



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0097239
(43) 공개일자 2022년07월07일

(51) 국제특허분류(Int. Cl.)
G06Q 50/10 (2012.01) G06F 16/33 (2019.01)
G06N 20/20 (2019.01) G06N 5/04 (2006.01)

(52) CPC특허분류
G06Q 50/10 (2015.01)
G06F 16/3347 (2019.01)

(21) 출원번호 10-2021-0177211(분할)
(22) 출원일자 2021년12월13일
심사청구일자 2021년12월13일

(62) 원출원 특허 10-2020-0187819
원출원일자 2020년12월30일
심사청구일자 2020년12월30일

(71) 출원인
주식회사 코어닷투데이
울산광역시 울주군 언양읍 유니스트길 50, 114동 806-8호

(72) 발명자
신재춘
울산광역시 북구 달천로 103-19 달천아이파크2차 204동 1104호

김대엽
경상북도 포항시 남구 오천읍 해병로241번길 5 제 일신천지타운

안소미
울산광역시 북구 상방로 12 한우리아파트 1-1103

(74) 대리인
서평강

전체 청구항 수 : 총 1 항

(54) 발명의 명칭 인공지능에 기반하여 시놉시스 텍스트를 분석하고 시청률을 예측하는 서버

(57) 요약

본 발명은 인공지능에 기반하여 시놉시스 텍스트를 분석하고 시청률을 예측하는 서버에 관한 것이다.

또한 본 발명은 시놉시스 텍스트 분석 및 시청률 예측 방법에 관한 것으로서, 더욱 상세하게는 드라마 콘텐츠를 포함한 시놉시스의 텍스트를 분석하여 콘텐츠에 맞는 캐릭터 및 배우를 추천하고 드라마의 시청률을 예측할 수 있도록 한 방법 및 이의 전자 장치에 관한 것이다.

대표도 - 도6

1	vector.most_similar("카리스마")	1	vector.most_similar("카리스마")
[('카리스마형', 0.811809778213501),		[('매력', 0.726121723651886),	
('카리스마넘치는', 0.8033837080001831),		('인간미', 0.6990047097206116),	
('카리스마만', 0.7894523739814758),		('연륜', 0.6927657723426819),	
('카리스마:', 0.7826545238494873),		('면모', 0.6809433102607727),	
('카리스마의', 0.7793933749198914),		('풍모', 0.680864691734314),	
('카리스마.', 0.7785433530807495),		('외모', 0.6766476631164551),	
('카리스마!', 0.7717763185501099),		('지략', 0.6704614162445068),	
('카리스마가', 0.7617940306663513),		('인품', 0.6683163642883301),	
('카리스마있고', 0.7511934638023376),		('성품', 0.6673632264137268),	
('카리스마적', 0.7474747896194458)]		('순둥이', 0.663332462310791)]	

(a)

(b)

(52) CPC특허분류

G06N 20/20 (2021.08)

G06N 5/04 (2013.01)

명세서

청구범위

청구항 1

프로세서; 및

하나 이상의 인스트럭션(instruction)을 저장하는 하나 이상의 메모리; 를 포함하고,

상기 프로세서는:

대상 드라마 콘텐츠의 시놉시스 텍스트 정보를 획득하고,

상기 시놉시스 텍스트 정보를 AI(artificial intelligence) 엔진 (engine)에 입력하도록 제어하고,

상기 시놉시스 텍스트 정보에 대한 상기 AI 엔진의 출력 동작에 기초하여, 상기 대상 드라마 콘텐츠를 위한 추천 캐릭터들에 대한 정보, 추천 배우들에 대한 정보, 추천 드라마들에 대한 정보 및 예측 시청률에 대한 정보를 출력(output)하도록 제어하고,

상기 AI 엔진의 출력 동작은:

상기 시놉시스 텍스트 정보를 벡터(vector)로 변환하는 임베딩(embedding)이 적용된 임베딩 된 시놉시스 텍스트 정보를 획득하는 동작;

상기 임베딩 된 시놉시스 텍스트 정보에 기초하여, 상기 대상 드라마 콘텐츠의 장르를 판단하는 동작;

상기 판단된 장르에 기초하여, 상기 추천 캐릭터들 및 상기 추천 배우들을 판단하는 동작; 및

상기 판단된 추천 배우들의 조합에 기초하여, 상기 추천 드라마들을 판단하는 동작;

상기 판단된 추천 드라마들에 기초하여, 상기 예측 시청률을 판단하는 동작; 을 포함하는 것을 특징으로 하며,

상기 시놉시스 텍스트 정보를 벡터(vector)로 변환하는 임베딩(embedding)이 적용된 임베딩 된 시놉시스 텍스트 정보를 획득하는 동작은,

상기 시놉시스 텍스트 정보를 상기 AI 엔진에 입력 가능하도록 임베딩 처리하여 벡터화하는 것을 의미하고,

상기 대상 드라마 콘텐츠의 장르를 판단하는 동작은, 상기 AI 엔진의 앙상블 학습(ensemble learning)에 기초하여 수행되고,

상기 앙상블 학습은, 모델(model)에 상기 AI 엔진이 수행 가능한 복수의 기계 학습(machine learning)들을 적용함에 기초하여 구성되며,

상기 앙상블 학습은 상기 복수의 기계 학습들을 통해 종래의 드라마 장르들을 처리함에 기초하여 수행되는, 서버.

발명의 설명

기술 분야

[0001] 본 발명은 인공지능에 기반하여 시놉시스 텍스트를 분석하고 시청률을 예측하는 서버에 관한 것이다.

[0002] 또한 본 발명은 시놉시스 텍스트 분석 및 시청률 예측 방법에 관한 것으로서, 더욱 상세하게는 드라마 콘텐츠를 포함한 시놉시스의 텍스트를 분석하여 콘텐츠에 맞는 캐릭터 및 배우를 추천하고 드라마의 시청률을 예측할 수 있도록 한 방법 및 이의 전자 장치에 관한 것이다.

배경 기술

[0003] 드라마는 TV, PC, 휴대폰, 스트리밍 서비스 등 콘텐츠를 재생하는 다양한 기기를 통해 다양한 시청자 층에 대하여 송출되고 있다. 그러나 이러한 드라마를 위한 시놉시스 텍스트 분석 시스템이 분석하는 시놉시스들의 수 및 앞으로 제공할 드라마를 위한 시놉시스의 수가 방대해지고 있으며, 시청자 층의 선호 역시 다각화 됨에 따라 중

래의 시놉시스 텍스트 분석 시스템이 드라마 제작을 위한 시놉시스를 분석하는데 어려움이 있으며, 그에 따른 시청률에 대한 분석 역시 예측력이 떨어지는 문제가 있다. 특히, 신규 드라마를 위한 시놉시스의 경우 사용자가 해당 시놉시스의 콘텐츠를 선호하는지 사용자의 성향이 파악되어 있지 않으며, 신규 드라마는 아직 아무도 소비한 적이 없기 때문에 기존에 방영했던 비슷한 취지의 드라마의 전례를 통해 시청률을 분석할 수 밖에 없다.

[0004] 한편, 머신 러닝을 기반으로 한 텍스트 분석은 다양한 소셜 네트워크 서비스(SNS), 클라우드 플랫폼 서비스 등에서 널리 이용되고 있으며, 텍스트의 흐름, 뉘앙스 또는 의도까지 인식하는 수준으로 발전하고 있다. 텍스트 인식을 기반으로 하는 인공지능(AI) 서비스도 현재 시장의 트렌드의 하나이다. 이러한 인공지능(AI) 기술은 현재 콘텐츠 사업 분야에도 적용되는 추세로서, 종래 텍스트 분석뿐 아니라 이를 이용하여 특정한 콘텐츠를 포함한 텍스트를 생성하는 수준까지 발전되고 있다.

발명의 내용

해결하려는 과제

[0005] 본 발명의 방법에 따르면 시놉시스 텍스트 분석 및 시청률 예측 시스템은 머신 러닝을 통해 다양한 시청자 층의 취향을 고려하여 드라마 콘텐츠의 적합 장르를 분석하고, 시놉시스 텍스트에 적합한 캐릭터 및 해당 캐릭터에 맞는 배우를 추천하여 객관적인 시청률을 도출할 수 있도록 한다. 이를 통해 드라마 담당자의 직관에만 의존해야 했던 기존 방식에서 벗어나 작가가 입력한 시나리오에 가장 적합한 배우를 추천하며, 이러한 배우의 조합에 어울리는 드라마를 제시하고 그 드라마의 성과를 보여줌으로써 예상되는 드라마 성과를 예측할 수 있다.

[0006] 본 발명에서 이루고자 하는 기술적 과제들은 이상에서 언급한 기술적 과제들로 제한되지 않으며, 언급하지 않은 또 다른 기술적 과제들은 아래의 기재로부터 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0007] 다양한 실시예들은, 시놉시스 텍스트 분석 및 시청률 예측을 위한 전자 장치의 동작 방법 및 이를 지원하는 전자 장치를 제공할 수 있다.

[0008] 다양한 실시예들에 따른 전자 장치의 동작 방법은: 대상 드라마 콘텐츠의 시놉시스 텍스트 정보를 획득하는 동작; 상기 시놉시스 텍스트 정보를 AI (artificial intelligence) 엔진 (engine)에 입력하는 동작; 및 상기 시놉시스 텍스트 정보에 대한 상기 AI 엔진의 출력 동작에 기초하여, 상기 대상 드라마 콘텐츠를 위한 추천 캐릭터들에 대한 정보, 추천 배우들에 대한 정보, 추천 드라마들에 대한 정보 및 예측 시청률에 대한 정보를 출력하는 동작을 포함할 수 있다.

[0009] 예시적 실시예에서, 상기 AI 엔진의 출력 동작은: 상기 시놉시스 텍스트 정보를 벡터(vector)로 변환하는 임베딩(embedding)이 적용된 임베딩 된 시놉시스 텍스트 정보를 획득하는 동작; 상기 임베딩 된 시놉시스 텍스트 정보에 기초하여, 상기 대상 드라마 콘텐츠의 장르를 판단하는 동작; 상기 판단된 장르에 기초하여, 상기 추천 캐릭터들 및 상기 추천 배우들을 판단하는 동작; 및 상기 판단된 추천 배우들의 조합에 기초하여, 상기 추천 드라마들을 판단하는 동작; 상기 판단된 추천 드라마들에 기초하여, 상기 예측 시청률을 판단하는 동작을 포함할 수 있다.

[0010] 예시적 실시예에서, 상기 대상 드라마 콘텐츠의 장르를 판단하는 동작은, 상기 AI 엔진의 앙상블 학습(ensemble learning)에 기초하여 수행되고, 상기 앙상블 학습은, 모델(model)에 상기 AI 엔진이 수행 가능한 복수의 기계 학습(machine learning)들을 적용함에 기초하여 구성되며, 상기 앙상블 학습은 상기 복수의 기계 학습들을 통해 종래의 드라마 장르들을 처리함에 기초하여 수행될 수 있다.

[0011] 예시적 실시예에서, 상기 추천 캐릭터들 및 상기 추천 배우들을 판단하는 동작은: 상기 판단된 장르에 출연한 배우들의 목록을 획득하는 동작; 상기 출연한 배우들이 연기했던 드라마 캐릭터들에 대한 제1 캐릭터 임베딩 정보들을 획득하는 동작; 상기 제1 캐릭터 임베딩 정보들과 상기 임베딩 된 시놉시스 텍스트 정보를 비교하여, 상기 임베딩 된 시놉시스 텍스트 정보와 유사한 제2 캐릭터 임베딩 정보들을 판단하는 동작; 및 상기 제2 캐릭터 임베딩 정보들에 해당하는 캐릭터들 및 상기 제2 캐릭터 임베딩 정보들에 해당하는 캐릭터들을 연기했던 배우들을 판단하는 동작을 포함할 수 있다.

[0012] 예시적 실시예에서, 상기 추천 드라마들을 판단하는 동작은: 상기 판단된 추천 배우들의 조합에 대한 임베딩 정보를 획득하는 동작; 상기 판단된 추천 배우들의 조합에 대한 임베딩 정보와 유사한 드라마 임베딩 정보들을 판

단하는 동작; 및 상기 유사한 드라마 임베딩 정보들에 해당하는 드라마들을 판단하는 동작을 포함할 수 있다.

- [0013] 예시적 실시예에서, 상기 예측 시청률에 대한 정보는 상기 추천 드라마들 중 유사성이 가장 높은 드라마의 시청률에 기반하여 구성될 수 있다.
- [0014] 예시적 실시예에서, 상기 AI 엔진은, 모델 (model)에 기계 학습 (machine learning)을 적용할 수 있도록 미리 구성되고, 상기 기계 학습은: - (a) 학습용 드라마 데이터를 처리함에 기초하여 획득된 훈련용 데이터 및 테스트용 데이터 중 훈련용 데이터에 기초하여 상기 모델을 학습; - (b) 상기 테스트용 데이터가 상기 모델로 입력됨에 대한 응답으로 출력되는 피드백 정보를 획득; - (c) 상기 피드백 정보에 기초하여 상기 모델을 업데이트; 및 - (d) 상기 (a) 내지 (c) 가 반복되되, 상기 (a) 내지 (c) 의 반복 횟수가 미리 설정된 임계값과 일치함에 기초하여 종료; 됨에 기초하여 수행될 수 있다.
- [0015] 다양한 실시예들에 따른 전자 장치는: 프로세서 (processor); 및 하나 이상의 인스트럭션 (instruction)을 저장하는 하나 이상의 메모리 (memory)를 포함할 수 있다.
- [0016] 예시적 실시예에서, 대상 드라마 콘텐츠의 시놉시스 텍스트 정보를 획득하는 동작; 상기 시놉시스 텍스트 정보를 AI (artificial intelligence) 엔진 (engine)에 입력하는 동작; 및 상기 시놉시스 텍스트 정보에 대한 상기 AI 엔진의 출력 동작에 기초하여, 상기 대상 드라마 콘텐츠를 위한 추천 캐릭터들에 대한 정보, 추천 배우들에 대한 정보, 추천 드라마들에 대한 정보 및 예측 시청률에 대한 정보를 출력하는 동작을 수행하도록 상기 프로세서를 제어할 수 있다.
- [0017] 예시적 실시예에서, 상기 AI 엔진의 출력 동작은: 상기 시놉시스 텍스트 정보를 벡터(vector)로 변환하는 임베딩(embedding)이 적용된 임베딩 된 시놉시스 텍스트 정보를 획득하는 동작; 상기 임베딩 된 시놉시스 텍스트 정보에 기초하여, 상기 대상 드라마 콘텐츠의 장르를 판단하는 동작; 상기 판단된 장르에 기초하여, 상기 추천 캐릭터들 및 상기 추천 배우들을 판단하는 동작; 및 상기 판단된 추천 배우들의 조합에 기초하여, 상기 추천 드라마들을 판단하는 동작; 상기 판단된 추천 드라마들에 기초하여, 상기 예측 시청률을 판단하는 동작을 포함할 수 있다.
- [0018] 예시적 실시예에서, 상기 대상 드라마 콘텐츠의 장르를 판단하는 동작은, 상기 AI 엔진의 앙상블 학습(ensemble learning)에 기초하여 수행되고, 상기 앙상블 학습은, 모델(model)에 상기 AI 엔진이 수행 가능한 복수의 기계 학습(machine learning)들을 적용함에 기초하여 구성되며, 상기 앙상블 학습은 상기 복수의 기계 학습들을 통해 종래의 드라마 장르들을 처리함에 기초하여 수행될 수 있다.
- [0019] 예시적 실시예에서, 상기 추천 캐릭터들 및 상기 추천 배우들을 판단하는 동작은: 상기 판단된 장르에 출연한 배우들의 목록을 획득하는 동작; 상기 출연한 배우들이 연기했던 드라마 캐릭터들에 대한 제1 캐릭터 임베딩 정보들을 획득하는 동작; 상기 제1 캐릭터 임베딩 정보들과 상기 임베딩 된 시놉시스 텍스트 정보를 비교하여, 상기 임베딩 된 시놉시스 텍스트 정보와 유사한 제2 캐릭터 임베딩 정보들을 판단하는 동작; 및 상기 제2 캐릭터 임베딩 정보들에 해당하는 캐릭터들 및 상기 제2 캐릭터 임베딩 정보들에 해당하는 캐릭터들을 연기했던 배우들을 판단하는 동작을 포함할 수 있다.
- [0020] 예시적 실시예에서, 상기 추천 드라마들을 판단하는 동작은: 상기 판단된 추천 배우들의 조합에 대한 임베딩 정보를 획득하는 동작; 상기 판단된 추천 배우들의 조합에 대한 임베딩 정보와 유사한 드라마 임베딩 정보들을 판단하는 동작; 및 상기 유사한 드라마 임베딩 정보들에 해당하는 드라마들을 판단하는 동작을 포함할 수 있다.
- [0021] 예시적 실시예에서, 상기 예측 시청률에 대한 정보는 상기 추천 드라마들 중 유사성이 가장 높은 드라마의 시청률에 기반하여 구성될 수 있다.
- [0022] 예시적 실시예에서, 상기 AI 엔진은, 모델 (model)에 기계 학습 (machine learning)을 적용할 수 있도록 미리 구성되고, 상기 기계 학습은: - (a) 학습용 드라마 데이터를 처리함에 기초하여 획득된 훈련용 데이터 및 테스트용 데이터 중 훈련용 데이터에 기초하여 상기 모델을 학습; - (b) 상기 테스트용 데이터가 상기 모델로 입력됨에 대한 응답으로 출력되는 피드백 정보를 획득; - (c) 상기 피드백 정보에 기초하여 상기 모델을 업데이트; 및 - (d) 상기 (a) 내지 (c) 가 반복되되, 상기 (a) 내지 (c)의 반복 횟수가 미리 설정된 임계값과 일치함에 기초하여 종료; 됨에 기초하여 수행될 수 있다.
- [0023] 상술한 본 개시의 다양한 실시예들은 본 개시의 바람직한 실시예들 중 일부에 불과하며, 본 개시의 다양한 실시예들의 기술적 특징들이 반영된 여러 가지 실시예들이 당해 기술분야의 통상적인 지식을 가진 자에 의해 이차상술할 상세한 설명을 기반으로 도출되고 이해될 수 있다.

발명의 효과

- [0024] 본 발명은 드라마 콘텐츠를 포함한 시놉시스의 텍스트를 분석하여 콘텐츠에 맞는 캐릭터 및 배우를 추천하고 드라마의 시청률을 예측할 수 있는 방법을 제시하여, 보다 적합한 드라마 구성 및 정확한 시청률 예측 결과를 확보할 수 있는 측면에서 기술적인 효과를 갖는다.
- [0025] 본 발명에서 얻을 수 있는 효과는 이상에서 언급한 효과들로 제한되지 않으며, 언급하지 않은 또 다른 효과들은 아래의 기재로부터 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

- [0026] 다양한 실시예들에 대한 이해를 돕기 위해 상세한 설명의 일부로 포함된, 첨부 도면은 다양한 실시예들을 제공하고, 상세한 설명과 함께 본 개시의 기술적 특징을 설명한다.
 - 도 1은 다양한 실시예들에 따른 시놉시스 텍스트 분석 및 시청률 예측을 위한 전자 장치의 동작 방법이 구현될 수 있는 시놉시스 텍스트 분석 및 시청률 예측 시스템을 설명하기 위한 도면이다.
 - 도 2는 다양한 실시예들에 따른 측정 장치 및/또는 사용자 장치의 구성을 도시한 도면이다.
 - 도 3은 다양한 실시예들에 따른 시청률 예측 및 출력을 위한 전자 장치의 동작 방법을 도시한 도면이다.
 - 도 4는 다양한 실시예들에 따른 시청률 예측 및 출력을 위한 UI/UX 일 예를 도시한 도면이다.
 - 도 5는 다양한 실시예들에 따른 시놉시스 텍스트 분석을 위한 전자 장치의 동작 방법을 도시한 도면이다.
 - 도 6은 단어 기반 단순 가공 형태의 임베딩 처리 예시 및 한글 자소 기반 형태소 가공 형태의 임베딩 처리 예시를 나타낸 도면이다.
 - 도 7은 추천된 배우들의 조합에 따른 드라마 추천의 예시를 나타낸 도면이다.
 - 도 8은 다양한 실시예들에 따른 시놉시스 텍스트 분석을 위한 AI (artificial intelligence) 엔진을 획득하는 과정을 도시한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0027] 이하의 실시예들은 다양한 실시예들의 구성요소들과 특징들을 소정 형태로 결합한 것들이다. 각 구성요소 또는 특징은 별도의 명시적 언급이 없는 한 선택적인 것으로 고려될 수 있다. 각 구성요소 또는 특징은 다른 구성요소나 특징과 결합되지 않은 형태로 실시될 수 있다. 또한, 일부 구성요소들 및/또는 특징들을 결합하여 다양한 실시예들을 구성할 수도 있다. 다양한 실시예들에서 설명되는 동작들의 순서는 변경될 수 있다. 어느 실시예의 일부 구성이나 특징은 다른 실시예에 포함될 수 있고, 또는 다른 실시예의 대응하는 구성 또는 특징과 교체될 수 있다.
- [0028] 도면에 대한 설명에서, 다양한 실시예들의 요지를 흐릴 수 있는 절차 또는 단계 등은 기술하지 않았으며, 당해 기술분야에서 통상의 지식을 가진 자의 수준에서 이해할 수 있을 정도의 절차 또는 단계는 또한 기술하지 아니하였다.
- [0029] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함(comprising 또는 including)"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에 기재된 "...부", "...기", "모듈" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다. 또한, "일(a 또는 an)", "하나(one)", "그(the)" 및 유사 관련어는 다양한 실시예들을 기술하는 문맥에 있어서(특히, 이하의 청구항의 문맥에서) 본 명세서에 달리 지시되거나 문맥에 의해 분명하게 반박되지 않는 한, 단수 및 복수 모두를 포함하는 의미로 사용될 수 있다.
- [0030] 이하, 다양한 실시예들에 따른 바람직한 실시 형태를 첨부된 도면을 참조하여 상세하게 설명한다. 첨부된 도면과 함께 이하에 개시될 상세한 설명은 다양한 실시예들의 예시적인 실시형태를 설명하고자 하는 것이며, 유일한 실시형태를 나타내고자 하는 것이 아니다.
- [0031] 또한, 다양한 실시예들에서 사용되는 특정(特定) 용어들은 다양한 실시예들의 이해를 돕기 위해서 제공된 것이

며, 이러한 특정 용어의 사용은 다양한 실시예들의 기술적 사상을 벗어나지 않는 범위에서 다른 형태로 변경될 수 있다.

[0033]

1. 시놉시스 텍스트 분석 및 시청률 예측 시스템 구현

[0034]

도 1은 다양한 실시예들에 따른 시놉시스 텍스트 분석 및 시청률 예측을 위한 전자 장치의 동작 방법이 구현될 수 있는 시놉시스 텍스트 분석 및 시청률 예측 시스템을 설명하기 위한 도면이다.

[0035]

도 1을 참조하면, 다양한 실시예들에 따른 시놉시스 텍스트 분석 및 시청률 예측 시스템은 다양한 종류의 전자 장치들에 구현될 수 있다. 예를 들어, 시놉시스 텍스트 분석 및 시청률 예측 시스템은 서버(100) 및/또는 사용자 단말(200)에 구현될 수 있다. 달리 말해, 서버(100) 및/또는 사용자 단말(200)은 각각의 장치에 구현된 시놉시스 텍스트 분석 및 시청률 예측 시스템을 기반으로, 본 개시의 다양한 실시예들에 따른 동작을 수행할 수 있다. 한편, 다양한 실시예들에 따른 시놉시스 텍스트 분석 및 시청률 예측 시스템은, 상기 도 1에 도시된 바에 국한되지 않고, 더 다양한 전자 장치와 서버들에 구현될 수도 있을 것이다.

[0036]

다양한 실시예들에 따른 서버(100)는, 복수 개의 사용자 단말(200)들과 무선 및/또는 유선 통신을 수행하며, 대량의 저장 용량을 갖는 데이터베이스를 포함하는 장치일 수 있다. 예를 들어, 서버(100)는 복수 개의 사용자 디바이스들과 연동되는 클라우드 디바이스(Cloud device)일 수 있다.

[0037]

다양한 실시예들에 따른 사용자 단말(200)은, 데스크탑 피시, 태블릿 피시, 모바일 단말 등의 개인 사용자에 의해 이용될 수 있는 장치일 수 있다. 이외에도 유사한 기능을 수행하는 다른 전자 장치들이 사용자 단말(200)로 이용될 수 있다.

[0039]

1.1. 기능적 구현

[0040]

다양한 실시예들에 따른 시놉시스 텍스트 분석 및 시청률 예측 시스템은 동작을 위한 다양한 모듈들을 포함할 수 있다. 시놉시스 텍스트 분석 및 시청률 예측 시스템에 포함된 모듈들은 시놉시스 텍스트 분석 및 시청률 예측 시스템이 구현되는(또는, 물리적 장치에 포함되는) 물리적 장치(예: 서버(100) 및/또는 사용자 단말(200))가 지정된 동작을 수행할 수 있도록 구현된 컴퓨터 코드 내지는 하나 이상의 인스트럭션(instruction)일 수 있다. 다시 말해, 시놉시스 텍스트 분석 및 시청률 예측 시스템이 구현되는 물리적 장치는 복수 개의 모듈들을 컴퓨터 코드 형태로 메모리에 저장하고, 메모리에 저장된 복수 개의 모듈들이 실행되는 경우 복수 개의 모듈들은 물리적 장치가 복수 개의 모듈들에 대응하는 지정된 동작들을 수행하도록 할 수 있다.

[0042]

1.2. 장치 구현

[0043]

도 2는 다양한 실시예들에 따른 서버 및/또는 사용자 디바이스의 구성을 도시한 도면이다.

[0044]

도 2를 참조하면, 서버(100) 및/또는 사용자 단말(200)은 입/출력부(210), 통신부(220), 데이터베이스(230) 및 프로세서(240)를 포함할 수 있다.

[0045]

입/출력부(210)는 사용자 입력을 받거나 또는 사용자에게 정보를 출력하는 각종 인터페이스나 연결 포트 등일 수 있다. 입/출력부(210)는 입력 모듈과 출력 모듈로 구분될 수 있는데, 입력 모듈은 사용자로부터 사용자 입력을 수신한다. 사용자 입력은 키 입력, 터치 입력, 음성 입력을 비롯한 다양한 형태로 이루어질 수 있다. 이러한 사용자 입력을 받을 수 있는 입력 모듈의 예로는 전통적인 형태의 키패드나 키보드, 마우스는 물론, 사용자의 터치를 감지하는 터치 센서, 음성 신호를 입력받는 마이크, 영상 인식을 통해 제스처 등을 인식하는 카메라, 사용자 접근을 감지하는 조도 센서나 적외선 센서 등으로 구성되는 근접 센서, 가속도 센서나 자이로 센서 등을 통해 사용자 동작을 인식하는 모션 센서 및 그 외의 다양한 형태의 사용자 입력을 감지하거나 입력받는 다양한 형태의 입력 수단을 모두 포함하는 포괄적인 개념이다. 여기서, 터치 센서는 디스플레이 패널에 부착되는 터치 패널이나 터치 필름을 통해 터치를 감지하는 압전식 또는 정전식 터치 센서, 광학적인 방식에 의해 터치를 감지하는 광학식 터치 센서 등으로 구현될 수 있다. 이외에도 입력 모듈은 자체적으로 사용자 입력을 감지하는 장치 대신 사용자 입력을 입력받는 외부의 입력 장치를 연결시키는 입력 인터페이스(USB 포트, PS/2 포트 등)의 형태로 구현될 수도 있다. 또 출력 모듈은 각종 정보를 출력해 사용자에게 이를 제공할 수 있다. 출력 모듈은 영상을 출력하는 디스플레이, 소리를 출력하는 스피커, 진동을 발생시키는 햅틱 장치 및 그 외의 다양한 형태의 출

력 수단을 모두 포함하는 포괄적인 개념이다. 이외에도 출력 모듈은 상술한 개별 출력 수단을 연결시키는 포트 타입의 출력 인터페이스의 형태로 구현될 수도 있다.

[0046] 일 예로, 디스플레이 형태의 출력 모듈은 텍스트, 정지 영상, 동영상을 디스플레이 할 수 있다. 디스플레이는 액정 디스플레이(LCD: Liquid Crystal Display), 발광 다이오드(LED: light emitting diode) 디스플레이, 유기 발광 다이오드(OLED: Organic Light Emitting Diode) 디스플레이, 평판 디스플레이(FPD: Flat Panel Display), 투명 디스플레이(transparent display), 곡면 디스플레이(Curved Display), 플렉시블 디스플레이(flexible display), 3차원 디스플레이(3D display), 홀로그래픽 디스플레이(holographic display), 프로젝터 및 그 외의 영상 출력 기능을 수행할 수 있는 다양한 형태의 장치를 모두 포함하는 광의의 영상 표시 장치를 의미하는 개념이다. 이러한 디스플레이는 입력 모듈의 터치 센서와 일체로 구성된 터치 디스플레이의 형태일 수도 있다.

[0047] 통신부(220)는 외부 기기와 통신할 수 있다. 따라서, 서버(100) 및/또는 사용자 단말(200)은 통신부를 통해 외부 기기와 정보를 송수신할 수 있다. 예를 들어, 서버(100) 및/또는 사용자 단말(200)은 통신부를 이용해 상호 간 통신을 수행하거나, 기타 외부 기기와 통신을 수행할 수 있다.

[0048] 여기서, 통신, 즉 데이터의 송수신은 유선 또는 무선으로 이루어질 수 있다. 이를 위해 통신부는 LAN(Local Area Network)를 통해 인터넷 등에 접속하는 유선 통신 모듈, 이동 통신 기지국을 거쳐 이동 통신 네트워크에 접속하여 데이터를 송수신하는 이동 통신 모듈, 와이파이(Wi-Fi) 같은 WLAN(Wireless Local Area Network) 계열의 통신 방식이나 블루투스(Bluetooth), 직비(Zigbee)와 같은 WPAN(Wireless Personal Area Network) 계열의 통신 방식을 이용하는 근거리 통신 모듈, GPS(Global Positioning System)과 같은 GNSS(Global Navigation Satellite System)을 이용하는 위성 통신 모듈 또는 이들의 조합으로 구성될 수 있다.

[0049] 데이터베이스(230)는 각종 정보를 저장할 수 있다. 데이터베이스는 데이터를 임시적으로 또는 반영구적으로 저장할 수 있다. 예를 들어, 서버(100)의 데이터베이스에는 서버(100)를 구동하기 위한 운영 프로그램(OS: Operating System), 웹 사이트를 호스팅하기 위한 데이터나 점자 생성을 위한 프로그램 내지는 어플리케이션(예를 들어, 웹 어플리케이션)에 관한 데이터 등이 저장될 수 있다. 또, 데이터베이스는 상술한 바와 같이 모듈들을 컴퓨터 코드 형태로 저장할 수 있다.

[0050] 데이터베이스(230)의 예로는 하드 디스크(HDD: Hard Disk Drive), SSD(Solid State Drive), 플래쉬 메모리(flash memory), 롬(ROM: Read-Only Memory), 램(RAM: Random Access Memory) 등이 있을 수 있다. 이러한 데이터베이스는 내장 타입 또는 탈부착 가능한 타입으로 제공될 수 있다.

[0051] 프로세서(240)는 서버(100) 및/또는 사용자 단말(200)의 전반적인 동작을 제어한다. 이를 위해 프로세서(240)는 각종 정보의 연산 및 처리를 수행하고 서버(100)의 구성요소들의 동작을 제어할 수 있다. 예를 들어, 프로세서(240)는 시놉시스 텍스트 분석 및 시청률 예측을 위한 프로그램 내지 어플리케이션을 실행시킬 수 있을 것이다. 프로세서(240)는 하드웨어 소프트웨어 또는 이들의 조합에 따라 컴퓨터나 이와 유사한 장치로 구현될 수 있다. 하드웨어적으로 프로세서(240)는 전기적인 신호를 처리하여 제어 기능을 수행하는 전자 회로 형태로 제공될 수 있으며, 소프트웨어적으로는 하드웨어적인 프로세서(240)를 구동시키는 프로그램 형태로 제공될 수 있다. 한편, 이하의 설명에서 특별한 언급이 없는 경우에는 서버(100) 및/또는 사용자 단말(200)의 동작은 프로세서(240)의 제어에 의해 수행되는 것으로 해석될 수 있다. 즉, 상술한 시놉시스 텍스트 분석 및 시청률 예측 시스템에 구현되는 모듈들이 실행되는 경우, 모듈들은 프로세서(240)가 서버(100) 및/또는 사용자 단말(200)를 이하의 동작들을 수행하도록 제어하는 것으로 해석될 수 있다.

[0052] 요약하면, 다양한 실시예들은 다양한 수단을 통해 구현될 수 있다. 예를 들어, 다양한 실시예들은 하드웨어, 펌웨어(firmware), 소프트웨어 또는 그것들의 결합 등에 의해 구현될 수 있다.

[0053] 하드웨어에 의한 구현의 경우, 다양한 실시예들에 따른 방법은 하나 또는 그 이상의 ASICs(application specific integrated circuits), DSPs(digital signal processors), DSPDs(digital signal processing devices), PLDs(programmable logic devices), FPGAs(field programmable gate arrays), 프로세서, 컨트롤러, 마이크로 컨트롤러, 마이크로 프로세서 등에 의해 구현될 수 있다.

[0054] 펌웨어나 소프트웨어에 의한 구현의 경우, 다양한 실시예들에 따른 방법은 이하에서 설명된 기능 또는 동작들을 수행하는 모듈, 절차 또는 함수 등의 형태로 구현될 수 있다. 예를 들어, 소프트웨어 코드는 메모리에 저장되어 프로세서에 의해 구동될 수 있다. 상기 메모리는 상기 프로세서 내부 또는 외부에 위치할 수 있으며, 이미 공지된 다양한 수단에 의해 상기 프로세서와 데이터를 주고받을 수 있다.

- [0056] 2. 다양한 실시예들에 따른 전자 장치의 동작
- [0057] 이하에서는, 상기와 같은 기술적 사상에 기반하여 다양한 실시예들에 대해 보다 상세히 설명한다. 이하에서 설명되는 다양한 실시예들에 대해서는 앞서 설명한 제 1 절의 내용들이 적용될 수 있다. 예를 들어, 이하에서 설명되는 다양한 실시예들에서 정의되지 않은 동작, 기능, 용어 등은 제 1 절의 내용들에 기반하여 수행되고 설명될 수 있다.
- [0059] 2.1 시청률 예측 및 출력 동작
- [0060] 이하의 설명에서는 사용자 단말(200)가 시청률 예측 및 출력 동작을 수행함을 전제로 다양한 실시예들에 대하여 설명이나, 다양한 실시예들에 따르면, 사용자 단말(200)로부터 상 드라마 콘텐츠의 시놉시스 텍스트 정보를 획득한 서버(100)가 시청률 예측 및 출력 동작을 수행할 수도 있다. 또는, 다양한 실시예들에 따르면, 하나의 서버(100) 또는 하나의 사용자 단말(200) 각각이 직접 대상 드라마 콘텐츠의 시놉시스 텍스트 정보를 획득(예를 들어, 사용자로부터 직접 입력)하여, 시청률 예측 및 출력 동작을 수행할 수도 있다. 또는, 다양한 실시예들에 따르면, 복수의 사용자 단말(200)들 및/또는 복수의 서버(100)들이 마련되어, 복수의 사용자 단말(200)들 및/또는 복수의 서버(100)들 간에 대상 드라마 콘텐츠의 시놉시스 텍스트 정보 송수신 및 시청률 예측 동작이 각 동작 별로 분배되어, 별개로 실시될 수도 있다.
- [0062] 도 3은 다양한 실시예들에 따른 시청률 예측 및 출력을 위한 전자 장치의 동작 방법을 도시한 도면이다.
- [0063] 다양한 실시예들에 따르면, 동작 S310에서, 사용자 단말(200)은 대상 드라마 콘텐츠의 시놉시스 텍스트 정보를 획득할 수 있다.
- [0064] 예를 들어, 상기 시놉시스 텍스트 정보는 상기 드라마 콘텐츠에 대한 줄거리 개요 또는 전체 내용을 간단히 정리한 텍스트일 수 있다.
- [0065] 예를 들어, 상기 시놉시스 텍스트 정보는 사용자가 사용자 단말(200)에 직접 입력할 수 있다.
- [0066] 예를 들어, 상기 시놉시스 텍스트 정보는 상기 드라마 콘텐츠를 대표하는 단어, 용어, 문구 기타 텍스트 정보를 포함할 수 있으며, 상기 시놉시스 텍스트 정보에 포함된 모든 텍스트는 AI(artificial intelligence) 엔진에 기반한 분석의 대상이 될 수 있다.
- [0067] 예를 들어, 사용자 단말(200)은 상기 시놉시스 텍스트 정보의 의미를 정확히 식별하기 위하여 상기 시놉시스 텍스트 정보 상의 오타 또는 띄어쓰기 등 기타 문법 오류가 있는지 판단할 수 있다.
- [0068] 예를 들어, 사용자 단말(200)은 상기 시놉시스 텍스트 정보 상에 오타 또는 띄어쓰기 등 기타 문법 오류가 존재하여 상기 시놉시스 텍스트 정보의 의미를 정확히 식별하지 못하는 경우, 사용자에게 문법 오류가 존재하는 부분을 표시하고 해당 부분에 대한 수정을 요청할 수 있다.
- [0069] 예를 들어, 사용자 단말(200)은 상기 시놉시스 텍스트 정보 상에 오타 또는 띄어쓰기 등 기타 문법 오류가 존재하더라도 상기 시놉시스 텍스트 정보의 의미를 정확히 식별할 수 있는 경우, 문법 오류에 대한 별도의 표시 또는 수정 요청 없이 시청률 예측을 위한 다음 동작을 수행할 수 있다.
- [0070] 예를 들어, 사용자 단말(200)은 상기 시놉시스 텍스트 정보 상에서 의미를 식별할 수 있는 단어, 용어, 문구 기타 텍스트 정보가 부족하다고 판단되는 경우, 사용자에게 텍스트 정보의 부족을 표시하고 추가적인 텍스트 정보의 입력을 요청할 수 있다.
- [0071] 예를 들어, 사용자 단말(200)은 상기 시놉시스 텍스트 정보가 특정 수의 글자 미만인 경우 또는 특정 개수의 단어 미만인 경우 사용자에게 텍스트 정보의 부족을 표시하고 추가적인 텍스트 정보의 입력을 요청할 수 있으며, 상기 특정 글자의 수 또는 상기 특정 단어의 개수는 사전에 사용자 단말(200)에 구성될 수 있다.
- [0072] 예를 들어, 사용자 단말(200)은 상기 시놉시스 텍스트 정보 상에서 의미를 식별할 수 있는 단어, 용어, 문구 기타 텍스트 정보가 과다하다고 판단되는 경우, 사용자에게 텍스트 정보가 충분함을 표시하고 입력된 텍스트 정보의 일부에 대한 삭제를 요청할 수 있다.

- [0073] 예를 들어, 사용자 단말(200)은 상기 시놉시스 텍스트 정보가 특정 수의 글자 초과인 경우 또는 특정 개수의 단어 초과인 경우 사용자에게 텍스트 정보의 과다를 표시하고 입력된 텍스트 정보의 일부에 대한 삭제를 요청할 수 있으며, 상기 특정 글자의 수 또는 상기 특정 단어의 개수는 사전에 사용자 단말(200)에 구성될 수 있다.
- [0074] 다양한 실시예들에 따르면, 동작 S320에서, 사용자 단말(200)은 시놉시스 텍스트 정보를 AI 엔진 (engine)에 입력할 수 있다.
- [0075] 예를 들어, 사용자 단말(200)은 상기 시놉시스 텍스트 정보를 AI 엔진에 입력 가능하도록 임베딩(embedding) 처리하여 벡터(vector)화 된 정보로서 AI 엔진에 상기 시놉시스 텍스트 정보를 입력할 수 있다.
- [0076] 예를 들어, 상기 시놉시스 텍스트 정보에 대한 임베딩(embedding) 처리는 자연어를 수치화 및 방향화하는 것을 의미할 수 있으며, 상기 자연어는 한국어, 영어 등 사용자 단말(200)이 인식 가능한 모든 언어를 포함할 수 있다.
- [0077] 예를 들어, AI 엔진에 대한 사용자 단말(200)의 상기 시놉시스 텍스트 정보 입력은 단어, 용어, 문구 등 의미를 파악할 수 있는 모든 단위에서 수행될 수 있다.
- [0078] 다양한 실시예들에 따르면, 동작 S330에서, 사용자 단말(200)은 상기 시놉시스 텍스트 정보에 대한 상기 AI 엔진의 출력 동작에 기초하여, 상기 대상 드라마 콘텐츠를 위한 추천 캐릭터들에 대한 정보, 추천 배우들에 대한 정보, 추천 드라마들에 대한 정보 및 예측 시청률에 대한 정보를 출력할 수 있다.
- [0079] 여기서, 상기 추천 캐릭터들에 대한 정보, 상기 추천 배우들에 대한 정보, 상기 추천 드라마들에 대한 정보 및 상기 예측 시청률에 대한 정보를 출력하기 위한 상기 AI 엔진의 출력 동작은, 후술할 도 5 내지 도 8의 동작에 따라 수행될 수 있다.
- [0081] 도 4는 다양한 실시예들에 따른 시청률 예측 및 출력을 위한 UI/UX 일 예를 도시한 도면이다. 예를 들어, 도 4의 예시는 사용자 단말(200)을 위한 UI/UX (user interface/user experience) 일 수 있다.
- [0082] 예를 들어, 도 4를 참조하면, '시놉시스*' 로 표시된 제1 영역(401)에는 대상 드라마 콘텐츠에 대한 텍스트가 놓여짐에 따라 시놉시스 텍스트 정보가 입력될 수 있으며, 사용자 단말(200)은, 제1 영역(401)에 시놉시스 텍스트 정보가 입력되면 이를 서버(100)로 송신하고 AI 엔진에 입력할 수 있다.
- [0083] 예를 들어, 사용자 단말(200)은, 시놉시스 텍스트 정보를 분석하여, 시놉시스 텍스트 정보에 대응하는 추천 캐릭터들 정보 및 추천 배우들 정보를 '주연1*', '주연2*', ..., '조연3*' 등의 카테고리로 표시되는 제2 영역(403)에 표시할 수 있다.
- [0084] 예를 들어, 사용자 단말(200)은, 사용자에게 제2 영역(403)의 각 카테고리에 복수의 추천 배우들을 표시(display)할 수 있다. 이 때, 각 카테고리는 추천 캐릭터들 정보에 따라 분류된 것일 수 있다.
- [0085] 예를 들어, 사용자 단말(200)은, 추천 캐릭터들 정보 및 추천 배우들 정보를 분석하여, 시놉시스 텍스트 정보에 대응하는 추천 드라마 정보 및 추천 드라마의 시청률 정보를 제3 영역(405)에 표시할 수 있다.
- [0086] 도 4의 사용자 단말(200)을 위한 UI/UX 및 사용자 디바이스의 동작은 후술할 도 5 내지 도 8에 개시된 서버(100) 및 AI 엔진의 동작과 연관되어 수행되며, 도 4의 UI/UX 예시는 본 발명의 개시를 위한 하나의 예시로서 본 발명의 다양한 실시예들이 도 4의 UI/UX 형태에 한정되지 않으며 본 발명의 다양한 실시예들을 구현할 수 있는 모든 형태의 UI/UX에 따라 수행될 수 있다.
- [0088] 2.2 시놉시스 텍스트 분석 동작
- [0089] 이하의 설명에서는 서버(100)가 시놉시스 텍스트 분석 동작을 수행함을 전제로 다양한 실시예들에 대하여 설명하나, 다양한 실시예들에 따르면, 서버(100)로부터 상 드라마 콘텐츠의 시놉시스 텍스트 정보를 획득한 사용자 단말(200)이 시청률 예측 동작을 수행할 수도 있다. 또는, 다양한 실시예들에 따르면, 하나의 서버(100) 또는 하나의 사용자 단말(200) 각각이 직접 대상 드라마 콘텐츠의 시놉시스 텍스트 정보를 획득(예를 들어, 사용자로부터 직접 입력)하여, 시놉시스 텍스트 분석 동작을 수행할 수도 있다. 또는, 다양한 실시예들에 따르면, 복수의 사용자 단말(200)들 및/또는 복수의 서버(100)들이 마련되어, 복수의 사용자 단말(200)들 및/또는 복수의 서버(100)들 간에 대상 드라마 콘텐츠의 시놉시스 텍스트 정보 송수신 및 시놉시스 텍스트 분석 동작이 각 동작

별로 분배되어, 별개로 실시될 수도 있다.

- [0091] 도 5는 다양한 실시예들에 따른 시놉시스 텍스트 분석을 위한 전자 장치의 동작 방법을 도시한 도면이다.
- [0092] 다양한 실시예들에 따르면, 동작 S510에서, 서버(100)는 임베딩 된 시놉시스 텍스트 정보를 획득할 수 있다.
- [0093] 예를 들어, 임베딩 된 시놉시스 텍스트 정보는 서버(100)가 시놉시스 텍스트 정보를 임베딩(embedding) 처리하여 벡터(vector)화한 것일 수 있다.
- [0094] 예를 들어, 임베딩 된 시놉시스 텍스트 정보는 사용자 단말(200)가 상기 시놉시스 텍스트 정보를 AI 엔진에 입력 가능하도록 임베딩(embedding) 처리하여 벡터(vector)화 한 것일 수 있다.
- [0095] 예를 들어, 상기 시놉시스 텍스트 정보에 대한 임베딩(embedding) 처리는 자연어를 수치화 및 방향화하는 것을 의미할 수 있으며, 상기 자연어는 한국어, 영어 등 사용자 단말(200)가 인식 가능한 모든 언어를 포함할 수 있다.
- [0096] 예를 들어, 시놉시스 텍스트 정보는 단순 단어 단위에서 임베딩 처리가 수행될 수 있다. 또는, 시놉시스 텍스트 정보는 형태소(morpheme) 단위의 임베딩 기법인 'FastText' 기법에 의해 임베딩 될 수 있으며, 'FastText' 기법에 따라 형태소 단위에서 임베딩 처리가 수행될 수 있다.
- [0097] 구체적인 일 예시로서, 도 6은 단어 기반 단순 가공 형태의 임베딩 처리 예시 및 한글 자소 기반 형태소 가공 형태의 임베딩 처리 예시를 나타낸 도면이다. 도 6(a)는 '카리스마'라는 단어가 시놉시스 텍스트에 포함되어 있는 경우 이를 단순 가공 형태의 임베딩 처리한 결과를 나타낸 도면으로서, '카리스마'라는 단어가 포함되어 활용되는 어절들을 유사성에 따라 리스트화 한 것이다. 한편, 도 6(b)는 '카리스마'라는 단어가 시놉시스 텍스트에 포함되어 있는 경우 이를 형태소 분석 가공 형태의 임베딩 처리한 결과를 나타낸 도면으로서, '카리스마'라는 단어와 의미가 유사한 단어들을 유사성에 따라 리스트화 한 것이다.
- [0098] 이 때, 동작 S510에 필요한 임베딩 처리 과정은 서버(100) 또는 사용자 단말(200)와 연동된 AI 엔진의 기계 학습(machine learning)에 기반한 임베딩 학습을 통해 수행될 수 있다.
- [0099] 예를 들어, 임베딩 학습을 수행하는 AI 엔진은 한글, 영어 등과 같은 자연어를 처리함에 기초하여 임베딩 용 데이터를 획득할 수 있으며, 획득한 임베딩 용 데이터를 형태소 단위로 분석하고 벡터화하여 'FastText' 모델을 학습할 수 있다.
- [0100] 다양한 실시예들에 따르면, 동작 S520에서, 서버(100)는 임베딩 된 시놉시스 텍스트 정보에 기초하여 대상 드라마 콘텐츠의 장르를 판단할 수 있다.
- [0101] 예를 들어, 서버(100)는 임베딩 된 시놉시스 텍스트 정보에 대한 컨텍스트(context) 벡터를 획득하여 대상 드라마 콘텐츠의 장르를 판단할 수 있다.
- [0102] 예를 들어, 임베딩 된 시놉시스 텍스트 정보에 대한 컨텍스트 벡터는 시놉시스 텍스트 정보에 포함된 모든 단어 정보들을 압축한 벡터이거나, 또는 임베딩 처리가 된 시놉시스 텍스트 정보에 포함된 모든 벡터들을 합친 벡터일 수 있다.
- [0103] 예를 들어 임베딩 된 시놉시스 텍스트 정보에 대한 컨텍스트 벡터는, 시놉시스 텍스트 정보에 포함된 내용을 바탕으로 드라마 장르를 지시하기 위한 벡터 정보일 수 있다.
- [0104] 이 때, 동작 S520에 필요한 장르 판단 과정은 서버(100) 또는 사용자 단말(200)와 연동된 AI 엔진의 기계 학습(machine learning)에 기반한 앙상블 학습(ensemble learning)을 통해 수행될 수 있다. 여기서, 앙상블 학습은 복수의 기계 학습 방법들을 사용하고 그 결과들을 결합하여 최종 결과를 도출해내는 학습 방법을 의미한다.
- [0105] 예를 들어, 앙상블 학습을 수행하는 AI 엔진은 AI 엔진 자신이 수행 가능한 복수의 기계 학습 모델들 중 일부를 선택하여 드라마 장르 판단에 대한 앙상블 학습을 수행할 수 있다.
- [0106] 예를 들어, 앙상블 학습을 수행하는 AI 엔진은 임베딩 된 시놉시스 텍스트 정보에 대한 컨텍스트 벡터를 정답 데이터로 학습하고, 복수의 기계 학습 모델들 중 성능이 우수하거나 정확도가 높은 3가지의 기계 학습 모델을 파악하여 앙상블 학습을 수행할 수 있다.
- [0107] 예를 들어, 앙상블 학습을 수행하는 AI 엔진은 Catboost Classifier, Ridge Classifier, Linear Discriminant

Analysis, Extreme Gradient Boosting, Light Gradient Boosting Machine, Logistic Regression, SVM-Linear Kernel, Random Forest Classifier, Extra Trees Classifier, Gradient Boosting Classifier, K Neighbors Classifier, Naïve Bayes, Decision Tree Classifier, Ada Boost Classifier 및/또는 Quadratic Discriminant Analysis와 같은 모델들을 활용할 수 있으며, 만약 앙상블 학습을 수행하는 AI 엔진이 Logistic Regression, Catboost Classifier 및 Light Gradient Boosting Machine 모델들이 우수한 성능을 가진다고 판단하거나 정확도가 높다고 판단되는 경우 해당 3가지의 모델을 통해 앙상블 학습을 수행할 수 있다.

- [0108] 예를 들어, 임베딩 된 시놉시스 텍스트 정보에 대한 앙상블 학습의 결과로서 판단된 장르는, 보편적으로 드라마 장르를 분류하는 여러 카테고리 중 하나로서 판단될 수 있다.
- [0109] 구체적으로, 서버(100)에는 보편적으로 드라마 장르를 분류하는 복수의 카테고리들이 미리 설정될 수 있다. 드라마 장르 분류에 대한 일 예시로서, 드라마 장르를 분류하는 복수의 카테고리들은 '가족:0', '경제:1', '로맨스/멜로:2', '리얼리티:3', '막장:4', '법률:5', '복수:6', '사극:7', '성장:8', '수사/범죄/스릴러/미스터리:9', '액션:10', '오피스:11', '음식/요리:12', '의학:13', '정치:14', '청춘:15', '코미디:16', '판타지:17', '풍자:18', '휴머니즘:19', 및 '힐링:20' 와 같이 구성될 수 있다.
- [0110] 만약 대상 드라마 콘텐츠에 대한 시놉시스 텍스트 정보가 “회사 가기 싫은 사람들의 아주 사소하고도 위대한 이야기. 이 시대의 평범한 직장인들을 위한 초밀착 리얼 오피스 드라마” 와 같이 입력되었다면, 이 시놉시스 텍스트 정보를 임베딩 처리하여 획득한 컨텍스트 벡터는 '직장인', '오피스'와 같은 용어들을 종합함에 따라 카테고리 '오피스:11'를 지시할 수 있으며, 그에 따라서 서버(100)는 대상 드라마 콘텐츠에 대한 장르를 오피스 드라마에 해당한다고 판단할 수 있다.
- [0111] 또는 만약 대상 드라마 콘텐츠에 대한 시놉시스 텍스트 정보가 “한국의 실리콘 벨리에서 성공을 꿈꾸며 스타트업에 뛰어들어 청춘들의 시작과 성장을 그린 드라마” 와 같이 입력되었다면, 이 시놉시스 텍스트 정보를 임베딩 처리하여 획득한 컨텍스트 벡터는 '성공', '꿈꾸며', '스타트업', '청춘', '성장'과 같은 용어들을 종합함에 따라 카테고리 '성장:8'을 지시할 수 있으며, 그에 따라서 서버(100)는 대상 드라마 콘텐츠에 대한 장르를 성장 드라마에 해당한다고 판단할 수 있다.
- [0112] 다양한 실시예들에 따르면, 동작 S530에서, 서버(100)는 대상 드라마 콘텐츠에 대해 판단된 장르를 기초로 하여, 대상 드라마 콘텐츠를 위해 추천할 캐릭터들 판단하고 추천 캐릭터들을 위해 추천할 배우들을 판단할 수 있다.
- [0113] 예를 들어, 서버(100)는 대상 드라마 콘텐츠에 대해 판단된 장르에 출연했던 배우들의 목록을 획득할 수 있으며, 해당 배우들이 연기했던 드라마 상 캐릭터들의 인물 설명을 임베딩 하여 캐릭터 임베딩 정보들을 획득할 수 있다. 즉, 서버(100)는 대상 드라마 콘텐츠에 대해 판단된 장르에 출연했던 배우들이 연기했던 드라마 상 캐릭터들을 캐릭터 벡터로서 계산할 수 있다.
- [0114] 예를 들어, 서버(100)는 계산한 캐릭터 벡터들과 기존에 획득했던 임베딩 된 시놉시스 텍스트 정보에 대한 컨텍스트 벡터 사이의 유사성을 비교하여 해당 시놉시스 텍스트 정보에 적합한 캐릭터들을 찾아내고, 해당 캐릭터들을 연기한 배우들을 판단할 수 있다.
- [0115] 이 때, 동작 S530에 필요한 추천 캐릭터 및 추천 배우 판단 과정은 서버(100) 또는 사용자 단말(200)와 연동된 AI 엔진의 기계 학습에 기반한 임베딩 학습을 통해 수행될 수 있다.
- [0116] 예를 들어, 임베딩 학습을 수행하는 AI 엔진은 드라마 콘텐츠에 대해 판단된 장르에 출연했던 배우들의 목록을 획득할 수 있으며, 해당 배우들이 연기했던 드라마 상 캐릭터들의 인물 설명을 'FastText' 모델에 따라 처리함에 기초하여 캐릭터 벡터들을 획득할 수 있다.
- [0117] 예를 들어, 서버(100)는 찾아낸 캐릭터들 배우들에 특성 문구를 추가하여 그 벡터를 더함으로써 더 적합한 캐릭터를 추천할 수도 있다. 구체적으로, 대상 드라마 콘텐츠에 대해 판단된 장르에 출연했던 배우들이 연기했던 드라마 상 캐릭터들에 대하여 계산한 캐릭터 벡터들에, 특성 문구가 임베딩 된 벡터를 더함으로써 유사성을 다시 판단하여 더 적합한 캐릭터를 추천할 수 있다. 일 예로, 특성 문구가 '카리스마 넘치는', '재벌가', 또는 '나쁜 남자'라면, 각 특성 문구가 임베딩 된 벡터를 계산하고 이미 획득한 캐릭터 벡터들과 합산하여 새로운 캐릭터 벡터를 획득할 수 있으며, 이를 다시 임베딩 된 시놉시스 텍스트 정보에 대한 컨텍스트 벡터와 유사성을 비교하여 더 적합한 캐릭터 및 그에 해당하는 배우를 추천할 수 있다.
- [0118] 다양한 실시예들에 따르면, 동작 S540에서, 서버(100)는 추천된 배우들의 조합을 위한 추천 드라마들을 판단할

수 있다.

- [0119] 예를 들어, 사용자는 서버(100)가 판단한 추천 배우들의 전부 또는 일부를 추천 드라마 판단을 위한 입력 값으로 설정할 수 있다.
- [0120] 예를 들어, 서버(100)는 추천된 배우들에 대해 임베딩 된 벡터들을 계산하고, 계산된 벡터들의 합을 추천 드라마의 벡터로서 활용할 수 있다. 즉, 추천된 배우들의 조합에 대해 임베딩 된 벡터를 추천 드라마의 벡터로서 활용할 수 있다.
- [0121] 예를 들어, 서버(100)는 문장, 문단 또는 문서 단위의 임베딩 기법인 'Doc2Vex' 기법을 통해 드라마들과 각 드라마들에 출연했던 배우들 간의 관계를 벡터로 임베딩 할 수 있다.
- [0122] 이 때, 동작 S540에 필요한 추천 캐릭터 및 추천 배우 판단 과정은 서버(100) 또는 사용자 단말(200)와 연동된 AI 엔진의 기계 학습에 기반한 임베딩 학습을 통해 수행될 수 있다.
- [0123] 예를 들어, 임베딩 학습을 수행하는 AI 엔진은 드라마들 및 각 드라마들에 출연했던 배우들에 대한 정보를 획득할 수 있으며, 각 드라마들을 'Doc2Vex' 모델에 따라 처리함에 기초하여 드라마 임베딩 벡터들을 획득할 수 있다. 획득한 드라마 임베딩 벡터들은, 추천된 배우들의 조합에 대해 임베딩 된 벡터와 비교하여 유사성이 판단될 수 있다.
- [0124] 구체적인 일 예시로서, 도 7은 추천된 배우들의 조합에 따른 드라마 추천의 예시를 나타낸 도면이다. 도 7(a)는 '조보아', '정연주', '박해진', '최광일' 및 '우정국' 배우들의 조합에 대하여 추천될 수 있는 드라마들을 유사성에 따라 리스트화 한 것이다. 한편, 도 7(b)는 다른 배우들은 동일하되 '조보아' 대신 '김혜리'를 입력 값 변경하여 달리 구성된 배우들의 조합에 대해 추천될 수 있는 드라마들을 유사성에 따라 리스트화 한 것으로서, 도 7(a)의 추천 드라마 리스트와는 일부 차이가 있음을 확인할 수 있다. '조보아' 배우에 대하여 임베딩 된 벡터가 '김혜리' 배우에 대하여 임베딩 된 벡터와 다르므로, 추천된 배우들에 대해 임베딩 된 벡터들을 합산한 결과도 달라지게 되어 추천 드라마 리스트 역시 차이가 발생하게 된다.
- [0125] 다양한 실시예들에 따르면, 동작 S550에서, 서버(100)는 추천된 드라마들을 바탕으로 예측 시청률 정보를 판단할 수 있다.
- [0126] 예를 들어, 상기 대상 드라마 콘텐츠에 예측 시청률 정보는 수치화된 정보로서 출력될 수 있으며, 백분율 (%) 단위로서 표시될 수 있다.
- [0127] 예를 들어, 예측 시청률 정보는 추천된 드라마들의 시청률을 바탕으로 구성되는 것일 수 있다.
- [0128] 예를 들어, 예측 시청률 정보는 추천된 드라마들 중 유사성이 가장 높은 드라마의 전 회 시청률 값들을 포함할 수 있다.
- [0129] 예를 들어, 예측 시청률 정보는 추천된 드라마들 중 유사성이 가장 높은 드라마의 전 회 평균 시청률 값들을 포함할 수 있다.
- [0130] 예를 들어, 예측 시청률 정보는 추천된 드라마들 중 유사성이 가장 높은 드라마의 전 회 시청률 값들을 포함하고, 전 회 시청률 값들 중 최저 시청률 값 및/또는 최고 시청률 값을 별도로 표시할 수 있다.
- [0131] 예를 들어, 예측 시청률 정보는 추천된 모든 드라마들의 시청률을 평균값으로 한 시청률 값을 포함할 수 있다.
- [0132] 예를 들어, 예측 시청률 정보는 추천된 드라마들 중 가장 높은 시청률을 기록했던 드라마의 시청률과 동일한 시청률 값을 포함할 수 있다.
- [0134] 2.3. AI 엔진 획득 동작
- [0135] 인공지능 시스템은 인간 수준의 지능을 구현하는 컴퓨터 시스템이며, 기존 Rule 기반 스마트 시스템과 달리 기계가 스스로 학습하고 판단하며 똑똑해지는 시스템이다. 인공지능 시스템은 사용할수록 인식이 향상되고 사용자 취향을 보다 정확하게 이해할 수 있게 되어, 기존 Rule 기반 스마트 시스템은 점차 딥러닝 기반 인공지능 시스템으로 대체되고 있다.
- [0136] 인공지능 기술은 기계 학습(딥러닝) 및 기계 학습을 활용한 요소 기술들로 구성된다.
- [0137] 기계 학습은 입력 데이터들의 특징을 스스로 분류/학습하는 알고리즘 기술이며, 요소 기술은 딥러닝 등의 기계

학습 알고리즘을 활용하여 인간 두뇌의 인지, 판단 등의 기능을 모사하는 기술로서, 언어적 이해, 시각적 이해, 추론/예측, 지식 표현, 동작 제어 등의 기술 분야로 구성된다.

- [0138] 인공지능 기술이 응용되는 다양한 분야는 다음과 같다. 언어적 이해는 인간의 언어/문자를 인식하고 응용/처리하는 기술로서, 자연어 처리, 기계 번역, 대화시스템, 질의 응답, 음성 인식/합성 등을 포함한다. 시각적 이해는 사물을 인간의 시각처럼 인식하여 처리하는 기술로서, 객체 인식, 객체 추적, 영상 검색, 사람 인식, 장면 이해, 공간 이해, 영상 개선 등을 포함한다. 추론 예측은 정보를 판단하여 논리적으로 추론하고 예측하는 기술로서, 지식/확률 기반 추론, 최적화 예측, 선호 기반 계획, 추천 등을 포함한다. 지식 표현은 인간의 경험정보를 지식데이터로 자동화 처리하는 기술로서, 지식 구축(데이터 생성/분류), 지식 관리(데이터 활용) 등을 포함한다. 동작 제어는 차량의 자율 주행, 로봇의 움직임을 제어하는 기술로서, 움직임 제어(항법, 충돌, 주행), 조작 제어(행동 제어) 등을 포함한다.
- [0139] 이하의 설명에서는 서버(100)가 AI 엔진 획득 동작을 수행함을 전제로 다양한 실시예들에 대하여 설명하였으나, 다양한 실시예들에 따르면, 서버(100) 외부의 다른 서버가 AI 엔진 획득 동작을 수행할 수도 있다. 또는, 다양한 실시예들에 따르면, 복수의 사용자 단말(200)들 및/또는 복수의 서버(100)들이 마련되어, 복수의 사용자 단말(200)들 및/또는 복수의 서버(100)들 간에 AI 엔진 획득 동작의 각 동작이 분배되어, 별개로 실시될 수도 있다. 후술하는 과정에 의해 획득되는 AI 엔진은, 전송한 도 3의 시청률 예측 및 출력 동작 및 도 5의 시놉시스 텍스트 분석 동작에서 언급되는 기계 학습(machine learning)에 활용될 수 있다.
- [0141] 도 8은 다양한 실시예들에 따른 시놉시스 텍스트 분석을 위한 AI (artificial intelligence) 엔진을 획득하는 과정을 도시한 도면이다.
- [0142] 도 8을 참조하면, 예를 들어, 서버(100)는, 종래의 드라마 장르들에 대한 데이터를 수집할 수 있다. 예를 들어, 종래의 드라마 장르들에 대한 데이터는, 일정 기간 동안 송출된 드라마들의 장르에 대한 데이터들이 수집될 수 있다. 예를 들어 드라마 장르들에 대한 데이터는 시청자의 성별, 시청자의 연령대, 드라마의 송출 시간대 등 다양한 특성들 중 하나 이상의 특성에 기초하여 수집될 수 있다.
- [0143] 예를 들어, 서버(100)는, 종래의 드라마들에 대한 데이터에 기초하여 학습 데이터를 획득할 수 있다. 예를 들어, 학습 데이터는 훈련용 데이터 및 테스트용 데이터를 포함할 수 있다.
- [0144] 예를 들어, 서버(100)는, 학습 데이터를 훈련용 데이터 및 테스트용 데이터로 분류할 수 있다. 예를 들어, 훈련용 데이터는 모델을 학습시키는데 활용되고, 테스트용 데이터는 학습된 모델을 검증 및 업데이트 하는데 활용될 수 있다.
- [0145] 예를 들어, 서버(100)는 학습된 모델에 기초하여 테스트용 데이터를 통해 피드백 정보를 획득할 수 있다.
- [0146] 예를 들어, 피드백 정보는, 서버(100) 관리자로부터 입력될 수 있다.
- [0147] 예를 들어, 서버(100)는, 피드백에 기초하여 학습된 모델을 업데이트할 수 있다.
- [0148] 예를 들어, 서버(100)는 피드백에 기초하여 학습된 모델의 특성을 수정하고, 수정된 모델에 기초하여 테스트용 드라마 데이터를 가공하고, 다시 피드백을 획득하여 업데이트 하는 과정을 반복할 수 있다.
- [0149] 예를 들어, 서버(100)는 반복 횟수가 미리 설정된 임계값 (예를 들어, 3-5회)에 도달하면, 동작을 종료하고 학습된 모델을 AI 엔진으로서 획득할 수 있다.
- [0151] 본 명세서와 도면에 개시된 본 발명의 실시 예들은 본 발명의 기술 내용을 쉽게 설명하고 본 발명의 이해를 돕기 위해 특정 예를 제시한 것일 뿐이며, 본 발명의 범위를 한정하고자 하는 것은 아니다. 즉 본 발명의 기술적 사상에 바탕을 둔 다른 변형 예들이 실시 가능하다는 것은 본 발명의 속하는 기술 분야에서 통상의 지식을 가진 자에게 자명한 것이다. 또한 상기 각각의 실시 예는 필요에 따라 서로 조합되어 운용할 수 있다. 예컨대, 본 발명의 모든 실시 예는 일부분들이 서로 조합되어 시스템에 의해 구현될 수 있다.
- [0152] 또한, 본 발명에 따른 시스템 등에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다.
- [0153] 이와 같이, 본 발명의 다양한 실시 예들은 특정 관점에서 컴퓨터 리드 가능 기록 매체(computer readable

recording medium)에서 컴퓨터 리드 가능 코드(computer readable code)로서 구현될 수 있다. 컴퓨터 리드 가능 기록 매체는 컴퓨터 시스템에 의해 리드될 수 있는 데이터를 저장할 수 있는 임의의 데이터 저장 디바이스이다. 컴퓨터 리드 가능 기록 매체의 예들은 읽기 전용 메모리(read only memory: ROM)와, 랜덤-접속 메모리(random access memory: RAM)와, 콤팩트 디스크- 리드 온니 메모리(compact disk-read only memory: CD-ROM)들과, 마그네틱 테이프(magnetic tape)들과, 플로피 디스크(floppy disk)들과, 광 데이터 저장 디바이스들, 및 캐리어 웨이브(carrier wave)들(인터넷을 통한 데이터 송신 등)을 포함할 수 있다. 컴퓨터 리드 가능 기록 매체는 또한 네트워크 연결된 컴퓨터 시스템들을 통해 분산될 수 있고, 따라서 컴퓨터 리드 가능 코드는 분산 방식으로 저장 및 실행된다. 또한, 본 발명의 다양한 실시 예들을 성취하기 위한 기능적 프로그램들, 코드, 및 코드 세그먼트(segment)들은 본 발명이 적용되는 분야에서 숙련된 프로그래머들에 의해 쉽게 해석될 수 있다.

[0154] 또한 본 발명의 다양한 실시 예들에 따른 장치 및 방법은 하드웨어, 소프트웨어 또는 하드웨어 및 소프트웨어의 조합의 형태로 실현 가능하다는 것을 알 수 있을 것이다. 이러한 소프트웨어는 예를 들어, 삭제 가능 또는 재기록 가능 여부와 상관없이, ROM 등의 저장 장치와 같은 휘발성 또는 비 휘발성 저장 장치, 또는 예를 들어, RAM, 메모리 칩, 장치 또는 집적 회로와 같은 메모리, 또는 예를 들어 콤팩트 디스크(compact disk: CD), DVD, 자기 디스크 또는 자기 테이프 등과 같은 광학 또는 자기적으로 기록 가능함과 동시에 기계(예를 들어, 컴퓨터)로 읽을 수 있는 저장 매체에 저장될 수 있다. 본 발명의 다양한 실시 예들에 따른 방법은 제어부 및 메모리를 포함하는 컴퓨터 또는 이와 같은 메모리 또는 컴퓨터를 포함한 차량 등에 의해 구현될 수 있고, 이러한 메모리는 본 발명의 실시 예들을 구현하는 명령들을 포함하는 프로그램 또는 프로그램들을 저장하기에 적합한 기계로 읽을 수 있는 저장 매체의 한 예임을 알 수 있을 것이다.

[0155] 따라서, 본 발명은 본 명세서의 청구항에 기재된 장치 또는 방법을 구현하기 위한 코드를 포함하는 프로그램 및 이러한 프로그램을 저장하는 기계(컴퓨터 등)로 읽을 수 있는 저장 매체를 포함한다. 또한, 이러한 프로그램은 유선 또는 무선 연결을 통해 전달되는 통신 신호와 같은 임의의 매체를 통해 전자적으로 이송될 수 있고, 본 발명은 이와 균등한 것을 적절하게 포함한다.

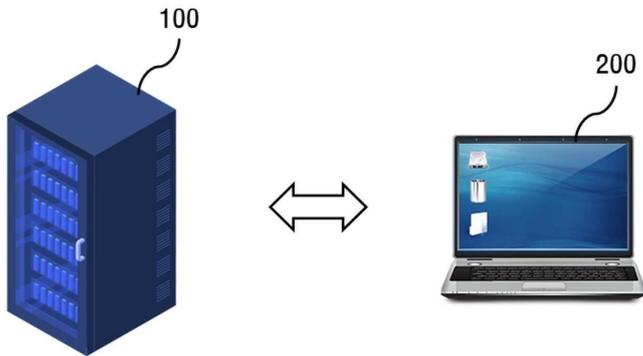
[0156] 상기에서는 본 발명의 실시 예들을 참조하여 설명하였지만, 본 명세서와 도면에 개시된 본 발명의 실시 예들은 본 발명의 기술 내용을 쉽게 설명하고 본 발명의 이해를 돕기 위해 특정 예를 제시한 것일 뿐이며 본 발명의 범위를 한정하고자 하는 것은 아니다. 또한 앞서 설명된 본 발명에 따른 실시 예들은 예시적인 것에 불과하며, 당해 분야에서 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 범위의 실시 예가 가능하다는 점을 이해할 것이다. 따라서, 본 발명의 진정한 기술적 보호 범위는 다음의 특허청구범위에 의해서 정해져야 할 것이다.

부호의 설명

- [0157] 100: 서버
- 200: 사용자 디바이스
- 210: 입/출력부 220: 통신부
- 230: 데이터베이스 240: 프로세서

도면

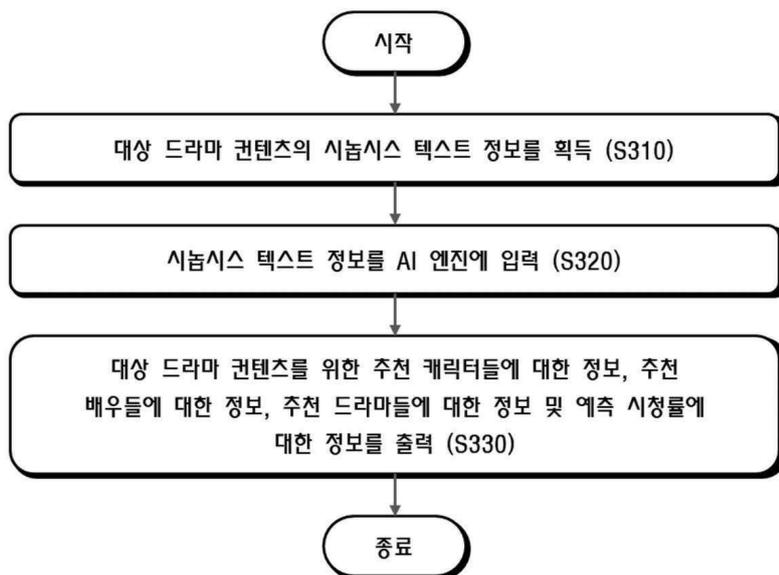
도면1



도면2



도면3



도면4

시놉시스*

현상의 백에 결함하지 않고 스스로 광과 시량을 이루기 위해 노력하는 청춘들의 성장 기록을 담은 드라마

401

개인 사항하기

시놉시스의 장르는 **미담연극** 으로 예측되었습니다.

이에 어울리는 배우는 아래와 같습니다.

주연 1* 김남길 **주연 2*** 박소담

△ 캐리스마 있는 △ 바보 같고 장난스러운

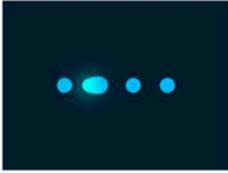
글로벌 캐릭터의 특성을 추가해 주세요

조연 1* 변우석 **조연 2*** 권수현 **조연 3*** 신애라

△ 날카로운 캐리스마 △ 아무런 △ 어릴 때 상처

글로벌 캐릭터의 특성을 추가해 주세요

글로벌 캐릭터의 특성을 추가해 주세요



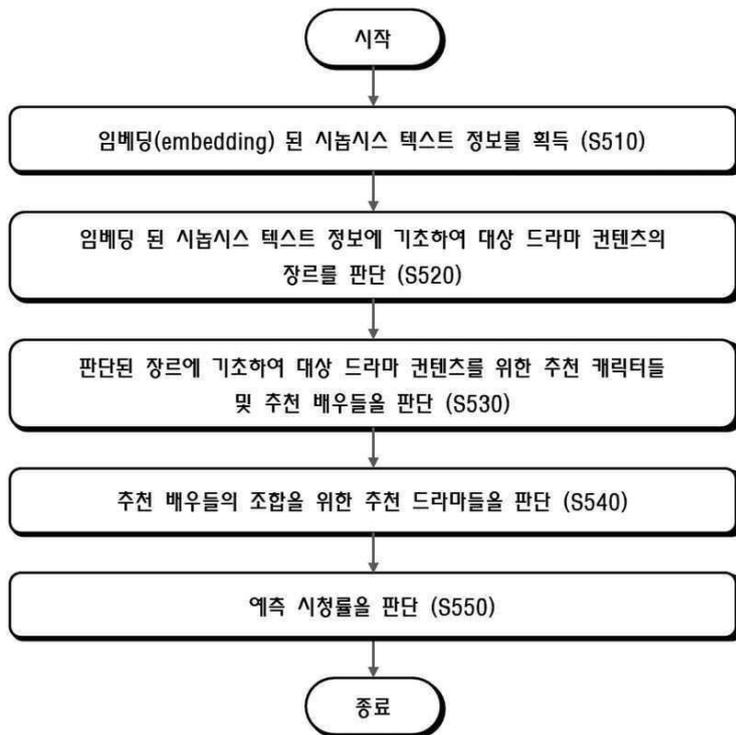
청춘기록 과 유사한 조합입니다.

청춘기록 시청률

회차	방송일	대한민국(전국)	서울(수도권)
제1회	9월 7일	6.362%	7.791%
제2회	9월 8일	6.802%	8.217%
제3회	9월 14일	7.200%	9.102%
제4회	9월 15일	7.823%	9.602%
제5회	9월 21일	7.801%	9.037%
제6회	9월 22일	6.995%	8.409%
제7회	9월 28일	7.730%	9.368%
제8회	9월 29일	7.734%	9.225%
제9회	10월 5일	7.404%	8.979%
제10회	10월 6일	8.238%	10.063%
제11회	10월 12일	8.262%	9.618%
제12회	10월 13일	7.798%	9.439%
제13회	10월 19일	7.760%	9.313%
제14회	10월 20일	7.772%	9.617%
제15회	10월 26일	7.625%	9.134%
제16회	10월 27일	8.740%	10.728%
평균 시청률		7.628%	9.227%

405

도면5



도면6

<pre>1 vector.most_similar("카리스마")</pre>	<pre>1 vector.most_similar("카리스마")</pre>
<pre>[('카리스마형', 0.811809778213501), ('카리스마넘치는', 0.8033837080001831), ('카리스마만', 0.7894523739814758), ('카리스마:', 0.7826545238494873), ('카리스마의', 0.7793933749198914), ('카리스마.', 0.7785433530807495), ('카리스마!', 0.7717763185501099), ('카리스마가', 0.7617940306663513), ('카리스마있고', 0.7511934638023376), ('카리스마적', 0.7474747896194458)]</pre>	<pre>[('매력', 0.726121723651886), ('인간미', 0.6990047097206116), ('연륜', 0.6927657723426819), ('면모', 0.6809433102607727), ('풍모', 0.680864691734314), ('외모', 0.6766476631164551), ('지략', 0.6704614162445068), ('인품', 0.6683163642883301), ('성품', 0.6673632264137268), ('순둥이', 0.663332462310791)]</pre>
(a)	(b)

도면7

```
1 get_sim_drama(doc2vec_model_100_5_1000, ['조보아', '정연주', '박해진', '최광일', '우정국'])
[(('D_포레스트', 0.6816328763961792),
('D_치즈인더트랩', 0.46372491121292114),
('D_탁구공', 0.4591841995716095),
('D_본 어게인', 0.4558047354221344),
('D_열혈장사꾼', 0.45390427112579346),
('D_소문난 칠공주', 0.4467722177505493),
('D_드라마 스페셜 - SLOW', 0.4371607005596161),
('D_드림하이 2', 0.43274176120758057),
('D_실종느와르 M', 0.42750146985054016),
('D_닥치고 꽃미남밴드', 0.4187033772468567)])
```

(a)

```
1 get_sim_drama(doc2vec_model_100_5_1000, ['김혜리', '정연주', '박해진', '최광일', '우정국'])
[(('D_포레스트', 0.5508899092674255),
('D_신데렐라와 네 명의 기사', 0.47394001483917236),
('D_열혈장사꾼', 0.4707231819629669),
('D_어머님은 내 며느리', 0.46129968762397766),
('D_치즈인더트랩', 0.4469233453273773),
('D_장미의 전쟁', 0.44375500082969666),
('D_정글피쉬 1', 0.43066078424453735),
('D_점프 2', 0.4224756956100464),
('D_엄마가 뽀났다', 0.4206192195415497),
('D_실종느와르 M', 0.4146650731563568)])
```

(b)

도면8

