

Conformal Prediction by Examples

2. ImageNet + APS: Adapting to classification difficulty

Byol Kim

August 19, 2025

지난 시간에는

분할 컨포멀 예측 워크플로

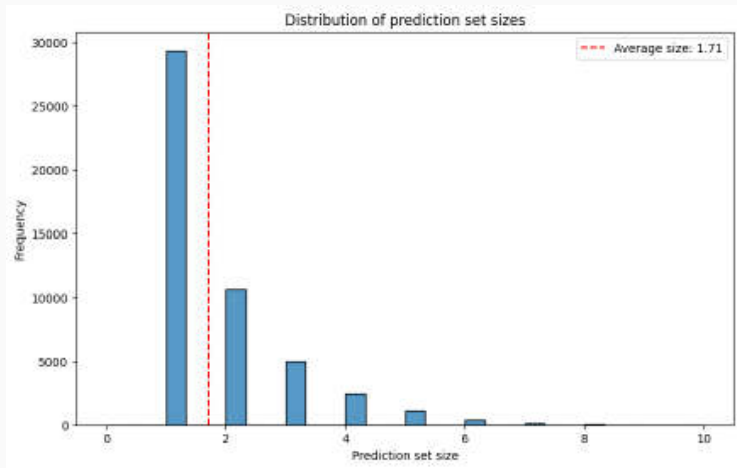
1. Score
2. Calibrate
3. Construct

바람직한 성질들

1. Validity ✓
2. Efficiency
3. Adaptivity

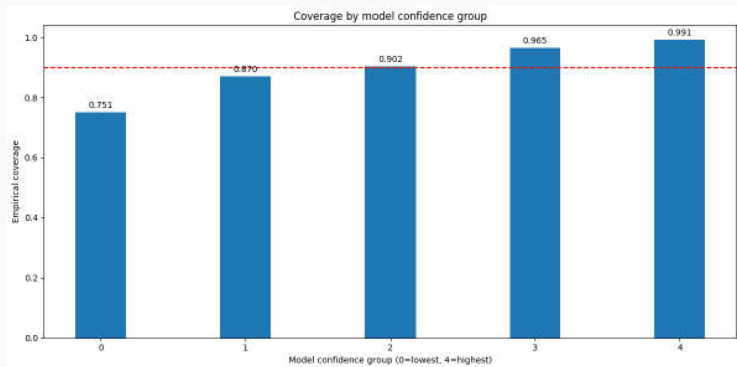
지난 시간에는

ImageNet 에 $s(x, y) = 1 - \hat{f}(x)_y$ 로 분할 컨포멀 예측을 적용한 결과



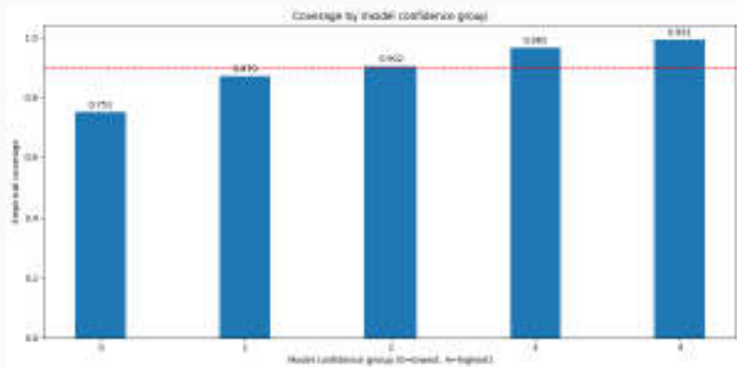
지난 시간에는

ImageNet 에 $s(x, y) = 1 - \hat{f}(x)_y$ 로 분할 컨포멀 예측을 적용한 결과



지난 시간에는

ImageNet 에 $s(x, y) = 1 - \hat{f}(x)_y$ 로 분할 컨포멀 예측을 적용한 결과



이번 시간

APS 논컨포미티 점수를 사용하여 분류 난이도에 적응

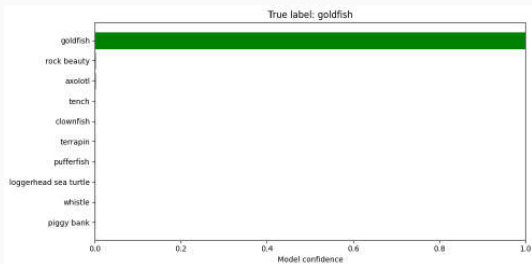
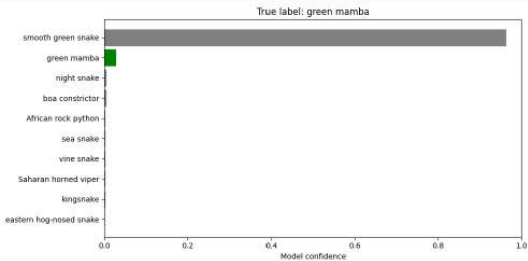
1. APS 논컨포미티 점수
2. 분할 컨포멀 예측 워크플로: 다음 실습의 경우
3. 주변 커버리지 보장

APS 논컨포미티 점수

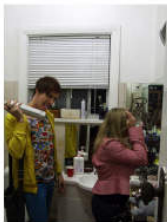
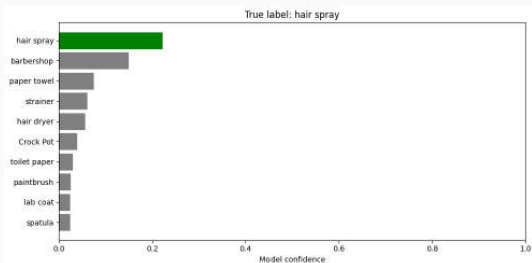
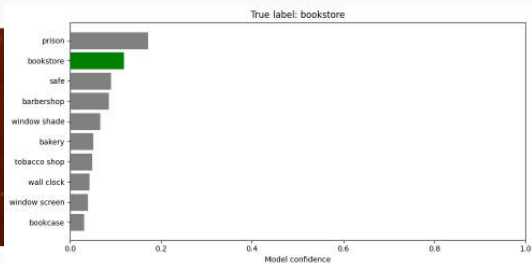
$$s(x, y) = 1 - \hat{f}(x)_y$$

단 하나의 소프트맥스 점수만을 활용

모델이 높은 확신을 보이는 이미지들



모델이 낮은 확신을 보이는 이미지들



만약 이미지 x 가 각각의 클래스에 속할 진짜 확률들

$$f^*(x)_k = \mathbb{P}\{Y = k | X = x\}$$

을 알고 있다면 구성할 수 있는 가장 이상적인 $100 \times (1 - \alpha)\%$ 예측집합은?

만약 이미지 x 가 각각의 클래스에 속할 진짜 확률들

$$f^*(x)_k = \mathbb{P}\{Y = k | X = x\}$$

을 알고 있다면 구성할 수 있는 가장 이상적인 $100 \times (1 - \alpha)\%$ 예측집합은?

1. 확률이 가장 높은 클래스부터 가장 낮은 클래스까지 클래스들을 정렬

$$f^*(x)_{\pi(x)_1} \geq f^*(x)_{\pi(x)_2} \geq \dots \geq f^*(x)_{\pi(x)_K}, \quad \pi(x) \in \mathfrak{S}_K$$

2. 확률이 가장 높은 클래스부터 순차적으로 확률의 누적합이 최초로 $1 - \alpha$ 를 넘을 때까지 클래스들을 추가

$$C^*(x) = \left\{ \pi(x)_1, \dots, \pi(x)_k \text{ s.t. } \sum_{j=1}^{k-1} f^*_{\pi(x)_j} < 1 - \alpha, \sum_{j=1}^k f^*_{\pi(x)_j} \geq 1 - \alpha \right\}$$

문제?

- 진짜 확률 $f^*(x)$ 를 모름

문제?

- 진짜 확률 $f^*(x)$ 를 모름
- $f^*(x)$ 를 $\hat{f}(x)$ 로 대체

문제?

- 진짜 확률 $f^*(x)$ 를 모름
- $f^*(x)$ 를 $\hat{f}(x)$ 로 대체
- 커버리지 보장 불가

문제?

- 진짜 확률 $f^*(x)$ 를 모름
- $f^*(x)$ 를 $\hat{f}(x)$ 로 대체
- 커버리지 보장 불가
- **컨포멀화 (conformalizing)** 문제의 구조에서 얻은 영감을 바탕으로 논컨포미티 점수를 설계, 분할 컨포멀 예측을 시행
 $\rightsquigarrow$ **APS(adaptive predictive sets)** 논컨포미티 점수

$$s(x, y) = s(x, y; f) = \sum_{k: f(x)_k > f(x)_y} f(x)_k$$

	높은 확신	낮은 확신
참 클래스가 최고 확률	0	0
그 외 클래스가 최고 확률	≈ 1	참 클래스의 순위에 따름

Romano et al. (2020) 가 제안

**분할 컨포멀 예측 워크플로: 다음 실
습의 경우**

일반적인 분할 컨포멀 예측 워크플로

주어진 것 사전 훈련된 예측 모델 $\hat{f}$

1. Score

$$s(x, y) = s(x, y; \hat{f})$$

2. Calibrate

$$S_i = s(X_i, Y_i), \quad i = 1, \dots, n$$

$\hat{q} = S_1, \dots, S_n$ 중 $\lceil (n+1)(1-\alpha) \rceil$ 번째 작은 수

$$= \frac{1}{n} \sum_{i=1}^n \delta_{S_i} \text{ 의 } \left(1 + \frac{1}{n}\right) (1-\alpha)\text{-분위수}$$

3. Construct

$$\hat{C}(x) = \{y : s(x, y) \leq \hat{q}\}$$

분할 컨포멀 예측 워크플로: APS 논컨포미티 점수의 경우

주어진 것 $N = 50000$ 개의 이미지-레이블 쌍과 각각의 $\hat{f}$ 값

1. Score

$$s(x, y) = \sum_{k: \hat{f}(x)_k > \hat{f}(x)_y} \hat{f}(x)_k$$

2. Calibrate

$$\begin{aligned} S_i &= s(X_i, Y_i), \quad i = 1, \dots, n \\ \hat{q} &= S_1, \dots, S_n \text{ 중 } \lceil (n+1)(1-\alpha) \rceil \text{ 번째 작은 수} \\ &= \frac{1}{n} \sum_{i=1}^n \delta_{S_i} \text{ 의 } \left(1 + \frac{1}{n}\right) (1-\alpha)\text{-분위수} \end{aligned}$$

3. Construct

$$\hat{C}(x) = \left\{ y : \sum_{k: \hat{f}(x)_k > \hat{f}(x)_y} \hat{f}(x)_k \leq \hat{q} \right\}$$

주변 커버리지 보장

주변 커버리지 보장

Q APS 논컨포미티 점수를 사용한 분할 컨포멀 예측은 주변 커버리지를 보장할까?

주변 커버리지 보장

Q APS 논컨포미티 점수를 사용한 분할 컨포멀 예측은 주변 커버리지를 보장할까?

핵심!!

$$\mathbb{P}\{S_{n+1} \leq S_{(k)}\} = \frac{k}{n+1}, \quad k = 1, \dots, n$$

- $S_1, \dots, S_n, S_{n+1}$ 모두 독립동일분포
- 모든 j 에 대해 S_{n+1} 가 $S_1, \dots, S_n, S_{n+1}$ 중 j 번째 작은 수일 확률 $= \frac{1}{n+1}$
- $S_{(0)} = -\infty, S_{(j)} = S_{(n;j)}$ 를 $S_1, \dots, S_n$ 중 j 번째 작은 수라 두면

$$\mathbb{P}\{S_{n+1} \leq S_{(k)}\} = \sum_{j=1}^k \mathbb{P}\{S_{(j-1)} \leq S_{n+1} \leq S_{(j)}\} = \sum_{j=1}^k \frac{1}{n+1} = \frac{k}{n+1}$$

이론적으로는 APS 논컨포미티 점수 사용에 영향을 받지 않음!!

요약

1. APS 논컨포미티 점수
2. 분할 컨포멀 예측 워크플로: 다음 실습의 경우
3. 주변 커버리지 보장

다음 실습 안내

코드를 완성하여 직접 APS 논컨포미티 점수에 의한 분할 컨포멀 예측을 구현

LABEL 논컨포미티 점수를 사용했을 때와 결과 비교

주의!! APS 논컨포미티 점수의 계산과 역변환 구현이 조금 까다로울수도