

설명 가능 인공지능 기법을 이용한 의류 분류에 관한 연구

최정훈, 서동준*
경북대학교

{cjh34544, *dongjunsuh}@knu.ac.kr

A Study on the Clothes Classification using Explainable Artificial Intelligence Method

Jeonghoon Choi, Dongjun Suh*
Kyungpook National University

요약

최근 인공지능은 컴퓨터 비전 분야에서 탁월한 성과를 보여주고 있다. 그러나 실제 분야의 응용에서는 모델의 결정에 대한 설명이 중요하다. 본 연구에서는 의류 분류 문제에 대한 해석 가능한 인공지능 기법의 적용을 제안하였다. Fashion MNIST 데이터셋을 사용하여 기본 CNN 모델을 학습하고, SHAP를 통해 모델의 의사 결정을 설명 가능하게 하였다. CNN 모델은 높은 분류 정확도를 달성하였고, SHAP 분석을 통해 각 특성이 분류에 미치는 상대적인 중요성을 시각적으로 확인하였으며, 향후 사용자 경험 및 피드백을 수용하여 XAI 기법을 보다 사용자 친화적으로 발전시키는 데 초점을 맞출 예정이다.

I. 서론

최근 인공지능은 다양한 분야에서 탁월한 성과를 보이고 있으며, 특히 컴퓨터 비전 분야에서는 의류 분류와 같은 과제에 있어서 놀라운 성능을 보여주고 있다. 그러나 이러한 복잡한 모델들은 종종 의사 결정 과정이 블랙박스처럼 불투명하게 이루어진다는 문제를 안고 있다. 특히 의류 분류와 같은 실제 응용에서는 모델의 결정에 대한 설명이 중요하다. 사용자 및 이해관계자들은 모델의 작동 원리를 이해함으로써 모델의 신뢰성을 높이고, 필요한 경우에는 모델의 예측을 수정할 수 있는 통찰력을 얻을 수 있다 [1].

본 연구에서는 XAI(eXplainable Artificial Intelligence) 기법을 의류 분류 문제에 적용하고자 한다. XAI는 모델의 의사 결정을 설명 가능하게 만드는 기법을 의미하며, 이는 사용자가 모델의 예측을 이해하고 신뢰할 수 있도록 도와준다 [2]. 본 논문에서는 의류 분류를 위한 딥러닝 모델인 CNN 모델에 대해 XAI 기법을 적용함으로써 모델의 투명성을 높이고, 모델이 어떻게 특정 의류를 분류하는지에 대한 해석 가능성을 제공하는 방법을 제안한다.

II. 본론

II-1. 배경 및 동기

의류 분류는 해당 분야의 실용적인 측면에서 중요한 문제이다. 특히 최근 온라인 쇼핑 산업의 급속한 성장과 함께 기업들은 정확한 의류 분류가 제공되어야 하고, 소비자는 온라인 쇼핑이나 오프라인 매장에서 적합한 제품

을 찾을 수 있어야 한다. 이러한 문제는 컴퓨터 비전 및 기계 학습 기술의 발전에 따라 해결 가능성이 높아졌다. 그러나 모델의 의사 결정에 대한 불투명성은 여전히 남아 있으며, 특히 의류 분류와 같이 사용자가 직접적으로 영향을 받는 응용 분야에서는 모델의 설명이 필수적이다.

XAI의 적용은 이러한 문제에 대한 해결방법으로 간주될 수 있다. 제안하는 기술은 고도로 복잡한 딥러닝 모델의 내부 작동 원리를 투명하게 설명함으로써 모델의 의사 결정을 실시간으로 이해할 수 있도록 도와준다. 최근 연구에서는 이미지 분류 작업에서의 XAI의 성과가 입증되었지만, 의류 분류의 적용에 대한 연구는 아직 제한적이다. 이에 본 연구에서는 XAI를 적용하여 모델의 투명성을 높이고, 사용자가 모델의 의사 결정을 이해하고 신뢰할 수 있도록 하는 방안을 제시한다.

II-2. 데이터셋

본 논문에서는 Fashion MNIST 데이터셋을 활용하여 의류 분류 모델을 학습하고 XAI 기법을 적용하였다. Fashion MNIST 데이터셋은 의류 관련 이미지를 10개의 클래스로 구분하는 데이터셋으로, 총 60000개의 훈련 이미지와 10000개의 테스트 이미지로 구성되어 있다 [3].

데이터 전처리로는 이미지 크기를 입력 층에 맞게 크기를 조절하고 픽셀 값들을 정규화하여 모델의 학습을 원활하게 하였다. 또한, 클래스 불균형을 해결하기 위해 데이터 각 라벨 별로 1대1 샘플링을 수행하였다. 이러한 전처리는 모델의 정확도 및 분석 속도 향상과 해석 가능성을 증진시키는 데 기여하였다. 그림 1은 데이터셋의 예시 이미지를 나타내었다.

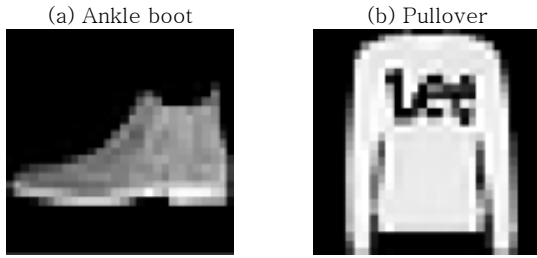


Figure 1 Example image of fashion MNIST dataset

II-3. 제안 기법

본 연구에서는 기본적인 CNN 을 의류 분류를 위한 모델로 선택하였다. CNN 은 이미지 처리에 탁월한 성능을 보이며, 의류 분류와 같은 컴퓨터 비전 작업에 적합한 모델이다.

CNN 아키텍처는 입력 이미지로부터 특징을 추출하고 분류를 수행하기 위한 다층 구조로 이루어져 있다. 본 연구에서는 2 개의 합성곱 레이어와 풀링 레이어가 번갈아 쌓인 구조를 사용하였다. 마지막으로 완전 연결 레이어를 통해 최종 의사 결정을 수행하였다.

최종 분류 결정에 대한 해석을 위해 SHAP(Shapley Additive exPlanations)를 선택하여 XAI 를 구현하였다. SHAP 는 Shapley value 의 개념을 활용하여 각 특성이 모델의 예측에 얼마나 기여하는지를 설명하는 방법 중 하나이다. Shapley value 는 공정 분배 게임 이론에서 나온 개념으로 주어진 게임에 참여한 플레이어들에게 상금을 공정하게 배분하는 하나의 방법이다. 대표적인 성질인 효율성에 관한 식 (1)은 다음과 같다.

$$\sum_{i=1}^n \varphi_i(\delta, N) = \delta(N) \quad (1)$$

효율성은 모든 플레이어가 게임에 참여한 경우에 받을 수 있는 총상금 $\delta(N)$ 으로 표현되며, 분배과정 φ 가 주어진 경우, 각 플레이어에게 주는 상금의 합은 $\delta(N)$ 과 같아야 공정함을 의미한다.

즉, SHAP 는 개별 예측값에 대한 각 변수들의 영향력을 모델 상관없이 균등하게 배분하는 방식이라 할 수 있다. SHAP 는 모델의 예측을 특성 별로 설명 가능한 형태로 분해하여 제시함으로써 모델의 블랙박스 특성을 극복하고, 의사 결정의 이해를 돕는다. 이미지 분류에서 SHAP 가 적용되는 방식은 라벨 별 이미지에 의도적으로 블러링(blurring)이나 마스킹(masking)을 삽입하여 전체 이미지의 라벨 정확도 변화를 보고 픽셀 별 변수 중요도를 평가한다.

모델 학습 후, SHAP 를 통해 각 의류 라벨 분류에 대한 특성의 상대적인 중요성을 해석하여 제시한다. 이를 통해 사용자 및 이해관계자들은 모델이 특정 의류 아이템을 분류하는데 어떤 특성을 고려하는지에 대한 직관을 얻을 수 있다.

II-4. 실험 및 고찰

본 논문에서 제안한 설명 가능 인공지능을 이용한 의류 분류의 성능을 검증하기 위해 CNN 모델을 Fashion MNIST 데이터셋으로 학습한 결과, 90% 이상의 정확도를 확인하였다. 분류 결과에 대해 설명력을 부여하기 위해 SHAP 분석 결과, 모델의 각 특성이 예측에 미치는 상대적인 중요성을 시각화하였다. 특히 어떤 의류 아이템이 분류되었을 때 어떤 부분이 중요한지를 시각적으로 확인할 수 있는 SHAP Summary Plot 을 통해 중요한 특성을 그림 2 에 나타내었다.

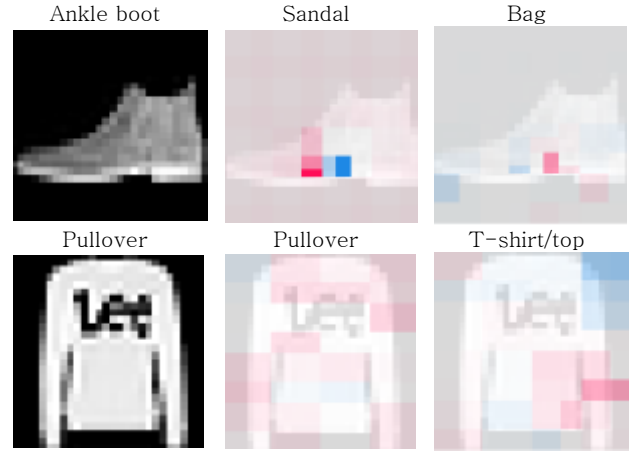


Figure 2 Example image of SHAP Summary plot

테스트 셋에서 2 개의 샘플을 가져온 후 설명 인자 개체에서 예측 확률 순서 별로 나타내었다. 붉은 부분은 해당 범주의 예측에 긍정적으로 기여한 부분을 나타내고, 파란 부분은 부정적으로 기여한 부분을 나타낸다.

첫번째 부츠 라벨 이미지는 샌들 라벨에 대해 신발이라는 카테고리에서 긍정적으로 예측한 부분이 많지만, 부정 영역이 존재하는 것으로 보아 분류상 오류가 난 것으로 설명할 수 있다. 두번째 풀오버 라벨 이미지는 정답 라벨로 예측해 냈으며 전반적으로 붉은 Shapley value 가 고르게 분포된 것을 확인하였다.

III. 결론

본 논문에서는 의류 분류에 XAI 기법을 적용하여 모델의 의사 결정을 설명 가능하게 만들었다. 이를 통해 모델의 의사 결정 해석에 도움이 되었고, 예측 결과를 더욱 신뢰할 수 있도록 하였다. 추후 연구에서는 의류 분야뿐만 아니라 다른 이미지 기반 작업에 대해서도 확장할 예정이며, 사용자 경험 및 피드백을 수용하여 XAI 기법을 보다 사용자 친화적으로 발전시키는 데 초점을 맞출 예정이다.

ACKNOWLEDGMENT

본 연구는 정부(과학기술정보통신부, 교육부)의 재원으로 한국연구재단의 지원을 받아 수행되었음 (과제번호: 2021R1A5A8033165, 2021R1I1A3049503)

참고 문헌

- [1] Arrieta et al. "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI." Information fusion 58: 82-115, 2020.
- [2] Adadi et al. "Peeking inside the black-box: a survey on explainable artificial intelligence (XAI)." IEEE access 6: 52138-52160, 2018.
- [3] Xiao et al. "Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms." arXiv preprint arXiv:1708.07747, 2017.