(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0092609
(43) 공개일자 2021년07월26일

(51) 국제특허분류(Int. Cl.)
G06N 20/00 (2019.01) *G06N 5/02* (2006.01)
(52) CPC특허분류
G06N 20/00 (2019.01)
G06N 5/025 (2019.01)
(21) 출원번호 10-2020-0006198
(22) 출원일자 2020년01월16일
심사청구일자 2020년01월16일

(71) 출원인
계명대학교 산학협력단
대구광역시 달서구 달구벌대로 1095, 계명대학교 산학협력관 201호(신당동)
(72) 발명자
고병철
대구광역시 수성구 범어천로 107 (범어동, 범어 STX KAN) 103동 701호
김상원
대구광역시 수성구 지범로22길 20 현대아파트, 105동 905호
(74) 대리인
김건우

전체 청구항 수 : 총 20 항

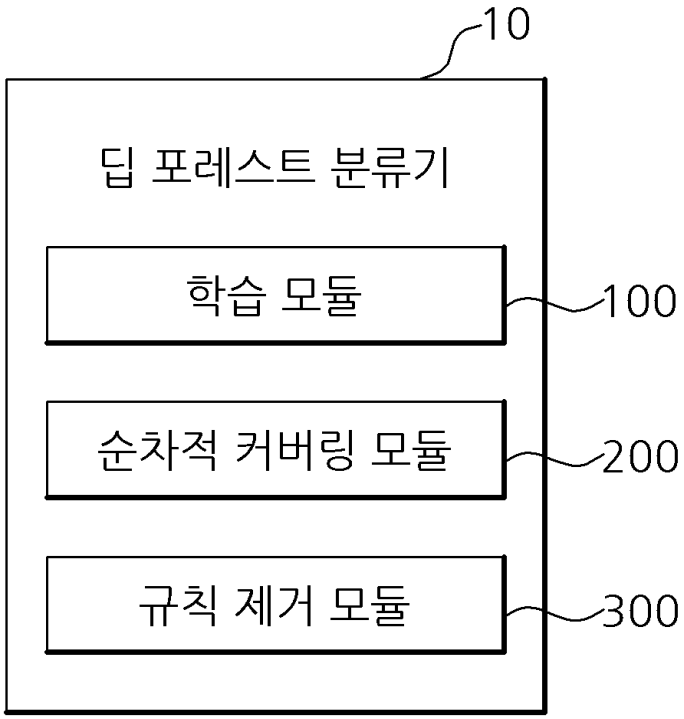
(54) 발명의 명칭 **해석할 수 있고 단순화된 딥 포레스트 분류 방법 및 분류기**

(57) 요 약

본 발명은 해석할 수 있고 단순화된 딥 포레스트 분류 방법에 관한 것으로서, 보다 구체적으로는 해석할 수 있고 단순화된 딥 포레스트 분류 방법으로서, (1) 복수의 랜덤 포레스트(Random Forest, RF)를 사용해 멀티 레이어로 구성된 완전히 학습된 딥 포레스트를 생성하는 단계; (2) 상기 딥 포레스트에 대해, 특징 기여도 및 특징 빈도에

(뒷면에 계속)

대 표 도 - 도1



기초한 순차적 커버링(sequential covering)을 통해 각각의 RF의 규칙들을 의사결정 셋에 저장하는 단계; 및 (3) 상기 의사결정 셋에서 불필요한 규칙을 제거하여 상기 딥 포레스트를 단순화하는 단계를 포함하는 것을 그 구성상의 특징으로 한다.

또한, 본 발명은 해석할 수 있고 단순화된 딥 포레스트 분류기에 관한 것으로서, 보다 구체적으로는 해석할 수 있고 단순화된 딥 포레스트 분류기로서, 복수의 랜덤 포레스트(Random Forest, RF)를 사용해 멀티 레이어로 구성된 완전히 학습된 딥 포레스트를 생성하는 학습 모듈; 상기 딥 포레스트에 대해, 특징 기여도 및 특징 빈도에 기초한 순차적 커버링(sequential covering)을 통해 각각의 RF의 규칙들을 의사결정 셋에 저장하는 순차적 커버링 모듈; 및 상기 의사결정 셋에서 불필요한 규칙을 제거하여 상기 딥 포레스트를 단순화하는 규칙 제거 모듈을 포함하는 것을 그 구성상의 특징으로 한다.

본 발명에서 제안하고 있는 해석할 수 있고 단순화된 딥 포레스트 분류 방법 및 분류기에 따르면, 의사결정 셋 형태로 완전히 학습된 딥 포레스트에 대해, 특징 기여도와 특징 빈도에 기초한 순차적 커버링(sequential covering)을 적용해 의사결정 셋에서 불필요한 규칙을 제거함으로써, 특징 기여도를 통해 규칙을 정량적으로 해석할 수 있고, 불필요한 규칙의 제거를 통해 강력한 성능을 유지하면서도 매개 변수와 규칙이 적은 단순화된 딥 포레스트를 구성할 수 있다.

이 발명을 지원한 국가연구개발사업

과제고유번호	1345301346
부처명	교육부
과제관리(전문)기관명	한국연구재단
연구사업명	지역대학우수과학자지원사업(1년~5년)
연구과제명	역전파가 필요 없는 심층 랜덤 네트워크 모델 개발과 응용 연구
기 여 율	1/1
과제수행기관명	계명대학교 산학협력단
연구기간	2019.06.01 ~ 2020.02.29

명 세 서

청구범위

청구항 1

해석할 수 있고 단순화된 딥 포레스트 분류 방법으로서,

(1) 복수의 랜덤 포레스트(Random Forest, RF)를 사용해 멀티 레이어로 구성된 완전히 학습된 딥 포레스트를 생성하는 단계;

(2) 상기 딥 포레스트에 대해, 특징 기여도 및 특징 빈도에 기초한 순차적 커버링(sequential covering)을 통해 각각의 RF의 규칙들을 의사결정 셋에 저장하는 단계; 및

(3) 상기 의사결정 셋에서 불필요한 규칙을 제거하여 상기 딥 포레스트를 단순화하는 단계를 포함하는 것을 특징으로 하는, 해석할 수 있고 단순화된 딥 포레스트 분류 방법.

청구항 2

제1항에 있어서, 상기 딥 포레스트는,

LMRF(lightweight multilayered random forest)인 것을 특징으로 하는, 해석할 수 있고 단순화된 딥 포레스트 분류 방법.

청구항 3

제1항에 있어서, 상기 단계 (2)에서는,

상기 RF의 모든 규칙을 추출할 때까지 상기 순차적 커버링을 반복하고, 상기 순차적 커버링에 의해 각 규칙의 최종 규칙 기여도에 따라 규칙이 재정렬되는 것을 특징으로 하는, 해석할 수 있고 단순화된 딥 포레스트 분류 방법.

청구항 4

제3항에 있어서, 상기 단계 (3)에서는,

상기 의사결정 셋에서, 미리 설정된 규칙의 비율에 따라 상기 재정렬된 규칙 중 불필요한 규칙을 제거하는 것을 특징으로 하는, 해석할 수 있고 단순화된 딥 포레스트 분류 방법.

청구항 5

제3항에 있어서, 상기 단계 (2)는,

(2-1) 상기 딥 포레스트의 레이어 및 RF의 의사결정 트리에서 특징 기여도를 계산하는 단계;

(2-2) 상기 계산된 특징 기여도를 이용해 규칙 기여도를 계산하는 단계;

(2-3) 상기 규칙 및 해당 규칙에 대해 계산된 규칙 기여도를 의사결정 셋에 저장하는 단계;

(2-4) 상기 의사결정 셋에서 특징 패턴을 추출하는 단계;

(2-5) 상기 규칙 기여도 및 특징 패턴을 이용해 최종 규칙 기여도를 산출하는 단계; 및

(2-6) 상기 최종 규칙 기여도에 따라 상기 의사결정 셋의 규칙들을 정렬하는 단계를 포함하는 것을 특징으로 하

는, 해석할 수 있고 단순화된 딥 포레스트 분류 방법.

청구항 6

제5항에 있어서, 상기 단계 (2-1)에서는,

특징 변수에 대하여 부모 노드와 자식 노드 사이의 클래스 확률 차이를 이용해 상기 특징 기여도를 계산하는 것을 특징으로 하는, 해석할 수 있고 단순화된 딥 포레스트 분류 방법.

청구항 7

제6항에 있어서, 상기 단계 (2-1)에서는,

i번째 규칙의 j번째 노드의 특징 기여도 feat.contrib(i,j)를 부모 (j-1) 노드와 자식 (j) 노드 사이의 클래스 확률 Pr={pr_1,…,$pr_{\#class}$}의 차이를 이용해 $feat.contrib(i,j) = (\mathbf{Pr}_{j-1} - \mathbf{Pr}_j)$ 에 의해 계산하는 것을 특징으로 하는, 해석할 수 있고 단순화된 딥 포레스트 분류 방법.

청구항 8

제5항에 있어서, 상기 단계 (2-2)에서는,

루트 노드에서 리프 노드까지의 경로를 따라 각 노드의 모든 특징 기여도를 합산하여 상기 규칙 기여도를 계산하는 것을 특징으로 하는, 해석할 수 있고 단순화된 딥 포레스트 분류 방법.

청구항 9

제8항에 있어서, 상기 단계 (2-2)에서는,

t번째 트리에서 P 심도(depth)로 구성된 i번째 규칙에 대한 규칙 기여도 r_i^t를 $\mathbf{r}_i^t = \frac{\sum_{j \in P} feat.contrib(i,j)}{\sum \#class \ \ of\ t\text{-}th\ tree}$ 에 의해 계산하는 것을 특징으로 하는, 해석할 수 있고 단순화된 딥 포레스트 분류 방법.

청구항 10

제5항에 있어서, 상기 단계 (2-5)에서는,

상기 의사결정 셋 dSet(v,l)의 최종 규칙 기여도 r^*_i를 규칙 기여도 r_i와 특징 x_j의 특징 패턴의 빈도 $fre_v^l(x_j)$의 가중 조합을 통해 $\mathbf{r}^*_i = \mathbf{r}_i \cdot \sum_{j \in \mathbf{r}_i} fre_v^l(x_j)$ 에 의해 산출하는 것을 특징으로 하는, 해석할 수 있고 단순화된 딥 포레스트 분류 방법.

청구항 11

해석할 수 있고 단순화된 딥 포레스트 분류기(10)로서,

복수의 랜덤 포레스트(Random Forest, RF)를 사용해 멀티 레이어로 구성된 완전히 학습된 딥 포레스트를 생성하는 학습 모듈(100);

상기 딥 포레스트에 대해, 특징 기여도 및 특징 빈도에 기초한 순차적 커버링(sequential covering)을 통해 각

각의 RF의 규칙들을 의사결정 셋에 저장하는 순차적 커버링 모듈(200); 및

상기 의사결정 셋에서 불필요한 규칙을 제거하여 상기 딥 포레스트를 단순화하는 규칙 제거 모듈(300)을 포함하는 것을 특징으로 하는, 해석할 수 있고 단순화된 딥 포레스트 분류기(10).

청구항 12

제11항에 있어서, 상기 딥 포레스트는,

LMRF(lightweight multilayered random forest)인 것을 특징으로 하는, 해석할 수 있고 단순화된 딥 포레스트 분류기(10).

청구항 13

제11항에 있어서, 상기 순차적 커버링 모듈(200)은,

상기 RF의 모든 규칙을 추출할 때까지 상기 순차적 커버링을 반복하고, 상기 순차적 커버링에 의해 각 규칙의 최종 규칙 기여도에 따라 규칙이 재정렬되는 것을 특징으로 하는, 해석할 수 있고 단순화된 딥 포레스트 분류기(10).

청구항 14

제13항에 있어서, 상기 규칙 제거 모듈(300)은,

상기 의사결정 셋에서, 미리 설정된 규칙의 비율에 따라 상기 재정렬된 규칙 중 불필요한 규칙을 제거하는 것을 특징으로 하는, 해석할 수 있고 단순화된 딥 포레스트 분류기(10).

청구항 15

제13항에 있어서, 상기 순차적 커버링 모듈(200)은,

(2-1) 상기 딥 포레스트의 레이어 및 RF의 의사결정 트리에서 특징 기여도를 계산하는 단계;

(2-2) 상기 계산된 특징 기여도를 이용해 규칙 기여도를 계산하는 단계;

(2-3) 상기 규칙 및 해당 규칙에 대해 계산된 규칙 기여도를 의사결정 셋에 저장하는 단계;

(2-4) 상기 의사결정 셋에서 특징 패턴을 추출하는 단계;

(2-5) 상기 규칙 기여도 및 특징 패턴을 이용해 최종 규칙 기여도를 산출하는 단계; 및

(2-6) 상기 최종 규칙 기여도에 따라 상기 의사결정 셋의 규칙들을 정렬하는 단계를 처리하는 것을 특징으로 하는, 해석할 수 있고 단순화된 딥 포레스트 분류기(10).

청구항 16

제15항에 있어서, 상기 단계 (2-1)에서는,

특징 변수에 대하여 부모 노드와 자식 노드 사이의 클래스 확률 차이를 이용해 상기 특징 기여도를 계산하는 것을 특징으로 하는, 해석할 수 있고 단순화된 딥 포레스트 분류기(10).

청구항 17

제16항에 있어서, 상기 단계 (2-1)에서는,

i번째 규칙의 j번째 노드의 특징 기여도 feat.contrib(i,j)를 부모 (j-1) 노드와 자식 (j) 노드 사이의 클래스 확률 Pr={pr_1,…,$pr_{\#class}$}의 차이를 이용해 $feat.contrib\ (i,j) = (\mathbf{Pr}_{j-1} - \mathbf{Pr}_j)$ 에 의해 계산하는 것을 특징으로 하는, 해석할 수 있고 단순화된 딥 포레스트 분류기(10).

청구항 18

제15항에 있어서, 상기 단계 (2-2)에서는,

루트 노드에서 리프 노드까지의 경로를 따라 각 노드의 모든 특징 기여도를 합산하여 상기 규칙 기여도를 계산하는 것을 특징으로 하는, 해석할 수 있고 단순화된 딥 포레스트 분류기(10).

청구항 19

제18항에 있어서, 상기 단계 (2-2)에서는,

t번째 트리에서 P 심도(depth)로 구성된 i번째 규칙에 대한 규칙 기여도 r_i^t를 $r_i^t = \frac{\sum_{j \in P} feat.contrib\ (i,j)}{\sum \#class\ of\ t-th\ tree}$ 에 의해 계산하는 것을 특징으로 하는, 해석할 수 있고 단순화된 딥 포레스트 분류기(10).

청구항 20

제15항에 있어서, 상기 단계 (2-5)에서는,

상기 의사결정 셋 dSet(v,l)의 최종 규칙 기여도 r^*_i를 규칙 기여도 r_i와 특징 x_j의 특징 패턴의 빈도 $fre_v^l(x_j)$의 가중 조합을 통해 $r^*_i = r_i \cdot \sum_{j \in r_i} fre_v^l(x_j)$ 에 의해 산출하는 것을 특징으로 하는, 해석할 수 있고 단순화된 딥 포레스트 분류기(10).

발명의 설명

기 술 분 야

[0001] 본 발명은 딥 포레스트 분류 방법 및 분류기에 관한 것으로서, 보다 구체적으로는 해석할 수 있고 단순화된 딥 포레스트 분류 방법 및 분류기에 관한 것이다.

배 경 기 술

[0002] 최근 딥 러닝은 이미지 분류, 음성인식, 주가예측, 자연어 채팅봇에 이르기까지 광범위하게 사용되는 기술로서, 신경망 및 랜덤 포레스트(Random Forest, RF)가 많이 사용된다. 그 중, 신경망은 그 특성상 설명력을 발휘해야 하는 근거들(즉, 설명력을 제공하는 근거인 해당 결과에 대한 결정 규칙들)이 특징 기반의 히든 레이어(hidden layer)들로 분산되어 흩어지기 때문에 설명력이 약하다는 단점이 있다.

[0004] 즉, DNN(Deep Neural Network) 기반의 딥 러닝은 주어진 데이터 세트에 대해 신경망 학습을 통해서 주어진 분류 문제에 대한 정확한 분류 결과를 제공 한다(예: Convolutional neural network). 이때, DNN은 분류의 정확도와 같이 학습의 결과가 우수하지만 해당 분류를 위해서 설정된 수학적 분류기들이 주어진 문제를 다수의 경계선으로 분리하여 주어진 데이터 들을 분류할 때의 경계선을 잘 모르게 되는 경우가 많고, 만약 알아낸다고 해도 비직교 경계선이 아니라서 말로 표현하기가 어렵다는 점이다.

[0006] 따라서 실제로 여러 문제에서 우수한 결과를 나타내는 딥 러닝에 설명력을 확보할 수 있는 신규의 방법 및 시스

템이 요구된다.

[0008] 한편, 본 발명과 관련된 선행기술로서, 공개특허 제10-2019-0061431호(발명의 명칭: 설명력이 있는 딥 러닝 방법 및 시스템, 공개 일자: 2019년 06월 05일) 등이 개시된 바 있다.

발명의 내용

해결하려는 과제

[0009] 본 발명은 기존에 제안된 방법들의 상기와 같은 문제점들을 해결하기 위해 제안된 것으로서, 의사결정 셋 형태로 완전히 학습된 딥 포레스트에 대해, 특징 기여도와 특징 빈도에 기초한 순차적 커버링(sequential covering)을 적용해 의사결정 셋에서 불필요한 규칙을 제거함으로써, 특징 기여도를 통해 규칙을 정량적으로 해석할 수 있고, 불필요한 규칙의 제거를 통해 강력한 성능을 유지하면서도 매개 변수와 규칙이 적은 단순화된 딥 포레스트를 구성할 수 있는, 해석할 수 있고 단순화된 딥 포레스트 분류 방법 및 분류기를 제공하는 것을 그 목적으로 한다.

과제의 해결 수단

[0010] 상기한 목적을 달성하기 위한 본 발명의 특징에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법은,

[0011] 해석할 수 있고 단순화된 딥 포레스트 분류 방법으로서,

[0012] (1) 복수의 랜덤 포레스트(Random Forest, RF)를 사용해 멀티 레이어로 구성된 완전히 학습된 딥 포레스트를 생성하는 단계;

[0013] (2) 상기 딥 포레스트에 대해, 특징 기여도 및 특징 빈도에 기초한 순차적 커버링(sequential covering)을 통해 각각의 RF의 규칙들을 의사결정 셋에 저장하는 단계; 및

[0014] (3) 상기 의사결정 셋에서 불필요한 규칙을 제거하여 상기 딥 포레스트를 단순화하는 단계를 포함하는 것을 그 구성상의 특징으로 한다.

[0016] 바람직하게는, 상기 딥 포레스트는,

[0017] LMRF(lightweight multilayered random forest)일 수 있다.

[0019] 바람직하게는, 상기 단계 (2)에서는,

[0020] 상기 RF의 모든 규칙을 추출할 때까지 상기 순차적 커버링을 반복하고, 상기 순차적 커버링에 의해 각 규칙의 최종 규칙 기여도에 따라 규칙이 재정렬될 수 있다.

[0022] 더욱 바람직하게는, 상기 단계 (3)에서는,

[0023] 상기 의사결정 셋에서, 미리 설정된 규칙의 비율에 따라 상기 재정렬된 규칙 중 불필요한 규칙을 제거할 수 있다.

[0025] 더욱 바람직하게는, 상기 단계 (2)는,

[0026] (2-1) 상기 딥 포레스트의 레이어 및 RF의 의사결정 트리에서 특징 기여도를 계산하는 단계;

[0027] (2-2) 상기 계산된 특징 기여도를 이용해 규칙 기여도를 계산하는 단계;

[0028] (2-3) 상기 규칙 및 해당 규칙에 대해 계산된 규칙 기여도를 의사결정 셋에 저장하는 단계;

[0029] (2-4) 상기 의사결정 셋에서 특징 패턴을 추출하는 단계;

[0030] (2-5) 상기 규칙 기여도 및 특징 패턴을 이용해 최종 규칙 기여도를 산출하는 단계; 및

[0031] (2-6) 상기 최종 규칙 기여도에 따라 상기 의사결정 셋의 규칙들을 정렬하는 단계를 포함할 수 있다.

[0033] 더더욱 바람직하게는, 상기 단계 (2-1)에서는,

[0034] 특징 변수에 대하여 부모 노드와 자식 노드 사이의 클래스 확률 차이를 이용해 상기 특징 기여도를 계산할 수 있다.

[0036] 더더더욱 바람직하게는, 상기 단계 (2-1)에서는,

[0037] i번째 규칙의 j번째 노드의 특징 기여도 feat.contrib(i,j)를 부모 (j-1) 노드와 자식 (j) 노드 사이의 클래스 확률 $Pr=\{pr_1,\cdots,pr_{\#class}\}$의 차이를 이용해 $feat.contrib\ (i,j) = \left(\mathbf{Pr}_{j-1} - \mathbf{Pr}_j\right)$ 에 의해 계산할 수 있다.

[0039] 더더욱 바람직하게는, 상기 단계 (2-2)에서는,

[0040] 루트 노드에서 리프 노드까지의 경로를 따라 각 노드의 모든 특징 기여도를 합산하여 상기 규칙 기여도를 계산할 수 있다.

[0042] 더더더욱 바람직하게는, , 상기 단계 (2-2)에서는,

[0043] t번째 트리에서 P 심도(depth)로 구성된 i번째 규칙에 대한 규칙 기여도 r_i^t를 $\mathrm{r}_i^t = \frac{\sum_{j\in P} feat.contrib\ (i,j)}{\sum \#class\ \ of\ t-th\ tree}$ 에 의해 계산할 수 있다.

[0045] 더더욱 바람직하게는, 상기 단계 (2-5)에서는,

[0046] 상기 의사결정 셋 dSet(v,l)의 최종 규칙 기여도 r^*_i를 규칙 기여도 r_i와 특징 x_j의 특징 패턴의 빈도 $fre_v^l(x_j)$의 가중 조합을 통해 $\mathrm{r}^*_i = \mathrm{r}_i \cdot \sum_{j\in \mathrm{r}_i} fre_v^l(x_j)$ 에 의해 산출할 수 있다.

[0048] 또한, 상기한 목적을 달성하기 위한 본 발명의 특징에 따른 해석할 수 있고 단순화된 딥 포레스트 분류기는,

[0049] 복수의 랜덤 포레스트(Random Forest, RF)를 사용해 멀티 레이어로 구성된 완전히 학습된 딥 포레스트를 생성하는 학습 모듈;

[0050] 상기 딥 포레스트에 대해, 특징 기여도 및 특징 빈도에 기초한 순차적 커버링(sequential covering)을 통해 각각의 RF의 규칙들을 의사결정 셋에 저장하는 순차적 커버링 모듈; 및

[0051] 상기 의사결정 셋에서 불필요한 규칙을 제거하여 상기 딥 포레스트를 단순화하는 규칙 제거 모듈을 포함하는 것을 그 구성상의 특징으로 한다.

[0053] 바람직하게는, 상기 딥 포레스트는,

[0054] LMRF(lightweight multilayered random forest)일 수 있다.

[0056] 바람직하게는, 상기 순차적 커버링 모듈은,

[0057] 상기 RF의 모든 규칙을 추출할 때까지 상기 순차적 커버링을 반복하고, 상기 순차적 커버링에 의해 각 규칙의 최종 규칙 기여도에 따라 규칙이 재정렬될 수 있다.

[0059] 더욱 바람직하게는, 상기 규칙 제거 모듈은,

[0060] 상기 의사결정 셋에서, 미리 설정된 규칙의 비율에 따라 상기 재정렬된 규칙 중 불필요한 규칙을 제거할 수 있다.

[0062] 더욱 바람직하게는, 상기 순차적 커버링 모듈은,

[0063] (2-1) 상기 딥 포레스트의 레이어 및 RF의 의사결정 트리에서 특징 기여도를 계산하는 단계;

[0064] (2-2) 상기 계산된 특징 기여도를 이용해 규칙 기여도를 계산하는 단계;

[0065] (2-3) 상기 규칙 및 해당 규칙에 대해 계산된 규칙 기여도를 의사결정 셋에 저장하는 단계;

[0066] (2-4) 상기 의사결정 셋에서 특징 패턴을 추출하는 단계;

[0067] (2-5) 상기 규칙 기여도 및 특징 패턴을 이용해 최종 규칙 기여도를 산출하는 단계; 및

[0068] (2-6) 상기 최종 규칙 기여도에 따라 상기 의사결정 셋의 규칙들을 정렬하는 단계를 처리할 수 있다.

[0070] 더더욱 바람직하게는, 상기 단계 (2-1)에서는,

[0071] 특징 변수에 대하여 부모 노드와 자식 노드 사이의 클래스 확률 차이를 이용해 상기 특징 기여도를 계산할 수 있다.

[0073] 더더더욱 바람직하게는, 상기 단계 (2-1)에서는,

[0074] i번째 규칙의 j번째 노드의 특징 기여도 feat.contrib(i,j)를 부모 (j-1) 노드와 자식 (j) 노드 사이의 클래스 확률 $Pr=\{pr_1,\cdots,pr_{\#class}\}$의 차이를 이용해 $feat.contrib\quad(i,j)=\left(\mathbf{Pr}_{j-1}-\mathbf{Pr}_j\right)$ 에 의해 계산할 수 있다.

[0076] 더더욱 바람직하게는, 상기 단계 (2-2)에서는,

[0077] 루트 노드에서 리프 노드까지의 경로를 따라 각 노드의 모든 특징 기여도를 합산하여 상기 규칙 기여도를 계산할 수 있다.

[0079] 더더더욱 바람직하게는, 상기 단계 (2-2)에서는,

[0080] t번째 트리에서 P 심도(depth)로 구성된 i번째 규칙에 대한 규칙 기여도 r_i^t를 $r_i^t=\frac{\sum_{j\in P} feat.contrib\quad(i,j)}{\sum \#class\quad of\ t-th\ tree}$ 에 의해 계산할 수 있다.

[0082] 더더욱 바람직하게는, 상기 단계 (2-5)에서는,

[0083] 상기 의사결정 셋 dSet(v,l)의 최종 규칙 기여도 r^*_i를 규칙 기여도 r_i와 특징 x_j의 특징 패턴의 빈도 $fre_v^l(x_j)$의 가중 조합을 통해 $r^*_i=r_i\cdot\sum_{j\in r_i} fre_v^l(x_j)$ 에 의해 산출할 수 있다.

발명의 효과

[0084] 본 발명에서 제안하고 있는 해석할 수 있고 단순화된 딥 포레스트 분류 방법 및 분류기에 따르면, 의사결정 셋 형태로 완전히 학습된 딥 포레스트에 대해, 특징 기여도와 특징 빈도에 기초한 순차적 커버링(sequential covering)을 적용해 의사결정 셋에서 불필요한 규칙을 제거함으로써, 특징 기여도를 통해 규칙을 정량적으로 해석할 수 있고, 불필요한 규칙의 제거를 통해 강력한 성능을 유지하면서도 매개 변수와 규칙이 적은 단순화된 딥 포레스트를 구성할 수 있다.

도면의 간단한 설명

[0085] 도 1은 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류기의 구성을 도시한 도면.

도 2는 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법의 흐름을 도시한 도면.

도 3은 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법의 단계 S100에서 생성된 딥 포레스트를 예를 들어 도시한 도면.

도 4는 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법의 단계 S100에서 생성된 딥 포레스트를 구성하는 RF를 예를 들어 도시한 도면.

도 5는 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법에서, 단계 S200의 세부적인 흐름을 도시한 도면.

도 6은 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법 및 분류기에서, 의사결정 셋 재정렬을 예를 들어 도시한 도면.

도 7은 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법 및 분류기에서, 특징 기여도 및 패턴에 기반한 전체 규칙 제거 알고리즘을 도시한 도면.

도 8 및 도 9는 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법 및 분류기를 사용한 IRIS 데이터 세트의 분류에서 특징 기여도를 도시한 도면.

도 10은 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법 및 분류기를 사용하여, ORL 데이터 세트의 특징 기여도에 기초한 시각적 주의 히트 맵(visual attention heat map)을 도시한 도면.

발명을 실시하기 위한 구체적인 내용

[0086] 이하, 첨부된 도면을 참조하여 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자가 본 발명을 용이하게 실시할 수 있도록 바람직한 실시예를 상세히 설명한다. 다만, 본 발명의 바람직한 실시예를 상세하게 설명함에 있어, 관련된 공지 기능 또는 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략한다. 또한, 유사한 기능 및 작용을 하는 부분에 대해서는 도면 전체에 걸쳐 동일한 부호를 사용한다.

[0088] 덧붙여, 명세서 전체에서, 어떤 부분이 다른 부분과 '연결' 되어 있다고 할 때, 이는 '직접적으로 연결' 되어 있는 경우뿐만 아니라, 그 중간에 다른 소자를 사이에 두고 '간접적으로 연결' 되어 있는 경우도 포함한다. 또한, 어떤 구성요소를 '포함' 한다는 것은, 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있다는 것을 의미한다.

[0090] 도 1은 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류기(100)의 구성을 도시한 도면이다. 도 1에 도시된 바와 같이, 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류기(10)는, 복수의 랜덤 포레스트(Random Forest, RF)를 사용해 멀티 레이어로 구성된 완전히 학습된 딥 포레스트를 생성하는 학습 모듈(100), 딥 포레스트에 대해, 특징 기여도 및 특징 빈도에 기초한 순차적 커버링(sequential covering)을 통해 각각의 RF의 규칙들을 의사결정 셋에 저장하는 순차적 커버링 모듈(200), 및 의사결정 셋에서 불필요한 규칙을 제거하여 딥 포레스트를 단순화하는 규칙 제거 모듈(300)을 포함하여 구성될 수 있다.

[0092] 도 2는 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법의 흐름을 도시한 도면이다. 도 2에 도시된 바와 같이, 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법은, 복수의 랜덤 포레스트를 사용해 멀티 레이어로 구성된 완전히 학습된 딥 포레스트를 생성하는 단계(S100), 딥 포레스트에 대해, 특징 기여도 및 특징 빈도에 기초한 순차적 커버링을 통해 각각의 RF의 규칙들을 의사결정 셋에 저장하는 단계(S200), 및 의사결정 셋에서 불필요한 규칙을 제거하여 딥 포레스트를 단순화하는 단계(S300)를 포함하여 구현될 수 있다.

[0094] 본 발명은 해석할 수 있고 단순화된 딥 포레스트 분류 방법 및 분류기에 관한 것으로서, 의사결정 셋 형태로 완전히 학습된 딥 포레스트에 대해, 특징 기여도와 특징 빈도에 기초한 순차적 커버링(sequential covering)을 적용해 의사결정 셋에서 불필요한 규칙을 제거함으로써, 특징 기여도를 통해 규칙을 정량적으로 해석할 수 있고, 불필요한 규칙의 제거를 통해 강력한 성능을 유지하면서도 매개 변수와 규칙이 적은 단순화된 딥 포레스트를 구성할 수 있다.

[0096] 또한, 본 발명은 새로운 ILM(Interpreting the learning model) 기반 해석 방법에 관한 것으로서, 본 발명을 통해 심층 학습 모델을 해석하고 단순화하여, 모델 구조의 중요한 특성을 유지하고 성능을 저하시키지 않으면서 규칙을 재정의할 수 있다. ILM 기반 접근을 통해, 특정 결정이 내려진 이유를 이해하고, 동일한 예측 성능을 유지하면서 그 결정을 설명할 수 있는 모델을 생성할 수 있다.

[0098] 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법 및 분류기는, 새로운 유형의 딥 모델인 딥 랜덤 포레스트(Random Forest, RF)를 사용하는데, 딥 RF는 여러 RF를 구성요소로 복수의 레이어에 연결하며, 학습 과정에서 역전파를 사용하지 않는다. 딥 RF는 일반적으로 DNN과 유사한 고성능을 달성할 수 있으나, 블랙박스 RF들로 구성되기 때문에 많은 규칙을 생성한다. 따라서 많은 수의 규칙이 딥 RF의 결과를 해석하는 데 중요한 장애물이 된다.

[0100] 이와 같은 문제를 해결하기 위해, 본 발명은 딥 RF가 훈련되고 여러 RF들을 사용해 멀티 레이어 네트워크가 생성된 후, 먼저 RF의 각 의사결정 트리(decision tree)의 예측을 수학적으로 정확한 특징 기여도(feature contribution)로 분해할 수 있다. 의사결정 트리의 개별 예측은 의사결정 경로(decision path)를 특징 당 단일 컴포넌트로 분류하여 설명할 수 있다. 이 과정은, 전체 RF 레이어의 모든 규칙을 레이어 별로 찾기 위해 반복적으로 적용될 수 있으며, RF의 분류 규칙 세트인 의사결정 셋(decision set 또는 decision rule set)에 저장될 수 있다.

[0102] 그 다음, 순차적 커버링(sequential covering)이 규칙 기여도 및 특징 패턴(특징의 빈도)의 조합에 기초하여, RF의 의사결정 셋의 규칙을 반복적으로 유지 및 제거할 수 있다. 이 정규화는 의사결정에 가장 큰 영향을 미치는 소수의 정제된 규칙만 유지하도록 할 수 있다. 순차적 커버링이 적용되면, 레이어 당 동일한 수의 의사결정 셋이 있지만, 성능 저하 없이 규칙과 특징의 수가 크게 줄어들게 된다.

[0104] 이하에서는, 해석할 수 있고 단순화된 딥 포레스트 분류 방법 및 분류기를 구성하는 각 구성요소에 대해 상세히 설명하도록 한다.

[0106] 단계 S100에서는, 학습 모듈(100)이, 복수의 랜덤 포레스트(Random Forest, RF)를 사용해 멀티 레이어로 구성된 완전히 학습된 딥 포레스트를 생성할 수 있다. 이때, 딥 포레스트는, LMRF(lightweight multilayered random forest)일 수 있다.

[0108] DNN은 인식률이 높지만, 매개 변수 조정이 필요한 지나치게 많은 하이퍼 매개 변수, 역전파에 의해 생성된 블랙

박스 모델, 높은 처리 비용, 및 많은 학습 데이터 양 등이 DNN의 해석에 상당한 부담이 된다. 이에 대한 대안적 접근으로, 여러 RF 또는 펀(ferns)으로 구성된 심층 앙상블 분류기가 연구되어 왔다.

[0110] 그중에서, LMRF(lightweight multilayered random forest) 모델은, 레이어-투-레이어 RF로 구성된다. DNN 레이어의 각 뉴런은 RF로 대체되며, 각 레이어는 여러 유형의 RF로 구성될 수 있다. 각 레이어는 균일한 RF 대신 무작위로 생성된 이종의 RF로 구성되어, 다양성을 높이면서도 일반성을 유지할 수 있다. 본 발명에서는, 변환된 특징 벡터를 결합하지 않고, 이전 레이어의 출력 특징만 다음 레이어의 새로운 입력 특징으로 사용하는 LMRF 모델을 설계하여 사용하였다.

[0112] 즉, 다양한 딥 앙상블 모델 중에서, RF 뉴런과 레이어의 수가 다른 방법보다 적기 때문에 실시간 시스템에 적용할 수 있는 LMRF에 규칙 제거 알고리즘을 적용하였다. 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법 및 분류기에 따른, 해석 가능하고 단순화된 LMRF(iLMRF)는 규칙 수가 크게 줄어들어도 성능이 유지되는 효과가 있다.

[0114] 도 3은 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법의 단계 S100에서 생성된 딥 포레스트를 예를 들어 도시한 도면이다. 도 3에 도시된 바와 같이, 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법의 단계 S100에서 생성된 딥 포레스트는, 각 RF가 많은 수의 의사결정 트리로 구성되어 있다.

[0116] 도 4는 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법의 단계 S100에서 생성된 딥 포레스트를 구성하는 RF를 예를 들어 도시한 도면이다. 즉, 도 4는 도 3에서 붉은색 타원형 파선으로 표시된 f번째 RF를 도시한 것으로서, RF는 특징 기여도에 기반하여 규칙을 분해함으로써 설명 및 해석을 할 수 있다.

[0118] 단계 S100에서는, 비 NN 스타일 딥 모델인 RF 앙상블을 기반으로 하는 LMRF를 학습할 수 있다. LMRF는 도 3에 도시된 바와 같이, RF $F_v^l \in \{1, \cdots, V\}$의 다중 레이어 $L^l(l \in \{1, \cdots, N\})$로 구성되며, 각 F_v^l은 다수의 의사결정 트리 t로 구성되고, 레이어 l에서 F_v^l의 t번째 트리는 $dt_{v,t}^l$로 표시된다.

[0120] 첫 번째 레이어에서 입력 벡터는 개별 RF, $\Phi^1 = [P(\Phi_1^1|F_1^1), P(\Phi_2^1|F_2^1), \cdots, P(\Phi_V^1|F_V^1)]$의 출력을 결합하여 변환된 특징 벡터 Φ^1로 인코딩될 수 있다. 두 번째 레이어에서 각 레이어 $L^l(l>1)$는 레이어 l-1의 인코딩 된 특징 벡터를 사용해 학습되며, 다음 레이어에 대한 새로운 특징 벡터 Φ^l를 생성하거나 최종 레이어에서 최종 클래스를 예측하는 데에 사용될 수 있다.

[0122] LMRF를 사용하면, DNN 레이어의 각 뉴런이 RF로 바뀌고 각 레이어는 여러 유형의 RF로 구성될 수 있다. 각 레이어는 다양성을 장려하고 일반성을 유지하기 위해 균일한 RF 대신 무작위로 생성된 이종의 RF로 구성될 수 있다. 레이어 확장 여부를 결정하기 위해 LMRF는 K-폴드 교차 검증을 사용하여, 과적합의 위험을 줄이면서 레이어 수 및 매개 변수를 자동으로 결정할 수 있다. LMRF가 K-폴드 교차 검증을 통해 수렴될 때, 최종 클래스 확률은 각 RF에서 예측된 클래스 확률 $[P(\Phi_1^N|F_1^N), \cdots, P(\Phi_V^N|F_V^N)]$을 평균하여 결정될 수 있고, 확률이 가장 높은 최종 클래스 레이블로 예측할 수 있다.

[0124] 단계 S200에서는, 순차적 커버링 모듈(200)이, 딥 포레스트에 대해, 특징 기여도 및 특징 빈도에 기초한 순차적 커버링(sequential covering)을 통해 각각의 RF의 규칙들을 의사결정 셋에 저장할 수 있다. 단계 S200에서는,

RF의 모든 규칙을 추출할 때까지 순차적 커버링을 반복하고, 순차적 커버링에 의해 각 규칙의 최종 규칙 기여도에 따라 규칙이 재정렬될 수 있다.

[0126] 순차적 커버링(sequential covering)은 전체 데이터셋을 포함하는 의사결정 셋을 생성하기 위해 개별적으로 단일 규칙을 반복적 학습하는 일반 규칙 유도 과정일 수 있다. 조밀하게 결합된 블랙박스 LMRF 모델이 구성된 후, 개별 RF의 규칙은 순차적 커버링 과정에 기초하여 의사결정 셋에 반복적으로 저장될 수 있다.

[0128] 단계 S200은, 도 3 및 도 4에 도시된 바와 같은 네트워크에서 특징 인코딩 레이어(첫 번째 레이어)를 제외한 캐스케이딩 레이어들로부터 수행될 수 있다. 여기서는, 블랙박스 구조로 구성된 iLMRF의 의사결정 프로세스가 두 가지 프로세스 즉, 특징 기여도 평가 및 중요하지 않은 규칙 제거를 통해 어떻게 설명될 수 있는지 보여줄 수 있다.

[0130] LMRF의 기본 단위 즉, 의사결정 트리는 최종값을 결정하는 의사결정 절차가 훈련된 노드가 나타내는 if-then 조건에 의존하기 때문에 규칙 기반 모델로 간주된다. 트리의 뿌리에서 잎까지의 각 경로는 일련의 예제를 분류하는 규칙에 해당한다. 인스턴스 X_n 및 해당 레이블 Y_n (데이터셋 $D=\{(X_1, Y_1), \cdots, (X_N, Y_N)\}$의 일부)이 루트 노드에 속하면, X_n은 학습 단계에서 결정된 특정 특징에 대한 임계값을 갖는 분할 함수(split function)를 만족하는 오른쪽 또는 왼쪽 자식 노드로 전달될 수 있다. 이러한 과정은, 주어진 데이터가 최적의 특징 공간을 생성하는 리프 노드에 도달할 때까지 반복될 수 있다. 노드는 특정 특징 인덱스 쌍, 임계값이 있는 분할 함수, 및 클래스 분포로 구성될 수 있으며, 의사결정 트리 당 전체 노드 체인(의사결정 경로 또는 규칙)이 의사결정 셋에 저장될 수 있다.

[0132] 본 발명의 단계 S200에서는, 순차적 커버링 알고리즘을 수정하여, 각 트리와 RF에서 최적의 규칙을 선택하였다. 분류 문제에서 특징 기여도(중요도)는 인스턴스가 특정 특징에 대해 분할될 때, 특징-특정 분포(feature-specific distribution)의 변경을 나타낼 수 있다.

[0134] 도 5는 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법에서, 단계 S200의 세부적인 흐름을 도시한 도면이다. 도 5에 도시된 바와 같이, 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법의 단계 S200는, 딥 포레스트의 레이어 및 RF의 의사결정 트리에서 특징 기여도를 계산하는 단계(S210), 계산된 특징 기여도를 이용해 규칙 기여도를 계산하는 단계(S220), 규칙 및 해당 규칙에 대해 계산된 규칙 기여도를 의사결정 셋에 저장하는 단계(S230), 의사결정 셋에서 특징 패턴을 추출하는 단계(S240), 규칙 기여도 및 특징 패턴을 이용해 최종 규칙 기여도를 산출하는 단계(S250) 및 최종 규칙 기여도에 따라 의사결정 셋의 규칙들을 정렬하는 단계(S260)를 포함하여 구현될 수 있다.

[0136] 단계 S210에서는, 딥 포레스트의 레이어 및 RF의 의사결정 트리에서 특징 기여도를 계산할 수 있다. 특징 기여도를 계산하기 위해, 의사결정 트리가 잎에 도달할 때까지 아래로 이동한다. 모든 특정 분할에 있어서, 분할을 결정하는 특징 변수의 특징 기여도는 부모 노드와 자식 노드 사이의 클래스 확률 차이로 정의될 수 있다.

[0138] 단계 S220에서는, 계산된 특징 기여도를 이용해 규칙 기여도를 계산할 수 있다. 즉, 최종 규칙 기여도를 얻기 위해, 데이터 인스턴스의 루트 노드에서 리프 노드까지의 경로를 따라 각 노드의 모든 특징 기여도를 합산할 수 있다. 이 알고리즘은 기여도 점수가 높은 최상의 규칙을 찾아 경로(규칙)를 순차적으로 추출할 수 있다.

[0140] t번째 트리에서 P 심도(depth)로 구성된 i번째 규칙에 대한 규칙 기여도 r_i^t는, 다음 수학식 1과 같이 계산될 수 있다.

수학식 1

[0141] $$r_i^t = \frac{\sum_{j \in P} feat.contrb \quad (i,j)}{\sum \#class \quad of \ t-th \ tree}$$

[0142] 여기서, i번째 규칙의 j번째 노드의 특징 기여도 feat.contrib(i,j)는 부모 (j-1) 노드와 자식 (j) 노드 사이의 클래스 확률 Pr={$pr_1,\cdots,pr_{\#class}$}의 차이를 이용해 다음 수학식 2와 같이 계산될 수 있다.

수학식 2

[0143] $$feat.contrb \quad (i,j) = \left(\mathbf{Pr}_{j-1} - \mathbf{Pr}_j\right)$$

[0145] 이러한 규칙 기여도는 t번째 트리의 모든 규칙 기여도의 합으로 나눔으로써 0과 1 사이의 값으로 정규화될 수 있다. r_i의 양수 또는 음수 값이 큰 것은, 여러 특징으로 구성된 규칙이 의사결정 클래스에 크기 기여한다는 것을 의미한다. 반대로, r_i의 양수 또는 음수 값이 작은 것은, 규칙이 의사결정 클래스에 약하게 기여함을 의미한다. 기여도에서 0값은 해당 특징이 의사결정 프로세스에 기여하지 않음을 의미한다.

[0147] 단계 S230에서는, 규칙 및 해당 규칙에 대해 계산된 규칙 기여도를 의사결정 셋에 저장할 수 있다.

[0149] 단계 S240에서는, 의사결정 셋에서 특징 패턴을 추출할 수 있다. 이때, 특징 패턴(feature pattern)의 pre-mining은 규칙 기여도의 가중치로 사용될 수 있다. 특징 패턴은 특징값의 발생 빈도일 수 있다(예를 들어, x=A). v번째 RF 및 l번째 레이어의 의사결정 셋 dSet(v,l)에서 모든 규칙으로부터 자주 발생하는 특징 패턴을 추출할 수 있다. 특징 패턴의 빈도는 의사결정 셋에서의 지원을 지반으로 다음 수학식 3과 같이 측정될 수 있다.

수학식 3

[0150] $$fre_v^l(x_j = A) = \frac{1}{|\mathrm{dSet}(v,l)|}\sum_{k \in \mathrm{dSet}(v,l)} I(x_j^{(k)} = A)$$

[0151] 여기서, |dSet|은 dSet(v,l)의 특징 카디널리티(cardinality)이고, fre(x_j=A)는 dSet(v,l)의 규칙에서 특징 패턴의 빈도의 정량화된 값이며, I는 인스턴스 k의 특징 xj가 레벨 A이면 1 값을 반환하고, 그렇지 않으면 0 값을 반환하는 인디케이터 함수이다. 특징 패턴은 dSet(v,l)에 포함된 모든 특징 중 겹치는 특징 수를 정규화한 것일 수 있다.

[0153] 단계 S250에서는, 규칙 기여도 및 특징 패턴을 이용해 최종 규칙 기여도를 산출할 수 있다. RF 레벨에서, 의사결정 셋 dSet(v,l)의 최종 규칙 기여도 r^*_i는 특징 기여도와 특징 패턴의 가중 조합을 통해 다음 수학식 4와 같이 추정될 수 있다.

수학식 4

[0154] $$r^{*}_{i} = r_i \cdot \sum_{j \in r_i} fre^{\,l}_{v}(x_j)$$

[0156] 수학식 4에서, 각 규칙 r_i에 포함된 특징 패턴이 높은 빈도를 갖는 경우, 최종 규칙 기여도 r^{*}_{i}는 특징 기여도에 비례하여 증가할 수 있다.

[0158] 단계 S260에서는, 최종 규칙 기여도에 따라 의사결정 셋의 규칙들을 정렬할 수 있다. 각 RF F_v^l에 대한 dSet(v,l)의 규칙은 최종 규칙 기여도에 따라 오름차순으로 정렬 및 재배열될 수 있다.

[0160] 단계 S200의 과정은 l번째 레이어의 RF를 다루는 모든 규칙을 추출할 때까지 반복될 수 있다.

[0162] 도 6은 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법 및 분류기에서, 의사결정 셋 재정렬을 예를 들어 도시한 도면이다. 초기 규칙은 IF 절의 특징 규칙과 THEN 절의 클래스 확률 쌍으로 구성될 수 있다. 그러나 본 발명의 순차적 커버링 과정을 통해 각 규칙의 최종 기여도에 따라 규칙이 재정렬될 수 있다.

[0164] 단계 S300에서는, 규칙 제거 모듈(300)이, 의사결정 셋에서 불필요한 규칙을 제거하여 딥 포레스트를 단순화할 수 있다. 보다 구체적으로, 단계 S300에서는, 의사결정 셋에서, 미리 설정된 규칙의 비율에 따라 재정렬된 규칙 중 불필요한 규칙을 제거할 수 있다.

[0166] 학습된 특징과 클래스 확률의 변화 사이의 상관관계를 나타내기 위해 규칙 기여도에 대한 특징 패턴을 가진 특징 기여도를 사용할 수 있다. 이 접근 방식은 어떤 특징, 규칙 및 RF가 LMRF의 예측 결과에 영향을 미치는지 이해하는데 도움을 줄 수 있다. 그러나 LMRF는 여러 블랙박스 RF로 구성되기 때문에 많은 수의 규칙을 생성한다. 따라서 중요하지 않은 규칙의 제거 프로세스를 통해 LMRF의 의사결정 프로세스를 설명할 수 있다.

[0168] dSet(v,l)의 약한 규칙은 주어진 최종 규칙 기여도 r^{*}_{i}에 따라 제거되며, 기여도가 높은 규칙만 유지될 수 있다. $List_{dSet}[l]$에 포함된 규칙은 사용자 사용자 입력에 따라 각 RF에 대해 동일한 속도로 제거하거나 필요한 정확도에 따라 각 RF에 대해 조정할 수 있다.

[0170] 도 7은 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법 및 분류기에서, 특징 기여도 및 패턴에 기반한 전체 규칙 제거 알고리즘을 도시한 도면이다. 도 7에 도시된 바와 같이, 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법에서는, 단계 S100에서 iLMRF의 학습을 완료한 후 테스트 데이터는 첫 번째 특징 인코더 레이어에 입력될 수 있다. 첫 번째 레이어의 출력이 연결되고, 첫 번째 레이어에 의해 생성된 클래스 벡터로 증강된 변환된 특징 벡터는 데이터가 최종 레이어에 매핑될 때까지 l번째 레이어의 리스트 dSet($List_{dSet}$)에 입력될 수 있다. 최종 레이어는 각 클래스의 확률값을 평균하고 확률값이 가장 높은 클래스를 최종 클래스로 결정할 수 있다.

[0172] 실험 1

[0174] 딥 포레스트 중 iLMRF 모델의 해석 및 단순화를 확인하고, 동일한 방식의 규칙 제거가 다른 심층 RF 및 DNN 기반 방식에 적용될 때의 성능을 비교하였다. 실험을 통해 단순화된 iLMRF는, 원래 LMRF뿐만 아니라 DNN 기반 알고리즘과 유사한 성능을 유지하면서도, 규칙의 상당 부분을 제거할 수 있음을 확인하였다.

[0176] 성능의 일관성을 입증하고 단순화된/압축된 iLMRF의 해석 가능성을 조사하기 위해 5개의 데이터 세트를 사용하여 테스트를 수행하였다.

[0178] CK+ 데이터 세트는 확장된 Cohn-Kanade 데이터 세트로서, FER(얼굴 표정 인식, Facial Expression Recognition)을 위한 공개 벤치마크 데이터 세트이다. MNIST(Modified National Institute of Standards and Technology) 데이터 세트는, 수기 이미지를 포함하며 머신러닝 분야의 평가에 널리 사용된다. IRIS 데이터 세트는, 3개의 홍채 종류 및 4차원 특징 벡터가 포함되며, 150개의 샘플로 구성된다. WDBC(The Wisconsin Diagnostic Breast Cancer) 데이터 세트는, 위스콘신 대학교 병원의 유방암 진단 결과를 제공한다. ORL(The Orivetti Research Lab) 데이터 세트는, AT & T Laboratories Cambridge에서 촬영한 얼굴 이미지 세트가 포함된다.

[0180] 실험 중에, CK+, IRIS, WDBC 데이터 세트는 특징 인코더 레이어(첫 번째 레이어) 없이 다른 유형의 특징 벡터들을 사용하고, MNIST, ORL 데이터 세트는 특징 추출을 위해 특징 인코더 레이어에 이미지가 입력된다.

[0182] 규칙 제거의 효율성을 검증하기 위해, 규칙의 비율(ratios of the rules)을 1.0에서 0.6까지 변경하면서, (1) 본 발명(iLMRF)과 대표적인 딥 랜덤 포레스트 기반 방법인 (2) gcForest (3) FTDRF를 비교하였다.

[0184] 다음 표 1은 CK+ 데이터 세트, 표 2는 MNIST 데이터 세트를 이용한 성능을 비교 결과를 각각 나타낸 것으로서, 정확도, 모델에 사용된 규칙 수, 매개 변수의 수 및 연산 수를 비교하여 표시하였다.

표 1

[0185]

Rule	Accuracy (%)			Rules (M)			# Param. (M)			# Op. (M)		
ratio	iLMRF	gcForest	FTDRF	iLMRF	gcForest	FTDRF	iLMRF	gcForest	FTDRF	iLMRF	gcForest	FTDRF
1.0	93.60	89.71	92.41	0.12	0.16	0.13	0.53	2.90	2.51	0.0060	0.0381	0.0233
0.9	92.86	90.00	92.15	0.11	0.15	0.12	0.51	2.78	2.39	0.0060	0.0381	0.0232
0.8	92.50	89.92	92.24	0.09	0.13	0.10	0.47	2.59	2.22	0.0059	0.0380	0.0231
0.7	91.87	89.92	92.18	0.08	0.12	0.09	0.44	2.38	2.03	0.0059	0.0379	0.0230
0.6	91.05	89.73	92.04	0.07	0.10	0.08	0.39	2.16	1.83	0.0058	0.0377	0.0228

표 2

[0186]

Rule	Accuracy (%)			Rules (M)			# Param. (M)			# Op. (M)		
ratio	iLMRF	gcForest	FTDRF	iLMRF	gcForest	FTDRF	iLMRF	gcForest	FTDRF	iLMRF	gcForest	FTDRF
1.0	97.98	98.77	98.57	0.08	0.94	0.26	2.12	26.42	7.17	0.0089	0.0852	0.0296
0.9	97.77	98.73	98.57	0.08	0.85	0.23	2.00	24.79	6.76	0.0088	0.0851	0.0294
0.8	97.41	98.74	98.57	0.07	0.76	0.21	1.86	22.98	6.28	0.0087	0.0850	0.0293
0.7	96.86	98.76	98.47	0.06	0.66	0.18	1.71	20.98	5.75	0.0087	0.0849	0.0292
0.6	96.00	98.75	98.39	0.05	0.57	0.16	1.54	18.86	5.19	0.0086	0.0850	0.0291

[0188] 보다 다양한 데이터 세트에서 알고리즘의 성능을 평가하기 위해, IRIS, WDBC, ORL 데이터 세트에서 동일한 실험

을 수행하였으며, 그 결과를 다음 표 3(IRIS 데이터 세트), 표 4(WDBC 데이터 세트), 표 5(ORL 데이터 세트)에 정확도 및 규칙 개수를 비교하여 표시하였다.

표 3

[0189]

Rule ratio	Accuracy (%)			# Rules (M)		
	iLMRF	gcForest	FTDRF	iLMRF	gcForest	FTDRF
1.0	100.00	98.00	100.00	0.0046	0.0362	0.0137
0.9	100.00	98.00	100.00	0.0046	0.0362	0.0137
0.8	100.00	98.00	100.00	0.0046	0.0329	0.0136
0.7	98.00	98.00	100.00	0.0039	0.0287	0.0120
0.6	100.00	98.00	100.00	0.0033	0.0248	0.0096

표 4

[0190]

Rule ratio	Accuracy (%)			# Rules (M)		
	iLMRF	gcForest	FTDRF	iLMRF	gcForest	FTDRF
1.0	96.49	95.21	97.34	0.0063	0.0450	0.0256
0.9	96.49	95.21	97.34	0.0063	0.0450	0.0247
0.8	96.49	95.21	97.34	0.0057	0.0385	0.0227
0.7	96.49	95.74	97.34	0.0050	0.0350	0.0199
0.6	96.49	95.74	96.81	0.0044	0.0298	0.0168

표 5

[0191]

Rule ratio	Accuracy (%)			# Rules (M)		
	iLMRF	gcForest	FTDRF	iLMRF	gcForest	FTDRF
1.0	97.50	97.50	90.00	0.0595	0.0957	0.1133
0.9	97.50	97.50	90.00	0.0543	0.0893	0.1045
0.8	97.50	97.50	90.00	0.0481	0.0786	0.0925
0.7	87.50	97.50	90.00	0.0422	0.0702	0.0809
0.6	87.50	97.50	90.00	0.0362	0.0594	0.0696

[0193] 표 1 내지 표 5에서 확인할 수 있는 바와 같이, 전체적으로 세 가지 방법 모두 많은 규칙을 제거하기는 하지만, gcForest 및 FTDRF는 여전히 iLMRF의 규칙보다 최소 1.6(gcForest of ORL)에서 최대 7.5배(gcForest of IRIS) 까지 많은 규칙을 포함하며, 정확성은 비슷하게 나타났다. 이러한 단순성은, 모델의 경량화의 지표가 될 수 있으며, 모델의 해석 가능성을 입증하기 위한 방법 중 하나이기도 하다.

[0195] 실험 2

[0197] 본 발명이 얼굴 표정을 효과적으로 인식하는지 여부를 테스트하기 위해 CK+ 데이터 세트에 대한 추가 실험을 수행하였다. 성능 비교를 위해 다른 최첨단 방법 즉, AlexNets 기반 FER 접근법, 변형 가능한 안면 행동 부분이 제한된 3D CNN 기반 FER 접근법(3DCNN-DAP), 다중 인셉션 레이어를 사용하는 DNN 기반 방법(Multiple

Inception), FER을 위한 LSTM을 가진 3D Inception-ResNet (3DIR) (3DIR with LSTM), 계층적 가중 RF(H-WRF)에 기초한 고속 FER, gcForest, FTDRF, 원래 LMRF를 본 발명과 비교하였다.

[0199] 다음 표 6은 CK+ 데이터 세트, 표 7은 MNIST 데이터 세트를 이용한 성능을 비교 결과를 각각 나타낸 것으로서, 정확도, 매개 변수의 수 및 연산 수를 비교하여 표시하였다.

표 6

[0200]

Methods	Acc. (%)	# Param.(M)	# Op. (M)
AlexNets [37]	92.2	62.3	720
3DCNN-DAP [38]	92.4	70	174
Multiple Inception [39]	93.2	12.36	23.7
3DIR with LSTM [40]	93.2	10.90	18.9
H-WRF [27]	92.6	0.25	0.0067
gcForest [7]	89.7	2.90	0.0381
FTDRF[8]	92.4	2.51	0.0233
LMRF (1.0)[24]	93.6	0.53	0.0060
iLMRF (0.8)	92.5	0.47	0.0059
iLMRF (0.7)	91.9	0.44	0.0059

표 7

[0201]

Methods	Acc. (%)	# Param. (M)	# Op.(M)
ResNet-101[41]	98.3	42	212
ShuffleNetV2[42]	97.0	1.3	2.7
MobileNetV2[43]	98.5	2.24	11.02
MobileNetV3[44]	98.6	1.66	16.59
gcForest [7]	98.8	26.42	0.085
FTDRF[8]	98.6	7.17	0.029
iLMRF (1.0)	98.0	2.12	0.0089
iLMRF (0.8)	97.4	1.86	0.0087
iLMRF (0.6)	96.0	1.54	0.0086

[0203] 두 실험을 통해 본 발명이 정확도 수준을 유지하면서 매개 변수 및 연산 수 측면에서 뛰어난 압축 성능을 도출할 수 있음을 확인할 수 있으며, 이를 통해 iLMRF 응용 프로그램의 범위를 저가형 시스템으로 확장할 수 있다.

[0205] 실험 3

[0207] 본 발명의 iLMRF가 해석 가능하다는 것을 나타내는 다른 예로서, 첫 번째 레이어의 두 RF 노드에서 클래스를 분류하는 데 사용되는 특징 기여도를 살펴볼 수 있다.

[0209] 도 8 및 도 9는 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법 및 분류기를 사용한 IRIS 데이터 세트의 분류에서 특징 기여도를 도시한 도면이다. 도 8은 첫 번째 레이어의 첫 번째 RF, 도 9는 첫 번째 레이어의 두 번째 RF의 특징 기여도를 각각 도시하였다. 도 8 및 도 9에 도시된 바와 같이, 분류를 위해 IRIS 데이터 세트를 사용하는 경우, petal length와 petal width는 첫 번째 RF에서 중요한 특징인 반면, 두 번째 RF에서는 petal length만이 중요한 특징으로 표시되었다.

[0211] 또한, ORL 데이터 세트를 사용하여 특징 기여도의 해석 가능성을 분석하였다. iLMRF는 히트 맵(heat map)과 유사하게 각 입력 변수에 관련성 점수를 할당할 때까지 로컬 특징 기여도를 사용하여 클래스 예측을 거꾸로 재분해한다.

[0213] 도 10은 본 발명의 일실시예에 따른 해석할 수 있고 단순화된 딥 포레스트 분류 방법 및 분류기를 사용하여, ORL 데이터 세트의 특징 기여도에 기초한 시각적 주의 히트 맵(visual attention heat map)을 도시한 도면이다. 히트 맵은 개인의 분류에 중추적인 픽셀을 식별할 수 있다. 도 10에서 입력 이미지에서 객체를 분류하는 데 유용한 특징 변수(픽셀)을 확인할 수 있다. 규칙 제거에 사용된 특징 기여도는 결과를 직관적으로 확인하고 해석할 수 있는 히트 맵을 제공함을 확인할 수 있다.

[0215] 본 발명에서 제안하고 있는 해석할 수 있고 단순화된 딥 포레스트 분류 방법 및 분류기에 따르면, 의사결정 셋 형태로 완전히 학습된 딥 포레스트에 대해, 특징 기여도와 특징 빈도에 기초한 순차적 커버링(sequential covering)을 적용해 의사결정 셋에서 불필요한 규칙을 제거함으로써, 특징 기여도를 통해 규칙을 정량적으로 해석할 수 있고, 불필요한 규칙의 제거를 통해 강력한 성능을 유지하면서도 매개 변수와 규칙이 적은 단순화된 딥 포레스트를 구성할 수 있다.

[0217] 이상 설명한 본 발명은 본 발명이 속한 기술분야에서 통상의 지식을 가진 자에 의하여 다양한 변형이나 응용이 가능하며, 본 발명에 따른 기술적 사상의 범위는 아래의 특허청구범위에 의하여 정해져야 할 것이다.

부호의 설명

[0218] 10: 딥 포레스트 분류기

100: 학습 모듈

200: 순차적 커버링 모듈

300: 규칙 제거 모듈

S100: 복수의 랜덤 포레스트를 사용해 멀티 레이어로 구성된 완전히 학습된 딥 포레스트를 생성하는 단계

S200: 딥 포레스트에 대해, 특징 기여도 및 특징 빈도에 기초한 순차적 커버링을 통해 각각의 RF의 규칙들을 의사결정 셋에 저장하는 단계

S210: 딥 포레스트의 레이어 및 RF의 의사결정 트리에서 특징 기여도를 계산하는 단계

S220: 계산된 특징 기여도를 이용해 규칙 기여도를 계산하는 단계

S230: 규칙 및 해당 규칙에 대해 계산된 규칙 기여도를 의사결정 셋에 저장하는 단계

S240: 의사결정 셋에서 특징 패턴을 추출하는 단계

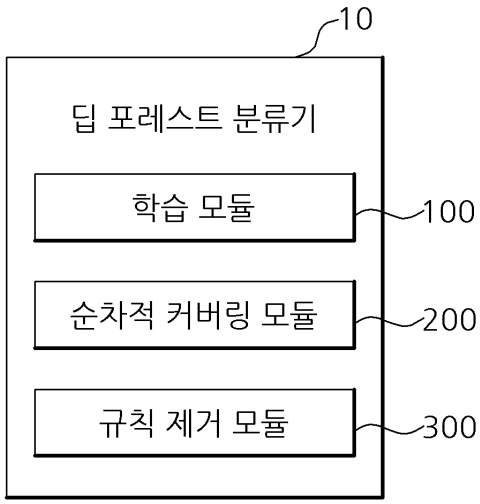
S250: 규칙 기여도 및 특징 패턴을 이용해 최종 규칙 기여도를 산출하는 단계

S260: 최종 규칙 기여도에 따라 의사결정 셋의 규칙들을 정렬하는 단계

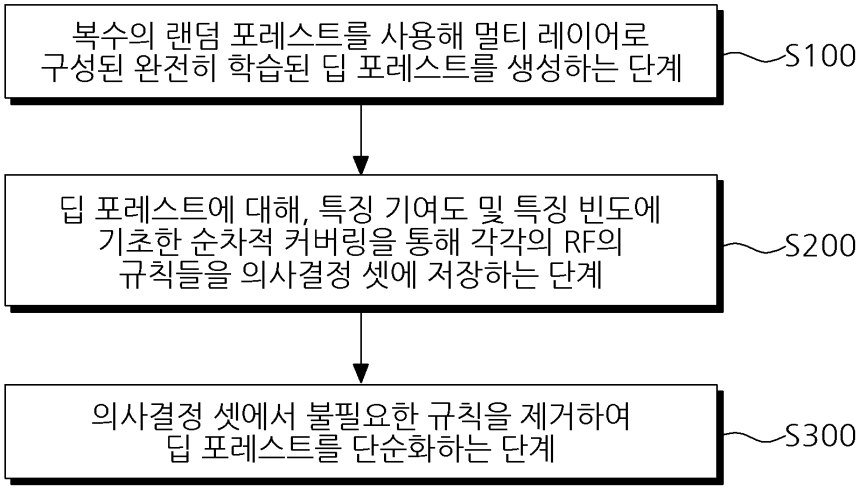
S300: 의사결정 셋에서 불필요한 규칙을 제거하여 딥 포레스트를 단순화하는 단계

도면

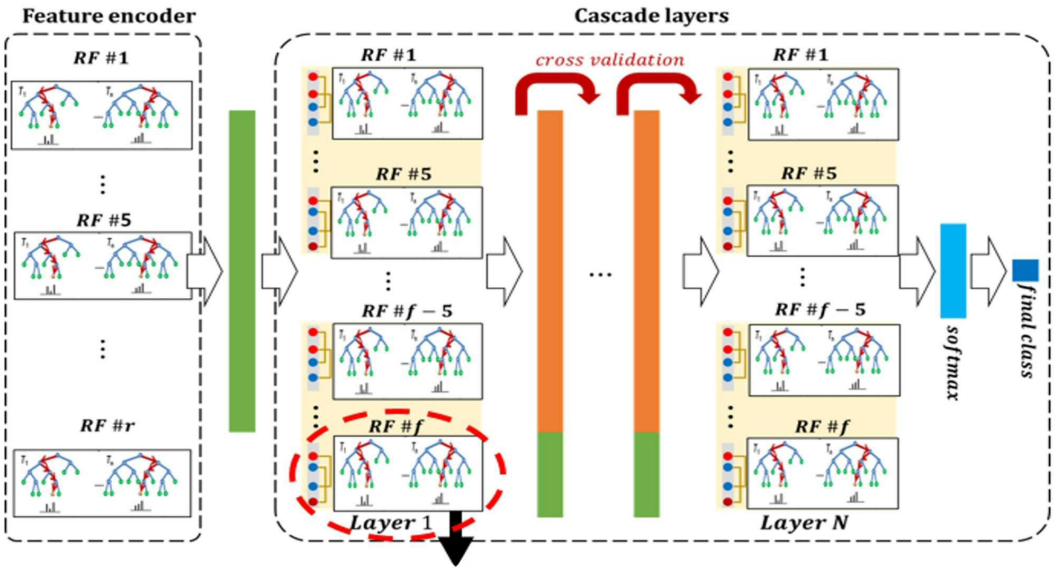
도면1


10
딥 포레스트 분류기
학습 모듈
100
순차적 커버링 모듈
200
규칙 제거 모듈
300

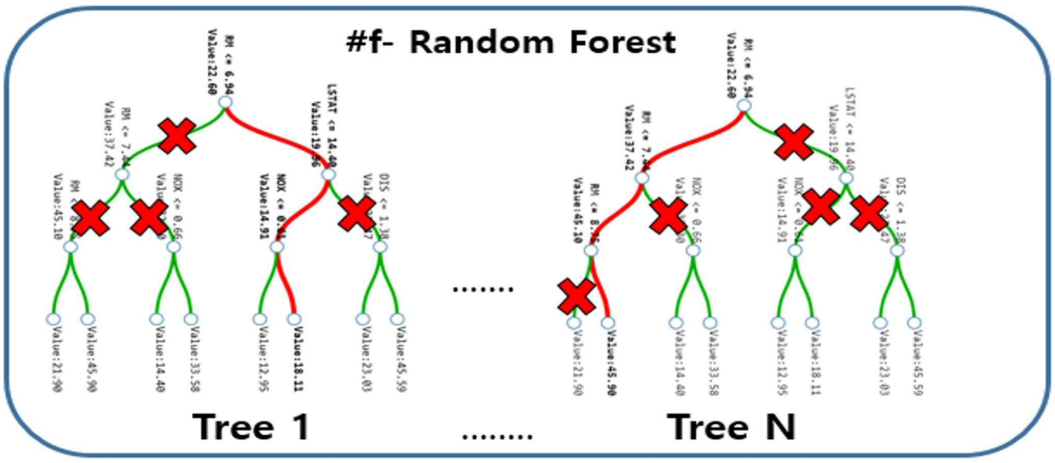
도면2


복수의 랜덤 포레스트를 사용해 멀티 레이어로 구성된 완전히 학습된 딥 포레스트를 생성하는 단계
S100
딥 포레스트에 대해, 특징 기여도 및 특징 빈도에 기초한 순차적 커버링을 통해 각각의 RF의 규칙들을 의사결정 셋에 저장하는 단계
S200
의사결정 셋에서 불필요한 규칙을 제거하여 딥 포레스트를 단순화하는 단계
S300

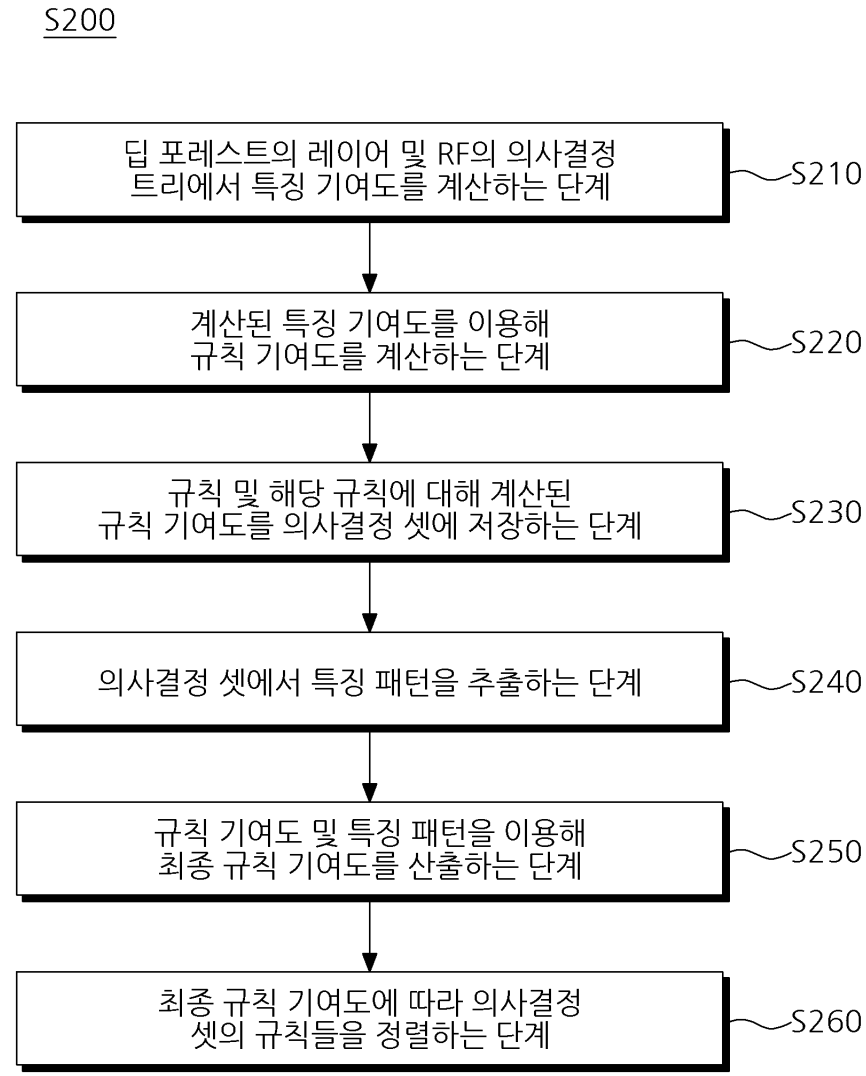
도면3


Feature encoder
Cascade layers
RF #1
RF #5
RF #r
RF #1
RF #5
RF #f − 5
RF #f
Layer 1
cross validation
RF #1
RF #5
RF #f − 5
RF #f
Layer N
softmax
final class

도면4


#f- Random Forest
Tree 1
.....
Tree N

도면5



도면6

Initial Rules of dSet (v, l)
Rule 1: $(x_2 > 2.59)$ *and* $(x_2 > 4.75)$ *and* $(x_0 <= 6.04)$ *and* $(x_3 > 1.84) \Rightarrow \{0\}$ [0, 0, 1]
Rule 2: $(x_2 > 2.59)$ *and* $(x_2 > 4.75)$ *and* $(x_0 <= 6.04)$ *and* $(x_3 <= 1.84) \Rightarrow \{-0.277\}$ [0, 0.5, 0.5]
Rule 3: $(x_3 <= 1.75)$ *and* $(x_3 > 0.7)$ *and* $(x_3 <= 1.55) \Rightarrow \{0.55\}$ [0, 1, 0]
.....
Reordered Rules of dSet (v, l)
Rule 1: $(x_3 > 0.7)$ *and* $(x_3 > 1.55)$ *and* $(x_2 <= 4.95) \Rightarrow \{0.83\}$ [0, 0.6, 0.4]
Rule 2: $(x_3 <= 1.75)$ *and* $(x_3 > 0.7)$ *and* $(x_3 <= 1.55) \Rightarrow \{0.55\}$ [0, 1, 0]
Rule 3: $(x_2 > 2.59)$ *and* $(x_2 > 4.75)$ *and* $(x_0 <= 6.04)$ *and* $(x_3 <= 1.84) \Rightarrow \{-0.277\}$ [0, 0.5, 0.5]

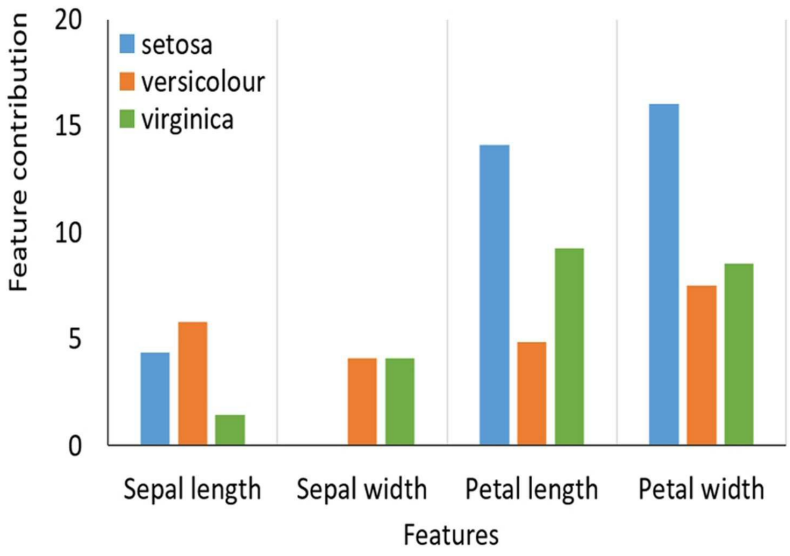
도면7

Algorithm 1: Rule elimination based on feature contribution and feature pattern

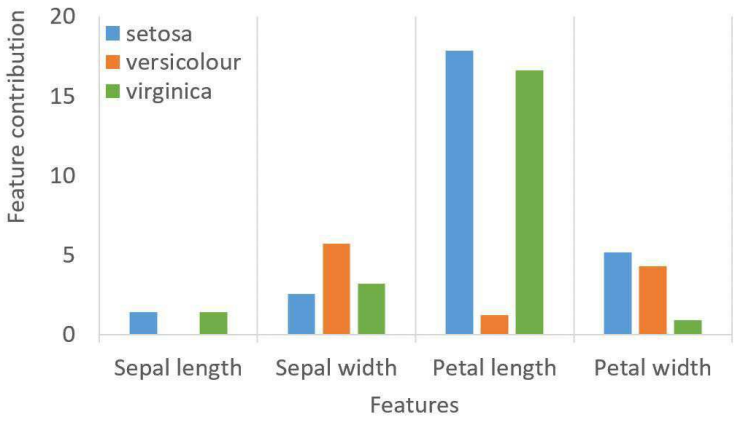
Input: The number of layers N, the number of RFs V, the number of trees T, random forest RF, list of dSets $\text{List}_{\text{dSet}}$

Start with an empty list of dSets, $\text{List}_{\text{dSet}} = \emptyset$
Learn LMRF
For each l layer**:**
  For each v RF:
    For each t tree:
      -Split a i-th rule from a decision tree
      -Calculate feature contribution of a i-th rule
      $feat.contrib\ (i,*)$
      -Calculate rule contribution for i-th rule r_i^t
      -Add rule and its r_i^t to $\text{dSet}(v,l)$
    End
    -Compute feature pattern $fre_v^l(x_j = A)$ by splitting rules in $\text{dSet}(v,l)$
    -Re-compute a new rule contribution $\text{r}^*{}_i$
    -Sort rules in $\text{dSet}(v,l)$ according to $\text{r}^*{}_i$
    -Add $\text{dSet}(v,l)$ to $\text{List}_{\text{dSet}}$ of l-th layer
      $\text{List}_{\text{dSet}}[l] = \text{List}_{\text{dSet}}[l] + \text{dSet}(v)$
  End
End
Output: The $\text{List}_{\text{dSet}}[l]$ consists of l layers

도면8



도면9


Feature contribution
20
15
10
5
0
setosa
versicolour
virginica
Sepal length
Sepal width
Petal length
Petal width
Features

도면10

