(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년12월01일
(11) 등록번호 10-2333505
(24) 등록일자 2021년11월26일

(51) 국제특허분류(Int. Cl.)
G10L 15/22 (2006.01) *G06F 16/00* (2019.01)
G06F 3/16 (2018.01) *G06F 40/20* (2020.01)
G10L 15/18 (2006.01)
(52) CPC특허분류
G10L 15/22 (2013.01)
G06F 16/3329 (2019.01)
(21) 출원번호 10-2017-7002966
(22) 출원일자(국제) 2015년07월02일
심사청구일자 2020년06월02일
(85) 번역문제출일자 2017년02월02일
(65) 공개번호 10-2017-0026593
(43) 공개일자 2017년03월08일
(86) 국제출원번호 PCT/US2015/038923
(87) 국제공개번호 WO 2016/004266
국제공개일자 2016년01월07일
(30) 우선권주장
14/323,050 2014년07월03일 미국(US)
(56) 선행기술조사문헌
JP2009110503 A

(73) 특허권자
마이크로소프트 테크놀로지 라이센싱, 엘엘씨
미국 워싱턴주 (우편번호 : 98052) 레드몬드 원 마이크로소프트 웨이
(72) 발명자
윌리암스 제이슨
미국 워싱턴주 98052-6399 레드몬드 원 마이크로소프트 웨이 엘씨에이 - 인터내셔날 패튼츠 (8/1172) 마이크로소프트 테크놀로지 라이센싱, 엘엘씨 내
츠바이그 제프리
미국 워싱턴주 98052-6399 레드몬드 원 마이크로소프트 웨이 엘씨에이 - 인터내셔날 패튼츠 (8/1172) 마이크로소프트 테크놀로지 라이센싱, 엘엘씨 내
(뒷면에 계속)
(74) 대리인
김태홍, 김진희

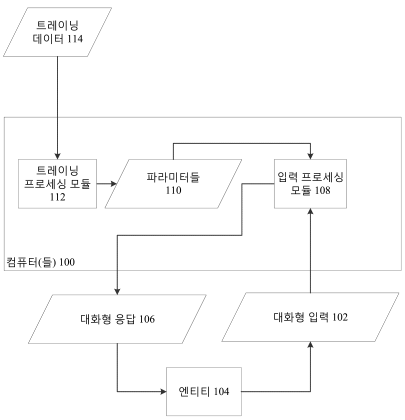
전체 청구항 수 : 총 20 항 심사관 : 임민섭

(54) 발명의 명칭 **소셜 대화형 입력들에 대한 컴퓨터 응답 생성**

(57) 요 약

사람과 컴퓨터 시스템 간의 대화형 상호작용은, 대화 유형에 의해 입력을 분류하고 대화 유형들에 대한 사람이 작성한 응답들을 제공하는 컴퓨터 시스템에 의해 제공될 수 있다. 입력 분류는 트레이닝된 2진 분류기(trained binary classifier)들을 사용하여 수행될 수 있다. 트레이닝은, 입력들을 대화 유형의 긍정적 또는 부정적 예시들 중 어느 하나로서 라벨링함으로써 수행될 수 있다. 대화형 응답들은 분류기들을 트레이닝할 때 사용된 입력들을 라벨링한 동일한 개인들에 의해 작성될 수 있다. 일부 경우들에서, 분류기들을 트레이닝하는 과정은 새로운 대화 유형의 제안을 초래시킬 수 있고, 이에 대해 사람 작성자들은 새로운 분류기에 대한 입력들을 라벨링하고, 해당 새로운 대화 유형에 대한 응답 콘텐츠를 작성할 수 있다.

대 표 도 - 도1



(52) CPC특허분류
G06F 3/167 (2013.01)
G06F 40/35 (2020.01)
G10L 15/1822 (2013.01)
G10L 2015/223 (2013.01)
G10L 2015/226 (2020.08)
(72) 발명자
라크쉬미라탄 아파르나
미국 워싱턴주 98052-6399 레드몬드 원 마이크로소프트 웨이 엘씨에이 - 인터내셔날 패튼츠 (8/1172) 마이크로소프트 테크놀로지 라이센싱, 엘엘씨 내
수아레즈 카를로스 가르시아 후라도
미국 워싱턴주 98052-6399 레드몬드 원 마이크로소프트 웨이 엘씨에이 - 인터내셔날 패튼츠 (8/1172) 마이크로소프트 테크놀로지 라이센싱, 엘엘씨 내

명 세 서

청구범위

청구항 1

대화형(conversational) 입력들을 대화 유형으로 분류하기 위한 컴퓨터 구현 분류기를 트레이닝하기 위한 프로세스에 있어서,

상기 분류기는, 대화형 입력을 나타내는 데이터를 수신하는 입력부와, 상기 대화형 입력이 상기 분류기에 의해 나타난 대화 유형과 어떻게 일치하는지를 표시하는 일치 출력을 제공하는 출력부를 갖고,

상기 분류기를 트레이닝하기 위한 프로세스는 프로세서와 컴퓨터 저장장치를 포함하는 프로세싱 시스템에 의해 수행되고,

상기 프로세스는,

상기 컴퓨터 저장장치 내로, 제1 세트의 대화형 입력들을 나타내고 상기 제1 세트의 대화형 입력들을 상기 대화 유형의 긍정적 예시들로서 표시하는 제1 라벨 데이터를 포함한 데이터와, 제2 세트의 대화형 입력들을 나타내고 상기 제2 세트의 대화형 입력들을 상기 대화 유형의 부정적 예시들로서 표시하는 제2 라벨 데이터를 포함한 데이터를 수신하는 단계;

상기 제1 세트의 대화형 입력들과 상기 제2 세트의 대화형 입력들 및 상기 제1 라벨 데이터와 상기 제2 라벨 데이터를 나타내는 데이터를 사용하여 상기 분류기를 구축하는 단계;

상기 컴퓨터 저장장치로부터, 제3 세트의 대화형 입력들을 나타내는 데이터에 액세스하는 단계;

상기 제3 세트의 대화형 입력들에서의 각각의 대화형 입력에 대해:

상기 분류기의 출력부로부터 일치 출력을 획득하기 위해 상기 분류기의 입력부에 상기 대화형 입력을 나타내는 데이터를 인가하는 단계, 및

상기 일치 출력이 상기 대화형 입력의 분류를 긍정적 예시와 부정적 예시 중 하나로서 표시하지 않는다라는 결정에 응답하여,

라벨링을 위해 상기 대화형 입력을 하나 이상의 사용자 컴퓨터를 통해 하나 이상의 개인에게 제시하는 단계,

상기 제시된 대화형 입력에 대한 제3 라벨 데이터를 상기 하나 이상의 개인으로부터 상기 하나 이상의 사용자 컴퓨터를 통해 수신하는 단계, 및

상기 수신된 제3 라벨 데이터를 상기 대화형 입력과 연관시켜서 상기 컴퓨터 저장장치에 저장하는 단계; 및

상기 제3 라벨 데이터를 갖는 상기 제3 세트의 대화형 입력들을 나타내는 데이터를 사용하여 상기 분류기를 재트레이닝하는 단계

를 포함하는 컴퓨터 구현 분류기를 트레이닝하기 위한 프로세스.

청구항 2

제1항에 있어서,

상기 하나 이상의 개인 중 하나 이상으로부터, 상기 하나 이상의 개인이 상기 라벨링을 수행했던 분류기와 연관된 대화 유형에 대한 대화형 응답들을 나타내는 데이터를 수신하고 이를 상기 컴퓨터 저장장치에 저장하는 단계

를 더 포함하는 컴퓨터 구현 분류기를 트레이닝하기 위한 프로세스.

청구항 3

제1항에 있어서,

상기 제3 세트의 대화형 입력들을 나타내는 데이터로부터 추가적인 대화 유형들을 식별하는 단계

를 더 포함하는 컴퓨터 구현 분류기를 트레이닝하기 위한 프로세스.

청구항 4

제3항에 있어서,

상기 식별하는 단계는,

상기 제3 세트의 대화형 입력들을 나타내는 데이터 내의 대화형 입력들을 클러스터링(clustering)하는 단계;

상이한 대화 유형을 각각의 클러스터에 할당하는 단계

를 포함한 것인 컴퓨터 구현 분류기를 트레이닝하기 위한 프로세스.

청구항 5

제3항에 있어서,

상기 식별하는 단계는,

상기 대화형 입력들을 복수의 분류기들에 인가하는 단계;

상기 복수의 분류기들 중 어느 것도 일치를 표시하지 않은 대화형 입력들을 식별하는 단계;

상기 식별된 대화형 입력들에 하나 이상의 새로운 대화 유형을 할당하는 단계

를 포함한 것인 컴퓨터 구현 분류기를 트레이닝하기 위한 프로세스.

청구항 6

제1항에 있어서,

분류기의 대화 유형은 대화 유형들의 계층구조(hierarchy) 내에 포함된 것인 컴퓨터 구현 분류기를 트레이닝하기 위한 프로세스.

청구항 7

제1항에 있어서,

상기 대화형 입력들에 대한 클래스 정보를 제공하기 위해 상기 제3 세트의 대화형 입력들을 나타내는 데이터를 자동으로 클러스터링하는 단계

를 더 포함하는 컴퓨터 구현 분류기를 트레이닝하기 위한 프로세스.

청구항 8

제1항에 있어서,

상기 분류기를 구축하는 단계는,

상기 프로세싱 시스템으로, 상기 분류기의 입력부에 인가된 데이터에 대해 상기 분류기에 의해 인가된 메트릭에 대한 파라미터들의 세트를 정의하는 단계

를 포함한 것인 컴퓨터 구현 분류기를 트레이닝하기 위한 프로세스.

청구항 9

제8항에 있어서,

상기 분류기에 의해 인가된 메트릭은, 상기 파라미터들의 세트에 따라 상기 분류기의 입력부에 인가된 데이터를 사용하여 유사성 메트릭 또는 거리 메트릭 또는 확률 메트릭을 계산하기 위해 프로세서에 의해 프로세싱된 컴퓨터 프로그램 명령어들을 포함한 것인 컴퓨터 구현 분류기를 트레이닝하기 위한 프로세스.

청구항 10

제1항에 있어서,

상기 제1 세트의 대화형 입력들, 상기 제2 세트의 대화형 입력들, 및 상기 제3 세트의 대화형 입력들에서의 대화형 입력을 나타내는 데이터는 상기 대화형 입력으로부터 도출된 복수의 특징들을 포함한 것인 컴퓨터 구현 분류기를 트레이닝하기 위한 프로세스.

청구항 11

컴퓨터 시스템에 있어서,

프로세서와 컴퓨터 저장장치를 포함한 프로세싱 시스템;

대화형 입력들을 대화 유형으로 분류하기 위한 분류기 - 상기 분류기는, 대화형 입력을 나타내는 데이터를 수신하는 입력부와, 상기 대화형 입력이 상기 분류기에 의해 나타난 대화 유형과 어떻게 일치하는지를 표시하는 일치 출력을 제공하는 출력부를 갖고,

상기 컴퓨터 저장장치는, 제1 세트의 대화형 입력들을 나타내고 상기 제1 세트의 대화형 입력들을 상기 대화 유형의 긍정적 예시들로서 표시하는 제1 라벨 데이터를 포함한 데이터와, 제2 세트의 대화형 입력들을 나타내고 상기 제2 세트의 대화형 입력들을 상기 대화 유형의 부정적 예시들로서 표시하는 제2 라벨 데이터를 포함한 데이터를 저장함 -;

상기 제1 세트의 대화형 입력들과 상기 제2 세트의 대화형 입력들 및 상기 제1 라벨 데이터와 상기 제2 라벨 데이터를 나타내는 데이터를 사용하여 상기 분류기를 구축하기 위한 트레이닝 모듈

을 포함하고,

상기 트레이닝 모듈은,

상기 컴퓨터 저장장치로부터, 제3 세트의 대화형 입력들을 나타내는 데이터에 액세스하고;

상기 제3 세트의 대화형 입력들에서의 각각의 대화형 입력에 대해:

상기 분류기의 출력부로부터 일치 출력을 획득하기 위해 상기 분류기의 입력부에 상기 대화형 입력을 나타내는 데이터를 인가하고,

상기 일치 출력이 상기 대화형 입력의 분류를 긍정적 예시와 부정적 예시 중 하나로서 표시하지 않는다라는 결정에 응답하여,

라벨링을 위해 상기 대화형 입력을 하나 이상의 사용자 컴퓨터를 통해 하나 이상의 개인에게 제시하고,

상기 제시된 대화형 입력에 대한 제3 라벨 데이터를 상기 하나 이상의 개인으로부터 상기 하나 이상의 사용자 컴퓨터를 통해 수신하고,

상기 수신된 제3 라벨 데이터를 상기 대화형 입력과 연관시켜서 상기 컴퓨터 저장장치에 저장하며;

상기 제3 라벨 데이터를 갖는 상기 제3 세트의 대화형 입력들을 나타내는 데이터를 사용하여 상기 분류기를 재트레이닝하도록 구성된 것인 컴퓨터 시스템.

청구항 12

제11항에 있어서,

상기 트레이닝 모듈은 또한, 상기 분류기를 구축하기 위해, 상기 분류기의 입력부에 인가된 데이터에 대해 상기 분류기에 의해 인가된 메트릭에 대한 파라미터들의 세트를 정의하도록 구성된 것인 컴퓨터 시스템.

청구항 13

제12항에 있어서,

상기 분류기에 의해 인가된 메트릭은, 상기 파라미터들의 세트에 따라 상기 분류기의 입력부에 인가된 데이터를 사용하여 유사성 메트릭 또는 거리 메트릭 또는 확률 메트릭을 계산하기 위해 프로세서에 의해 프로세싱된 컴퓨터 프로그램 명령어들을 포함한 것인 컴퓨터 시스템.

청구항 14

제11항에 있어서,

상기 제1 세트의 대화형 입력들, 상기 제2 세트의 대화형 입력들, 및 상기 제3 세트의 대화형 입력들에서의 대화형 입력을 나타내는 데이터는 상기 대화형 입력으로부터 도출된 복수의 특징들을 포함한 것인 컴퓨터 시스템.

청구항 15

제11항에 있어서,

상기 트레이닝 모듈은 또한, 상기 하나 이상의 개인 중 하나 이상으로부터, 상기 하나 이상의 개인이 상기 라벨링을 수행했던 분류기와 연관된 대화 유형에 대한 대화형 응답들을 나타내는 데이터를 수신하고 이를 상기 컴퓨터 저장장치에 저장하도록 구성된 것인 컴퓨터 시스템.

청구항 16

제12항에 있어서,

상기 트레이닝 모듈은 또한, 상기 제3 세트의 대화형 입력들을 나타내는 데이터를 프로세싱한 것에 기초하여 추가적인 대화 유형을 식별하도록 구성된 것인 컴퓨터 시스템.

청구항 17

제16항에 있어서,

추가적인 대화 유형을 식별하기 위해, 상기 트레이닝 모듈은 또한,

상기 제3 세트의 대화형 입력들을 나타내는 데이터에 기초하여 상기 제3 세트의 대화형 입력들 내의 대화형 입력들을 클러스터들 내로 클러스터링하며;

대화 유형을 각각의 클러스터에 할당하도록 구성된 것인 컴퓨터 시스템.

청구항 18

제16항에 있어서,

추가적인 대화 유형을 식별하기 위해, 상기 트레이닝 모듈은 또한,

상기 제3 세트의 대화형 입력들을 나타내는 데이터를 복수의 분류기들에 인가하고;

상기 복수의 분류기들 중 어느 것도 일치를 표시하지 않은 대화형 입력들을 식별하며;

상기 식별된 대화형 입력들에 하나 이상의 새로운 대화 유형을 할당하도록 구성된 것인 컴퓨터 시스템.

청구항 19

제11항에 있어서,

분류기의 대화 유형은 대화 유형들의 계층구조 내에 포함된 것인 컴퓨터 시스템.

청구항 20

제11항에 있어서,

상기 트레이닝 모듈은 또한, 상기 대화형 입력들에 대한 클래스 정보를 제공하기 위해 상기 제3 세트의 대화형 입력들을 자동으로 클러스터링하도록 구성된 것인 컴퓨터 시스템.

발명의 설명

기 술 분 야

[0001] 본 발명은 소셜 대화형 입력들에 대한 컴퓨터 응답 생성에 관한 것이다.

배 경 기 술

[0002] 컴퓨터가 개인들로부터의 대화형(conversational) 입력들, 즉 대화형인 자연 언어(natural language) 입력에 대해 적절한 대화형 응답들을 제공함으로써 컴퓨터의 응답들이 흥미로운 사용자 경험을 생성하게 하도록 하는 것은 해결과제이다. 이러한 해결과제의 일부는 자연 언어에 내재된 다양한 변이들을 고려하여, 입력을 이해하는 것이다. 이러한 해결과제의 또다른 부분은 입력에 대한 그 어떠한 응답도 적절하도록 보장하는 것이다. 단어 선택, 어조, "개성" 또는 인상과의 일관성과 같이, 응답이 적절한 것으로 간주되는지 여부, 및 응답이 재미있고, 재치있고, 흥미로우며, 불쾌하지 않은지 등의 여부에 많은 요소들이 영향을 미친다.

[0003] 일부 시스템들에서, 이러한 해결과제는 정규 표현들을 사용하고, 사람이 작성한 콘텐츠를 패턴들과 연관시키는 것과 같이, 수동으로 이러한 패턴들을 규정함으로써 해결된다. 입력이 정규 표현과 일치하면, 일치하는 정규 표현과 연관된 콘텐츠가 응답으로서 제공될 수 있다. 이러한 시스템들은 정규 표현들을 작성하고 분석하는데 있어서 전문적인 기술을 가진 사람들(일반적으로 컴퓨터 프로그래머), 및 응답에 대한 적절한 콘텐츠를 작성하는 기술을 가진 사람들(일반적으로 콘텐츠 편집자)에 의해 개발된다.

발명의 내용

해결하려는 과제

[0004] 일부 시스템들에서는, 대용량 데이터 세트로부터 대화들을 마이닝(mining)하여 자동으로 응답들을 생성함으로써 이러한 해결과제가 해결된다. 입력이 하나 이상의 이전 대화들의 입력과 일치하는 경우, 이전 대화(들)의 해당 입력에 대한 임의의 응답들에 기초하여 자동으로 응답이 생성될 수 있다. 이러한 시스템에서, 자동 생성된 응답은 컴퓨터가 이 응답을 생성하기 전에 사람에 의해 점검되지 않으므로, 부적절한 응답을 생성할 위험이 높아진다.

과제의 해결 수단

[0005] 본 요약은 아래의 상세한 설명에서 보다 상세하게 설명되는 개념들의 선택을 단순한 형태로 소개하기 위해 제공된 것이다. 본 요약은 청구된 발명내용의 중요하거나 또는 필수적인 특징들을 식별시키려는 의도는 없으며, 또한 그 범위를 제한시키려는 의도도 없다.

[0006] 사람과 컴퓨터 시스템 간의 대화형 상호작용은, 대화 유형에 의해 입력을 분류하고 대화 유형들에 대한 사람이 작성한 응답들을 제공하는 컴퓨터 시스템에 의해 제공될 수 있다. 입력 분류는 트레이닝된 2진 분류기(trained binary classifier)들을 사용하여 수행될 수 있다. 트레이닝은, 입력들을 대화 유형의 긍정적 또는 부정적 예시들 중 어느 하나로서 라벨링함으로써 수행될 수 있다. 대화형 응답들은 분류기들을 트레이닝할 때 사용된 입력들을 라벨링한 동일한 개인들에 의해 작성될 수 있다. 일부 경우들에서, 분류기들을 트레이닝하는 프로세스는 새로운 대화 유형의 제안을 초래시킬 수 있고, 이에 대해 사람 작성자들은 새로운 분류기에 대한 입력들을 라벨링하고, 해당 새로운 대화 유형에 대한 응답 콘텐츠를 작성할 수 있다.

[0007] 본 발명은 컴퓨터 시스템으로서, 그러한 컴퓨터 시스템의 임의의 개별 컴포넌트로서, 그러한 컴퓨터 시스템 또는 그러한 컴퓨터 시스템의 임의의 개별 컴포넌트에 의해 수행되는 프로세스로서, 또는 컴퓨터들에 의해 프로세싱될 때, 이들 컴퓨터들이 그러한 컴퓨터 시스템 또는 그러한 컴퓨터 시스템의 임의의 개별 컴포넌트를 제공하도록 이들 컴퓨터들을 구성시키는 컴퓨터 프로그램 명령어들이 저장되어 있는 컴퓨터 저장장치를 포함하는 제조물품으로서 구현될 수 있다.

[0008] 아래의 설명에서는, 본 기술의 일부를 형성하는 첨부 도면들을 참조하며, 이 도면들에서는 본 기술의 예시적인 특정 구현예들이 실례로서 도시되어 있다. 본 발명개시의 범위를 벗어나지 않고서 다른 실시예들이 이용될 수 있고 구조적 변경들이 취해질 수 있다는 것이 이해된다.

도면의 간단한 설명

[0009] 도 1은 컴퓨터 시스템이 대화형 입력들에 대한 응답 생성을 지원하는 예시적인 애플리케이션 환경의

블록도이다.

도 2는 입력 프로세싱 모듈의 예시적인 구현을 설명하는 데이터 흐름도이다.

도 3은 입력 프로세싱 모듈의 예시적인 동작을 설명하는 흐름도이다.

도 4는 트레이닝 프로세스의 예시적인 구현을 설명하는 흐름도이다.

도 5는 트레이닝 프로세싱 모듈의 예시적인 구현의 데이터 흐름도이다.

도 6은 이러한 시스템의 컴포넌트들이 구현될 수 있는 예시적인 컴퓨터의 블록도이다.

발명을 실시하기 위한 구체적인 내용

[0010] 다음의 섹션은 대화형 입력들에 대한 응답 생성을 지원하는 컴퓨터의 예시적인 동작 환경을 설명한다.

[0011] 도 1을 참조하면, 컴퓨터(100)는 엔티티(104)로부터 대화형 입력(102)을 수신한다. 대화형 입력(102)은 입력의 수신자에게 의미를 전달하도록 의도된 데이터를 포함한다. 대화형 입력은 임의의 다양한 디바이스들 및 프로세싱을 통해 수신될 수 있다. 대화형 입력의 간단한 예시는 "오늘 시애틀의 날씨는 어떻습니까?"이다.

[0012] 대화형 입력(102)에 응답하여, 컴퓨터(100)는 대화형 응답(106)을 생성한다. 대화형 응답(106)은 대화형 입력(102)을 제공했던 엔티티(104)에게 의미를 전달하도록 의도된 데이터를 포함한다. 대화형 응답(106)은 임의의 다양한 디바이스들 및 프로세싱을 통해 엔티티(104)에 전달될 수 있다. 대화형 응답의 간단한 예시는 "오늘 시애틀에서는 비가 내리고 있습니다"이다.

[0013] 대화형 응답을 생성하기 위해, 컴퓨터(100)는 입력 프로세싱 모듈(108)을 포함하며, 그 예시적인 구현이 도 2 및 도 3과 관련하여 이하에서 보다 상세하게 설명된다.

[0014] 입력 프로세싱 모듈(108)은 트레이닝 프로세스를 사용하여 설정된 파라미터들(110)을 갖는다. 트레이닝 프로세스는 트레이닝 프로세싱 모듈(112)을 통해 구현된다. 트레이닝 프로세싱 모듈은 트레이닝 데이터(114)를 수신하는 입력부들 및 파라미터들(110)을 제공하는 출력부들을 갖는다. 트레이닝 프로세싱 모듈의 예시적인 구현은 도 4 및 5와 관련하여 아래에서 보다 상세하게 설명된다.

[0015] 도 1에서, 컴퓨터(100)는 입력 프로세싱 모듈(108) 및 트레이닝 프로세싱 모듈(112) 둘 다를 포함한다. 컴퓨터(100)는 하나 이상의 구별적인 범용 컴퓨터 또는 다른 구별적인 컴퓨팅 디바이스들을 포함할 수 있으며, 이들 각각은 하나 이상의 프로세서들 또는 프로세싱 디바이스들을 포함할 수 있다. 따라서, 모듈들(108, 112)은 동일한 컴퓨팅 디바이스 상에 있거나 또는 별개의 컴퓨팅 디바이스들 상에 있을 수 있거나, 또는 그 일부분들은 동일한 컴퓨팅 디바이스 상에 있을 수 있다. 각각의 모듈(108, 112)은 하나 이상의 컴퓨팅 디바이스들 상에 있을 수 있다.

[0016] 도 1에서, 하나의 구현예에서, 트레이닝 프로세싱 모듈(112)은 입력 프로세싱 모듈의 파라미터들을 직접 업데이트할 수 있다. 다른 구현예에서, 트레이닝 프로세싱 모듈은 입력 프로세싱 모듈과는 별개의 업데이트된 파라미터들을 생성할 수 있고, 그 후, 업데이트된 파라미터들을 때때로 입력 프로세싱 모듈에서 전송할 수 있다.

[0017] 일부 구현예들에서, 트레이닝 프로세싱 모듈에 의한 트레이닝은 한 번 발생할 수 있다. 다른 구현예들에서, 트레이닝은 진행중인 프로세스로서 수행될 수 있고, 입력 프로세싱 모듈의 파라미터들은 주기적으로와 같이, 때때로 업데이트될 수 있다.

[0018] 트레이닝 프로세싱 모듈은 상이한 컴퓨팅 디바이스들 상의 복수의 상이한 입력 프로세싱 모듈들에 의해 사용되고, 이에 따라, 이에 의해 공유되는 하나의 파라미터 세트를 생성할 수 있다. 복수의 상이한 트레이닝 프로세싱 모듈들은, 상이한 입력 프로세싱 모듈들에 상관없이 또는, 예를 들어, 최종 사용자 또는 상이한 최종 사용자들의 상이한 상황들에 대해, 하나의 입력 프로세싱 모듈로부터 선택되고 이에 의해 사용될 상이한 파라미터 세트를 생성할 수 있다.

[0019] 도 1은 또한 대화형 입력들(102)을 수신하고 대화형 응답들(106)을 엔티티(104)에 제공하는 컴퓨터(100)를 도시한다. 엔티티(104)는 일반적으로 개인이다.

[0020] 대화형 입력들(102)은 일반적으로 자연 언어로 된 하나 이상의 단어들 또는 구문들이다. 다른 입력들은 비제한적인 예시로서, 수학 기호, 이모티콘 등과 같이, 자연 언어의 일부로서 공식적으로 인식되지 않을 수 있지만 의미를 전달하기 위해 통상적으로 사용되는 기호를 포함할 수 있다. 대화형 입력과 관련된 다른 데이터는 입력

디바이스로부터의 제스처 데이터, 착용식 센서 및 환경 센서로부터의 센서 데이터, 이미지 데이터, 비디오 데이터 및/또는 오디오 데이터를 포함할 수 있으며, 이들로부터는 다양한 메타데이터도 추출될 수 있다. 엔티티의 제스처, 음성의 어조, 표정, 가능한 감정 등과 관련된 메타데이터가 유용될 수 있다.

[0021] 이러한 입력들은 다양한 방법으로 제공될 수 있다. 예를 들어, 엔티티는 키보드로 타이핑할 수 있다. 엔티티는 마이크로폰으로 말하고, 컴퓨터(100)는 음성 인식을 수행할 수 있다. 엔티티는 이미지 또는 비디오 데이터를 제공할 수 있고, 컴퓨터(100)는 이미지 인식 또는 다른 이미지 프로세싱을 수행할 수 있다. 엔티티는 텍스트 데이터를 제공할 수 있고, 컴퓨터(100)는 텍스트를 프로세싱하여 단어 및 구문을 추출할 수 있다.

[0022] 대화형 응답들(106)은 일반적으로 자연 언어로 된 하나 이상의 단어들 또는 구문들이다. 다른 입력들은 비제한적인 예시로서, 수학 기호, 이모티콘 등과 같이, 자연 언어의 일부로서 공식적으로 인식되지 않을 수 있지만 의미를 전달하기 위해 통상적으로 사용되는 기호를 포함할 수 있다. 사용자에 대한 응답의 전달과 관련하여 다양한 목적으로 다른 데이터가 또한 생성될 수 있다.

[0023] 이러한 대화형 출력은 다양한 방법으로 제공될 수 있다. 예를 들어, 응답들은 디스플레이 스크린 상에서 텍스트로서 디스플레이될 수 있다. 응답들은 스피커 또는 헤드폰에서 재생되는 오디오 데이터를 제공하기 위해 음성 생성을 사용하여 프로세싱될 수 있다. 응답들은 텍스트에만 국한되지 않는다면, 다른 이미지 데이터로 변환될 수 있고, 그 이미지 데이터는 디스플레이될 수 있다. 예를 들어, 이러한 이미지 데이터는 엔티티의 제스처 및 표정을 보여주는 엔티티의 아바타를 포함할 수 있다. 사용자에 의해 선택된 원하는 상호작용 포맷에 의존하여 (예컨대, 점자, 인쇄 또는 다른 유형들의 출력 디바이스들과 같은) 다른 출력 포맷들이 생성될 수 있다.

[0024] 컴퓨터(100) 또는 그 일부는 엔티티(104)에 의해 사용될 수 있는 디바이스(도시되지 않음)의 일부일 수 있거나, 또는 엔티티(104)에 의해 사용되는 디바이스로부터 분리될 수 있다.

[0025] 예를 들어, 엔티티는 마이크로폰, 스피커, 및 가능하게는 터치 스크린 디스플레이 상의 키보드와 함께, 휴대 전화와 같은 핸드 헬드 디바이스를 사용할 수 있다. 이러한 핸드 헬드 디바이스는 일반적으로 컴퓨터를 포함한다. 하나의 구현예에서, 핸드 헬드 디바이스 상의 컴퓨터는 입력 프로세싱 모듈(108)을 포함할 수 있다. 다른 구현예에서, 핸드 헬드 디바이스는 대화형 입력(102)을 다른 컴퓨팅 디바이스 상의 입력 프로세싱 모듈에 전달할 수 있으며, 이어서, 다른 컴퓨팅 디바이스는 대화형 응답(106)을 핸드 헬드 디바이스에 제공할 수 있다. 핸드 헬드 디바이스와 다른 컴퓨팅 디바이스 사이의 연결은 공용 컴퓨터 네트워크, 셀룰러 전화 연결 등과 같은 임의의 종류의 데이터 통신 네트워크를 통해 이루어질 수 있다. 핸드 헬드 디바이스와 마찬가지로, 손목밴드, 시계, 이어피스 또는 안경의 형태를 취하는 착용식 컴퓨팅 디바이스가 또한 이와 유사한 구성을 가질 수 있다.

[0026] 다른 예로서, 엔티티는 마이크로폰 및 스피커, 및 가능하게는 다른 입력 및 출력 디바이스들을 가질 수 있는 자동차 또는 다른 차량 내에 있을 수 있다. 차량에는 차량 내의 다양한 디바이스들로부터의 입력을 프로세싱하고 이 디바이스들에게 출력을 제공하는 온보드 컴퓨터가 포함될 수 있다. 하나의 구현예에서, 온보드 컴퓨터는 입력 프로세싱 모듈(108)을 포함할 수 있다. 다른 구현예에서, 온보드 컴퓨터는 셀룰러 전화 연결 또는 위성 통신 연결 등을 통해 온보드 컴퓨터와 무선 통신하는 다른 컴퓨팅 디바이스 상의 입력 프로세싱 모듈에 입력 데이터를 전달할 수 있다. 이어서, 다른 컴퓨팅 디바이스는 대화형 응답(106)을 온보드 컴퓨터에 전송할 수 있다.

[0027] 다른 예로서, 엔티티는 게임 콘솔, 셋탑 박스, 스마트 텔레비전 또는 다른 디스플레이, 및 스마트 폰, 태블릿 컴퓨터 및/또는 원격 제어기와 같은 하나 이상의 핸드 헬드 디바이스들을 비제한적인 예시로서 비롯하여 많은 디바이스들을 포함할 수 있는 홈 엔터테인먼트 시스템을 사용하는 가정 환경 내에 있을 수 있다. 이러한 디바이스들 중 임의의 하나 이상은 이러한 환경 내의 엔티티로부터 대화형 입력을 수신할 수 있는 마이크로폰, 키보드, 또는 터치 스크린 디스플레이와 같은 하나 이상의 입력 디바이스들을 포함할 수 있다. 동일한 환경 내에 있는 이들 디바이스들 중 임의의 하나 이상은 또한 하나 이상의 출력 디바이스들을 포함할 수 있으며, 이를 통해 대화형 응답은 엔티티에게 제공될 수 있다. 하나 이상의 이들 디바이스들 상의 컴퓨터는 입력 프로세싱 모듈(108)을 포함할 수 있다. 이 컴퓨터는 입력 디바이스들 및/또는 출력 디바이스들 중 어느 하나를 포함하는 디바이스와는 상이한 디바이스 내에 있을 수 있다. 다양한 디바이스들 간의 연결은 비제한적인 예시로서, 개인 근거리 네트워크, 무선 통신 연결, 셀룰러 전화 연결 등을 비롯하여, 임의의 종류의 데이터 통신 네트워크를 통해 이루어질 수 있다.

[0028] 이제까지 예시적인 동작 환경들을 설명하여 왔으며, 이제부터는 도 1에서의 입력 프로세싱 모듈(108)의 예시적인 구현예의 보다 상세한 설명을 도 2 및 도 3과 관련하여 설명할 것이다.

[0029] 입력 프로세싱 모듈은 각각 상이한 대화 유형을 위한 복수의 분류기들(200)을 포함한다. 각각의 분류기는 입력

데이터로부터 도출된 특징 데이터(202)를 수신하는 입력부를 갖는다. 각각의 분류기는 분류기의 입력부에 인가된 특징 데이터가 분류기에 대한 대화 유형과 일치하는지 여부를 나타내는 일치 데이터(204)를 제공하는 출력부를 갖는다. 각각의 분류기는, 다른 분류기들의 파라미터들과 구별되며 트레이닝 프로세스에 대한 응답으로 설정되는 파라미터들을 갖는다.

[0030] 의사 결정 트리, 신경망, 벡터 머신 등과 같은 임의의 다양한 유형의 패턴 분류기들이 입력 프로세싱 모듈에서 사용될 수 있다. 일반적으로, 패턴 분류기는, 복수의 특징들에 의해 정의된 입력을 수신하고, 일치를 위해 분류기가 트레이닝되었던 패턴들과 상기 입력이 얼마나 잘 일치하는지를 나타내는 출력을 제공한다. 일반적으로, 출력은 0 내지 1과 같이, 정의된 범위 내의 점수이며, 이 점수는 유사성 또는 거리 메트릭(distance metric) 또는 확률 측정의 형태를 사용하여 계산된다. 유사성 또는 거리 또는 확률 메트릭의 파라미터들과 같은 분류기들의 파라미터들은 트레이닝 프로세스를 통해 설정된다. 하나의 예시적인 구현예에서, 분류기는 로지스틱 회귀 분류(logistic regression classification) 방법을 사용하는 2진 분류기를 사용하여 구현될 수 있다.

[0031] 분류기들에 의해 사용되는 특징들은 일반적으로 표준형(canonical form)이며, 입력은 이러한 표준형으로 변환된다. 대화형 입력의 경우, 이러한 특징들은 대화형 입력 내의 단어 형태, 해당 단어들의 순서, 해당 단어들의 n 그램(n-gram)과 같은 정보, 및 대화형 입력 소스, 이용가능한 임의의 감정적 또는 어조 신호들, 환경 데이터, 관련 키워드 등와 같은 기타 관련 정보를 포함할 수 있다. 예를 들어, 특징들은 존재하는 모든 단어들, 단어들의 n 그램들(예컨대, 또는 2개 또는 3개 단어 시퀀스들, 예를 들어, "how old are you?"의 경우, 예시적인 n 그램은 how_old, old_are, are_you 등임), 단어의 발언(utterance) 길이, 각 단어의 스피치의 일부(예컨대, "old"는 형용사임), 비슷한 단어 세트를 설명하는 임베딩(embedding)(숫자), 입력을 설명하는 구문 분석 트리 등을 포함할 수 있다.

[0032] 선택기(206)는 분류기(200)로부터의 일치 데이터(204)를 수신하도록 연결된 입력부들을 갖는다. 선택기는 분류기들로부터 수신된 일치 데이터(204)에 기초하여, 각각의 대화 유형에 대한 대화형 응답들의 컬렉션(212)으로부터, 대화 응답(208)을 제공하는 출력부를 갖는다. 대화형 응답(208)은 다양한 방식으로 생성될 수 있다.

[0033] 하나의 예시적인 구현예에서, 분류기에 의해 출력된 일치 데이터(204)는 분류기에 의해 계산된 점수이고, 선택기(206)는 최상의 점수를 생성하는 분류기에 대응하는 대화 유형을 선택한다. 다른 예로서, 최상의 점수의 범위 내에서 점수들을 생성하는 모든 분류기들이 고려될 수 있다. 대화 유형들의 계층 구조가 있는 경우에는 여러 분류기들을 선택하는 것이 특히 유용하다. 보다 일반적인 유형의 분류기 및 보다 구체적인 유형의 분류기 둘 다가 적절한 일치 점수들을 생성하는 경우, 보다 구체적인 대화 유형을 갖는 분류기에 기초하여 대화 유형이 선택될 수 있다. 모든 점수들에 임계값을 적용하여 점수들이 대화형 입력의 존재를 나타낼만큼 충분히 높도록 보장할 수 있다. 이어서, 선택기는 선택된 대화 유형에 대해 저장된 응답들 중에서 대화형 응답을 선택하기 위해 대화형 응답들의 컬렉션(212)에 액세스한다. 세트로부터의 대화형 응답의 선택은 랜덤 방식, 의사 랜덤 방식, 준 랜덤 선택, 라운드 로빈 선택, 최저 사용 빈도(least recently used), 응답들의 피드백 및 등급 등과 같은 다양한 방식으로 수행될 수 있다.

[0034] 일반적으로, 각 대화 유형에 대한 대화형 응답들의 컬렉션(212)은 그러한 서비스들에 대한 콘텐츠를 생성하는 경험을 가진 개인들에 의해 생성된다. 아래에서 보다 자세하게 설명되는 바와 같이, 이러한 사람이 작성한 응답들은 분류기들의 입력들에 라벨들을 제공함으로써 분류기들의 트레이닝에 또한 참여하는 개인들에 의해 작성될 수 있다. 대화형 입력에 적용된 라벨들은 대화형 입력이 긍정적 또는 부정적 대화 유형 예시인지 여부를 나타낸다. 컬렉션(212)은 각각의 대화형 응답을 대화 유형과 연관시키는 방식으로 대화형 응답들을 저장한다. 따라서, 컬렉션(212)은 대화 유형이 주어지면 대화형 응답의 액세스 준비 및 선택을 허용하는 임의의 다양한 방식으로 저장될 수 있는 대화형 응답/대화 유형 쌍들을 나타내는 데이터 세트이다. 대화형 응답들은 일반적으로 일관된 저장 포맷을 제공하기 위해 길이가 특정 문자 개수로 제한될 수 있는 문자열을 사용하여 표현된다. 대화 유형은 영숫자 식별자와 같이, 직접적으로 또는 간접적으로 해당 대화 유형의 분류기와 연관될 수 있는 임의의 식별자로 표현될 수 있다.

[0035] 분류기들의 세트에 의해 출력된 일치 데이터(204)에 기초하여 컬렉션(212)으로부터 대화형 응답을 선택하는데 임의의 다양한 다른 기술들이 사용될 수 있다.

[0036] 일부 구현예들에서, 대화 상태(214)를 기술하는 정보가 저장될 수 있고, 이것은 이어서 선택기(206) 또는 분류기(200)에 의해 사용될 수 있다. 대화 상태는 하나 이상의 이전 대화 입력들, 대화형 응답들 및 대화 유형을 기술하는 데이터일 수 있다. 이러한 정보는 선택기(206)에 의해 제공되는 것으로서 예시되어 있지만, 분류기(200) 또는 대화 상태를 추적하도록 특별히 설계된 시스템의 다른 컴포넌트에 의해 제공될 수 있다. 예시된 이

러한 정보는 분류기들 중 하나에 제공되지만, 모든 분류기들, 분류기들의 서브 세트에 제공될 수 있거나, 또는 분류기 어느 것에도 제공되지 않을 수 있다. 일부 구현예들에서, 대화 상태는 분류기들에 인가되는 특징(202)으로서 취급될 수 있다. 현재 프로세싱되는 대화형 입력에 대한 대화 유형의 분류는 대화 상태(214)에 의해 영향을 받을 수 있다. 대안적으로, 또는 조합하여, 현재 프로세싱되는 대화형 입력에 대한 대화 유형 및 대화형 응답의 선택은 대화 상태(214)에 의해 영향을 받을 수 있다. 실용적인 예로서, 대화 입력이 "필라델피아의 오늘 날씨는 어떻습니까?"인 경우, 날씨 관련 대화형 응답이 생성된다. 다음 대화 입력이 "내일은 어떤가요?"이면, 날씨 관련 대화형 응답이 생성되는 것이 보장되도록 대화 상태 정보(214)가 사용될 수 있다. 예를 들어, 대화 유형 계층 구조에 대한 정보를 대화 상태 정보와 결합하여 이러한 선택을 행할 수 있다.

[0037] 도 3의 흐름도는 이러한 예시적인 입력 프로세싱 모듈의 동작을 설명한다.

[0038] 도 3에서, 동작은 대화형 입력을 수신하는 것으로 시작한다(단계 300). 대화형 입력으로부터 도출된 특징들은 복수의 분류기들의 입력부들에 인가된다(단계 302). 일치 데이터가 복수의 분류기들로부터 수신되고(304), 이 데이터는 대화 유형을 선택하는데 사용된다(단계 306). 그 후, 선택된 대화 유형에 기초하여 대화형 응답이 출력된다(단계 308).

[0039] 이제까지 입력 프로세싱 모듈을 설명하여 왔으며, 이제부터는 분류기들을 위한 트레이닝 프로세스의 예시적인 구현예를 보다 상세하게 설명할 것이다.

[0040] 하나의 구현예에서, 유한적 세트의 알려진 대화 유형들 및 대응하는 예시들로 시작하는 종래의 분류기를 위한 종래의 트레이닝 프로세스가 사용된다. 랜덤 샘플링 또는 일부 다른 접근법을 사용하여, 예시들이 분류기에 인가되고, 분류기의 파라미터들은 분류기의 출력과 분류기에 인가된 예시들 간의 오류들에 기초하여 조정된다. 분류기는 주어진 대화 유형에 대해 트레이닝된 후, 대화형 응답들을 생성하는데 사용될 수 있다.

[0041] 하나의 구현예에서, 분류기는 활성 트레이닝 프로세스를 사용하여 트레이닝될 수 있는 2진 분류기일 수 있다. 이러한 프로세스에서, 트레이닝은 세트, 예를 들어 10개의 대화형 입력의 긍정적 예시들과 세트, 예를 들어, 10개의 부정적 예시들을 선택함으로써 시작된다(단계 400). 긍정적 예시는 분류기의 대화 유형과 일치하는 대화형 입력이다. 예를 들어, 분류기를 위한 대화 유형이 "장소의 오늘 날씨에 관한 질문"인 경우, "오늘 시애틀의 날씨는 어떻습니까?"는 긍정적 예시이며, "당신은 몇살입니까?"는 부정적 예시이다. 2진 분류기를 사용하는 하나의 구현예에서, "0" 또는 "1"의 라벨이 예시적인 발언문들에 적용되는데, 여기서 "1"은 분류기가 이 발언문을 발화해야 함을 나타내며, "0"은 분류기가 발화하지 않아야 함을 나타낸다.

[0042] 분류기는 종래 기술을 사용하여 긍정적 및 부정적 예시들의 세트를 사용하는 분류기의 초기 파라미터들을 설정함으로써 초기화된다(단계 402). 초기화된 분류기가 주어지면, 분류기에 대화형 입력들(각각을 트레이닝 입력이라고 칭함)의 트레이닝 세트를 인가함으로써 트레이닝이 수행된다(단계 404). 트레이닝 세트는 서비스와 관련하여 저장된 실제 대화형 입력들과 같은, 다양한 소스들로부터 수집될 수 있다. 이 예시에서, 트레이닝 입력들은 실제 대화형 입력들로부터 도출된 특징들이다. 각각의 트레이닝 입력은 분류기로부터의 점수를 초래시킨다. 결과가 명백하게 긍정적 또는 부정적 예시인 경우, 즉 점수 범위의 어느 극값의 임계량 내에 있는 경우, 트레이닝 입력을 자동으로 표시할 수 있지만; 이러한 자동 라벨링은 오류를 생성하고 이 오류를 증강시킬 수 있어서 바람직하지 않다. 그렇지 않고, 긍정적 일치도 아니고 부정적 일치도 아닌 모호한 결과들을 갖는 임의의 트레이닝 입력들이 이 분류기에 대한 대화 유형의 긍정적 또는 부정적 예시로서 라벨링을 위해 개인에게 제시된다(단계 406). 그 후, 분류기는 이들 추가적인 트레이닝 입력들의 라벨링을 사용하여 업데이트된다(단계 408). 그 후, 본 프로세스는 단계 404로의 화살표에 의해 표시된 바와 같이 1회 이상 반복될 수 있다.

[0043] 도 5는 도 4에서 설명된 동작에서 사용되는 트레이닝 프로세싱 모듈의 예시적인 구현의 데이터 흐름도이다. 도 5에서, 대화형 입력들의 컬렉션(500)이 컴퓨터로 판독가능한 저장장치 내에 저장된다. 각각의 저장된 입력을 트레이닝 입력이라고 칭한다. 트레이닝 입력은 라벨 및 대응하는 대화 유형과 연관될 수 있다. 여러 트레이닝 유형들을 사용하여, 트레이닝 입력은 여러 라벨들과 연관될 수 있다. 트레이닝 모듈(502)은 대화형 입력들로부터 도출된 특징들(504) 및 라벨들(506)을 수신한다. 이것들은 대화 유형에 대한 분류기(510)에 대한 파라미터(508)들을 설정하는데 사용된다.

[0044] 일반적으로 라벨들은, 트레이닝 입력들을 검토하고, 이 입력들을 대화 유형의 긍정적 또는 부정적 예시로서 라벨링하는 개인(512)들에 의해 입력된다. 개인(512)에게는 사용자 디바이스 상의 사용자 인터페이스(516)를 통해 트레이닝 입력 및 대화 유형의 표시(참조번호 514에서 도시됨)가 제시된다. 개인은 사용자 인터페이스(516)를 통해 해당 트레이닝 입력에 대한 라벨(518)을 제공한다.

[0045] 개인(512)은 또한 사용자 인터페이스(516) 또는 다른 사용자 인터페이스를 통해서든, 동일하거나 상이한 사용자 디바이스 상에서든, 대화 유형에 대한 대화형 응답(520)들을 제공할 수 있다. 이들 대화형 응답들은 도 2에서 도시된 바와 같이, 입력 프로세싱 모듈 내에 저장된다.

[0046] 동일한 개인들이 대화 응답들과 대화형 입력들의 라벨들 둘 다를 준비하게 함으로써, 본 시스템은 주어진 입력들에 적절한 응답들을 가질 확률을 높일 수 있다. 이러한 능력은 라벨링된 입력들을 사용하여 트레이닝된 대화 유형에 의한 분류기들과, 대화 유형에 의해 선택된 대화형 응답들을 가짐으로써 제공된다. 달리 말하면, 대화형 응답들을 생성하는 개인들은 대화형 입력들에 대한 라벨들을 제공하는 개인과 동일하기 때문에 컴퓨터는 대화형 입력들에 대한 적절한 응답들을 제공한다. 라벨링된 대화형 입력들의 긍정적 및 부정적 예시들에 의해 트레이닝된 대화 유형별 분류기들을 사용함으로써, 콘텐츠 편집자는 라벨러와 편집자 둘 다일 수 있다.

[0047] 대화형 입력들의 대규모 코퍼스(corpus)를 사용하여, 컴퓨터는 다양한 형태의 클러스터링 알고리즘들을 사용하여 가능한 대화 유형들을 식별할 수 있다. 트레이닝을 위한 샘플들은 발언문들이 배치된 대화 유형들에 기초하여 선택될 수 있으며, 새로운 대화 유형들이 또한 식별될 수 있다. 클러스터링 알고리즘은 유사한 대화형 입력들을 세트로 그룹화하여, 유사한 대화형 입력들을 갖는 각각의 세트를 갖는 여러가지 상이한 세트들을 초래시킬 수 있다. 한 세트로부터의 샘플들은 해당 세트에 대한 대응하는 대화 유형에 대한 긍정적 예시들로서 사용될 수 있으며, 다른 세트들로부터의 샘플들은 부정적 예시들로서 사용될 수 있다.

[0048] 하나의 예시적인 구현예로서, 자동 클러스터링은, 예를 들어, 백 오브 워드 벡터(bag-of-word vector)를 사용하여 발언문들의 벡터 공간 표현에 k 평균(k-mean)을 적용하거나, 또는 LSA(latent semantic analysis) 또는 LDA(latent dirichlet allocation)을 사용하여 발언문들의 투영들에 k 평균을 적용하는 것을 포함한다. 다른 예시적인 구현예에서, 기존의 분류기가 사용될 수 있고, 분류기가 발화하는 발언문들, 즉, 알려진 대화 유형의 발언문들은 제외되며, 분류기로 하여금 발화하게 하지 않는 나머지 발언문은 새로운 대화 유형과 연관될 수 있다.

[0049] 구현예에서, 분류기들은 또한 소셜 대화를 위한 단어들 및 구문들의 클래스들을 포함하도록 확장될 수 있다. 위에서 설명한 바와 같이, 대화 유형들은 계층 구조로 그룹화될 수 있다. 이러한 계층 구조는 대화형 입력들과 연관된 키워드 또는 대화형 입력들의 세트들로부터 도출된 다른 클래스 정보에 기초할 수 있다. 이러한 클래스들은, 예를 들어, 자동 클러스터링을 사용하여 자동으로 식별될 수 있다. 특정 예시로서, 자동 클러스터링은 단어를 연속 공간 내에 투영하고 그 연속 공간에서 클러스터링함으로써 수행될 수 있다. 예를 들어, 신경망으로 학습된 LSA, LDA, 또는 단어 임베딩을 사용하여 투영을 수행한 다음, k 평균을 사용하여 클러스터링을 수행할 수 있다. 클래스들을 도입하기 위한 다른 방법은 콘텐츠 편집자가 예시적인 단어들/구문들을 제공하고, 클래스를 자동으로 확장하는 것이다. 예를 들어, 콘텐츠 편집자가 "출생하다", "성장하다", "어린 시절"을 제공한 경우, "~에서 자라났다", "나는 ~ 출신이다", "~학교에 다녔다"와 같은 더 많은 용어들이 자동으로 추가될 수 있다. (LSA, LDA, 또는 단어 임베딩과 같은) 클러스터링 기술들의 반 지도형(semi-supervised) 변형들이 또한 사용될 수 있다.

[0050] 하나의 양태에서, 대화형 입력은 대화형 입력으로부터 도출된 특징들을 복수의 분류기들에 인가하는 것에 기초하여 대화형 응답을 선택하기 위한 수단을 사용하여 프로세싱된다. 각각의 분류기는 대화 유형과 연관된다. 대화형 응답들이 또한 대화 유형들과 연관된다.

[0051] 다른 양태에서, 대화형 입력은 메모리 내로의 대화형 입력을 나타내는 입력 데이터를 수신함으로써 프로세싱될 수 있다. 입력 데이터로부터 도출된 특징 데이터는 복수의 분류기들에 인가될 수 있으며, 각각의 분류기는 복수의 대화 유형들 중의 대화 유형을 나타낸다. 분류기들은 입력된 특징 데이터에 응답하여 일치 데이터를 출력한다. 대화형 응답은 분류기들에 의해 출력된 일치 데이터에 따라 복수의 대화 유형들 중 하나 이상의 대화 유형들에 따라 선택된다.

[0052] 이 섹션에서의 앞서 설명한 임의의 양태들에서, 특징들은 대화형 입력으로부터 생성된 n 그램들을 포함할 수 있다. 이러한 특징들은 대화형 입력과 관련된 감정을 나타내는 데이터를 포함할 수 있다. 이러한 특징들은 대화형 입력의 주제에 대한 하나 이상의 클래스들을 나타내는 데이터를 포함할 수 있다.

[0053] 이 섹션에서의 앞서 설명한 임의의 양태들에서, 특징 데이터가 2진 분류기와 연관된 대화 유형의 긍정적 예시인 것으로서 라벨링된 대화형 입력과 일치할 확률을, 특징 데이터에 응답하여 출력되는 일치 데이터가 나타내도록, 각각의 분류기는 2진 분류기일 수 있다.

[0054] 이 섹션에서의 앞서 설명한 임의의 양태들에서, 특징 데이터는 대화형 입력 및 대화형 입력과 연관된 다른 정보

로부터 도출될 수 있다.

[0055] 이 섹션에서의 앞서 설명한 임의의 양태들에서, 대화형 응답은 복수의 대화 유형들에 대한 복수의 대화형 응답들을 저장하는 컬렉션으로부터 선택될 수 있다.

[0056] 이 섹션에서의 앞서 설명한 임의의 양태들에서, 분류기들에 의해 출력된 일치 데이터에 따라 하나 이상의 대화 유형들이 복수의 대화 유형들 중에서 선택될 수 있다. 대화형 응답은 선택된 하나 이상의 대화 유형들에 따라 선택될 수 있다. 하나 이상의 대화 유형들은 분류기들에 의해 출력된 일치 데이터 중에서의 점수들에 기초하여 선택될 수 있다.

[0057] 이 섹션에서의 앞서 설명한 임의의 양태들에서, 각 분류기는 점수를 출력할 수 있다. 점수는 임의의 값 범위 내에 있을 수 있으며, 어떤 값들은 분류기에 의해 표현된 대화 유형과 특징들의 일치를 나타내며, 어떤 값들은 일치가 없음을 나타낸다. 일치를 나타내는 값들은 범위 내 최대값과 같이, 범위의 극값일 수 있다. 점수는, 이와 달리, 대화 유형과 일치할 확률을 나타낼 수 있다.

[0058] 이 섹션에서의 앞서 설명한 임의의 양태들에서, 대화 상태 정보가 유지될 수 있다. 대화형 응답은 또한 유지된 대화 상태 정보에 기초해서 선택될 수 있다. 대화형 응답은 또한 유지된 상태 정보 및 대화 유형 계층 구조 정보에 기초하여 선택될 수 있다.

[0059] 이 섹션에서의 앞서 설명한 임의의 양태들에서, 대화형 입력의 엔티티로부터의 입력 및 대화형 응답의 해당 엔티티로의 출력은, 대화형 입력으로부터 특징들을 생성하고 이들 특징들을 분류기들에 인가하며 대화형 응답을 선택하는 다른 디바이스들 및 컴퓨터들로부터 분리된 하나 이상의 디바이스들 상에서 수행될 수 있다. 그러한 경우, 입력들을 수신하는 디바이스는 대화형 입력들에 관한 정보를 전송하기 위해 하나 이상의 컴퓨터들에 연결된다. 마찬가지로, 대화형 응답들에 관한 정보를 갖는 전송들을 수신하는 디바이스는 그러한 전송들을 수신하기 위해 하나 이상의 컴퓨터에 연결된다. 이러한 각각의 디바이스 또는 컴퓨터는 메모리, 입력 데이터가 수신되도록 해주는 입력부, 입력 데이터를 프로세싱하도록 구성된 프로세서, 및 출력 데이터가 다른 디바이스 또는 컴퓨터에 제공될 수 있도록 해주는 출력부를 포함한다.

[0060] 이 섹션에서의 앞서 설명한 임의의 양태들에서, 대화형 입력을 제공하고 대화형 응답을 수신하기 위해 엔티티에 의해 사용되는 디바이스 또는 디바이스들은, 비제한적인 예시로서, 이동 전화기, 태블릿 컴퓨터, 핸드 헬드 컴퓨팅 디바이스, 오토모바일 기반 컴퓨터, 랩탑 컴퓨터, 휴대용 컴퓨터, 서버 컴퓨터, 노트북 컴퓨터, 슬레이트 컴퓨터, 데스크탑 컴퓨터, 홈 엔터테인먼트 시스템, 게임 콘솔, 케이블 위성 또는 다른 텔레비전 수신을 위한 셋톱 박스, 텔레비전, 스마트 디스플레이 디바이스 등을 비롯한, 복수의 디바이스들 중 임의의 디바이스일 수 있다.

[0061] 전술한 양태들 중 어느 하나에서, 분류기는 대화형 입력들을 대화 유형으로 분류하도록 트레이닝될 수 있다. 이러한 트레이닝 프로세스는 제1 개수의 대화형 입력들을 대화 유형의 긍정적 예시로서 라벨링하고, 제2 개수의 대화형 입력들을 대화 유형의 부정적 예시로서 라벨링함으로써 구현될 수 있다. 분류기는 라벨링된 대화형 입력들을 사용하여 구축된다. 대화형 입력들의 코퍼스가 액세스된다. 코퍼스로부터의 대화형 입력들은 분류기로부터의 각각의 대화형 입력에 대한 출력을 획득하기 위해 분류기에 인가된다. 추가적인 대화형 입력들은 대화형 입력들에 대한 분류기로부터의 출력들에 기초한 라벨링을 위해 하나 이상의 개인들에게 제시된다. 제시된 추가적인 대화형 입력들에 대해 하나 이상의 개인들로부터 라벨들이 수신된다. 분류기는 추가적인 라벨링된 대화형 입력들을 사용하여 재트레이닝된다.

[0062] 분류기들의 트레이닝을 수반하는 본 섹션의 전술한 양태들 중 어느 하나의 양태에서, 분류기들은 재트레이닝된 분류기를 사용하여, 인가 단계, 제시 단계, 수신 단계, 및 재트레이닝 단계를 반복하는 프로세스를 사용하여 트레이닝될 수 있다.

[0063] 전술한 양태들 중 어느 하나에서, 하나 이상의 개인들이 라벨링을 수행했던 분류기와 연관된 대화 유형에 대한 대화형 응답들이 라벨링을 수행했던 개인들로부터 수신될 수 있다.

[0064] 전술한 양태들 중 어느 하나에서, 추가적인 대화 유형들이 대화형 입력들의 코퍼스로부터 식별될 수 있다. 예를 들어, 코퍼스 내의 대화형 입력들을 클러스터링하고, 대화 유형을 클러스터들에 할당할 수 있다. 다른 예로서, 대화형 입력들은 복수의 분류기들에 인가될 수 있다. 복수의 분류기들 중 어느 것도 일치를 나타내지 않는 대화형 입력들이 식별된다. 식별된 대화 입력들에 하나 이상의 새로운 대화 유형들이 할당된다.

[0065] 전술한 양태들 중 어느 하나에서, 라벨링을 위한 대화형 입력들을 제시하는 것은 분류기로의 추가적인 대화형

입력이 긍정적 일치도 아니고 부정적 일치도 아님을 분류기의 출력이 나타내는지 여부에 따라 제시될 추가적인 대화형 입력들을 선택하는 것을 포함할 수 있다.

[0066] 전술한 양태들 중 어느 하나에서, 분류기의 대화 유형은 대화 유형들의 계층 구조 내에 포함될 수 있다. 임계값 위의 점수를 갖는 분류기들에 기초하여 대화 유형이 선택될 수 있으며, 보다 구체적인 대화 유형이 더 일반적인 대화 유형에 우선해서 선택된다. 대화형 입력들의 코퍼스를 자동으로 클러스터링하면 대화형 입력들에 대한 클래스 정보를 제공할 수 있다.

[0067] 이 섹션에서의 앞서 설명한 임의의 양태들은 하나 이상의 컴퓨터 시스템들 내에서, 그러한 컴퓨터 시스템의 임의의 개별 컴포넌트 또는 컴포넌트들로서, 그러한 컴퓨터 시스템들 중 하나 이상에 의해 수행되는 프로세스 또는 그러한 컴퓨터 시스템의 임의의 개별 컴포넌트에 의해 수행되는 프로세스, 또는 하나 이상의 컴퓨터들에 의해 프로세싱될 때, 이들 컴퓨터들이 그러한 컴퓨터 시스템 또는 그러한 컴퓨터 시스템의 임의의 개별 컴포넌트를 제공하도록 이들 컴퓨터들을 구성시키는 컴퓨터 프로그램 명령어들이 저장되어 있는 컴퓨터 저장장치를 포함하는 하나 이상의 제조 물품들로서 구현될 수 있다.

[0068] 도 1 내지 도 5와 관련하여 설명된 예시적인 구현예에서는, 두 개의 애플리케이션들이 존재할 수 있는데, 그 하나는 분류기들을 구축하기 위한 것이며(예를 들어, 트레이닝 프로세싱 모듈), 다른 하나는 대화 입력들에 응답하기 위해 트레이닝된 분류기를 사용하는 것이다(예를 들어, 입력 프로세싱 모듈). 이러한 각각의 애플리케이션들은 해당 애플리케이션의 상이한 컴포넌트들을 구현하는 여러 애플리케이션들로 구성될 수 있다. 각각의 애플리케이션은 별개의 플랫폼(예컨대, 컴퓨터 상에서 실행되는 애플리케이션들에 의해 컴퓨터의 자원들에 대한 액세스를 관리하는 운영 체제가 있는 컴퓨터) 상에서 실행되도록 설계될 수 있다. 컴퓨터는 아래의 도 6과 관련하여 설명된 대로 구현된다.

[0069] 도 6은 도 1 내지 도 5의 시스템의 다양한 컴포넌트들이 구현될 수 있는 예시적인 컴퓨터를 예시한다. 컴퓨터는 임의의 다양한 범용 또는 특수 목적 컴퓨팅 하드웨어 구성들일 수 있다. 사용될 수 있는 컴퓨터 유형들의 몇가지 예시들은, 비제한적인 예시로서, 개인용 컴퓨터, 게임 콘솔, 셋톱 박스, 핸드 헬드 또는 랩톱 디바이스들(예컨대, 미디어 플레이어, 노트북 컴퓨터, 태블릿 컴퓨터, 휴대폰, 개인 정보 단말기, 음성 녹음기), 서버 컴퓨터, 멀티프로세서 시스템, 마이크로프로세서 기반 시스템, 프로그램가능한 가전 제품, 네트워크 PC, 미니 컴퓨터, 메인 프레임 컴퓨터, 및 상기 유형들의 컴퓨터 또는 디바이스 중 임의의 것을 포함하는 분산 컴퓨팅 환경 등을 포함한다.

[0070] 도 6을 참조하면, 예시적인 컴퓨터(600)는 적어도 하나의 프로세싱 유닛(602) 및 메모리(604)를 포함한다. 컴퓨터는 다중 프로세싱 유닛(602)을 가질 수 있다. 프로세싱 유닛(602)은 서로 독립적으로 동작하는 하나 이상의 프로세싱 코어들(미도시됨)을 포함할 수 있다. 그래픽 프로세싱 유닛(620)과 같은, 추가적인 코프로세싱 유닛이 또한 컴퓨터 내에 존재할 수 있다. 메모리(604)는 (동적 랜덤 액세스 메모리(DRAM) 또는 다른 랜덤 액세스 메모리 디바이스와 같은) 휘발성, (판독 전용 메모리, 플래시 메모리 등과 같은) 비휘발성, 또는 이들의 일부 조합일 수 있다. 이 메모리 구성은 점선(606)에 의해 도 6에서 도시된다. 컴퓨터(600)는 비제한적인 예시로서, 자기적 기록 또는 광학적 기록 디스크 또는 테이프를 비롯하여, 추가적인 저장장치(착탈식/비착탈식)를 포함할 수 있다. 이러한 추가적인 저장장치는 도 6에서 착탈식 저장장치(608)와 비착탈식 저장장치(610)로 나타난다. 도 6에서의 다양한 컴포넌트들은, 일반적으로, 하나 이상의 버스들(630)과 같은 상호연결 메커니즘에 의해 상호연결된다.

[0071] 컴퓨터 저장 매체는 컴퓨터에 의해 어드레스가능한 물리적 저장 위치들에서 데이터가 저장되고 이 위치들로부터 데이터가 검색될 수 있는 임의의 매체이다. 컴퓨터 저장 매체는 휘발성 및 비휘발성 메모리, 및 착탈식 및 비착탈식 저장 매체를 포함한다. 메모리(604, 606), 착탈식 저장장치(608), 및 비착탈식 저장장치(610)는 모두 컴퓨터 저장 매체의 예시들이다. 컴퓨터 저장 매체의 일부 예시들은, RAM, ROM, EEPROM, 플래쉬 메모리 또는 다른 메모리 기술, CD-ROM, 디지털 다용도 디스크(DVD) 또는 다른 광학적 또는 자기광학적 기록 저장 디바이스, 자기 카세트, 자기 테이프, 자기 디스크 저장장치 또는 다른 자기 저장장치 디바이스들이다. 컴퓨터 저장 매체 및 통신 매체는 상호 배타적인 매체 카테고리들이다.

[0072] 컴퓨터(600)는 또한 컴퓨터가 통신 매체를 통해 다른 디바이스들과 통신할 수 있게 해주는 통신 연결부(들)(612)을 포함할 수 있다. 통신 매체는 일반적으로, 유선 또는 무선 물질을 통해 반송파 또는 다른 전송 메커니즘과 같은 변조된 데이터 신호를 전파함으로써 유선 또는 무선 물질을 통해 컴퓨터 프로그램 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터를 전송한다. "변조된 데이터 신호"의 용어는 신호 내 정보를 인코딩하여 신호의 수신 디바이스의 구성 또는 상태를 변경하는 방식으로 하나 이상의 신호 특성이 설정 또는 변경된

신호를 의미한다. 비제한적인 예시로서, 통신 매체는 유선 네트워크 또는 직접적 유선 연결부와 같은 유선 매체, 및 음향, RF, 적외선, 및 이와 다른 무선 매체와 같은 무선 매체를 포함한다. 통신 연결부(612)는 통신 매체와 인터페이스하여 통신 매체를 통해 데이터를 송신하고 통신 매체로부터 데이터를 수신하는, 네트워크 인터페이스 또는 무선 송신기와 같은 디바이스들이다.

[0073] 컴퓨터(600)는 키보드, 마우스, 펜, 카메라, 터치 입력 디바이스 등과 같은 다양한 입력 디바이스(들)(614)을 가질 수 있다. 디스플레이, 스피커, 프린터 등과 같은 출력 디바이스(들)(616)이 또한 포함될 수 있다. 이러한 모든 디바이스들은 본 업계에서 잘 알려진 것이며 여기서는 길게 논의될 필요가 없다. 입력 및 출력 디바이스들은 도 6에서의 컴퓨터의 다양한 컴포넌트들을 포함하는 하우징의 일부일 수 있거나, 또는 그 하우징으로부터 분리될 수 있고 직렬 버스, 무선 통신 연결부 등과 같은 다양한 연결 인터페이스들을 통해 컴퓨터에 연결될 수 있다. 다양한 입력 및 출력 디바이스들은 내추럴 사용자 인터페이스(natural user interface; NUI)로 구현될 수 있으며, 이 NUI는 마우스, 키보드, 원격 제어 등과 같은 입력 디바이스들에 부과된 인공적인 제약들로부터 벗어나서, 사용자가 디바이스와 "내추럴" 방식으로 상호작용하게 해주는 임의의 인터페이스 기술이다.

[0074] NUI 방법들의 예시들에는 음성 인식, 터치 및 스타일러스 인식, (스크린 상과 스크린에 인접한 곳에서의) 제스처 인식, 에어 제스처, 머리와 안구 추적, 구두 및 음성, 시각, 터치, 제스처들, 및 머신 지능에 의존한 것들이 포함되며, 터치 민감 디스플레이, 음성 및 스피치 인식, 의도 및 목표 이해, (입체 카메라 시스템, 적외선 카메라 시스템, 다른 카메라 시스템, 및 이들의 조합들과 같은) 깊이 카메라들을 이용한 모션 제스처 검출, 가속도계 또는 자이로스코프를 이용한 모션 제스처 검출, 안면 인식, 삼차원 디스플레이, 머리, 안구, 및 시선 추적, 몰입형 증강 현실 및 가상 현실 시스템(이 모든 것들은 더 내추럴한 인터페이스를 제공한다) 뿐만이 아니라, 전기장 감지 전극들(EEG 및 관련 방법들)을 이용한 두뇌 활동을 감지하기 위한 기술들의 이용을 포함할 수 있다.

[0075] 다양한 저장장치(610), 통신 연결부(612), 출력 디바이스(616) 및 입력 디바이스(614)는 하우징 내에서 컴퓨터의 나머지와 통합될 수 있거나, 또는 컴퓨터 상의 입력/출력 인터페이스 디바이스들을 통해 연결될 수 있으며, 이 경우 참조번호들(810, 812, 814, 816)은 경우에 따라 디바이스로의 연결을 위한 인터페이스 또는 디바이스 자체를 나타낼 수 있다.

[0076] 상기 도 1 내지 도 5에서 설명한 바와 같은 시스템의 각각의 컴포넌트(이것은 "모듈" 또는 "엔진" 등으로 칭해질 수도 있음)는 컴퓨터 상에서 동작하며, 하나 이상의 컴퓨터들의 하나 이상의 프로세싱 유닛들 및 하나 이상의 프로세싱 유닛들에 의해 프로세싱되는 하나 이상의 컴퓨터 프로그램들을 사용하여 구현될 수 있다. 컴퓨터 프로그램은 프로그램 모듈들과 같은, 컴퓨터로 실행가능한 명령어들 및/또는 컴퓨터로 해석되는 명령어들을 포함하며, 이 명령어들은 이상의 컴퓨터들 내의 하나 이상의 프로세싱 유닛들에 의해 프로세싱된다. 일반적으로, 이러한 명령어들은, 프로세싱 유닛에 의해 프로세싱될 때, 프로세싱 유닛으로 하여금 데이터에 대한 연산을 수행하게 하거나 또는 프로세서 또는 컴퓨터로 하여금 다양한 컴포넌트들 또는 데이터 구조들을 구현하게 하도록 명령하는 루틴, 프로그램, 오브젝트, 컴포넌트, 데이터 구조 등을 정의한다. 이러한 컴포넌트들은 저장장치 또는 메모리 내의 데이터에 액세스하고 저장장치 또는 메모리 내에 데이터를 저장함으로써 입력 및 출력을 갖는다.

[0077] 이 컴퓨터 시스템은 통신 네트워크를 통해 링크된 여러 컴퓨터들에 의해 연산들이 수행되는 분산형 컴퓨팅 환경들에서 실시될 수 있다. 분산형 컴퓨팅 환경에서, 컴퓨터 프로그램들은 로컬 및 원격 컴퓨터 저장 매체 둘 다에 위치할 수 있다.

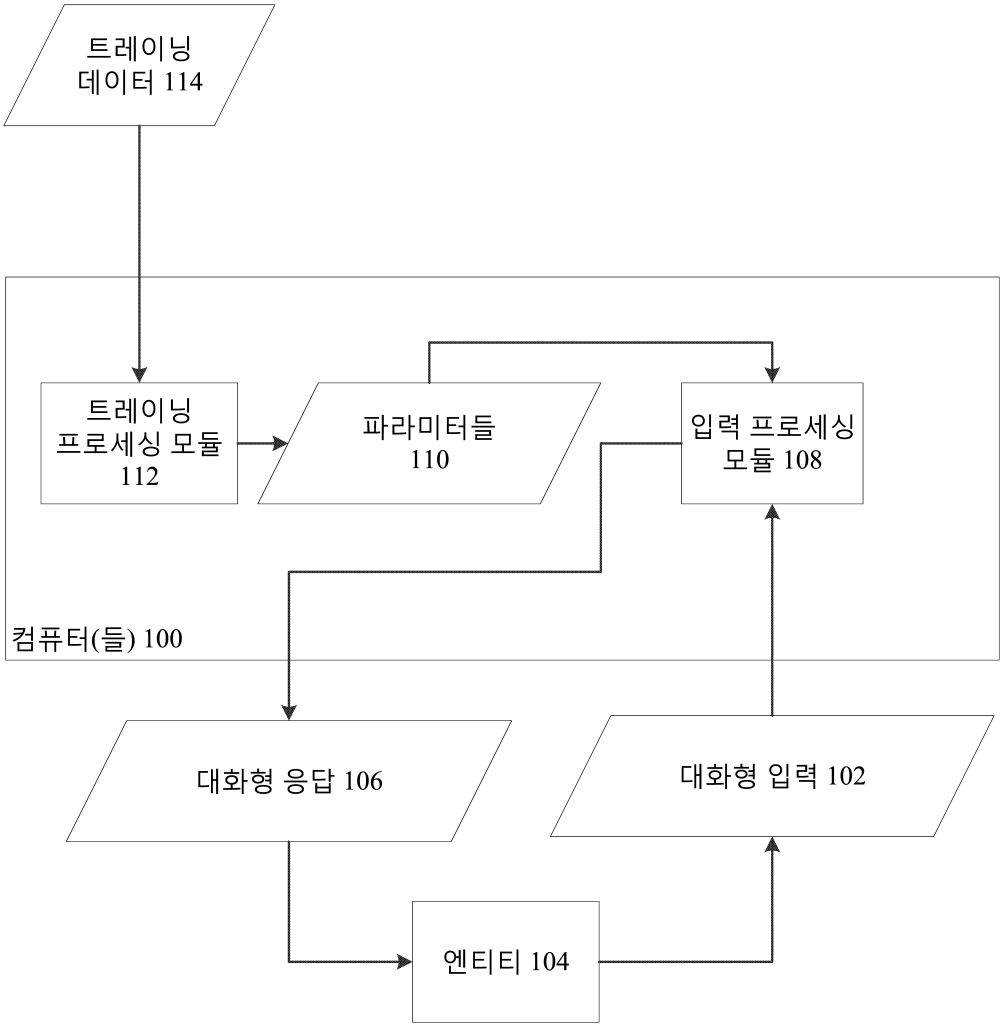
[0078] 대안적으로, 또는 추가적으로, 여기서 설명된 다양한 하나 이상의 컴포넌트들의 기능은 적어도 부분적으로, 하나 이상의 하드웨어 로직 컴포넌트들에 의해 수행될 수 있다. 예를 들어, 비제한적인 예시로서, 이용될 수 있는 예시적인 유형들의 하드웨어 로직 컴포넌트들은 FPGA(Field-programmable Gate Array), ASIC(Program-specific Integrated Circuit), ASSP(Program-specific Standard Product), SOC(System-on-a-chip system), CPLD(Complex Programmable Logic Device) 등을 포함한다.

[0079] 첨부된 청구범위의 전제부에서의 "제조 물품", "프로세스", "머신" 및 "물질 조성"의 용어들은 청구범위를, 35 U.S.C. §101에서의 이들 용어들의 사용에 의해 정의된 특허가능한 발명내용의 범위에 속하는 것으로 간주되는 발명내용으로 제한되도록 의도된 것이다.

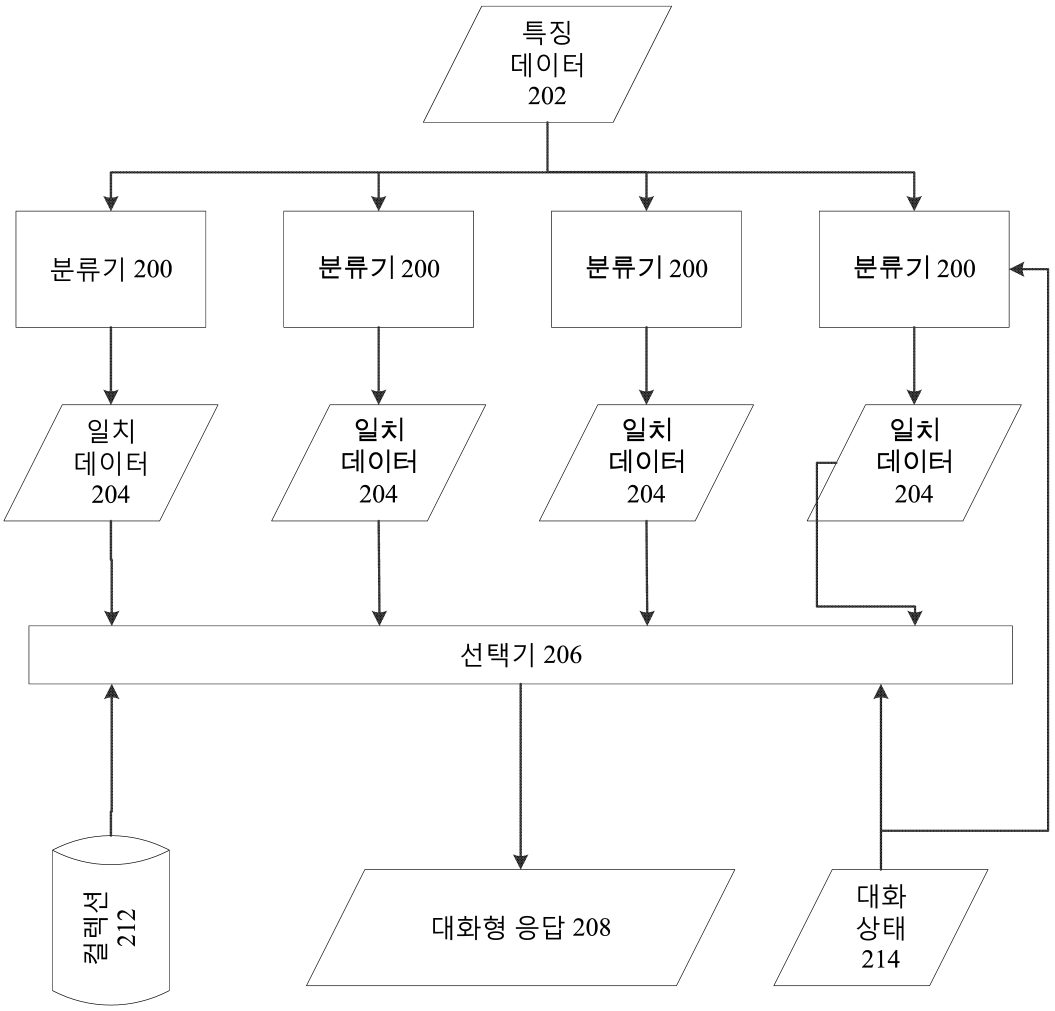
[0080] 첨부된 청구범위에서 정의된 발명내용은 상술한 특정 구현예로 반드시 제한될 필요는 없다는 것을 이해해야 한다. 상술한 특정 구현예는 단지 예시로서 개시된 것일 뿐이다.

도면

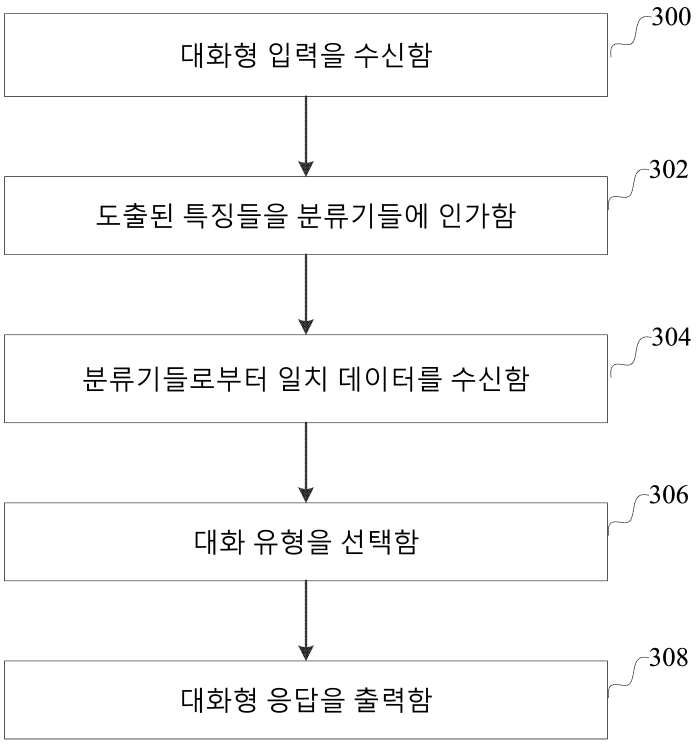
도면1


트레이닝
데이터 114
트레이닝
프로세싱 모듈
112
파라미터들
110
입력 프로세싱
모듈 108
컴퓨터(들) 100
대화형 응답 106
대화형 입력 102
엔티티 104

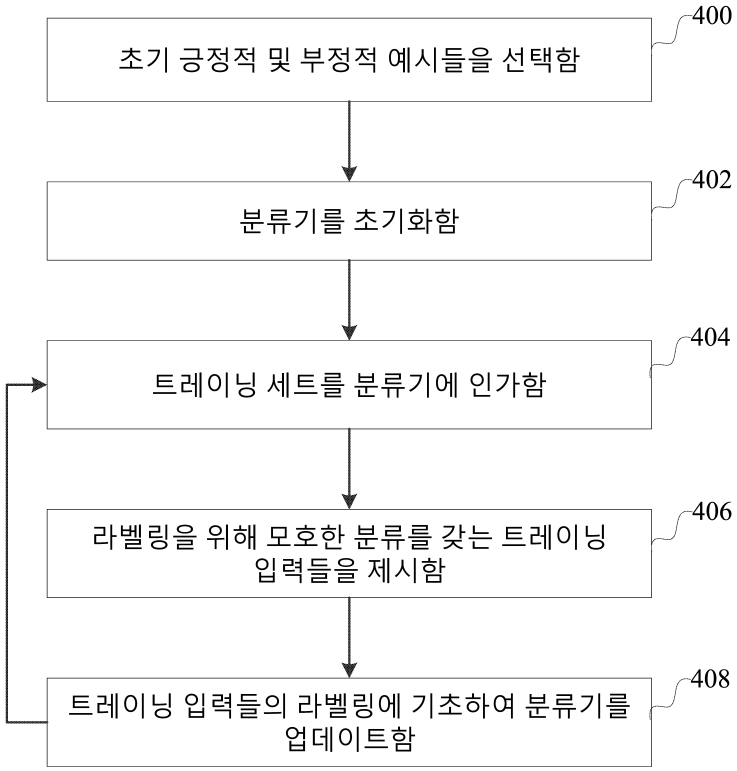
도면2


특징
데이터
202
분류기 200
분류기 200
분류기 200
분류기 200
일치
데이터
204
일치
데이터
204
일치
데이터
204
일치
데이터
204
선택기 206
컬렉션
212
대화형 응답 208
대화
상태
214

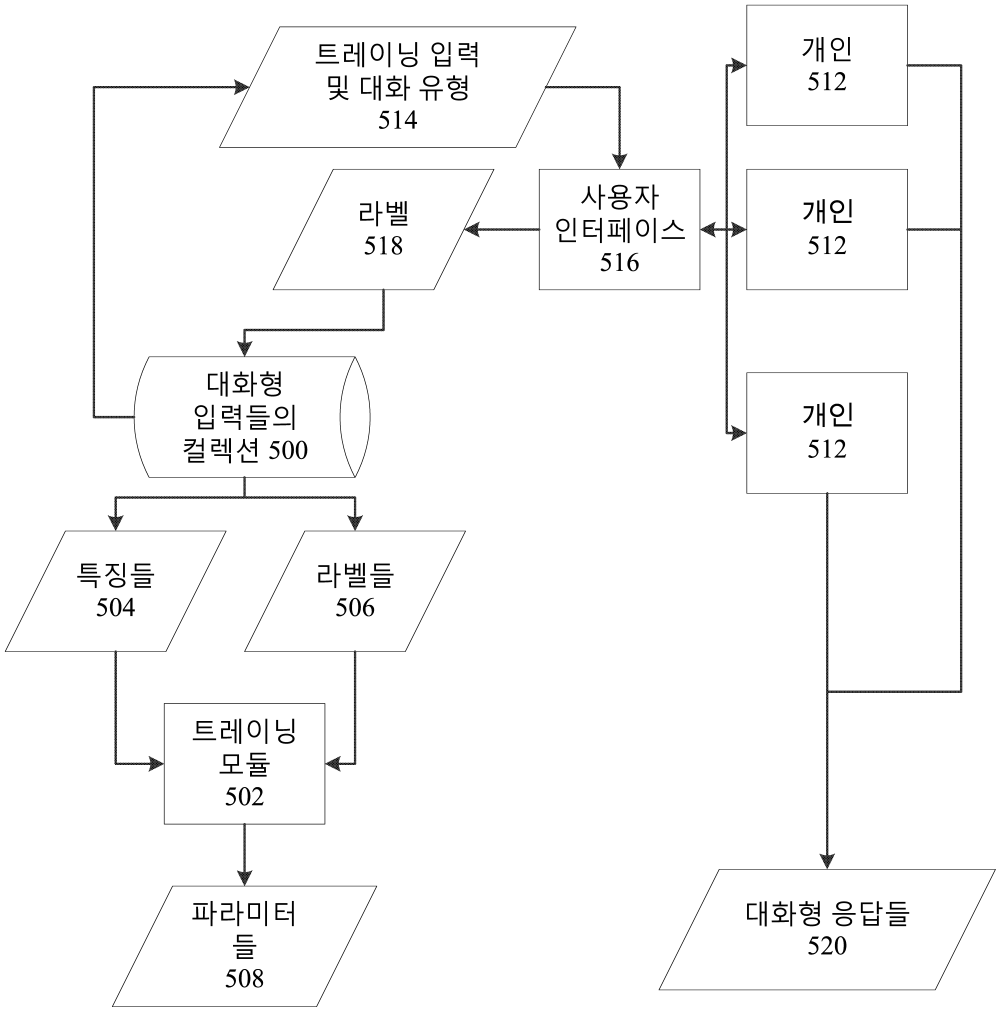
도면3


대화형 입력을 수신함
300
도출된 특징들을 분류기들에 인가함
302
분류기들로부터 일치 데이터를 수신함
304
대화 유형을 선택함
306
대화형 응답을 출력함
308

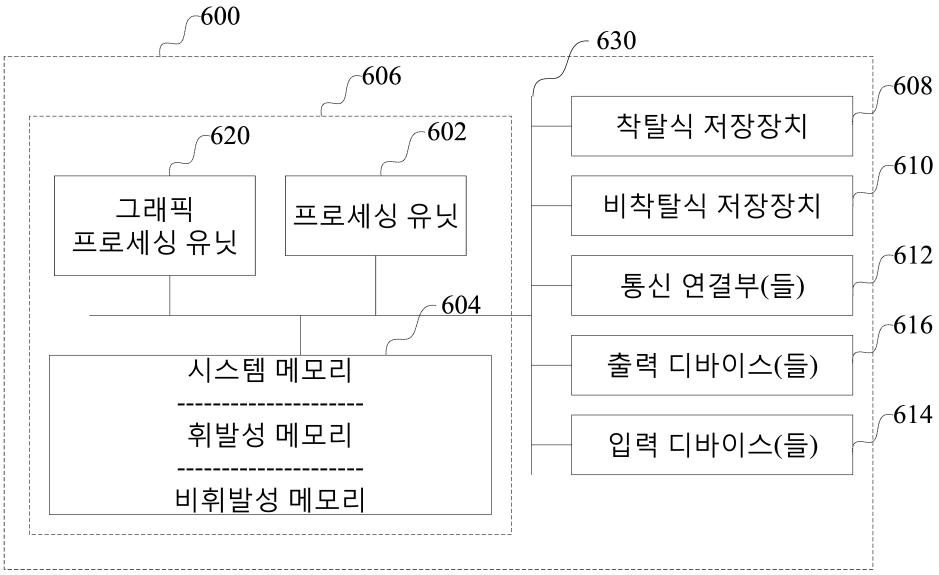
도면4


초기 긍정적 및 부정적 예시들을 선택함
400
분류기를 초기화함
402
트레이닝 세트를 분류기에 인가함
404
라벨링을 위해 모호한 분류를 갖는 트레이닝 입력들을 제시함
406
트레이닝 입력들의 라벨링에 기초하여 분류기를 업데이트함
408

도면5


트레이닝 입력
및 대화 유형
514
라벨
518
사용자
인터페이스
516
개인
512
개인
512
개인
512
대화형
입력들의
컬렉션 500
특징들
504
라벨들
506
트레이닝
모듈
502
파라미터
들
508
대화형 응답들
520

도면6


600
630
606
608
620
602
610
그래픽
프로세싱 유닛
프로세싱 유닛
착탈식 저장장치
비착탈식 저장장치
612
통신 연결부(들)
604
616
시스템 메모리
휘발성 메모리
비휘발성 메모리
출력 디바이스(들)
614
입력 디바이스(들)