

인공지능2 - 03 베이즈 정리

Bayes' theorem

(주)씨에스리 기술연구소 전해경 이사
(hgchon@cslee.co.kr)

세상의 불확실성을 반영한 “주관적 확률이론” 베이즈 정리

<p>Concept</p>	<p>(정의) 이전의 경험과 현재의 증거를 토대로 어떤 사건의 확률의 정확도를 보정하는 방식으로 추론하는 주관적 확률이론 (특징) 불확실성 추론, 사후 확률</p> $P(A B) = \frac{P(A) P(B A)}{P(B)} = \frac{P(A)P(B A)}{P(A)P(B A) + P(A^c)P(B A^c)}$
<p>KeyWord</p>	<p>주관적 확률이론, 불확실성 추론, 사전확률, 사후확률(수정된 확률),</p>

주관적 확률 이론인 베이즈 정리가 나오게 된 배경

주사위를 던졌을 때 1 이 나올 확률은? 네~ 초등학교도 다 아는 1/6 이 맞습니다.

그런데, 현실에서는 어떤 사람이 주사위를 던지면 던지는 면이 항상 그대로 나오는 경우가 종종 생깁니다. 그럼 이 경우에도 1 이 나올 확률이 1/6 일까요? 이 경우 이 사람이 주사위를 던졌을 때 ‘주사위 윗면이 그대로 나온 통계(확률)’과 1 이 윗면으로 올 확률을 통해서 보정을 해본다면 좀 더 현실적인 확률이 나올 수 있지 않을까요?

고전적 확률이론은 “주사위를 무한대만큼 던지는 실험을 했을 때”라는 가정하에 동등한 비율로 확률이 도출된다는 동등분포이론(베르누이 무차별원칙)을 기반으로 하는데, 실제 현실에서는 무한대만큼 실험을 할 수도 없고, 불확실성도 많기 때문에 그대로 적용하기에는 한계가 있습니다.

그래서 확률의 정확도를 높이기 위한 시도로 ‘상대적 빈도 확률이론’, ‘선행적 확률이론’, ‘주관적 확률이론’ 등이 발전되어 나오게 되었습니다. 베이즈 정리(Bayes' theorem)는 이 중에서 ‘주관적 확률이론’의 대표적인 이론입니다. 베이즈 정리에 의하면 하나의 가설을 정하고 그 가설을 믿는 정도를 ‘사전적 확률’로 정하고 과거의 통계와 측정된 데이터 등을 반영하여 우리의 믿음(사전확률)을 보정한 ‘사후적 확률’을 구하는 통계 이론입니다.

그렇다면 왜 인공지능과 기계학습에서 베이즈 정리를 알아야 하는가?

인간의 인지과정을 살펴보면 인간은 사건에 대한 판단을 내릴 때 과거의 경험과 지식을 바탕으로 ‘~일 것이다’라는 가설을 세우고 새로운 정보가 생기면 계속 지식을 업데이트해서 판단의 확실성을 더해갑니다. 즉, 인간이 가설을 세우고, 여러 정보를 이용해서 가설의 확률을 계속 보정해 나간다는 점에서 **인간이 생각하고 판단하는 인지의 방법과 베이즈의 사상이 매우 유사합니다.** 이러한 이유로 인공지능 및 기계학습 분야에서 베이즈 정리를 기초로 하는 기법들을 발전시켜왔고 많이 사용되고 있습니다.

I. 데이터를 이용한 사후확률의 추정, 베이즈 정리의 개요

가. 베이즈 정리 정의

- 사전적 확률을 구한 후 확률에 영향을 미치는 변수의 확률을 반영하여 사후 확률을 구하는 방법으로 확률의 정확도를 높이고자 하는 주관적 확률 이론

예를 들어 설명하면

- 가설: 화성에는 생명체가 있다(H)
- 특정인 i가 이 가설 H를 선형적으로 믿는 정도를 $P(H)$ 라고 함 (← 사전적 확률)
- 생명체 탐지기 x가 화성에 보내져서 화성에 생명체가 있는가, 없는가를 탐지한다고 하자. 그런데 탐지기 x는 완전한 기계가 아니어서 오류를 범할 수 있다.
- 이 탐지기 x가 측정한 데이터 D의 신뢰도는 다음과 같다.
 - ① 실제로 생명체가 있는데(H) 생명체가 있다고 확률을 $P(D|H)$
 - ② 생명체가 없는데(\bar{H}) 생명체가 있다고 할 확률(오류율)을 $P(D|\bar{H})$ 라 하고 D를 data report라고 하자.
- 이 탐지기를 실제로 화성에 보냈더니 생명체가 있다는 보고(D)를 보내왔다

이 보고를 받고 특정인 i는 자신이 가지고 있는 H에 대한 믿음을 수정하게 될 것이다. 이 수정된 가설을 $P(H|D)$ 라는 새확률(사후적 확률)로 표시하고 이 확률을 구하는 것이 베이즈 정리의 핵심이다

나. 베이즈 정리 특징

- 불확실성 추론 : 직접 관측으로는 쉽게 얻어낼 수 없는 현상 추론에 용이
- 사후확률 : 사전지식 $P(A|B)$ 으로부터 사후확률 $P(B|A)$ 를 추론

II. 데이터를 이용한 사후확률의 추정 방법

가. 이전의 경험과 현재의 증거를 토대로 어떤 사건의 확률을 추정하는 예시

[예제] 진단검사 문제(어떤 사람이 병에 걸렸을 확률)

1. 처음에는 어떤 사람이 병에 걸려있을 확률에 대해 아는 것이 없어, 전 세계 인구 일반이 해당 질병에 걸릴 확률 유병률 1% 가정(=선형적 확률)
2. 정확도가 90%인 검사를 받았더니 '양성판정'을 받음
3. 양성 판정을 받은 새로운 사실을 토대로 이 사람이 실제로 병에 걸려있을 확률(=사후적 확률)을 추정

나. 예제의 간단한 수학적 표현 (베이즈 정리)

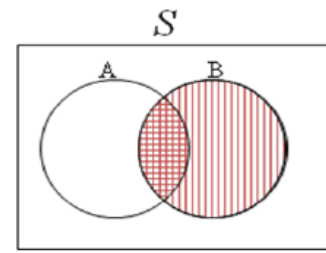
- 병에 걸렸다(A), 양성판정(B)

	(사건 A)의 사후 확률	∝	(사건 A)의 사전 확률	×	가능성(정확도)
수학적 표현	$P(A B)$		$P(A)$		$P(B A)$
예(진단검사)	$P(\text{병에 걸렸다} \text{양성 판정})$		$P(\text{병에 걸렸다})$		$P(\text{양성 판정} \text{병에 걸렸다})$

(간단한 확률 개념) 사건 B가 일어났다는 가정 하에 사건 A가 일어날 확률(A와 B의 교집합 확률)

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, P(B) > 0$$

$$\therefore P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$$



다. 전확률(Total Probability) 이론을 적용한 베이즈 정리

- 위의 예제는 영향을 미치는 인자가 1개일 경우로 사건이 2개로 간단하게 표현이 되지만, 현실에서는 영향을 미치는 인자를 최대한 많이 반영한다면 좀 더 정확한 사후 확률을 구할 수 있게 됨 → 전확률 이론

1) 전확률(Total Probability) 이론: 서로 배반인(서로 겹치는 경우가 없는) 사건 A_1, A_2, \dots, A_n 이 어떤 시행의 결과로 일어나고, 이들 사이의 완전계(A_1, A_2, \dots, A_n 전체의 합집합=전체집합)를 이루어서 임의의 사건B가 일어난다고 보는 이론

(전확률 이론의 수학적 표현)

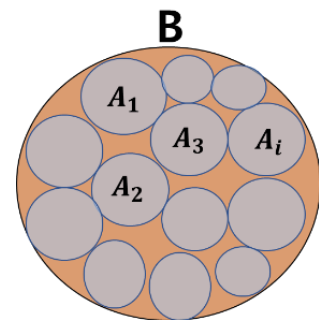
$$P(B) = P(B \cap A_1) + P(B \cap A_2) + P(B \cap A_3) + \dots + P(B \cap A_n)$$

수학적 표현으로 정리하면,

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \dots + P(B|A_n)P(A_n)$$

$$\therefore P(B) = \sum_{i=1}^n P(B \cap A_i) = \sum_{i=1}^n P(B|A_i)P(A_i)$$

라는 일반식으로 정리될 수 있음



2) 전확률(Total Probability)을 적용한 베이즈 정리 도출

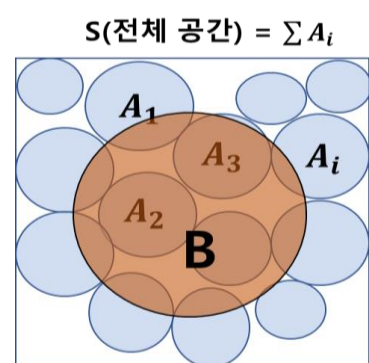
B를 표본공간 S 내의 임의의 사건이라 하고 서로 배반인 사건 A_1, A_2, \dots, A_n 의 합(Union)이 전체공간 S라 하면, B 사건 근거로 A_j 사건이 발생할 확률은 다음과 같음

(전확률 이론을 적용한 베이즈 정리) B의 관측값이 주어진 경우 A_j 사건의 사후 확률

$$P(A_j|B) = \frac{P(B \cap A_j)}{P(B)} = \frac{P(B|A_j)P(A_j)}{P(B)}$$

$$= \frac{P(B|A_j)P(A_j)}{\sum_i P(B|A_i)P(A_i)}$$

라는 일반식으로 정리될 수 있음



III. 인지과학 및 인공지능에서의 베이지 정리의 활용

- 베이지정리는 처음에는 아무 정보가 없던 상태에서 정보를 받아들이고, 측정된 데이터를 기반으로 기존 사전확률을 업데이트하고 사후확률의 정확도를 높여나가는 점에서 인간의 사고 과정을 표현하는데 적합
- 그러나 기존에는 사전분포와 가능도가 잘 매칭되지 않는 문제와 사후분포를 사람의 손으로 계산하는데 어려움으로 활용에 어려움이 있었으나 최근에는 컴퓨터 계산 능력의 발전으로 활용도가 높아짐
- 인공지능 및 기계학습 분야에서 알고리즘 설계시에 베이지 정리를 기초로 알고리즘 업그레이드 적용
- 최근에는 복잡한 베이지 알고리즘이 아닌 단순화시킨 베이지인 ‘나이브 베이지(Naïve Bayes)’ 알고리즘도 1950년대 이후 광범위하게 연구되고 있으며, 텍스트 분류 및 스팸메일 분류 등에 이용되고 있음

활용분야	설명
패턴인식 활용	- 통계적 모델을 기반으로 패턴 분류 - 베이지 정리를 이용하여 패턴 분류하여 인식
스팸메일 필터링	- 베이지 규칙을 이용하여 전자편지에 들어있는 일반 낱말 개수와 스팸 낱말 개수를 비교 하여, 스팸일 가능성이 얼마나 되는지를 알려주는 확률 값을 계산하여 필터링
기타활용	- 임상실험, 기상예보, 품질관리 등 다양한 분야에서 활용

“끝”

Contents connect communications!!

아이리포에 오시면 더 많은 지식을 가져가실 수 있습니다.

아이리포 온라인 : <http://www.ilifo.co.kr>

아이리포 지덤시리즈 : <http://www.jidum.com>

아이리포 IT지식창고 : <https://www.ilifo.co.kr/boards/knowledge>

아이리포 기술사/감리사 카페 : <http://cafe.naver.com/itlf>

서울시 마포구 상암동 1610번지, DDMC 3층 아이리포 교육센터

TEL: 02-303-9997 | MAIL: edu@ilifo.co.kr