

## ✓ 텍스트 분류

- 텍스트 문서나 문장을 사전 정의된 클래스나 카테고리로 분류하는 머신 러닝의 한 유형
- 텍스트의 내용과 의미를 분석한 다음, 텍스트 라벨링을 사용하여 가장 적절한 라벨을 할당함
- 텍스트 분류의 실제 적용에는 정서 분석(리뷰에서 긍정적이거나 부정적인 감정 결정), 스팸 탐지(예: 정크 이메일 발견) 및 주제 분류(예: 뉴스 기사를 관련 주제로 구성)가 포함됨
- 텍스트 분류는 컴퓨터가 대량의 비정형 텍스트를 이해하고 구성할 수 있도록 함으로써 자연어 처리(NLP)에서 중요한 역할을 함
- 이를 통해 콘텐츠 필터링, 추천 시스템, 고객 피드백 분석 등의 작업이 단순화됨

### • 텍스트 분류 유형

#### ◦ 텍스트 정서 분석

- 텍스트에서 표현된 정서나 감정을 파악하여 일반적으로 긍정적, 부정적 또는 중립으로 분류합니다. 이는 제품 리뷰, 소셜 미디어 게시물, 고객 피드백을 분석하는 데 사용됨
- 텍스트 정서 분석과 관련된 독성 감지는 온라인에서 공격적이거나 유해한 언어를 식별 - 이는 온라인 커뮤니티의 관리자가 온라인 토론, 댓글 또는 소셜 미디어 게시물에서 존중받는 디지털 환경을 유지하는 데 도움이 됨

#### ◦ 의도 인식

- 사용자의 텍스트 입력 이면에 있는 목적(또는 의도)을 이해하는 데 사용되는 텍스트 정서 분석의 또 다른 하위 집합임
- 챗봇과 가상 비서는 종종 의도 인식을 사용하여 사용자 쿼리에 응답함

#### ◦ 이진 분류

- 텍스트를 두 클래스 또는 카테고리 중 하나로 분류함

- 일반적인 예로는 이메일이나 메시지와 같은 텍스트를 스팸 또는 적절한 카테고리로 분류하여 원치 않는 잠재적으로 유해한 콘텐츠를 자동으로 필터링하는 스팸 메일 분류하기가 있음

#### ◦ 다중 클래스 분류

- 텍스트를 3개 이상의 개별 클래스 또는 카테고리로 분류함
- 이를 통해 뉴스 기사, 블로그 게시물, 연구 논문 등의 콘텐츠에서 정보를 더 쉽게 구성하고 검색할 수 있음
- 다중 클래스 분류와 관련된 주제 분류는 문서나 기사를 사전 정의된 주제나 테마로 그룹화함
  - 예를 들어, 뉴스 기사는 정치, 스포츠, 연예 등의 주제로 분류될 수 있음

#### ◦ 언어 식별

- 텍스트가 작성되는 언어를 파악함
- 이는 다국어 컨텍스트 및 언어 기반 애플리케이션에 유용함

#### ◦ 개체명 인식

- 사람 이름, 조직, 위치, 날짜 등 텍스트 내에서 명명된 개체를 식별하고 분류하는 데 중점을 둠

#### ◦ 질문 분류

- 예상 답변 유형을 기준으로 질문을 분류하는 작업을 다루며, 이는 검색 엔진 및 질문 답변 시스템에 유용할 수 있음

## ✓ AI 모델 > 분류 - 텍스트 모델

- 학습 데이터로 입력한 텍스트를 클래스로 직접 분류하고 학습시키면 새롭게 입력되는 텍스트를 분류할 수 있는 모델을 만들 수 있음

- [텍스트의 의미가 아니라 형태가 얼마나 비슷한지를 기준으로 분류하는 모델임](#)
- 모델이 학습할 텍스트는 직접 작성하거나 txt, csv 파일로 업로드할 수 있음
  - txt, csv 파일로 업로드하면 이미 입력된 내용의 맨 마지막에 추가됨
  - 텍스트 모델은 고급 모드(학습 조건 설정 및 학습 차트 표시)를 지원하지 않음



- 데이터 입력 팝업 창을 열고, 입력한 데이터를 학습한 모델로 분류

학습한 모델로 분류하기



- 이 블록의 값 블록(value)에 입력한 텍스트를 학습한 모델로 분류

엔트리

을(를) 학습한 모델로 분류하기



- 입력한 데이터를 학습한 모델로 분류한 결과를 가져오는 값 블록
- 입력한 데이터와 가장 비슷한 클래스의 이름을 가져와요. 만일 데이터를 입력하지 않았다면 아무 것도 가져오지 않음

분류 결과

- 입력한 데이터가 분류된 클래스에 대한 신뢰도를 가져오는 값 블록
- 만일 데이터를 입력하지 않았다면 0 을 가져옴

클래스 ▼

에 대한 신뢰도

- 입력한 데이터가 선택한 클래스와 가장 비슷하다면 참, 아니라면 거짓으로 판단하는 블록
- 만일 데이터를 입력하지 않았다면 항상 거짓을 가져옴

분류 결과가

클래스 ▼

인가?

## 분류: 텍스트 모델 학습하기

모델 학습은 인터넷이 연결되어 있어야 정상적으로 동작합니다. [튜토리얼 보기](#)

### 텍스트분류

#### 데이터 입력

모델이 학습할 데이터를 입력합니다.

긍정

5개



모델이 학습할 텍스트 데이터를 아래에 작성해 주세요.  
클래스당 5개 이상의 데이터를 입력해야 합니다.  
각각의 데이터는 쉼표로 구분합니다. (예: 맛있다, 맛있어, 맛있네)

기쁘다. 좋다. 즐겁다, 뿌듯하다. 행복하다. 날아갈 것 같다, 개운하

부정

25개



모델이 학습할 텍스트 데이터를 아래에 작성해 주세요.  
클래스당 5개 이상의 데이터를 입력해야 합니다.  
각각의 데이터는 쉼표로 구분합니다. (예: 맛있다, 맛있어, 맛있네)

간절하다, 간질간질하다, 갈구하다, 갈등하다, 갈팡질팡하다, 갈피를

#### 학습

입력한 데이터로 모델을 학습합니다.

학습하기

모델 학습을 완료했습니다.

100%

#### 결과

학습한 모델의 결과를 확인합니다.

결과 확인하기

분류한 클래스

부정



- [28-텍스트분류-인공지능](#)

