 통계기법을 발생 오류를

예측하는 분야에만 활용하고 있을 뿐, 행동 예측이나 시점별 예측에 대한 활용을 다루지 않는다.

선행기술문헌

특허문헌

[0004] (특허문헌 0001) 대한민국 공개특허 제10-2011-0043277호

발명의 내용

해결하려는 과제

[0005] 본 발명이 해결하고자 하는 과제는 과업 목표 달성을 위해 취할 시점별 활동 및 이벤트 속성값 예측과 최종 성과 유형별 근접도 산출 결과를 제공할 수 있는 진단 장치 및 진단 방법을 제공하는 것이다.

과제의 해결 수단

[0006] 본 발명에 따른 목표 달성 진단 장치는 프로세스 목표, 최종성과 유형, 속성, 수행기간 및 평가시점을 정의하고, 정의된 목표, 최종성과 유형, 속성, 수행기간 및 평가시점에 부합되는 하나 이상의 과거 프로세스 인스턴스의 속성값들을 수집 또는 확인하며, 속성값들로부터 이벤트 프로파일을 추출하는 전처리부 및 추출된 이벤트 프로파일에 기초하여 사전확률 및 사후확률을 산출하고, 산출된 사전확률 및 사후확률을 이용하여 정의된 평가시점에서의 진행 중인 프로세스 인스턴스의 시점별 최종 성과 유형별 근접도를 산출하는 근접도 산출부를 포함하며, 진행 중인 프로세스 인스턴스의 활동 및 활동에 따른 이벤트 속성값을 통해 후속 평가시점의 활동 및 그 활동에 따른 이벤트값들을 예측하는 시점별 예측부를 더 포함할 수 있다.

[0007] 프로세스 인스턴스의 속성은 환경 속성, 객체 속성, 활동 속성 및 이벤트 속성 중에서 적어도 하나 이상을 포함할 수 있다. 전처리부는 과거 프로세스 인스턴스들에 대해 활동 속성과 이벤트 속성 중 적어도 하나의 속성에 기초하여 군집분석을 실시하여 군집(cluster)을 형성하고, 각 군집으로부터 군집(cluster) 내 프로세스 인스턴스의 속성값들이 형성하는 하나 이상의 이벤트 프로파일을 추출한다. 이 군집분석을 '이벤트 프로파일 군집분석'이라고 하고, 그 결과물인 군집을 '프로파일 군집'이라 명명한다.

[0008] 근접도 산출부는 추출된 하나 이상의 이벤트 프로파일에 대해 프로세스 인스턴스가 각 이벤트 프로파일에 속할 사전확률을 산출하고, 산출된 사전확률에 우도(Likelihood) 함수를 적용하여 사후확률을 산출하며, 산출된 사전확률 및 사후확률에 베이저언 정리(Bayesian Theorem)를 적용하여 프로세스 인스턴스의 시점별 최종 성과 유형별 근접도를 산출한다. 그리고, 근접도 산출부는 프로세스 목표를 달성하기 위해 필요한 수행기간이 다양하게 발생할 수 있는 경우, 모든 발생 가능한 수행기간 각각에 대해 각 수행기간을 가진 프로세스에 속하는 과거 프로세스 인스턴스의 속성값들을 수집 또는 확인하여 그 수행기간에 대응하는 최종 성과 유형별 근접도와 그 수행기간이 발생할 확률을 산출하고, 각 수행기간에 대응하는 최종 성과 유형별 근접도와 그 수행기간이 발생할 확률에 기초하여 최종성과 유형의 발생 기대확률을 산출한다.

[0009] 본 발명에 따른 목표 달성 진단 방법은 프로세스 목표, 최종성과 유형, 속성, 수행기간 및 평가시점을 정의하고, 정의된 목표, 최종성과 유형, 속성, 수행기간 및 평가시점에 부합되는 하나 이상의 과거 프로세스 인스턴스로부터 속성 값을 수집 또는 확인한다. 그리고, 프로세스 인스턴스의 속성 값으로부터 이벤트 프로파일을 추출한다. 프로세스 인스턴스의 속성 값으로부터 이벤트 프로파일이 추출되면, 추출된 이벤트 프로파일에 기초하여 사전확률 및 사후확률을 산출한다. 다음으로, 산출된 사전확률 및 사후확률을 이용하여 정의된 평가시점에서의 프로세스 인스턴스의 시점별 최종 성과 유형별 근접도를 산출한다.

[0010] 프로세스 인스턴스의 속성은 환경 속성, 객체 속성, 활동 속성 및 이벤트 속성 중에서 적어도 하나 이상을 포함할 수 있다. 과거 프로세스 인스턴스들의 속성 값들로부터 이벤트 프로파일을 추출하는 단계는 활동 속성과 이벤트 속성 중 적어도 하나의 속성에 기초하여 이벤트 프로파일 군집분석을 통해 과거 프로세스 인스턴스들로 군집(cluster)을 형성하고, 형성된 각 군집(cluster)에 속한 과거 프로세스 인스턴스들의 속성 값들에 기초하여 군집 각각에서 하나 이상의 이벤트 프로파일을 추출한다.

[0011] 그리고, 추출된 이벤트 프로파일에 기초하여 사전확률 및 사후확률을 산출하는 단계는 추출된 하나 이상의 이벤트 프로파일에 대해 프로세스 인스턴스가 각 이벤트 프로파일에 속할 사전확률을 산출하고, 산출된 사전확률에

우도(Likelihood) 함수를 적용하여 사후확률을 산출한다. 시점별 최종 성과 유형별 근접도를 산출하는 단계는 산출된 사전확률 및 사후확률에 베이저언 정리(Bayesian Theorem)를 적용하여 프로세스 인스턴스의 시점별 최종 성과 유형별 근접도를 산출한다.

발명의 효과

[0012] 본 발명에 따른 목표 달성 진단 장치 및 방법은 프로세스의 진행 방향을 예측하여 조기에 대응이 가능하게 한다. 그리고, 다양한 공정 프로세스에서, 통계적 공정관리(Statistical Process Control, SPC) 시스템에 반영되어 원가 절감 및 품질 개선이 가능하며, 마트/병원/은행 프로세스 등에서 진행 상황에 따라 조기 대응을 가능하게 한다. 또한, 개인 생활 모니터링 및 진단에 적용되어 향후 개선 방향 설정에 도움이 될 수 있다.

도면의 간단한 설명

[0013] 도 1은 본 발명에 따른 목표 달성 진단 장치의 일 실시예를 나타내는 구성도이다.
 도 2는 본 발명의 일 실시예에 따른 목표 달성 진단 장치의 처리 과정을 나타내는 흐름도이다.
 도 3a는 본 발명의 일 실시예에 따른 목표 달성 진단 장치의 취업 프로세스 다이어그램의 일례를 나타내는 도면이다.
 도 3b는 도 3a의 취업 프로세스 다이어그램의 일례에서 이벤트 속성값을 그래프로 나타낸 도면이다.
 도 4a 및 도 4b는 본 발명의 일 실시예에 따른 목표 달성 진단 장치의 복수 이벤트 분석 과정을 설명하기 위한 도면이다.
 도 5는 본 발명의 일 실시예에 따른 목표 달성 진단 방법의 일 실시예를 나타내는 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0014] 이하, 본 발명의 실시예를 첨부된 도면들을 참조하여 상세하게 설명한다. 본 명세서에서 사용되는 용어 및 단어들은 실시예에서의 기능을 고려하여 선택된 용어들로서, 그 용어의 의미는 발명의 의도 또는 관례 등에 따라 달라질 수 있다. 따라서 후술하는 실시예에서 사용된 용어는, 본 명세서에 구체적으로 정의된 경우에는 그 정의에 따르며, 구체적인 정의가 없는 경우는 당업자들이 일반적으로 인식하는 의미로 해석되어야 할 것이다.

[0015] 본 발명에 따른 목표 달성 진단 장치는 프로세스 진행 방향 및 결과 예측에 대한 우도 함수(Likelihood Function) 기반 베이저언(Bayesian) 확률 산출 방법을 이용한다. 이를 위해 본 발명에 따른 목표 달성 진단 장치는 진단 대상이 되는 프로세스 인스턴스의 시점별 활동 및 이벤트를 예측하고, 그 프로세스 인스턴스의 시점별 최종 성과 유형별 근접도를 산출할 수 있다. 프로세스는 미리 정해진 목표를 달성하기 위한 활동의 집합으로 구성되어 있다. 프로세스는 여러 가지의 결과들 중 하나로 종료된다. 프로세스 예시로 반도체 공정 프로세스, 은행 입금 프로세스, 다이어트 프로세스, 범죄 프로세스, 인생 프로세스, 취업 프로세스, 졸업 프로세스, 취업 프로세스 등이 있다. 프로세스 중 바람직하지 않은 상황을 조기에 발견함으로써 프로세스의 성과를 향상시키고 자원의 손실을 줄일 수 있다. 또한 다가올 미래 시점에 어떤 활동을 취할지를 예측할 수 있어, 선제 대응할 수 있다.

[0016] 도 1은 본 발명에 따른 목표 달성 진단 장치의 일 실시예를 나타내는 구성도이다.

[0017] 도 1을 참조하면, 본 발명에 따른 목표 달성 진단 장치(100)는 프로세스 인스턴스가 각각의 평가시점에 이를 때마다 프로세스 인스턴스의 후속 시점별 활동을 예측하고, 프로세스 인스턴스의 최종 성과 유형별 근접도를 산출한다.

[0018] 표 1은 본 발명에서 사용되는 용어를 정의한 것이다.

표 1

[0019] 발명의 용어 정의

프로세스 (process)	특정 목표를 달성하기 위해 행해지는 일련의 활동들(activity)
프로세스 인스턴스 (process instance)	프로세스 실행 사례. 예를 들어, 은행 입금 프로세스의 인스턴스는 한 고객이 은행을 방문하여 입금을 시도하여 그 업무를 종료하고 은행을 나가는 것을 의미한다. 즉, 입금을 하고자 방문하는 모든 고객이 각각 하나의 인스턴스를 생성하게 된다.
이벤트 (event)	프로세스가 수행될 때, 그 프로세스 인스턴스의 각 활동(activity)에 의해서 생성되는 결과(outcome). 예를 들어, 은행 입금 프로세스에서 은행원이 입금액을 전산 입력하면 입금액과 시간이 DB에 기록되고, 예금 잔액이 갱신되는 이벤트가 발생된다.
이벤트 프로파일 (event profile)	프로세스 인스턴스가 완료되기까지 생성되는 일련의 이벤트들의 값들이 형성하는 양상의 개요. 예를 들어 은행 입금 프로세스에서 분석 대상이 되는 이벤트들이 방문시간, 입금액, 대기시간, 고객 연령, 입금 완료율이라고 하면, 이벤트 프로파일은 다음과 같을 수 있다. 오전 10시와 11시 사이에 방문한 남자 고객들은 입금액이 10,000,000원과 20,000,000원 사이에 속하고, 대기시간이 10분과 20분 사이에 속하고, 연령이 40세와 49세 사이에 속하고, 입금을 완료하고 돌아가는 비율이 80%와 90% 사이이다. 또한 이벤트 프로파일은 다음과 같이 분포로 표현될 수 있다. 예를 들어, 오전 10시와 11시 사이에 방문한 남자 고객들의 입금액은 평균이 15,000,000원이고 표준편차가 1,500,000원인 정규분포를 따르고, 대기시간은 평균 대기시간이 15분인 포아송분포를 따르고, 입금을 완료하고 돌아가는 비율은 평균 비율이 85%이고 표준편차가 6%인 정규분포를 따른다.

[0020] 프로세스는 미리 정해진 목표를 달성하기 위한 활동의 집합으로 구성될 수 있다. 프로세스는 여러 가지의 결과들 중 어떤 하나의 결과로 종료될 수 있다. 프로세스의 일례로 반도체 공정 프로세스, 은행 입금 프로세스, 다이어트 프로세스, 범죄 프로세스, 인생 프로세스, 취업 프로세스, 졸업 프로세스, 관광 프로세스 및 취업 프로세스 등을 포함할 수 있다. 본 발명에 따른 목표 달성 진단 장치(100)는 프로세스 수행 과정에서 바람직하지 않은 상황을 초기에 발견함으로써 프로세스의 성과를 향상시키고 자원의 손실을 줄일 수 있다. 또한, 본 발명에 따른 목표 달성 진단 장치(100)는 다가올 미래 시점에 어떤 활동을 취할지를 예측할 수 있어, 선제 대응할 수 있다.

[0021] 본 발명에 따른 목표 달성 진단 장치(100)는 다양한 프로세스에 활용될 수 있다. 예를 들어, 식이조절 프로세스의 경우, 정해진 목표는 감량 무게가 될 수 있고, 활동은 식단 조절, 운동, 대중 교통 이용, 수면 시간 조절, 섭취 칼로리 조절, 다이어트에 긍정적인 모임 참여, 다이어트에 부정적인 모임 참여, 다이어트에 긍정적인 주변 사람과 교류 및 다이어트에 부정적인 주변 사람과 교류 등을 포함할 수 있다. 그리고, 프로세스 최종 성과는 프로세스의 목표에 따라 목표 감량 무게 이상 감량, 목표 감량 무게 이하 감량, 체중 유지, 체중 증가, 체중 감소 등으로 나타날 수 있다. 프로세스와 관련된 활동들의 일부는 시간적 측면에서 순차적, 병행, 선택적 또는 일정 주기 반복 형태로 수행될 수 있다. 그리고, 프로세스의 종료시점은 프로세스 인스턴스 마다 상이할 수 있다.

[0022] 또 다른 프로세스의 예로 취업 프로세스를 들 수 있다. 정해진 목표는 취업일 수 있고, 활동은 매 학기 학교 과목 수강, 학원 과목 수강, 공모전 참여, 자격증 시험 응시, 해외 어학 연수 수행, 인턴십 수행 및 현장 연수 활동 수행 등을 포함할 수 있다. 그리고, 정해진 목표에 대한 예측 결과는 대기업 취업, 중소기업 취업, 상위 교육 기관 진학, 취업 실패 및 졸업 유예 중 하나로 나타날 수 있다. 이때 최종 성과 유형을 취업과 미취업으로 이분해서 정의할 수도 있다.

[0023] 프로세스의 목표를 달성하기 위해 프로세스 인스턴스 내에서 수행되는 활동의 조합은 상이할 수 있다. 또한 프로세스의 목표를 달성한 프로세스 인스턴스들은 수행했던 동일 활동에서 발생한 이벤트들에서 서로 상이한 값들을 가질 수 있다. 예를 들어, 취업을 목표로 대학생에 의해 수행되는 취업 프로세스에서 그 목표를 달성한 두 학생의 평균 평점이 각각 3.2와 4.2이고 공모전 수상 경력이 1회와 5회일 수 있다.

[0024] 평가 대상 객체는 정해진 프로세스 수행 절차에 따라 하나 이상의 활동을 완수하여 프로세스 수행기간 이후에 프로세스 성과 중 하나를 획득하여 평가 대상이 될 수 있는 객체를 의미한다. 예를 들어, 취업을 목표로 대학생에 의해 수행되는 취업 프로세스에서는 프로세스 수행 주체인 대학생이 평가 대상 객체가 된다. 그리고, 평가

대상 프로세스 인스턴스는 정해진 프로세스 수행 절차에 따라 활동을 완수하여 수행기간 이후에 프로세스 성과 중 하나를 획득하여 평가 대상이 될 수 있는 프로세스 인스턴스를 의미한다.

[0025]

본 발명에 따른 목표 달성 진단 장치(100)는 프로세스 인스턴스의 시점별 최종 성과 유형별 근접도를 산출한다. 시점별 최종 성과 유형별 근접도는 프로세스 종료시점 이전의 평가시점에서 예측하는 해당 프로세스 인스턴스의 최종 성과 유형별 예상 발생 확률로 정의된다. 후술하는 도 3a 및 도 3b에서는 4년의 진행기간을 가지며 단위시간을 1년으로 정하고 매 학년의 말일을 평가시점으로 정의한 취업 프로세스의 예를 들어 설명한다. 최종 성과 유형이 취업과 미취업이라고 할 때, 하나의 객체가 프로세스를 시작해서 2년이 완료된 시점인 두 번째 평가시점에 있을 때, 그 시점로부터 2년 뒤인 프로세스의 종료시점에 발생할 것으로 예상되는 각 최종 성과 유형의 확률인 '예상 취업 확률'과 '예상 미취업 확률' 이 해당 프로세스 인스턴스의 두 번째 평가시점에서의 최종 성과 유형별 근접도이다.

[0026]

본 발명에 따른 목표 달성 진단 장치(100)는 전처리부(10) 및 근접도 산출부(30)를 포함한다.

[0027]

전처리부(10)는 평가 대상 객체에 대해 예측하고자 하는 목표를 달성하기 위한 프로세스의 목표, 최종 성과 유형, 속성, 수행기간과 평가시점을 정의한다. 본 발명에서 프로세스는 예측하고자 하는 목표(특정 목표)를 달성하기 위해 행해지는 일련의 활동을 나타낸다. 프로세스의 수행 기간은 프로세스 인스턴스의 시점별 최종 성과 유형별 근접도 산출의 대상 기간이 되며, 단위 시간으로 표시될 수 있다. 단위 시간은 시간(Hour), 일(Day), 주(Week) 및 월(Month)과 같이 시간과 관련된 단위를 사용할 수 있으며, 목적과 대상에 따라 달라질 수 있다.

[0028]

그리고 전처리부(10)는 누적된 통계 데이터를 분석하여 프로세스 목표, 최종성과 유형, 속성, 평가기간 및 평가시점을 정의할 수 있다. 예를 들어, 취업자에 관한 통계 데이터를 분석하여 특정 과목 이수 여부, 해외 어학 연수 여부, 자격증 교육 참가 여부, 공모전 참가 여부, 인턴 수행 여부 등을 프로세스의 활동 속성으로 정의하며, 학년별 학점, 어학점수, 자격증, 수상경력 및 인턴경험 횟수 등을 프로세스의 이벤트 속성으로 정의하며, 대기업 취직, 중소기업 취직 및 공무원 취직 등을 프로세스 최종 성과 유형으로 정의할 수 있다. 또한 전처리부(10)는 사전에 프로세스 목표, 최종성과 유형, 속성, 수행기간 및 평가시점 등과 관련된 데이터를 입력 받아 프로세스 목표, 최종성과 유형, 속성, 수행기간 및 평가시점 을 정의할 수 있다.

[0029]

예를 들어, 취업 프로세스의 경우 전처리부(10)는 대학 졸업에 필요한 4년의 기간을 고려하여 단위 시간을 1년으로 정의하고, 4개의 단위시간에 해당하는 프로세스 수행기간과 4년 대학 생활 동안 매 학년의 마지막 날을 평가 시점으로 설정할 수 있다. 이는 입학해서 대학생활 4년 후 졸업하는 학생에 대해 입학 후 4년 동안 평가시점에 이를 때마다 프로세스 종료시점인 졸업시점에서 해당 학생의 취업 여부를 예측하고자 하는 의도를 나타낸다. 본 발명에 따른 목표 달성 진단 장치(100)는 프로세스 수행기간이 4년인 경우에는 프로세스 시작 후 4년이 되는 프로세스 종료시점 이전에 해당 프로세스를 구성하고 있는 '활동'을 취해서 종료시점에서 평가할 수 있는 프로세스 성과를 생성한 프로세스 인스턴스들에 한해서 이벤트 프로파일 군집분석의 대상으로 삼는다.

[0030]

다음으로, 전처리부(10)는 정의된 목표, 최종성과 유형, 속성, 수행기간 및 평가시점에 부합하는 하나 이상의 과거 프로세스 인스턴스로부터 속성 값을 수집 또는 확인한다. 본 발명에 따른 목표 달성 진단 장치(100)에서 분석에 사용되는 프로세스 인스턴스의 속성 값(Attribute Value)는 과거, 현재 및 미래 속성을 포함할 수 있다. 그리고, 전처리부(10)는 분석의 대상이 되는 프로세스의 일부 속성으로만 분석을 수행할 수 있다. 분석 과정에서 미래 속성의 값은 예측되거나, 이미 현재 시점에서 확정된 값(예를 들어, 이번 학기 15학점 수강 이후에 다음 평가시점에서 프로세스 인스턴스의 누적 이수 학점, 다음 평가시점에 시행되는 정책 관련 값 등)들로 사용될 수 있으며, 예측이나 확정이 불가능한 미래 속성에 대해서는 무시할 수 있는 값(null)이 할당될 수 있다. 미래 속성은 후술하는 시점별 예측부(50)에서 추가적으로 설명한다.

[0031]

그리고, 프로세스 인스턴스의 속성은 환경 속성, 객체 속성, 활동 속성 및 이벤트 속성으로 분류될 수 있다. 여기서 객체는 프로세스 인스턴스의 수행 주체를 의미한다. 예를 들어, 취업 프로세스 인스턴스에서 객체는 취업이라는 목표 달성의 주체가 되는 학생이 된다. 환경 속성은 객체가 처해 있는 환경을 설명하기 위해 사용되는 속성이고, 객체 속성은 객체를 설명하기 위해 사용되는 속성이다. 그리고, 활동 속성은 프로세스 인스턴스의 활동과 관련된 속성이다. 이벤트 속성은 프로세스 인스턴스의 활동 수행으로 발생한 이벤트에 의해 유발되는 프로세스 인스턴스의 상태와 관련된 속성으로, 환경에서 활동을 취했을 때 나타나는 결과에 대한 속성이다. 평가시점별로 다른 값을 가지는 속성도 있다. 예를 들어, 1년 전의 평균 평점은 평가시점을 기준으로 정해지는 상대적인 값이다. 표 2는 프로세스 인스턴스의 속성 값의 일례를 나타낸다.

표 2

[0032]

프로세스 인스턴스의 속성 값의 일례

속성 분류	프로세스 인스턴스의 속성 값의 일례
환경 속성	대한민국, 지난 연도별 취업률, 당해 연도 대학 졸업생 수, 당해 연도 기업 채용 인원 수, 누적 실업자 수, 후속 연도 예상 대학 졸업생 수, 후속 연도 예상 연도 기업 채용인원 수, 후속 연도 예상 누적 실업자 수 등
객체 속성	성별, 나이, 고향, 부모님 이름, 부모님 학력, 부모님 고향, 과거 거주지, 현재 거주지, 미래 예상 거주지, 과거 학력, 현재 학력, 미래 학력, 과거 가족관계, 현재 가족관계, 미래 가족관계, 1년 전 학년, 현재 학년, 1년 후 학년, 인구통계학적 특성, 사회통계학적 특성 성격 특성, 행동특성 등
활동 속성	지난 달 제품 구매 여부, 이번 달 제품 구매 여부, 다음 달 제품 구매 여부, 지난 학기 과목 수강 여부, 현재 학기 과목 수강 여부, 다음 학기 과목 수강 여부, 어제 운동 여부, 현재 운동 여부, 내일 운동 여부, 어제 식사 여부, 오늘 식사 여부, 내일 식사 여부, 지난 학기 학습 여부, 이번 학기 학습 여부, 다음 학기 학습 여부.
이벤트 속성	두 달 전까지 보유한 구매 제품 목록, 한 달 전까지 보유한 구매 제품 목록, 현재 보유한 구매 제품 목록, 한 달 후에도 보유 예상되는 구매 제품 목록, 한 학기 전 수강 과목 목록, 현재 학기 수강 과목 목록, 다음 학기 수강 예상 과목 목록, 어제 운동시간, 현재 운동시간, 내일 예상 운동시간, 어제 칼로리 섭취량, 현재 칼로리 섭취량, 내일 예상 칼로리 섭취량, 지난 학기 총 학습시간, 이번 학기 총 학습시간, 다음 학기 예상 학습시간, 1년 전의 평균 평점, 한 학기 전의 평균 평점, 현재 평균 평점, 다음 학기 예상 평균 평점, 1년 전까지 체중, 현재 체중, 한 달 후의 예상 체중, 지난 평가시점 성과 달성 유무, 현 평가시점 성과 달성 유무, 다음 평가시점 예상된 성과 달성 유무, 지난 평가시점 달성된 성과 유형, 현 평가시점 달성된 성과 유형, 다음 평가시점 달성될 예상 성과 유형, 평가시점에 사용했던 검색 키워드 목록, 현 평가시점에 사용했던 검색 키워드 목록, 다음 평가시점에 사용할 것으로 예상되는 검색 키워드 목록 등.

[0033]

전처리부(10)는 분석 대상이 된 과거 프로세스 인스턴스의 속성 값을 대상으로 이벤트 프로파일(Event Profile, EP) 추출을 위한 군집분석을 실시한다. 전처리부(10)는 분석 대상이 된 과거 프로세스 인스턴스의 속성 값에 대해 하나 이상의 활동 속성 및 하나 이상의 이벤트 속성 중 적어도 하나에 기초하여 군집을 형성한다. 본 발명에서는 이 군집을 '프로파일 군집'이라고 명명한다. 여기서 군집은 군집분석에 사용된 속성 측면에서 유사한 프로세스 인스턴스들의 그룹을 의미한다. 예를 들어, 매 평가시점에서 획득된 평균 평점이라는 속성을 기준으로 수행된 대학생 취업 프로세스에 대한 군집분석에서 다음 세 가지와 같은 군집을 형성할 수 있다. 평균 평점이 지속적으로 상승하는 추세를 보이는 프로세스 인스턴스들로 구성된 그룹, 평균 평점이 지속적으로 하락하는 추세를 보이는 프로세스 인스턴스들로 구성된 그룹, 평균 평점이 일정 변동 폭을 벗어나지 않고 유지되는 프로세스 인스턴스들로 구성된 그룹 등 세 가지 군집이 형성될 수 있다.

[0034]

그리고, 전처리부(10)는 군집분석 시에 사용된 하나 이상의 속성의 값들을 시간 순차적으로 나열한다. 시간 순차적으로 나열된 하나 이상의 속성의 값들에서 각 프로파일 군집에 속한 프로세스 인스턴스들은 유사성을 띤다. 전처리부(10)는 각 프로파일 군집 내 프로세스 인스턴스들의 속성 값들에 기초하여 집단 각각에서 하나 이상의 이벤트 프로파일을 추출한다. 표 1에서 정의한 바와 같이, 이벤트 프로파일은 프로세스 인스턴스가 완료되기까지 생성되는 일련의 이벤트 속성값들이 형성하는 양상의 개요이다.

[0035]

예를 들어, 평균 평점이 지속적으로 상승하는 추세를 보이는 프로세스 인스턴스들로 구성된 군집의 이벤트 프로파일은 해당 군집 내 프로세스 인스턴스들의 1학년 말에 획득된 평균 평점은 2.0~3.1 구간을, 2학년 말에 획득된 평균 평점은 3.3~3.5 구간을, 3학년 말에 획득된 평균 평점은 3.7~3.8 구간을, 4학년 말에 획득된 평균 평점은 4.0~4.2 구간을 형성하는 것으로 정의될 수 있다. 전처리부(10)에서 이벤트 프로파일 군집분석을 수행하고, 이벤트 프로파일 군집분석 실시로 도출된 프로파일 군집 각각에서 이벤트 프로파일을 추출하는 과정은 후술하는 도 3a 및 도 3b에서 실시예를 들어 구체적으로 설명한다. 그리고, 전처리부(10)에서 추출된 이벤트 프로파일은 군집도 산출부(30)로 전달된다.

[0036]

표 3은 일 실시예로 군집도 산출부(30)에서 평가시점마다 단 하나의 이벤트값을 획득하는 프로세스 인스턴스의 최종 성과 유형별 군집도를 산출하는 과정에서 사용되는 변수의 정의를 나타낸다. 다시 말해, 평가시점의 개수와 같은 수의 프로세스 이벤트 속성에 기초하여 최종 성과 유형별 군집도를 산출하며, 평가시점마다 해당 이벤트 속성에 대한 값을 하나 획득한다. 단, 표 3의 평가시점마다 하나의 이벤트값을 획득하는 프로세스는 하나의 일례 일뿐, 이로써 한정되는 것은 아니며 평가시점에서 하나 이상의 속성값들이 획득될 수 있다.

표 3

변수 정의

용어	정의
L	프로젝트 수행기간
I	현재 진행 중인 프로세스 인스턴스
n	프로세스 인스턴스 I에서 생성되는 각 활동의 이벤트값이 형성하는 이벤트 프로파일의 총 수
m	최종 성과 유형의 총 수
A_t	t 번째 활동. t 는 정수이며 $1 \leq t \leq L$
e_t	t 번째 이벤트. t 는 정수이며 $1 \leq t \leq L$
O_i	i 번째 최종 평가 유형. i 는 정수이며 $1 \leq i \leq m$
EP_j	j 번째 이벤트 프로파일 (EP). j 는 정수이며 $1 \leq j \leq n$
ATP	비정상 종료 확률(Abnormal Termination Probability)
t	t 번째 평가시점. t 는 정수이며 $1 \leq t \leq L$
d_t	t 번째 평가시점의 활동에 의해 발생된 이벤트의 값.
D_t	t 번째 평가시점까지 프로세스 인스턴스 I에 의해 발생한 활동에 따른 이벤트들의 값들로 구성된 집합 $D_t = \{d_1, d_2, \dots, d_t\}$ 예를 들어, 매 학년말을 평가시점으로 하고 평균 평점이라는 한 가지 이벤트만으로 구성된 대학생의 취업 프로세스에서 3번째 평가시점에 획득될 수 있는 D_t 는 1학년말 평균 평점과 2학년말 평균 평점과 3학년말 평균 평점을 원소로 하는 벡터로 정의될 수 있다.

[0037]

근접도 산출부(30)는 시점별 최종 성과 유형별 근접도 산출을 위해서 사전확률을 산출한다. 먼저, 근접도 산출부(30)는 전처리부(10)에서 추출된 하나 이상의 이벤트 프로파일에 대해 프로세스 인스턴스가 각 이벤트 프로파일에 속할 사전확률을 계산한다.

수학식 1

$$P(I \in EP_k) = \frac{n_k}{n} = \frac{n_k}{N}$$

[0039]

[0040]

근접도 산출부(30)는 수학식 1을 통해 사전확률을 산출할 수 있다. 수학식 1에서 N 은 분석 대상이 되는 프로세스 인스턴스(I)의 총 수이며, n_k 는 k 번째 이벤트 프로파일(EP)에 속하는 프로세스 인스턴스의 수이고, $P(I \in EP_k)$ 는 프로세스 인스턴스(I)가 k 번째 이벤트 프로파일에 속할 사전확률이다. 프로세스가 수행됨에 따라 해당 프로세스 인스턴스가 t 번째 평가시점까지 매 평가시점마다 획득한 이벤트 속성값이 벡터 D_t 를 형성할 때, 전처리부(10)는 수학식 2를 통해 해당 프로세스 인스턴스가 k 번째 이벤트 프로파일에 속할 사후확률을 산출한다.

수학식 2

$$\begin{aligned} P(I \in EP_k | D_t) &= \frac{P(I \in EP_k \cap D_t)}{P(D_t)} \\ &= \frac{P(D_t | I \in EP_k) \times P(I \in EP_k)}{P(D_t)} \\ &= \frac{P(D_t | I \in EP_k) \times P(I \in EP_k)}{\sum_{j=1}^n P(D_t | I \in EP_j) \times P(I \in EP_j)} \end{aligned}$$

[0041]

[0042]

수학식 2의 계산 과정은 다음과 같다. 프로세스 인스턴스가 j번째 이벤트 프로파일(EP_j)에 속할 사전확률인

$$P(I \in EP_j)$$

를 모든 가능한 j값에 대해 산출한다. j는 정수이며, $1 \leq j \leq n$ 을 만족한다.

[0043]

프로세스 인스턴스 I가 j 번째 이벤트 프로파일(EP_j)에 속하는 경우에 D_t 가 형성될 확률인 $P(D_t | I \in EP_j)$ 는 수학식 3과 같이 우도(likelihood) 함수를 이용해서 산출할 수 있으며, 모든 가능한 j값에 대해 산출한다.

수학식 3

$$P(D_t | I \in EP_j) = \prod_{r=1}^t f(d_r | I \in EP_j)$$

[0044]

[0045]

수학식 3에서 $f(d_r | I \in EP_j)$ 는 프로세스 인스턴스 I가 j 번째 이벤트 프로파일(EP_j)에 속하는 경우에 평가시점 r에서 이벤트 속성값으로 d_r 이 발생할 확률이며 이와 같은 확률을 평가시점 t를 포함한 그 이전 시점에 대해 각각 계산하여 모두 곱하는 방식으로 $P(D_t | I \in EP_j)$ 를 산출한다. 베이저언 정리(Bayesian Theorem)에 기초해

서 상술한 과정을 통해 산출된 확률 $P(I \in EP_j)$ 및 $P(D_t | I \in EP_j)$ 를 이용하여 수학식 2의 $P(I \in EP_k | D_t)$ 를 산출할 수 있으며, 모든 가능한 k값에 대해 산출한다. k는 정수이며, $1 \leq k \leq n$ 을 만족한다.

[0046]

사전확률 및 사후확률이 산출되면, 근접도 산출부(30)는 수학식 1 내지 수학식 3에 의해 산출된 사전확률 및 사후확률을 이용해서 t번째 평가시점에서 프로세스 인스턴스의 최종 성과 유형별 근접도를 산출한다. 예를 들어, 취업 프로세스의 최종 성과를 취업 성공을 의미하는 정상 종료와 취업 실패를 의미하는 비정상 종료를 두 가지 유형으로 분류할 경우에 프로세스 인스턴스의 평가시점별 최종 성과 유형별 근접도 중 하나는 비정상 종료 확률로서 프로세스 종료시점에서 취업 실패 예상 확률이 될 수 있다. 여기서 비정상 종료 확률에 초점을 맞추는 이유는 프로세스 진행 중 바람직하지 않은 상황을 조기에 발견함으로써 프로세스의 성과를 향상시키고 자원의 손실을 줄이기 위함이다.

[0047]

수학식 4는 i번째 최종 성과 유형을 O_i 라 할 때, 평가시점 t에서 n개의 이벤트 프로파일 중 하나에 반드시 속하는 프로세스 인스턴스의 최종 성과 유형별 근접도를 산출하는 일반화된 수학식을 나타낸다. 산출부(30)는 수학식 4를 이용하여 프로세스 인스턴스의 시점별 최종 성과 유형별 근접도를 산출할 수 있다.

수학식 4

$$P(O_i|D_t) = \sum_{j=1}^n P(I \in EP_j|D_t)P(O_i|I \in EP_j \cap D_t)$$

수학식 4에서 평가시점 t까지 획득된 이벤트값들이 벡터 D_t 를 형성하는 경우에 프로세스의 종료시점에서 최종 성과가 i번째 유형에 속할 확률인 $P(O_i|D_t)$ 는 평가시점 t까지 획득된 이벤트값들이 벡터 D_t 를 형성하는 경우에 프로세스 인스턴스 I가 j 번째 이벤트 프로파일(EP_j)에 속하는 확률인 $P(I \in EP_j|D_t)$ 와 평가시점 t까지 획득된 이벤트값들이 벡터 D_t 를 형성하고 프로세스 인스턴스 I가 j 번째 이벤트 프로파일(EP_j)에 속하는 경우에 프로세스의 종료시점에서 최종 성과가 i번째 유형인 O_i 에 속할 확률인 $P(O_i|I \in EP_j \cap D_t)$ 를 곱한 값을 모든 최종 성과 유형에 대해 구하고 그 곱셈의 결과값들을 모두 합산하여 산출한다. 수학식 4를 유도하는 세부 과정은 후술하는 도 3a 내지 도 3b에서 실시예를 들어 설명하도록 한다.

그리고, 근접도 산출부(30)는 프로세스 수행기간이 다양하게 발생할 수 있는 경우 최종성과 유형별 기대확률을 산출할 수 있다. 진단 목적에 따라 프로세스 수행기간이 다양한 경우가 있다. 예를 들어, 대학 입학 후 4번째 평가시점(졸업시점)에 취업 지원이 이루어지는 학생이 있는가 하면, 대학 입학 후 5번째 평가시점에 취업 지원이 이루어지는 학생도 있다. 이와 같은 경우, 최종성과 유형별 기대 확률을 다음과 같이 산출할 수 있다. 설명의 이해를 돕기 위해, L은 어느 특정 평가시점과 정확히 일치한다고 가정한다.

L의 수행기간을 가진 프로세스에 속하는 과거 프로세스 인스턴스들의 데이터를 수집하여, 상술한 과정을 거쳐 산출된 프로세스 인스턴스의 t번째 평가시점에서 i번째 최종 성과 유형별 근접도를 $P(O_{i,L}|D_{t,L})$ 라 한다. 근접도 산출부(30)는 i번째 최종 성과 유형의 근접도 $P(O_{i,L}|D_{t,L})$ 에 기초하여 i번째 최종성과 유형의 발생 기대확률을 수학식 5를 이용하여 산출할 수 있다.

수학식 5

$$\sum_{L=T}^{\infty} P(\text{project_length} = L | L \geq T) * P(O_{i,L}|D_{t,L})$$

수학식 5에서 T는 최소 프로젝트 수행기간을 나타내며, $P(\text{project_length} = L | L \geq T)$ 는 T번째 평가시점에서 종료되는 프로세스 또는 그 이후로 더 진행되는 프로세스 인스턴스 중에서 프로세스의 수행기간이 L일 확률을 의미한다.

본 발명에 따른 목표 달성 진단 장치(100)는 시점별 예측부(50)를 더 포함할 수 있다. 시점별 예측부(50)는 프로세스 인스턴스의 속성 값이 확보된 가장 최근 평가시점 t 이후 미래의 후속 평가시점에 대한 프로세스 인스턴스의 시점별 활동 속성값을 예측할 수 있으며, 프로세스 인스턴스의 시점별 이벤트 속성값을 예측할 수 있다. 본 발명에 따른 목표 달성 진단 장치(100)는 프로세스 인스턴스의 속성 값으로부터 추출된 이벤트 프로파일과 프로세스 인스턴스의 기획된 속성값들과 예측된 속성값들로 구성된 '확장된 벡터 D_t '를 형성하여 상술한 확률을 산출함으로써, 평가시점에서의 시점별 최종 성과 유형별 근접도를 산출할 수 있다.

시점별 예측부(50)는 수학식 6을 통해 미래 후속 시점의 활동에 대한 활동 속성값과 이벤트에 대한 이벤트 속성값에 대한 예측값들을 제시할 수 있다.

수학식 6

$$s = \arg \max_j P(I \in EP_j | D_t)$$

[0056]

[0057]

수학식 6은 가장 최근 평가시점 t 까지 획득된 이벤트 속성값들이 벡터 D_t 를 형성하는 경우에 프로세스 인스턴스 I 가 각 이벤트 프로파일에 속할 확률을 구해서 가장 높은 확률을 갖게 하는 이벤트 프로파일 유형을 파악하기 위한 것이다. 예를 들어, $P(I \in EP_j | D_t)$ 의 값이 $j=2$ 일 때 가장 크다면, s 는 2를 값으로 갖는다. 이는 해당 프로세스 인스턴스가 후속 시점에서 획득하게 될 이벤트 속성값들이 2번째 이벤트 프로파일에 가장 근접할 것이라는 가정의 근거가 되며, 그 가정 아래 후속 시점에서 이벤트의 값에 대한 예측값을 결정함에 있어 2번째 이벤트 프로파일에 의해 정의된 해당 후속 시점에서의 이벤트 속성값의 양상을 참고할 수 있다. 예를 들어, 수학식 6에 의해 3번째 평가 시점에 도달한 프로세스 인스턴스가 후속 평가시점에서 2번째 이벤트 프로파일에 가장 근접한 이벤트 속성값들을 발생시킬 것이라는 가정이 도출되고, 2번째 이벤트 프로파일에 의하면 아직 도달하지 않은 4번째 평가시점에서의 이벤트 속성값들이 평균이 50이고 표준편차가 5인 정규분포를 구성한다고 정의되어 있다면, 4번째 평가시점에 대한 이벤트 예측값으로 그 정규분포의 평균인 50을 사용하여 최종 성과 유형별 근접도를 산출할 수 있다.

[0058]

프로세스 인스턴스의 시점별 활동 예측은 현 평가시점 이후의 후속 평가시점까지 행해지는 활동 목록 예측으로 과목 수강 여부 예측, 공모전 참가 여부 예측, 자격증 시험 응시 여부 예측 등을 포함할 수 있다. 그리고, 프로세스 인스턴스의 시점별 활동의 이벤트 속성값 예측은 현 평가시점 이후의 후속 평가시점까지 행해지는 활동에 의해 생성될 수 있는 이벤트 속성값으로 수강 과목 목록 예측, 참여 공모전 목록 예측, 응시 자격증 시험 목록 예측 등을 포함할 수 있다. 이와 같이 다가올 미래 시점에 프로세스에 의해 어떤 활동이 취해질지 예측할 수 있어, 선제 대응할 수 있다.

[0059]

목표 달성 진단 장치(100)의 평가 시점은 사전에 고정된 시점이 될 수 있고, 특정 이벤트가 발생하는 시점이 될 수 있다. 예를 들어, 고정된 평가 시점은 평가시점을 단위시간과 일치시켜, 각 단위시간의 종료 시점을 평가시점으로 하는 것이다.

[0060]

도 2는 본 발명의 일 실시예에 따른 목표 달성 진단 장치의 처리 과정을 나타내는 흐름도이다.

[0061]

도 1 및 도 2를 참조하면, 먼저 전처리부(10)는 예측하고자 하는 목표를 달성하기 위한 프로세스의 목표, 최종 성과 유형, 속성, 수행기간 및 평가시점을 정의한다(S201). 본 발명에서 프로세스는 예측하고자 하는 목표(특정 목표)를 달성하기 위해 행해지는 일련의 활동을 나타낸다. 프로세스의 수행 기간은 프로세스 인스턴스의 최종 성과 유형별 근접도 산출의 대상 기간이 되며, 단위 시간으로 표시될 수 있다.

[0062]

다음으로, 전처리부(10)는 정의된 목표, 최종성과 유형, 속성, 수행기간 및 평가시점에 부합하는 하나 이상의 과거 프로세스 인스턴스로부터 속성 값을 수집 또는 확인한다(S202). 본 발명에 따른 목표 달성 진단 장치(100)에서 분석에 사용되는 프로세스 인스턴스의 속성 값(Attribute Data)은 과거, 현재 및 미래 속성을 포함할 수 있으며, 분석의 대상이 되는 프로세스의 일부 속성만으로도 분석을 수행할 수 있다. 분석 과정에서 미래 속성의 값은 예측되거나, 이미 현재 시점에서 확정된 값들로 사용될 수 있다. 그리고, 프로세스 인스턴스의 속성은 환경 속성, 객체 속성, 활동 속성 및 이벤트 속성으로 분류될 수 있다.

[0063]

프로세스 인스턴스의 속성 값이 수집되면, 전처리부(10)는 이벤트 프로파일 추출을 위한 군집분석을 수행하기 이전에 프로세스 인스턴스의 세분화(Segmentation)를 위한 군집분석(Cluster Analysis)을 수행할 수 있다(S203). 이 군집분석을 '세분화 군집분석'이라고 하고, 그 결과물인 군집을 '세분화 군집'이라 명명한다. 전처리부(10)는 수집된 프로세스 인스턴스의 속성 값을 대상으로 이벤트 프로파일 군집분석을 실시하기 이전에, 세분화 군집분석을 수행하여, 유사한 환경 속성값이나 유사한 객체 속성값을 지닌 프로세스 인스턴스들로 군집을 형성하게 하고, 이 세분화 군집들에 대해 각각 이벤트 프로파일 군집분석을 수행하여 이벤트 프로파일을 도출함으로써 도출된 이벤트 프로파일 간의 상이성을 증대시킬 수 있다. 하나 이상의 활동 속성 및 하나 이상의 이벤트 속성 중에서 적어도 하나의 속성에 기초하여 수행되는 이벤트 프로파일 군집분석과 달리, 세분화를 위한 군집분석을 위해 전처리부(10)는 하나 이상의 환경 속성 및 객체 속성 중에서 적어도 하나의 속성에 기초하여 군

집분석(Cluster Analysis)를 실시하여 하나 이상의 군집(cluster)을 형성한다.

[0064] 세분화 군집분석은 과거 프로세스 인스턴스들로부터 이벤트 프로파일을 추출하기 전에 과거 프로세스 인스턴스들을 환경 속성이나 활동 속성을 기준으로 유사성을 판단하여 유사한 인스턴스들로 군집을 형성하게 한다. 세분화 군집분석은 다양한 속성의 환경 및 객체에 의해 생성된 프로세스 인스턴스들이 섞여 있는 전체 프로세스 인스턴스 그룹으로부터 이벤트 프로파일을 추출하지 않고 유사한 속성의 환경 및 객체에 의해 생성된 프로세스 인스턴스들로 군집을 형성하게 하여 세분화한 후에, 각 세분화 군집으로부터 이벤트 프로파일을 추출하는 것이 더 효과적이고 효율적이라는 점에 착안하여 수행된다. 전처리부(10)의 세분화 군집분석 과정은 이벤트 프로파일 군집분석을 수행하기 이전에 필요에 따라 선택적으로 수행될 수 있다.

[0065] 세분화 군집 내 과거 프로세스 인스턴스의 속성 값을 대상으로 이벤트 프로파일 군집분석을 실시한다(S204). 전처리부(10)는 S203단계에서 형성된 세분화 군집 내 과거 프로세스 인스턴스의 속성 값에 대해 하나 이상의 활동 속성 및 하나 이상의 이벤트 속성 중 적어도 하나에 기초하여 군집(cluster)을 형성한다. 그리고, 전처리부(10)는 군집분석 시에 사용된 하나 이상의 속성의 값들을 시간 순차적으로 나열한다. 시간 순차적으로 나열된 하나 이상의 속성의 값들에서 각 프로파일 군집에 속한 프로세스 인스턴스들은 유사성을 띤다. 전처리부(10)는 각 프로파일 군집 내 프로세스 인스턴스들의 속성 값들에 기초하여 집단 각각에서 하나 이상의 이벤트 프로파일을 추출한다(S205). S205 단계의 수행 대상인 프로파일 군집은 S204 단계의 이벤트 프로파일 군집분석에 의해 형성된 프로파일 군집을 의미한다. 그리고, 전처리부(10)에서 추출된 이벤트 프로파일은 근접도 산출부(30)로 전달된다(S206).

[0066] 전처리부(10)에서 이벤트 프로파일이 추출되면, 근접도 산출부(30)는 시점별 최종 성과 유형별 근접도 산출을 위해서 사전확률을 산출한다(S207). 근접도 산출부(30)는 도 1에 대한 설명에 기재된 수학적 식 1을 통해 사전확률을 산출할 수 있다. 그리고, 프로세스가 수행됨에 따라 해당 프로세스 인스턴스가 t번째 평가시점까지 획득한 이벤트 속성값들이 벡터 D_t 를 형성(S208)할 때, 근접도 산출부(30)는 도 1에서 설명한 수학적 식 2를 통해 해당 프로세스 인스턴스가 k번째 이벤트 프로파일에 속할 사후확률을 산출한다(S209).

[0067] 사전확률 및 사후확률이 산출되면, 근접도 산출부(30)는 산출된 사전확률 및 사후확률을 이용해서 t번째 평가시점에서 프로세스 인스턴스의 시점별 최종 성과 유형별 근접도를 산출한다(S210). 근접도 산출부(30)는 도 1에서 설명한 수학적 식 4를 이용하여 프로세스 인스턴스의 시점별 최종 성과 유형별 근접도를 산출할 수 있다.

[0068] 도 3a는 본 발명의 일 실시예에 따른 목표 달성 진단 장치의 취업 프로세스 다이어그램의 일례를 나타내는 도면이다.

[0069] 도 3b는 도 3a의 취업 프로세스 다이어그램의 일례에서 이벤트 속성값을 그래프로 나타낸 도면이다.

[0070] 도 3a 및 도 3b를 참조하면, 취업 프로세스를 실시예로 하여 목표 달성 진단 장치(100)의 취업 프로세스를 설명한다. 진단 목표가 취업이라고 가정하면, 목표 달성 진단 장치(100)는 최종 성과 유형을 취업과 미취업으로 정의하고, 활동 속성으로 대학생활 1학년 때 과목 수강 여부, 대학 2학년 때 과목 수강 여부, 대학 3학년 때 과목 수강 여부, 대학 4학년 때 과목 수강 여부를 정의하고, 이벤트 속성으로 대학 1학년 평균 평점, 대학 2학년 평균 평점, 대학 3학년 평균 평점, 대학 4학년 평균 평점으로 정의하고, 대학 졸업에 필요한 4년의 기간을 고려하여 단위 시간을 1년으로 정의하고, 4개의 단위시간에 해당하는 프로세스 수행기간과 각 학년의 말일을 평가 시점으로 설정할 수 있다. 이는 학생이 입학 후 향후 4년까지 취업률을 분석하고자 하는 의도를 나타낸다. 목표 달성 진단 장치(100)는 프로세스 수행기간이 4년인 경우에는 4년이 되는 평가시점 이전에 '활동'을 취해서 총 4번의 평가시점에서 이벤트 속성값을 획득하고 그에 따른 프로세스 성과를 확인할 수 있는 과거 프로세스 인스턴스들에 한해서 세분화 군집분석과 이벤트 프로파일 군집분석의 대상으로 삼는다.

[0071] 그리고, 목표 달성 진단 장치(100)는 정의된 목표, 최종성과 유형, 속성, 수행기간 및 평가시점에 부합하는 하나 이상의 과거 프로세스 인스턴스들로부터 속성 값을 수집한다. 분석 과정에서 미래 속성의 값은 예측되거나, 이미 현재 시점에서 확정된 값으로, 이번 학기 15학점 수강 이후에 다음 평가시점에서 프로세스 인스턴스의 누적 이수 학점, 다음 평가시점에 시행되는 정책 관련 값 등이 사용될 수 있으며, 예측이나 확정이 불가한 미래 속성에 대해서는 무시할 수 있는 값(null)이 할당될 수 있다.

[0072] 도 1 및 도 2에 기재된 목표 진단 과정에 따라, 취업 프로세스에서 객체인 '학생'의 이벤트 속성인 '대학 1학년 말 평균 평점', '대학 2학년 말 평균 평점', '대학 3학년 말 평균 평점', '대학 4학년 말 평균 평점'에 기초해서 전처리부(10)가 이벤트 프로파일 군집분석을 실시하여 3개의 군집(cluster)을 획득한 경우, 각 프로파일 군집(cluster)이 표 4와 같이 3개의 이벤트 프로파일을 형성한 것으로 가정한다. 표 4의 일례에서, 평점 평균은 4.5

점 만점으로 환산하여 산출된다. 표 4의 일례에서는 4번의 평가시점이 있고, 각 평가시점은 각 학년 2학기의 종료일이다. 그리고, 프로세스 인스턴스의 활동은 해당 학년의 과목 수강이고, 활동에 따라 생성되는 이벤트의 값은 해당 학년말의 평균 평점이다.

표 4

이벤트 속성값의 일례

	Event1	Event2	Event3	Event4	ATP
EP 1	$N(3.2, 0.01^2)$	$N(3.3, 0.01^2)$	$N(3.1, 0.01^2)$	$N(3.5, 0.01^2)$	20%
EP 2	$N(2.5, 0.02^2)$	$N(3.5, 0.02^2)$	$N(2.5, 0.02^2)$	$N(4.0, 0.02^2)$	30%
EP 3	$N(3.5, 0.01^2)$	$N(2.5, 0.01^2)$	$N(2.2, 0.01^2)$	$N(1.9, 0.01^2)$	90%

도 3b에 개시된 그래프는 표 4의 제1 이벤트 프로파일(EP 1), 제2 이벤트 프로파일(EP 2) 및 제3 이벤트 프로파일(EP3)을 그래프 형태로 나타낸 것이다. 4개의 이벤트 속성인 '대학 1학년말 평균 평점', '대학 2학년말 평균 평점', '대학 3학년 평균 평점', 대학 4학년말 평균 평점'에 기초해서 이벤트 프로파일이 도출되었으며, 프로파일 군집 내 과거 프로세스 인스턴스들의 각 이벤트 속성의 값들은 정규분포를 형성한다. 도 3b에 개시된 각 그래프에서 평가시점에 대응하는 이벤트 속성값이 형성하는 정규분포의 평균을 시간 순차적으로 연결하고 각 평균을 중심으로 정규분포 곡선을 그려 시간에 따른 프로세스 인스턴스의 이벤트 속성값의 변화 양상을 간략히 표시했다. 제1 이벤트 프로파일(EP 1)은 제1 EP 그래프(301), 제2 이벤트 프로파일(EP 2)은 제2 EP 그래프(302), 제3 이벤트 프로파일(EP3)은 제3 EP 그래프(303)과 대응한다.

상기 예시로 든 간단한 취업 프로세스를 다이어그램 형태로 표시하면, 도 3a와 같이 표시될 수 있다. 도 3의 취업 프로세스 다이어그램(300)은 4개의 활동(Activity, A_1, A_2, A_3, A_4)으로 구성되어 있으며, 직렬 프로세스(Serial Process) 형태를 가진다. 그리고, 각각의 활동(A_1, A_2, A_3, A_4)은 하나의 이벤트(e_1, e_2, e_3, e_4)만을 생성한다. 또한, 프로세스에는 3가지 유형의 이벤트 프로파일이 존재한다.

A_i 는 i 학년의 활동이고, e_i 는 i 번째 학년 후의 평균 평점이다. 표 4에서, Event1은 이벤트 속성으로 1학년 2학기의 종료일까지의 평균 평점을 의미하고, Event2는 이벤트 속성으로 2학년 2학기의 종료일의 평균 평점을 의미한다. 그리고, 비정상 종료 확률(Abnormal Termination Probability)인 ATP는 취업 프로세스가 4년 뒤에 비정상적인 성과로 종료될 확률로서, 도 3a 및 도 3b의 일례에서는 취업 실패로 가정한다. 따라서, 표 4에서, 제1 이벤트 프로파일(EP 1)에 속하는 과거 프로세스 인스턴스들의 취업 실패 비율이 20%라는 것을 나타낸다.

제1 이벤트 프로파일(EP 1)에서 Event3의 값은 $N(3.1, 0.01^2)$ 분포를 형성한다. 즉, 3학년 2학기의 종료일인 평가시점에서, 제1 이벤트 프로파일(EP 1)에 속하는 과거 프로세스 인스턴스들의 평균평점의 분포는 평균이 3.1이고 표준편차가 0.01인 정규분포를 가진다. 현실적으로 기존에 알려진 분포들 중 하나를 따르지 않을 경우에는 경험분포를 사용하여 표현할 수 있다.

목표 달성 진단 장치(100)는 정상 종료와 비정상 종료로만 최종 성과 유형을 구분하는 취업 프로세스 일 실시예의 t 번째 평가시점에서 프로세스 인스턴스의 비정상 종료 확률인 취업 실패의 근접도는 도 1에서 설명한 수학식 4에 기초하여, 수학식 7에 의해 같이 산출될 수 있다.

수학식 7

$$P(AT|D_t) = P(I \in EP_1|D_t)P(AT|I \in EP_1 \cap D_t) + P(I \in EP_2|D_t)P(AT|I \in EP_2 \cap D_t) + P(I \in EP_3|D_t)P(AT|I \in EP_3 \cap D_t)$$

수학식 8 내지 수학식 11은 수학식 7의 유도 과정을 나타낸다.

수학식 8

$$\begin{aligned} P(AT) &= P(I \in EP_1)P(AT|I \in EP_1) + P(I \in EP_2)P(AT|I \in EP_2) \\ &\quad + P(I \in EP_3)P(AT|I \in EP_3) \\ &= P(I \in EP_1) \frac{P(AT \cap (I \in EP_1))}{P(I \in EP_1)} + P(I \in EP_2) \frac{P(AT \cap (I \in EP_2))}{P(I \in EP_2)} \\ &\quad + P(I \in EP_3) \frac{P(AT \cap (I \in EP_3))}{P(I \in EP_3)} \\ P(AT) &= P(AT \cap (I \in EP_1)) + P(AT \cap (I \in EP_2)) + P(AT \cap (I \in EP_3)) \end{aligned}$$

수학식 9

$$\begin{aligned} P(AT \cap D_i) &= P(AT \cap D_i \cap (I \in EP_1)) \\ &\quad + P(AT \cap D_i \cap (I \in EP_2)) \\ &\quad + P(AT \cap D_i \cap (I \in EP_3)) \end{aligned}$$

수학식 10

$$\begin{aligned} \frac{P(AT \cap D_i)}{P(d_i)} &= \frac{P(AT \cap D_i \cap (I \in EP_1))}{P(d_i)} \\ &\quad + \frac{P(AT \cap D_i \cap (I \in EP_2))}{P(d_i)} \\ &\quad + \frac{P(AT \cap D_i \cap (I \in EP_3))}{P(d_i)} \end{aligned}$$

수학식 11

$$\begin{aligned} P(AT|D_i) &= P(I \in EP_1|D_i)P(AT|I \in EP_1 \cap D_i) \\ &\quad + P(I \in EP_2|D_i)P(AT|I \in EP_2 \cap D_i) \\ &\quad + P(I \in EP_3|D_i)P(AT|I \in EP_3 \cap D_i) \end{aligned}$$

또 다른 실시예로서, 유사한 방법으로 특정 최종 성과 유형인 중소기업 취업(중취) 근접도는 수학식 12와 같이 산출될 수 있다.

수학식 12

$$\begin{aligned} P(\text{중취}|D_i) &= P(I \in EP_1|D_i)P(\text{중취}|I \in EP_1 \cap D_i) \\ &\quad + P(I \in EP_2|D_i)P(\text{중취}|I \in EP_2 \cap D_i) \\ &\quad + P(I \in EP_3|D_i)P(\text{중취}|I \in EP_3 \cap D_i) \end{aligned}$$

도 3a 및 도 3b의 목표 달성 진단 장치(100)는 취업 프로세스를 정상 종료와 비정상 종료의 두 가지 최종 성과 유형으로 분류하였으나 이는 하나의 실시예로서 두 가지 최종 성과 유형만을 갖는 프로세스로 한정하는 것은 아니다.

도 4a 및 도 4b는 본 발명의 일 실시예에 따른 목표 달성 진단 장치의 복수 이벤트 분석 과정을 설명하기 위한

도면이다.

- [0089] 도 4a 및 도 4b를 참조하면, 도 4a 및 도 4b의 일례에서는 하나의 활동에 의해 생성되는 하나의 이벤트에 대해서만 설명하였으나, 본 발명에 따른 목표 달성 진단 장치(100)는 각 활동에 의해 생성되는 하나의 이벤트 속성뿐만 아니라 둘 이상의 이벤트 속성에 대해서도 목표 달성 진단이 가능하다. 도 4a에서는 매 평가시점에서 이벤트 속성값으로 평균 평점뿐만 아니라, 자격증 수 및 TOEIC 점수 등 총 3가지의 이벤트 속성값을 획득할 수 있다.
- [0090] 도 4a 및 도 4b는 평가시점 별로 이벤트 속성값들의 변화를 나타내고 있다. (t-1)번째 평가시점에서 '평균 평점' 이벤트 속성의 값은 3.3, '토익 점수' 이벤트 속성의 값은 800점, '자격증 점수' 이벤트 속성의 값은 4로 나타난다. 목표 달성 진단 장치(100)는 세 이벤트 속성을 동시에 활용하여 목표 달성 여부를 진단할 수 있다. 도 4a 및 도 4b는 평가시점별로 이벤트 속성값의 변화를 나타내고 있다. 이와 같이, 특정 평가시점에서 획득된 둘 이상의 이벤트 속성값들로 벡터를 형성하고, 이러한 벡터들이 평가시점에 걸쳐서 하나의 이벤트 프로파일을 형성할 수 있다. 예를 들어, 프로세스의 목표가 취업일 경우, 목표 달성 진단 장치(100)는 평균 평점, 토익 점수 및 자격증 수의 세 이벤트 속성을 모두 고려하여 취업 실패 확률을 산출할 수 있다. 본 발명에 따른 목표 달성 진단 장치(100)는 둘 이상의 이벤트 속성값이 한 시점에서 발생하는 경우에도 차원 수정과 같은 약간의 수정을 통해 적용할 수 있으며, 한 시점에서 단일 이벤트 속성값이 획득되는 경우에 대해 산출 가능한 모든 종류의 확률을 둘 이상의 이벤트가 한 시점에서 발생하는 경우에도 동일하게 산출할 수 있다.
- [0091] 도 1 내지 도 4b에서 설명된 목표 달성 진단 장치(100)는 실시예를 들어 설명한 취업 확률뿐만 아니라 다양한 분야에 응용이 가능하다. 목표 달성 진단 장치(100)는 은행 입금 프로세스, 다이어트 프로세스, 범죄 프로세스, 인생 설계 프로세스, 취업 프로세스, 졸업 프로세스, 반도체 공정 프로세스 및 발명 프로세스와 같이 다양한 분야에 응용할 수 있다.
- [0092] 예를 들어, 사람의 방문 장소를 분석하여 발명 프로세스를 진단할 수 있다. 프로세스의 수행기간을 5년으로 보고, 매월 1일을 평가시점으로, 장소 방문을 활동으로 보며, 해당 달에 방문한 장소의 유형(음식 관련 장소, 건강 관련 장소, 스트레스 관련 장소 등)을 이벤트 속성값으로 정의하고, 발명되는 병의 유형을 프로세스의 최종성과 유형으로 설정하여 분석을 진행할 수 있다.
- [0093] 도 5는 본 발명의 일 실시예에 따른 목표 진단 달성 방법을 나타내는 흐름도이다.
- [0094] 도 5를 참조하면, 본 발명의 일 실시예에 따른 목표 진단 달성 방법은 예측하고자 하는 목표를 달성하기 위한 프로세스의 목표, 최종 성과 유형, 속성, 수행기간 및 평가시점을 정의한다(501). 본 발명에서 프로세스는 예측하고자 하는 목표(특정 목표)를 달성하기 위해 행해지는 일련의 활동을 나타낸다. 프로세스의 수행 기간은 프로세스 인스턴스의 최종 성과 유형별 근접도 산출의 대상 기간이 되며, 단위 시간으로 표시될 수 있다. 평가시점은 객체에 대한 시점별 최종 성과 유형별 근접도 산출하기 위한 시점으로, 활동에 의해 이벤트 속성값이 존재하는 과거 및 현재 시점이 평가시점이 될 수 있다. 또한, 평가시점은 과거 및 현재 시점뿐만 아니라, 미래의 시점이 될 수 있다.
- [0095] 다음으로, 정의된 목표, 최종성과 유형, 속성, 수행기간 및 평가시점에 부합하는 하나 이상의 과거 프로세스 인스턴스로부터 속성 값을 수집 또는 확인한다(502). 본 발명에 따른 목표 달성 진단 방법에서 분석에 사용되는 프로세스 인스턴스의 속성은 과거, 현재 및 미래 속성을 포함할 수 있으며, 분석의 대상이 되는 프로세스의 일부 속성만으로도 분석을 수행할 수 있다. 분석 과정에서 미래 속성의 값은 예측되거나, 이미 현재 시점에서 확정된 값들로 사용될 수 있다. 그리고, 프로세스 인스턴스의 속성은 환경 속성, 객체 속성, 활동 속성 및 이벤트 속성으로 분류될 수 있다.
- [0096] 프로세스 인스턴스의 속성 값이 수집되면, 이벤트 프로파일 추출을 위한 군집분석을 수행하기 이전에 프로세스 인스턴스의 세분화를 위한 군집분석을 수행할 수 있다(503). 이 군집분석을 '세분화 군집분석'이라고 하고, 그 결과물인 군집을 '세분화 군집'이라 명명한다. 수집된 프로세스 인스턴스의 속성 값을 대상으로 이벤트 프로파일 군집분석을 실시하기 이전에, 세분화 군집분석을 수행하여, 유사한 환경 속성값이나 유사한 객체 속성값을 지닌 프로세스 인스턴스들로 군집을 형성하게 하고, 이 세분화 군집들에 대해 각각 이벤트 프로파일 군집분석을 수행하여 이벤트 프로파일을 도출함으로써 도출된 이벤트 프로파일 간의 상이성을 증대시킬 수 있다.
- [0097] 하나 이상의 환경 속성 및 하나 이상의 객체 속성 중에서 적어도 하나의 속성에 기초하여 군집분석(Cluster Analysis)을 실시하여 하나 이상의 군집(cluster)을 형성한다. 세분화를 위한 군집분석 과정은 이벤트 프로파일 에 대한 군집분석을 수행하기 이전에 필요에 따라 선택적으로 수행될 수 있다.

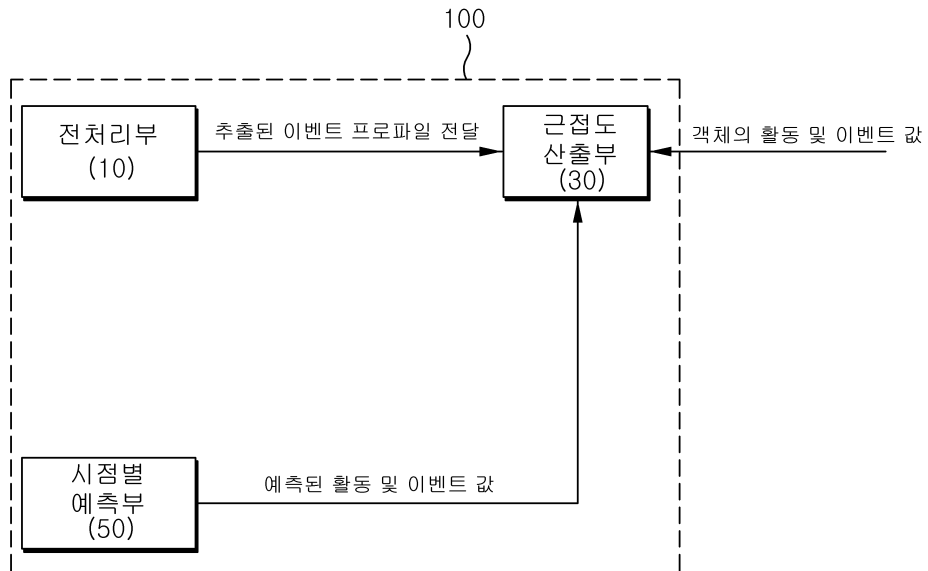
- [0098] 다음으로, 503단계에서 형성된 각 세분화 군집 내 과거 프로세스 인스턴스의 속성 값을 대상으로 이벤트 프로파일 군집분석을 실시한다(504). 이 군집분석을 '이벤트 프로파일 군집분석'이라고 하고, 그 결과물인 군집을 '프로파일 군집'이라 명명한다. 세분화 군집이 수행되었을 경우에 프로파일 군집분석은 세분화 군집을 대상으로 수행한다. 분석 대상인 과거 프로세스 인스턴스의 속성 값에 대해 하나 이상의 활동 속성 및 하나 이상의 이벤트 속성 중에서 하나 이상에 기초하여 군집(cluster)을 형성한다. 그리고, 군집분석 시에 사용된 하나 이상의 속성의 값들을 시간 순차적으로 나열한다. 시간 순차적으로 나열된 하나 이상의 속성의 값들에서 각 군집에 속한 프로세스 인스턴스들은 유사성을 띤다. 다음으로 각 프로파일 군집 내 프로세스 인스턴스들의 속성 값들에 기초하여 집단 각각에서 하나 이상의 이벤트 프로파일을 추출한다 (505).
- [0099] 이벤트 프로파일이 추출되면, 시점별 최종 성과 유형별 근접도 산출을 위해서 진행 중인 프로세스 인스턴스가 각 이벤트 프로파일에 속할 사전확률을 산출한다(506). 사전확률은 도 1에서 설명한 수학적 식 1을 통해 산출될 수 있다. 그리고, 프로세스가 수행됨에 따라 해당 프로세스 인스턴스가 t번째 평가시점까지 획득한 이벤트 속성값들이 벡터 D_t 를 형성(507)할 때, 도 1에서 설명한 수학적 식 2를 통해 해당 프로세스 인스턴스가 k번째 이벤트 프로파일에 속할 사후확률을 산출한다(508).
- [0100] 사전확률 및 사후확률이 산출되면, 산출된 사전확률 및 사후확률을 이용해서 t번째 평가시점에서 프로세스 인스턴스의 시점별 최종 성과 유형별 근접도를 산출한다(509). 프로세스 인스턴스의 시점별 최종 성과 유형별 근접도는 도 1에서 설명한 수학적 식 4를 이용하여 산출될 수 있다.
- [0101] 상술한 내용을 포함하는 본 발명은 컴퓨터 프로그램으로 작성이 가능하다. 그리고 상기 프로그램을 구성하는 코드 및 코드 세그먼트는 당분야의 컴퓨터 프로그래머에 의하여 용이하게 추론될 수 있다. 또한, 상기 작성된 프로그램은 컴퓨터가 읽을 수 있는 기록매체 또는 정보저장매체에 저장되고, 컴퓨터에 의하여 관독되고 실행함으로써 본 발명의 방법을 구현할 수 있다. 그리고 상기 기록매체는 컴퓨터가 관독할 수 있는 모든 형태의 기록매체를 포함한다.
- [0102] 이상 바람직한 실시예를 들어 본 발명을 상세하게 설명하였으나, 본 발명은 전술한 실시예에 한정되지 않고, 본 발명의 기술적 사상의 범위 내에서 당분야에서 통상의 지식을 가진자에 의하여 여러 가지 변형이 가능하다.

부호의 설명

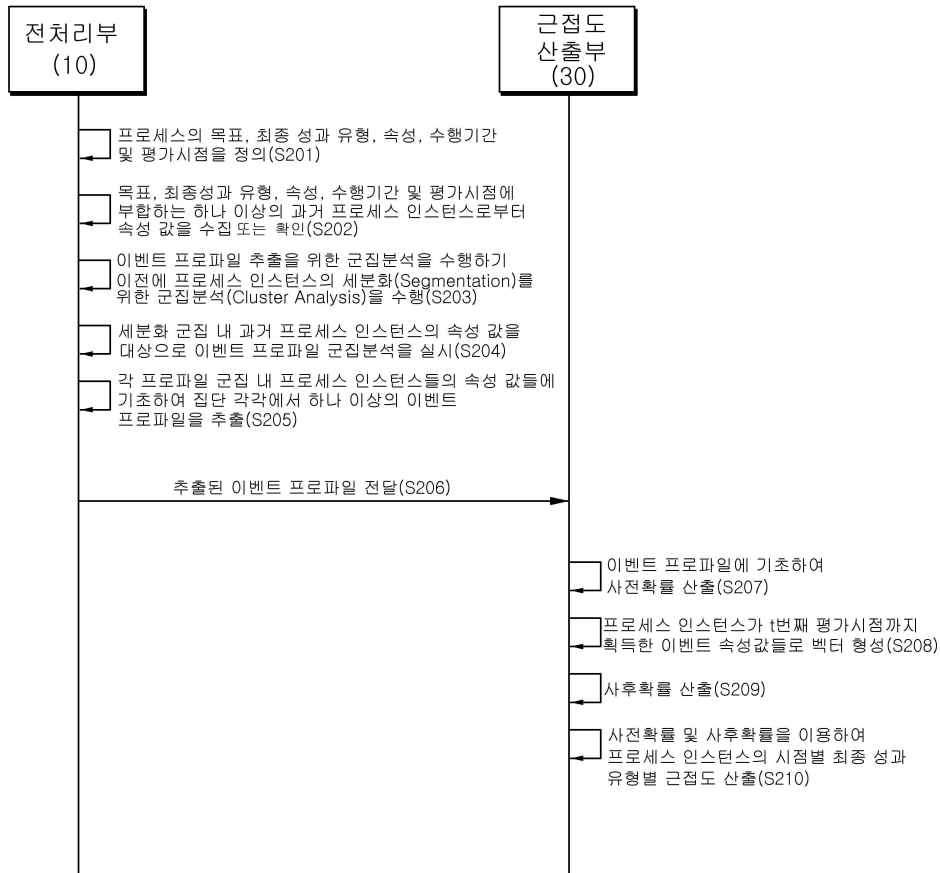
- [0103] 10: 전처리부
30: 근접도 산출부
50: 시점별 예측부
301: 제1 EP 그래프
302: 제2 EP 그래프
303: 제3 EP 그래프

도면

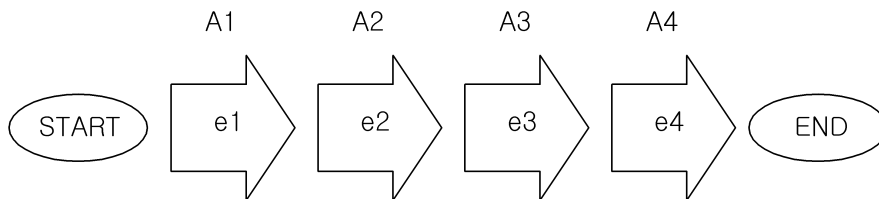
도면1



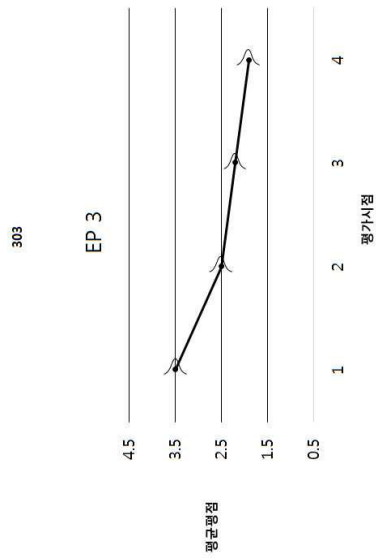
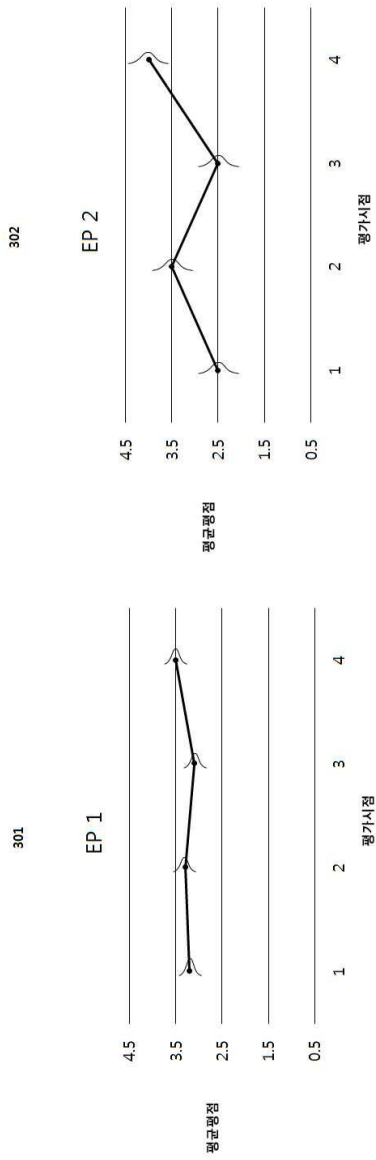
도면2



도면3a



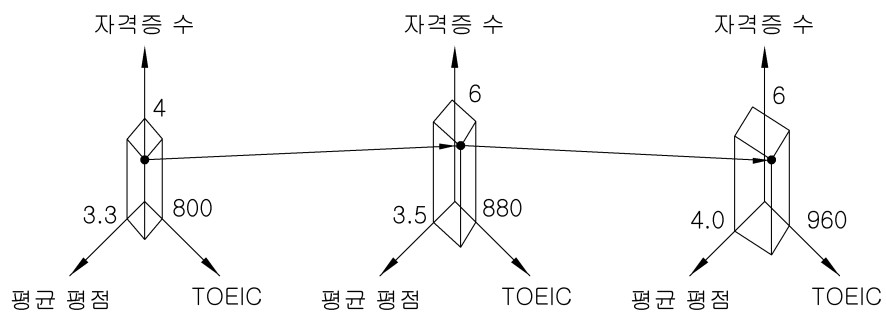
도면3b



도면4a

(t-1)번째 평가시점	t번째 평가시점	(t+1)번째 평가시점
평균평점: 3.3 토익점수: 800 자격증 수: 4개	평균평점: 3.5 토익점수: 880 자격증 수: 6개	평균평점: 4.0 토익점수: 960 자격증 수: 6개

도면4b



도면5

