



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2023년08월24일
(11) 등록번호 10-2570966
(24) 등록일자 2023년08월22일

(51) 국제특허분류(Int. Cl.)
G06Q 30/02 (2023.01) G06N 3/04 (2023.01)
G06N 3/08 (2023.01)

(52) CPC특허분류
G06Q 30/0202 (2023.01)
G06N 3/04 (2023.01)

(21) 출원번호 10-2021-0000132

(22) 출원일자 2021년01월04일

심사청구일자 2021년01월04일

(65) 공개번호 10-2022-0098424

(43) 공개일자 2022년07월12일

(56) 선행기술조사문헌

KR1020190134934 A*

Jonghoon Jin 외 2명, 'Flattened Convolutional Neural Networks for Feedforward Acceleration', Neural and Evolutionary Computing, 2015년 11월, pp.1-11*

'tensorflow의 tf.concat에 대해서 알아보기(axis = 0, 1에 따른 변화)', 티스토리(2019.03.18. 게재)*

'밀바닥 딥러닝_7장 합성곱 신경망(CNN)', 티스토리(2019.01.27. 게재)*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

연세대학교 원주산학협력단

강원도 원주시 흥업면 연세대길 1

(72) 발명자

안재준

서울특별시 동작구 보라매로5가길 7 캐릭터그린빌 1903호

최세환

서울특별시 동대문구 한천로8가길 17-1

(뒷면에 계속)

(74) 대리인

오영진

전체 청구항 수 : 총 3 항

심사관 : 권태현

(54) 발명의 명칭 **고객의 미래 행동 예측 장치**

(57) 요약

본 발명의 일 실시예에 따른 고객의 미래 행동 예측 장치는, 고객의 시간별 데이터 및 행동 데이터를 포함하는 로테이터에 기초하여 상기 고객의 향후 행동을 예측할 수 있는 고객의 미래 행동 예측 장치에 관한 것으로, 상기 로테이터를 입력받는 입력모듈; 상기 입력모듈로부터 전달받은 로테이터를 가공하여 시간×행동으로 정의된 2차원 매트릭스를 생성하는 데이터 전처리모듈; 상기 2차원 매트릭스에 HCNN(Horizontal Convolution Neural Network)을 적용하여 상기 고객의 시간별 행동특징인 제1 특징 매트릭스를 추출하는 제1 특징 추출모듈; 및 상기 2차원 매트릭스에 VCNN(Vertical Convolution Neural Network)을 적용하여 상기 고객의 행동별 시간특징인 제2 특징 매트릭스를 추출하는 제2 특징 추출모듈;을 포함한다..

대표도 - 도2

100
200
300
400
500
600

(52) CPC특허분류

G06N 3/08 (2023.01)

(72) 발명자

임진영

강원도 원주시 흥업면 북원로 1412-11 휴먼스빌
305호

장인호

강원도 원주시 예술관길 15 원주 삼성아파트 3동
303호

박성중

경기도 고양시 일산동구 산두로 54 정발마을3단지
건영빌라 302동 104호

이 발명을 지원한 국가연구개발사업

과제고유번호	1711118846
과제번호	2020-51-0194
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	개인기초연구(과기정통부)(R&D)
연구과제명	정형 시계열데이터 분석을 위한 합성곱 기반의 새로운 딥러닝 모형 개발
기 여 율	1/1
과제수행기관명	연세대학교 원주산학협력단
연구기간	2020.06.01 ~ 2021.02.28

명세서

청구범위

청구항 1

고객의 시간별 데이터 및 행동 데이터를 포함하는 로데이터에 기초하여 상기 고객의 향후 행동을 예측할 수 있는 고객의 미래 행동 예측 장치에 있어서,

상기 로데이터를 입력받는 입력모듈;

상기 입력모듈로부터 전달받은 로데이터를 가공하여 시간×행동으로 정의된 2차원 매트릭스를 생성하는 데이터 전처리모듈;

상기 2차원 매트릭스에 HCNN(Horizontal Convolution Neural Network)을 적용하여 상기 고객의 시간별 행동특징인 제1 특징 매트릭스를 추출하는 제1 특징 추출모듈; 및

상기 2차원 매트릭스에 VCNN(Vertical Convolution Neural Network)를 적용하여 상기 고객의 행동별 시간특징인 제2 특징 매트릭스를 추출하는 제2 특징 추출모듈;

을 포함하고,

상기 제1 특징 추출모듈은, 상기 2차원 매트릭스에 Horizontal 1D 컨벌루션을 수행하여 제1 특징맵을 생성한 후, 상기 제1 특징맵에 concatenation을 적용하여 제1 특징 매트릭스를 생성하고,

상기 제2 특징 추출모듈은, 상기 2차원 매트릭스에 Vertical 1D 컨벌루션을 수행하여 제2 특징맵을 생성한 후, 상기 제2 특징맵에 concatenation을 적용하여 제2 특징 매트릭스를 생성하고,

상기 제1 특징 매트릭스인 horizontal feature matrix는 time x filter 형태로 시간별 특징을 표현하고,

상기 제2 특징 매트릭스인 vertical feature matrix는 filter x variable 형태로 시간에 따른 각 변수의 특징을 지니고,

상기 제1 특징 추출모듈 및 상기 제2 특징 추출모듈은 행과 열에 대한 공간정보를 유지하기 위해 concatenation을 적용하는 고객의 미래 행동 예측 장치.

청구항 2

삭제

청구항 3

삭제

청구항 4

청구항 1에 있어서,

상기 제1 특징 매트릭스 및 제2 특징 매트릭스에 평탄화(flatten)를 적용하는 평탄화 모듈; 및

상기 평탄화 모듈의 출력을 입력받아 상기 고객의 미래 행동을 추론하는 결정모듈;

을 더 포함하는 것을 특징으로 하는 고객의 미래 행동 예측 장치.

청구항 5

청구항 4에 있어서,

상기 결정모듈의 활성화 함수(activation function)은 시그모이드 함수(sigmoid function)인 것을 특징으로 하

는 고객의 미래 행동 예측 장치.

발명의 설명

기술 분야

[0001] 본 발명은 고객의 미래 행동 예측 장치에 관한 것으로, 구체적으로 2차원 정형시계열 데이터분석을 위한 합성곱 기반의 딥러닝 시스템을 활용하여 고객의 미래 행동 예측을 수행하는 장치에 관한 것이다.

배경 기술

[0003] 고객의 미래 행동 예측은 기업이 사전에 일어날 위험을 방지할 수 있게 도와주며, 다양한 마케팅 전략을 개발할 수 있게 해주는 중요한 방법으로, 미래 행동에는 이탈, 구매 및 방문 유무 등 다양한 분야가 존재한다.

[0004] 이탈 분야의 경우, 기존 고객을 유지하는 것은 새로운 고객을 유치시키는 것 보다 더욱 어렵고 많은 비용이 필요하기 때문에, 고객 이탈 예측은 중요한 분야로 인식된다.

[0005] 통신사에서는 기존 고객이 이탈하면 경쟁사를 이용할 확률이 높기 때문에 상당한 피해와 함께 통신 시장의 입지가 줄어들게 되며, 게임사는 기존 고객의 이탈이 단순히 사용자 수가 줄어드는 것에 그치지 않고, 함께 게임을 즐기는 다른 고객에게 악영향을 주어 수익에 상당한 영향을 미치게 된다.

[0006] 고객의 구매 및 방문 분야는 사전에 행동의 유무를 예측했을 때, 비용 절감과 기회비용을 증가시킬 수 있으며, 특히 유통업체에서는 고객의 방문 가능성을 예측하여 프로모션 비용을 크게 절감하고 투자 수익률을 높일 수 있으며, 나아가 방문 가능성이 낮은 고객들에게는 특별한 이벤트를 통하여 방문을 유도할 수 있다.

[0007] 이처럼 고객의 미래 행동 예측은 높은 이윤을 도출 가능한 중요한 요소이지만, 고객 데이터는 대부분 방대한 크기, 높은 차원, 불균형 비율 그리고 2차원 형태인 구조적 문제들을 지니고 있으며, 이러한 데이터 자체가 내포하고 있는 문제들이 고객의 미래 행동 예측을 한층 더 어렵게 만들고 있다.

[0008] 따라서 현재까지도 고객의 미래 행동 예측에 대한 연구가 활발히 진행되고 있는 상황이며, 고객의 미래 행동 예측에 대한 연구는 주로 고객의 행동과 시간 변화에 집중하고 있다.

[0009] 고객의 사소한 행동은 풍부한 잠재적 의미를 내포하고 있으며, 통화내역, 회선정보, 계정정보, 인구통계 프로파일, 불만사항 등 사소한 고객의 행동을 변수로 이용해 높은 고객 이탈 예측의 정확도를 보였다.

[0010] 고객의 시간 정보는 과거부터 현재까지의 패턴을 포함하고 있기 때문에, 미래 행동에 많은 영향을 끼치고 행동에 대한 포괄적 이해를 제공한다.

[0011] 고객의 행동과 시간 정보는 논의된 바와 같이 고객의 미래 행동의 예측에 있어 중요한 요소들이다.

[0012] 이러한 요소들을 반영하고 구조적 문제를 해결하기 위해 최근에는 딥러닝(Deep Learning)을 이용한 연구가 활발히 진행중이나, 이러한 고객의 미래 행동 예측의 경우 시계열적 특징이 추가된 2차원 정형시계열 데이터를 갖고 있으므로, 이에 대한 분석이 다수 요구된다.

[0013] 그러나 종래의 모델로는 입력 형태가 1차원으로 고정되어 있어 2차원 정형시계열 데이터에 적합하지 않고, 유일하게 2차원 형태가 가능한 RNN 계열은 최근 정보에 치우치는 경향이 있고, CNN 계열은 구조적 오류로 인해 정형시계열 데이터에 적합하지 않다는 문제점이 있다.

[0014] 한편, 하기 선행기술문헌은 고객별 정보분석을 통한 예측 전자상거래 서비스 방법에 관한 내용을 개시하고 있을 뿐 본 발명의 기술적 요지는 개시하고 있지 않다.

선행기술문헌

특허문헌

[0016] (특허문헌 0001) 대한민국 공개특허공보 제2000-0054695호

발명의 내용

해결하려는 과제

- [0017] 본 발명의 일 실시예에 따른 고객의 미래 행동 예측 장치는 전술한 문제점을 해결하기 위하여 다음과 같은 해결 과제를 목적으로 한다.
- [0018] 시계열 정보가 추가된 2차원 정형데이터 분석에 활용이 가능한 딥러닝 모델을 제안하고, 이를 이용한 고객의 미래 행동의 예측을 정확하고 신속하게 진행할 수 있는 고객의 미래 행동 예측 장치를 제공하는 것이다.
- [0019] 본 발명의 해결과제는 이상에서 언급된 것들에 한정되지 않으며, 언급되지 아니한 다른 해결과제들은 아래의 기재로부터 당해 기술분야에 있어서의 통상의 지식을 가진 자에게 명확하게 이해되어 질 수 있을 것이다.

과제의 해결 수단

- [0021] 본 발명의 일 실시예에 따른 고객의 미래 행동 예측 장치는, 고객의 시간별 데이터 및 행동 데이터를 포함하는 로테이터에 기초하여 상기 고객의 향후 행동을 예측할 수 있는 고객의 미래 행동 예측 장치에 관한 것으로, 상기 로테이터를 입력받는 입력모듈; 상기 입력모듈로부터 전달받은 로테이터를 가공하여 시간×행동으로 정의된 2차원 매트릭스를 생성하는 데이터 전처리모듈; 상기 2차원 매트릭스에 HCNN(Horizontal Convolution Neural Network)을 적용하여 상기 고객의 시간별 행동특징인 제1 특징 매트릭스를 추출하는 제1 특징 추출모듈; 및 상기 2차원 매트릭스에 VCNN(Vertical Convolution Neural Network)를 적용하여 상기 고객의 행동별 시간특징인 제2 특징 매트릭스를 추출하는 제2 특징 추출모듈;을 포함한다.
- [0022] 상기 제1 특징 추출모듈은, 상기 2차원 매트릭스에 Horizontal 1D 컨벌루션을 수행하여 제1 특징맵을 생성한 후, 상기 제1 특징맵에 concatenation을 적용하여 제1 특징 매트릭스를 생성하는 것이 바람직하다.
- [0023] 상기 제2 특징 추출모듈은, 상기 2차원 매트릭스에 Vertical 1D 컨벌루션을 수행하여 제2 특징맵을 생성한 후, 상기 제2 특징맵에 concatenation을 적용하여 제2 특징 매트릭스를 생성하는 것이 바람직하다.
- [0024] 상기 제1 특징 매트릭스 및 제2 특징 매트릭스에 평탄화(flatten)를 적용하는 평탄화 모듈; 및 상기 평탄화 모듈의 출력을 입력받아 상기 고객의 미래 행동을 추론하는 결정모듈;을 더 포함하는 것이 바람직하다.
- [0025] 상기 결정모듈의 활성화 함수(activation function)은 시그모이드 함수(sigmoid function)인 것이 바람직하다.

발명의 효과

- [0027] 본 발명의 일 실시예에 따른 고객의 미래 행동 예측 장치는, 1D-convolution과 concatenation을 통해 기존 딥러닝 방법들의 구조적 특징에 기인하여 발생하였던 문제점을 해결할 수 있는 효과를 기대할 수 있다.
- [0028] 특히 본 발명에서 적용되는 1D convolution은 일반적인 CNN에서 사용하는 2D convolution보다 연산량이 적기 때문에 분석 시간을 줄이는 측면에서 매우 중요한 의미를 가지며, 이를 토대로 기존 딥러닝 방법들의 문제점 해결 방법에 대하여 새로운 방향성을 제시할 수 있는 효과를 기대할 수 있다.
- [0029] 본 발명의 효과는 이상에서 언급된 것들에 한정되지 않으며, 언급되지 아니한 다른 효과들은 아래의 기재로부터 당해 기술분야에 있어서의 통상의 지식을 가진 자에게 명확하게 이해되어질 수 있을 것이다.

도면의 간단한 설명

- [0031] 도 1은 본 발명의 일 실시예에 따른 고객의 미래 행동 예측 장치에 적용되는 1D convolution layer의 개념을 설명하기 위한 이미지이다.
- 도 2는 본 발명의 일 실시예에 따른 고객의 미래 행동 예측 장치를 간략히 도시한 블록도이다.
- 도 3은 본 발명의 일 실시예에 따른 고객의 미래 행동 예측 장치에 적용되는 HV-Net의 전반적인 프로세스를 설명하기 위한 이미지이다.
- 도 4는 본 발명의 일 실시예에 따른 고객의 미래 행동 예측 장치에서 제1 특징 추출모듈 및 제2 특징 추출모듈에서 진행되는 1D convolution 및 concatenation의 과정을 설명하기 위한 이미지이다.
- 도 5는 본 발명의 일 실시예에 따른 제1 특징 추출모듈 및 제2 특징 추출모듈의 세부 구성을 도시한 개념도이다.
- 도 6은 본 발명의 일 실시예에 따른 고객의 미래 행동 예측 장치에 적용되는 모델의 정확도의 평가시 활용되는

confusion matrix이다.

도 7은 본 발명의 일 실시예에 따른 고객의 미래 행동 예측 장치에서 적용되는 모델과 다른 모델과의 성능차이를 확인하기 위한 도표이다.

발명을 실시하기 위한 구체적인 내용

- [0032] 첨부된 도면을 참조하여 본 발명에 따른 바람직한 실시예를 상세히 설명하되, 도면 부호에 관계없이 동일하거나 유사한 구성 요소는 동일한 참조 번호를 부여하고 이에 대한 중복되는 설명은 생략하기로 한다.
- [0033] 또한, 본 발명을 설명함에 있어서 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다. 또한, 첨부된 도면은 본 발명의 사상을 쉽게 이해할 수 있도록 하기 위한 것일 뿐, 첨부된 도면에 의해 본 발명의 사상이 제한되는 것으로 해석되어서는 아니 됨을 유의해야 한다.
- [0035] 본 발명의 일 실시예에 따른 고객의 미래 행동 예측 장치를 설명하기에 앞서 먼저 본 발명에 적용되는 기술적 특징인 Convolution layer 및 Latent Semantic Analysis(LSA)에 대하여 설명하도록 한다.
- [0036] CNN은 데이터를 작은 영역으로 나누어 부분적인 특징을 추출하여 공간 정보를 반영하는 모델로, 이미지 인식에서 뛰어난 성능을 보이는 모델이며, 이러한 CNN의 핵심은 공간적 정보를 추출하는 convolution layer이다.
- [0037] Convolution layer는 filter와 activation function으로 이루어지는데, Filter란 서로 다른 가중치로 구성된 matrix로, 이미지 공간에서 다양한 특징을 추출한다.
- [0038] Activation function(활성화 함수)은 추출된 특징을 비선형 형태로 표현하기 위한 방법으로, 주로 sigmoid, tanh, ReLU가 사용된다.
- [0039] Convolution layer의 방법은 크게 2-D 그리고 1-D convolution이 존재하는데, 두 방법의 차이는 반영하는 공간이다.
- [0040] 도 1(a)는 kernel size 2×2 , stride는 1인 2-D convolution 적용 과정을 도시하고 있으며, filter가 왼쪽 상단에서 시작해 오른쪽 하단까지 이동하며, Kernel size는 일반적으로 가로와 세로의 길이가 동일하게 적용되며, 대부분의 이미지 인식에 사용되는 형태이지만 많은 연산 시간이 요구된다.
- [0041] 도 1(b)는 kernel size는 2×5 , stride는 1인 1-D convolution의 연산과정으로, 1-D convolution의 Filter size는 2-D convolution과는 다르게 행과 열의 길이가 동일하지 않다.
- [0042] Filter는 위에서부터 순차적으로 하나의 row씩 이동해 feature map을 추출하며, 이를 통하여 2-D convolution의 문제점인 많은 computing 시간을 1-D convolution을 적용함으로써 해결할 수 있다.
- [0043] 한편, Singular value decomposition(SVD)는 선형대수에서 널리 사용되는 개념으로, 최근 빅데이터 분석에서 데이터 축약 및 노이즈 제거 방법으로 다양하게 활용된다.
- [0044] SVD 방법은 $m \times n$ 행렬인 A 를 하기 수식(1)로 분해하며, 여기에서 U 와 V 는 A 의 행 및 열에 대한 잠재적 의미를 표현한다.
- [0046] <수식 (1)>

$$A = U \Sigma V^T$$

상기 수식 (1)에서 Σ 는 $m \times n$ 의 고윳값 행렬이며, Σ 행렬의 k 개 값을 사용하여 새로운 근사 A_k 행렬을 도출하는데, 여기에서 A_k 는 A 의 노이즈 정보를 줄이고, 중요한 특징만을 부각하고, SVD는 LSA의 기반이 되는 중요한 방법론이다.

LSA는 데이터의 행과 열에 대한 잠재적 의미를 추론하여 분석하는 방법으로, LSA 방법을 통해 도출된 추상화 vector는 높은 유사도를 갖게 되며, LSA의 L1과 L2는 하기 수식 (2)로부터 도출된다.

<수식 (2)>

$$A_K = U_K \Sigma_K V_K^T$$

[0055] 상기 수식 (2)의 A_k 우변에 V_k 를 곱하면 $L1$, A_k 좌변에 U_k 를 곱하면 $L2$ 가 하기 수식 (3) 및 수식 (4)와 같이 생성되며, 여기에서 $L1$ 은 행의 중요 특징을 담고 있으며, $L2$ 는 열에 대한 중요 특징을 지닌 정보 행렬이다.

[0057] <수식 (3)>

$$U_k \sum K = L1$$

[0058]

[0060] <수식 (4)>

$$\sum K V_{KT} = L2$$

[0061]

[0063] 이하, 본 발명의 일 실시예에 따른 고객의 미래 행동 예측 장치에 대하여 도 2 내지 도 7을 참조하여 설명하도록 한다.

[0064] 본 발명의 일 실시예에 따른 고객의 미래 행동 예측 장치는 고객의 시간별 데이터 및 행동 데이터를 포함하는 로데이터(Raw Data)에 기초하여 고객의 향후 행동을 예측할 수 있는 고객의 미래 행동 예측 장치에 관한 것으로, 도 2에 도시된 바와 같이 입력모듈(100), 데이터 전처리모듈(200), 제1 특징 추출모듈(300), 제2 특징 추출모듈(400), 평탄화 모듈(500) 및 결정모듈(600)을 포함하도록 구성된다.

[0065] 입력모듈(100)은 벡터 형태인 로데이터(Raw Data)를 입력받는 기능을 수행하고, 데이터 전처리모듈(200)은 도 3에 도시된 바와 같이 벡터 형태인 로데이터를 가공하여 시간×행동으로 정의된 2차원 매트릭스를 생성하는 기능을 수행한다.

[0066] 제1 특징 추출모듈(300)은 2차원 매트릭스에 HCNN(Horizontal Convolution Neural Network)을 적용하여 상기 고객의 시간별 행동특징인 제1 특징 매트릭스를 추출하는 기능을 수행하고, 제2 특징 추출모듈(400)은 상기 2차원 매트릭스에 VCNN(Vertical Convolution Neural Network)를 적용하여 상기 고객의 행동별 시간특징인 제2 특징 매트릭스를 추출하는 기능을 수행한다.

[0067] 평탄화 모듈(500)은 도 3에 도시된 바와 같이 제1 특징 매트릭스(300) 및 제2 특징 매트릭스(400)에 평탄화(flatten)를 적용하여 제1 특징 추출모듈(300) 및 제2 특징 추출모듈(400)에서 각각 추출된 행동과 시간의 잠재적 의미를 반영하게 된다.

[0068] 마지막으로 결정모듈(600)은 고객의 미래 행동을 예측하는 기능을 수행한다.

[0069] 이하 제1 특징 추출모듈(300) 및 제2 특징 추출모듈(400)의 특징을 좀 더 구체적으로 설명하도록 한다.

[0070] 일반적인 CNN의 구조는 convolution layer에서 feature map을 생성하고, 생성된 feature map에 다시 convolution layer를 적용함으로써 검사 차원을 축소하고, 여러 공간정보가 혼합된 의미를 추출하게 된다.

[0071] 본 발명의 일 실시예에 따른 제1 특징 추출모듈(300) 및 제2 특징 추출모듈(400)은 상술한 LSA와 같이 행과 열에 대한 정보를 유지하기 위하여 concatenation 방법을 도입한다.

[0072] 즉, 도 4에 도시된 바와 같이 1D Convolution을 이용하여 행동과 시간의 잠재적 의미를 생성하고, concatenation을 사용하여 공간정보를 유지한다.

[0073] $time \times variable$ matrix에 $1 \times variable$ kernel, stride는 1인 1-D convolution을 적용해 n개의 feature map이 추출되고, 추출된 feature map size는 $time \times 1$ 이고, 추출된 n개의 feature map을 concatenation하여 $time \times n$ matrix를 생성하며, 생성된 matrix의 각 행은 raw matrix의 행에 대한 latent semantic을 의미한다.

[0074] 이와 같은 과정을 반복하여 행의 위치 정보는 유지하면서 유의미한 의미를 추출하며, 최종적으로 추출된 feature map은 공간 전체가 아닌 horizontal 혹은 vertical의 유의미한 정보를 표현한다.

[0075] 고객의 시간과 행동의 관계를 반영하기 위해 HCNN이 적용된 제1 특징 추출모듈(300)과 VCNN이 적용된 제2 특징 추출모듈(400)을 동시 학습하되, 개별적으로 학습하는 경우 관계를 반영하기 힘들기 때문에, 동일한 loss function 하에 모델을 생성하였으며, 각 module의 1-D convolution의 activation function은 tanh로 설정하였다.

- [0076] 도 5(1)은 제1 특징 추출모듈(300)의 HCNN의 구조로 시간별 특징을 추출하기 위한 module이다.
- [0077] Horizontal 1-D convolution에서 추출된 feature map에 concatenation을 적용하여 horizontal feature matrix를 생성하고, 생성된 matrix는 $\text{time} \times \text{filter}$ 형태를 지니며, 시간별 특징을 표현한다.
- [0078] matrix의 value는 raw 데이터의 행에 대한 정보를 나타내며, 이는 LSA의 L1과 같다.
- [0079] 제2 특징 추출모듈(400)의 VCNN은 시간에 따른 행동의 변화를 표현하는 module이며, 여러 변수를 혼합하여 정보를 추출하는 것이 아닌, 단일 variable의 고유 정보를 추론한다.
- [0080] 도 5(2)는 VCNN의 구조로, vertical 1D convolution과 concatenation으로 구성되며, 생성된 vertical feature matrix는 $\text{filter} \times \text{variable}$ 형태이며, 시간에 따른 각 변수의 유의미한 의미를 지닌다.
- [0081] 추출된 matrix의 value는 raw 데이터의 열에 대한 정보를 지니며, 이는 LSA의 L2와 같다.
- [0082] 상술한 평탄화 모듈(500)은 서로 다른 의미를 추론하는 HCNN과 VCNN을 결합하여 고객의 미래 행동 예측을 위해 flatten을 적용하며, 구체적으로 horizontal feature matrix와 vertical feature matrix에 flatten을 적용하여 각기 다르게 생성된 정보를 온전히 반영할 수 있으며, 이를 통해 생성된 요소들을 활용하여 선택적으로 혼합이 가능하며, 여러 경우의 수가 도출 가능하게 된다.
- [0083] 다음으로 결정모듈(600)이 dense layer를 통해 고객의 미래 행동을 예측하는데, 이때 dense layer의 활성화함수(activation function)는 sigmoid 함수로 설정하는 것이 바람직하다.
- [0084] 상기와 같은 과정으로 HV-Net은 2D 고객 데이터의 구조적 문제 및 높은 차원 문제를 해결하고, 고객의 시간과 행동 정보를 이용하여 미래 행동을 예측할 수 있다.
- [0085] 한편, L사의 구매이력 데이터를 적용하여 본 발명의 일 실시예에 따른 고객의 미래 행동 예측 장치의 성능을 실험하였다.
- [0086] 구매이력 데이터는 고객의 시간과 행동 정보를 담고 있으며, 데이터에는 총 15216명의 고객이 존재하며, 1년 동안의 상품거래이력이 포함되어 있다.
- [0087] 행동은 구매일자, 거래상품종류, 상품 수, 상품 금액 등 총 12개가 존재하며, Raw 데이터는 sparse 성질을 지니고 있어, 월별로 통합 변환하여 실험을 진행하였다.
- [0088] 실험 목적은 고객의 구매이력 데이터를 이용하여 미래 행동인 방문 유무를 예측하는 것으로, 연구결과를 검증하기 위해 training set과 test set의 실험 기간을 다르게 설정하였다.
- [0089] Training set기간은 1월부터 10월까지, test set은 2월부터 11월까지로 설정하였으며, Target은 다음 달에 해당 고객의 방문 여부로, 방문 고객은 1, 방문하지 않은 고객은 0으로 설정하였다.
- [0090] Training set과 test set 모두 target ratio는 8:2로 나타났으며, 두 클래스 간의 차이가 상당한 불균형 형태를 보였다.
- [0091] 불균형 데이터를 이용한 모델의 경우, 분석 결과가 다수 클래스로 치우치는 문제가 발생하기 때문에, 이 문제를 해결하기 위해 under sampling method를 적용했다.
- [0092] Under sampling는 imbalanced data를 해결하기 위한 방법 중 하나로 다수 클래스의 일부만 sampling하여 데이터의 균형을 맞추는 방법이다.
- [0093] 실험 방식은 under sampling ratio가 0.2인 경우 target ratio가 1:1이기 때문에 0.2부터 1.0까지 0.1씩 증가시켜 총 10개의 case를 설정하였으며, Under sample된 training set을 이용해 제안 모델인 HV-Net과 benchmark model들을 학습시키고, 최종적으로 test set을 이용해 성과평가를 하였다.
- [0094] 고객의 미래 행동 예측을 평가하기 위한 지표로 recall과 specificity를 이용하였으며, 본 발명에서의 recall은 방문한 고객에 대한 정확도이며, Specificity는 방문하지 않은 고객의 정확도이다.
- [0095] 각 지표는 도 6 및 하기 수식 (5), (6)이다.

[0097] <수식 (5)>

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

[0098]

[0100] <수식 (6)>

$$Specificity = \frac{True\ Negative}{True\ Negative + False\ Positive}$$

[0101]

[0103] Accuracy를 통한 모델 평가는 imbalanced data에서는 적합하지 않아, 두 행동에 대한 정확한 결과를 볼 수 없기 때문에, Under sampling case마다 recall과 specificity를 통해 결과를 비교한다.

[0104] 도 7은 test set에 대한 모델별 recall과 specificity 결과로서, 도출된 결과를 통해 본 발명에서 제안한 HV-Net이 다른 bench mark 모델보다 더 좋은 성능을 나타내고 있음을 알 수 있다.

[0105] recall의 결과를 보면 대부분의 모델이 sample ratio가 증가하면서 동일하게 증가하고 있는데, 이는 imbalanced data의 성질에 영향을 받은 결과로 판단된다.

[0106] ratio 0.2일 때, 제안 모델을 제외한 모델들이 대략 60에 성능을 지니고 있으며, ANN의 경우, ratio 1.0 일 때, recall이 99.4로 100에 가까운 성능을 보임으로써 다른 모델과 비교했을 때, imbalance 영향을 가장 많이 받아 과적합 되었으며, 반대로 CNN이 90.2로 가장 낮은 성능을 지녔다.

[0107] ANN, SVM, LSTM은 공통적으로 ratio 0.4 이상에서 80 이상의 성능을 ratio 1.0에서는 95 이상의 성능을 나타내는데, 이는 과하게 학습데이터에 적합된 결과이다.

[0108] specificity의 경우, ratio 0.2에서 SVM이 91.8로 가장 높으며, ANN이 75.4로 가장 낮으며, specificity는 ratio가 증가함에 따라 계속 해서 감소하며, ratio 1.0에서는 HV-Net은 46.1로 가장 높은 성능을 지녔다.

[0109] recall과 specificity를 같이 봤을 때, HV-Net이 모든 ratio에서 꾸준히 높은 결과를 보였다.

[0110] HV-Net이 유일하게 ratio 0.2 부터 0.5까지 두 지표 모두 70 이상을 유지하였으며, SVM의 경우 ratio 0.3일 때는 높은 성능을 보였지만, 0.4부터 과적합 되어 낮은 성능을 나타냈고, CNN도 이와 비슷하다.

[0111] 추가로 진행한 HV-Net의 단일 모델인 HCNN과 VCNN의 결과는 다른 성향을 보여주고 있는데, HCNN의 결과는 과적합 형태를 보이고 있으며, VCNN의 경우는 제안모델과 유사한 결과를 보인다.

[0112] 이는 여러 variable의 혼합된 정보가 아닌 각 변수의 고유정보를 추출한 접근이 효과가 있음을 시사하며, 또한 VCNN에서 추출한 고유정보와 HCNN에서 추출한 시간별 정보의 결합이 더 좋은 성능을 도출한다는 것이 입증되었다.

[0113] 2-D convolution을 적용한 CNN이 다른 bench mark 모델에 비해 높았고, 이는 2차원 고객 data에 대한 적합성이 다른 모델 보다 뛰어났음을 의미하지만, 구조적 결합으로 인해 HV-Net 보다는 낮은 성능을 보인 것으로 판단된다.

[0115] 본 명세서에서 설명되는 실시예와 첨부된 도면은 본 발명에 포함되는 기술적 사상의 일부를 예시적으로 설명하는 것에 불과하다. 따라서 본 명세서에 개시된 실시예들은 본 발명의 기술적 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이므로, 이러한 실시예에 의하여 본 발명의 기술 사상의 범위가 한정되는 것이 아님은 자명하다. 본 발명의 명세서 및 도면에 포함된 기술적 사상의 범위 내에서 당해 기술분야에 있어서의 통상의 지식을 가진 자가 용이하게 유추할 수 있는 변형 예와 구체적인 실시예는 모두 본 발명의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

부호의 설명

[0117] 100: 입력모델

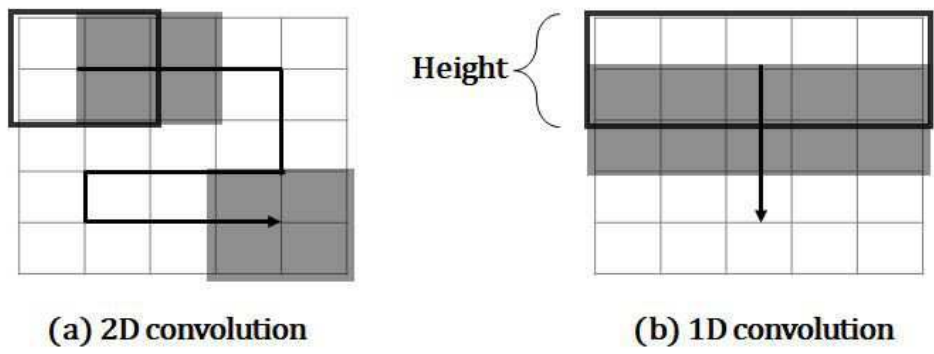
200: 데이터 전처리모델

300: 제1 특징 추출모델

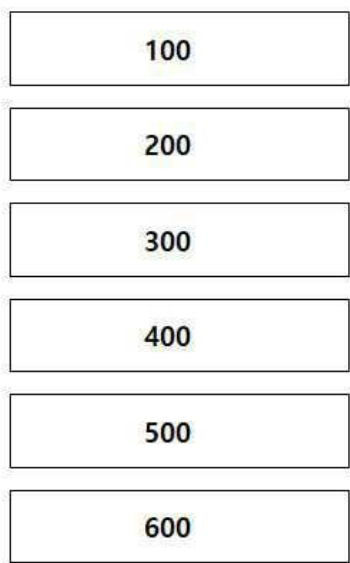
400: 제2 특징 추출모듈
500: 평탄화 모듈
600: 결정모듈

도면

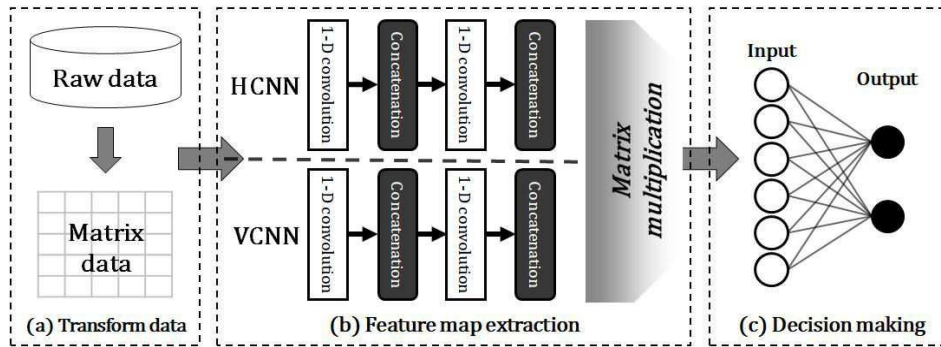
도면1



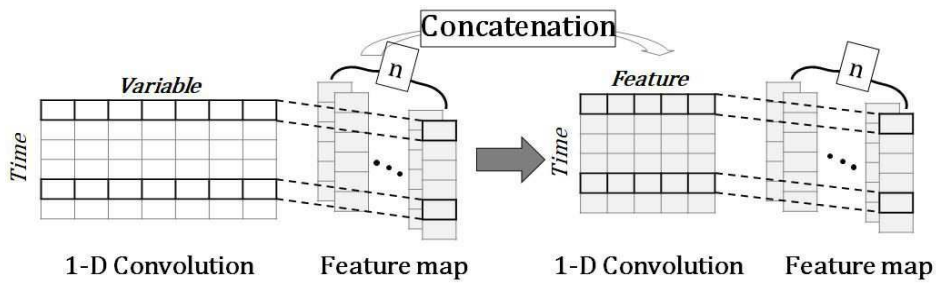
도면2



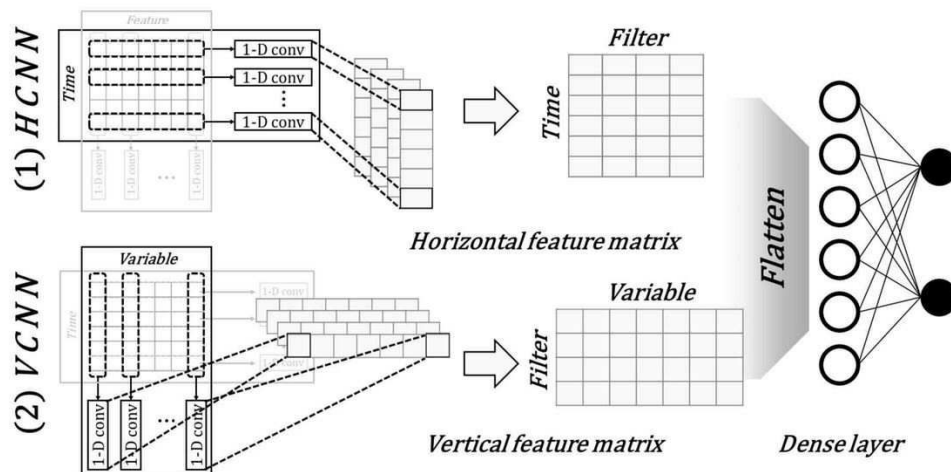
도면3



도면4



도면5



도면6

		Predicted	
		Visit	No-Visit
Actual	Visit	True Positive	False Positive
	No-Visit	False Negative	True Negative

도면7

Model	Indicators	Under smaple ratio(%)								
		0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
ANN	recall	60.5	75.8	83.5	89.3	93.8	96	98.3	99.3	99.4
	specificity	75.4	59.1	46.9	33.6	20.6	13.3	5.2	2.6	1.7
SVM	recall	61.5	72.1	83.4	88	91.3	93.1	94.4	95.4	95.5
	specificity	91.8	70.3	61.5	54.4	44.6	37.2	31.9	27.4	24.3
LSTM	recall	67.2	78.1	81.2	85.2	88	89.3	93.1	94	95.4
	specificity	77.7	67.5	60.7	56.5	51.8	48.5	39.3	35.2	26.7
CNN	recall	65.5	69.9	75.8	81.2	82.6	84.9	85.7	88.8	90.2
	specificity	77.3	76.2	67.3	54.8	51.6	50.9	49.9	43.1	36
HCNN	recall	68.6	74	79.9	81.1	84.6	88.8	89.5	91.7	94.4
	specificity	76.1	73.3	62.6	59	55.4	43.8	39.7	34.4	27.1
VCNN	recall	63.6	69.7	77.2	80.6	82.5	88.3	89.2	92.2	93.3
	specificity	80.2	76.7	69.6	66.5	60.4	55.6	55	47.2	41.4
HV-Net	recall	70.5	77.4	79.7	83.8	85.5	87.1	88.7	92.2	93.6
	specificity	82.8	80.9	75.2	70.8	63.7	60.6	61.5	52.9	46.1