Multi-Output Conformal Prediction

2026. 09. 18

Data Mining & Quality Analytics Lab.

이 용 우

발표자 소개

❖ 이 용 우 (Yongwoo Lee)

- 고려대학교 일반대학원 산업경영공학과 재학
- Data Mining & Quality Analytics Lab. (김성범 교수님)
- 석사 과정 3학기차 (2025. 09 ~)

❖ Research Interest

- Tabular Data Deep Learning
- Uncertainty Quantification

❖ Contact

- ywlee1@korea.ac.kr

Contents

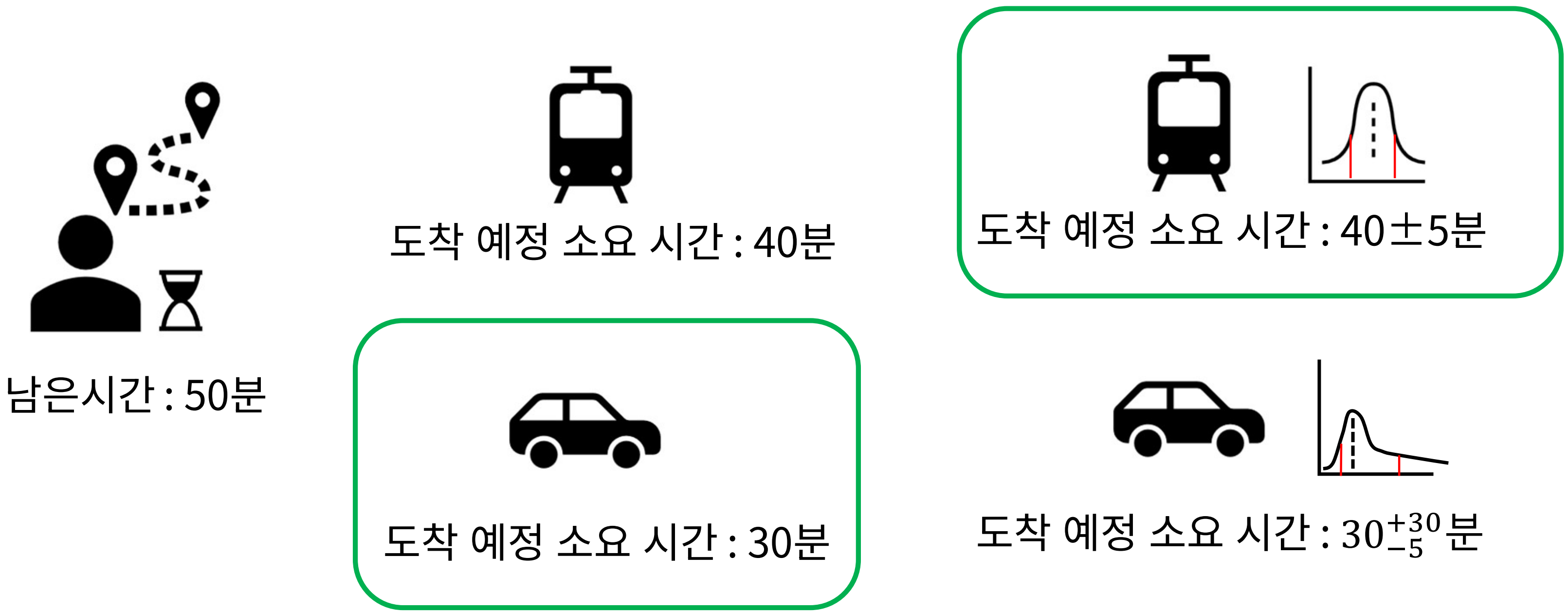
Introduction

Introduction

Background

❖ 점예측(Point-Prediction)만으로 충분한가?

- 점예측만으로는 개별 예측의 불확실성 크기를 알 수 없어 개별 예측의 위험도(Tail Risk)를 제어하지 못함
- 허용 가능한 위험 범위와 비교하기 위해 통계적으로 보장하는 예측구간/영역이 필요.

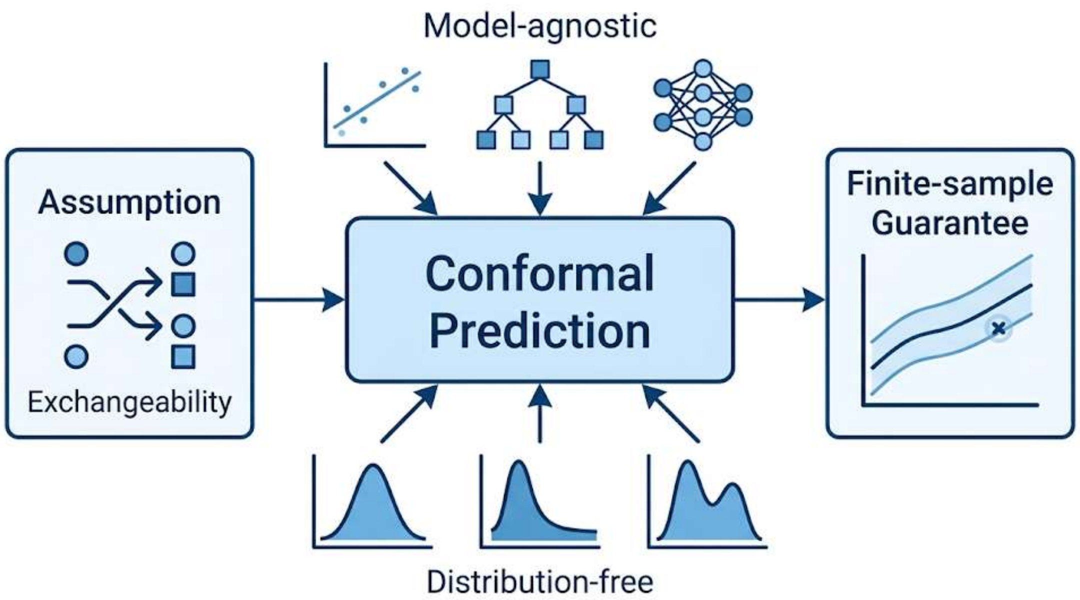


Introduction

Conformal Prediction

❖ Conformal Prediction이란?

- conform = 부합하다, 동조하다. → conformity: 새 데이터가 기존 데이터에 얼마나 잘 동조/부합하는가
- 교환가능성(exchangeability)하에, 예측모델 상관없이(Model-Agnostic),분포 상관없이(distribution-free) 유한 샘플에 대해 동조성(conformity)을 근거로 보장된 예측영역/구간을 만드는 절차

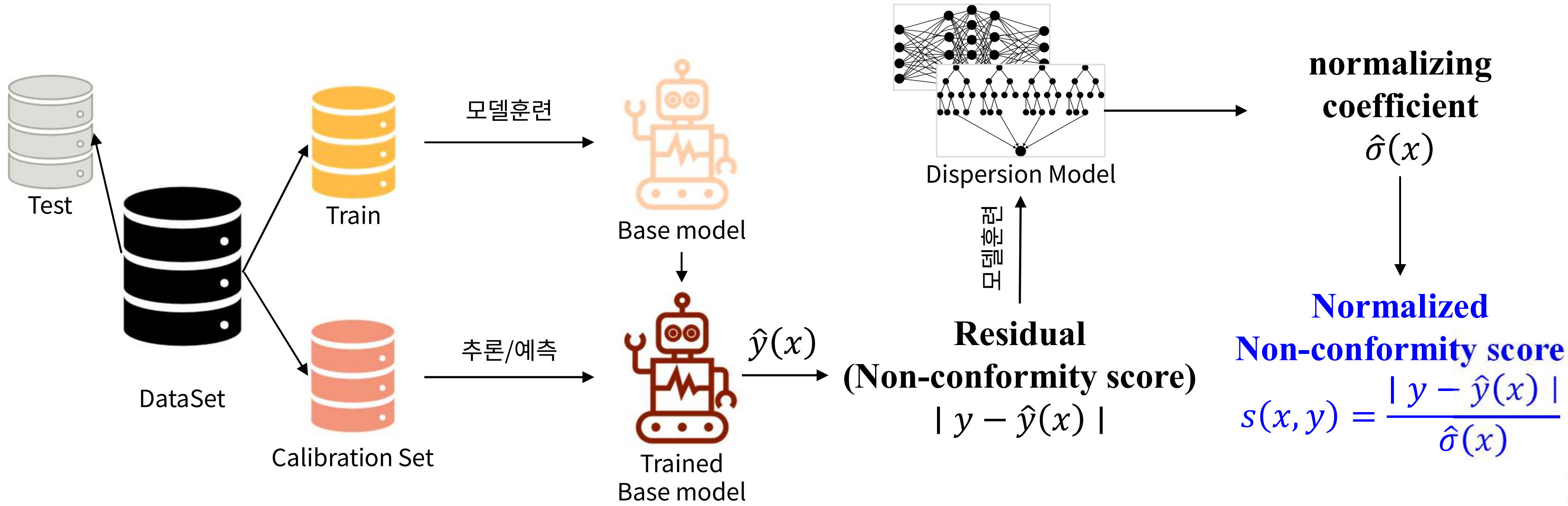


Introduction

Conformal Prediction

❖ 예측 영역을 어떻게 만드는가? (기본 아이디어 : Locally Adaptive CP)

- 보정 데이터(Calibration Set)의 비동조성 점수(Non-conformity score)를 정렬해 임계값 $\hat{q}$ 하나를 뽑는다
- 새로운 입력에 대해 비동조성 점수가 이 임계값 $\hat{q}$ 이하가 되는 정답의 범위를 역으로 계산해 예측 구간/영역을 만든다



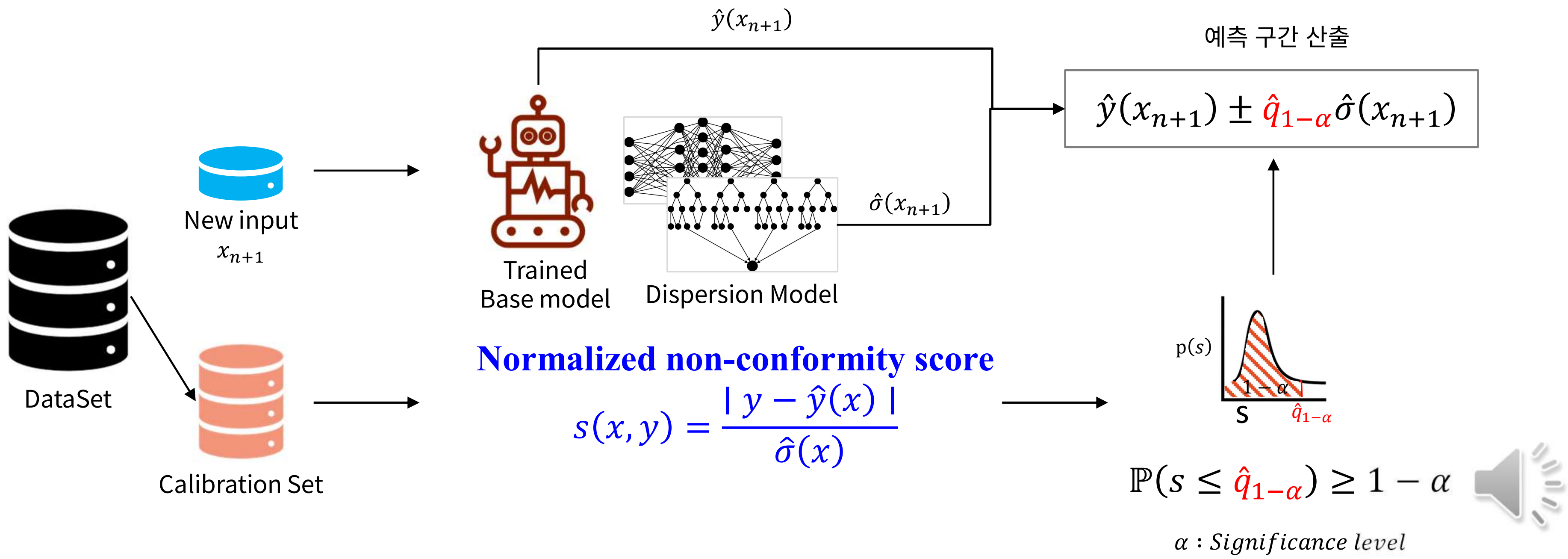
Papadopoulos, Harris, and Haris Haralambous. "Reliable prediction intervals with regression neural networks." Neural Networks 24.8 (2011): 842-851.

Introduction

Conformal Prediction

❖ 예측 영역을 어떻게 만드는가? (기본 아이디어 : Locally Adaptive CP)

- 보정 데이터(Calibration Set)의 비동조성 점수(Non-conformity score)를 정렬해 임계값 $\hat{q}$ 하나를 뽑는다
- 새로운 입력에 대해 비동조성 점수가 이 임계값 $\hat{q}$ 이하가 되는 정답의 범위를 역으로 계산해 예측 구간/영역을 만든다



Papadopoulos, Harris, and Haris Haralambous. "Reliable prediction intervals with regression neural networks." Neural Networks 24.8 (2011): 842-851.

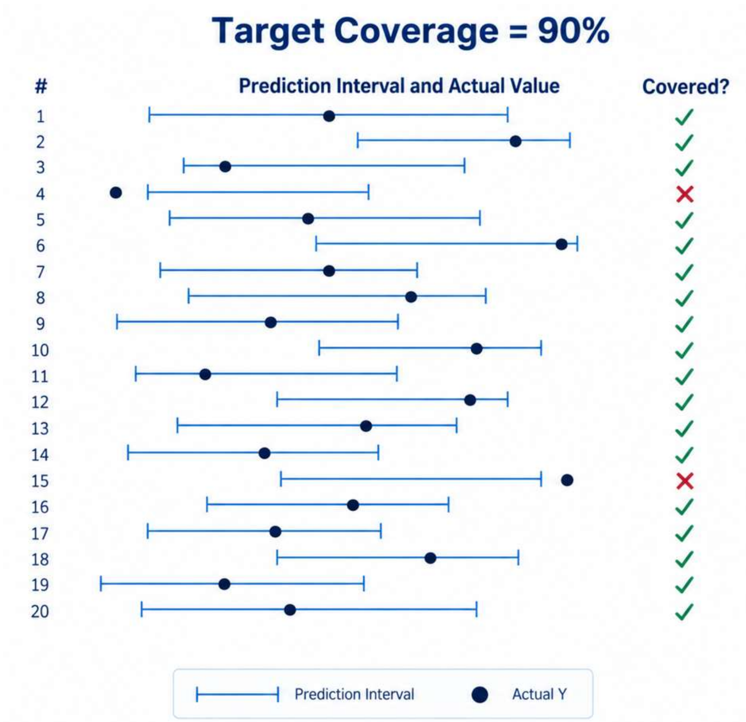
Introduction

Conformal Prediction

❖ 어떤 예측영역이 좋은 예측영역인가? - Validity

- Validity: 정답이 예측영역안에 포함(cover)될 확률을 목표한 확률 이상으로 실제로 달성하는 성질
- Prediction Set의 신뢰수준을 정량적으로 보장하는 것이 CP의 핵심

도착 예정 소요 시간 : ?



18/20 Covered = 90%

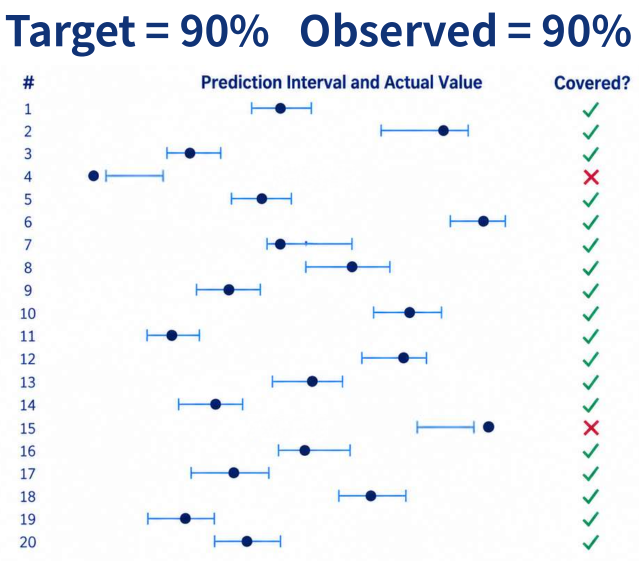
Target = 90% Observed = 90%

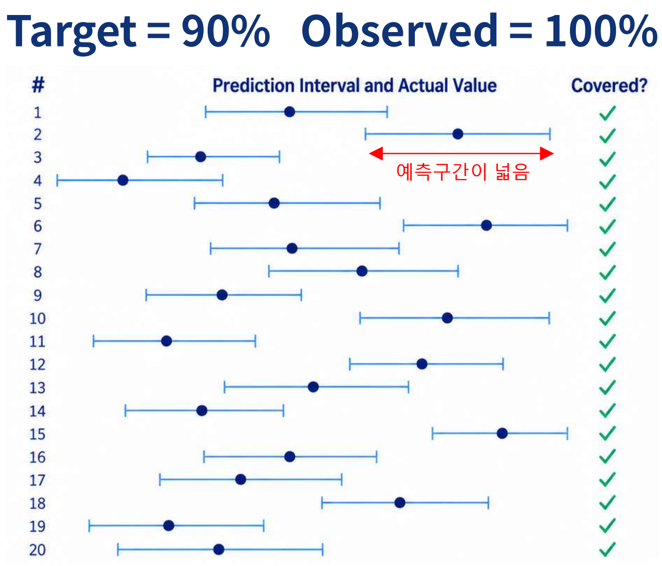
Introduction

Conformal Prediction

❖ 어떤 예측영역이 좋은 예측영역인가? - Efficiency

- 항상 (−∞, ∞)를 내놓으면 coverage는 100%임.
- Validity만으로는 무의미 — 무한 구간은 언제나 유효하지만 쓸모가 없음
- Efficiency: Validity를 유지하면서 예측영역이 작은 정도





Introduction

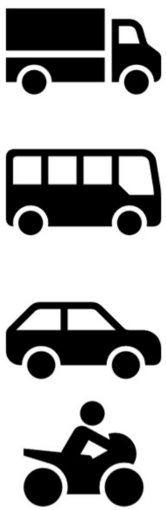
Multi-Output Conformal Prediction

❖ 만약 출력변수가 다변량이라면?

- 출력마다 90%를 맞춰도, 전부 맞을 확률은 90%가 아니다
- 이를 억지로 보장하려 들면 예측 영역이 비대해져 무의미해진다

각 출력간 의존관계를 어떻게 정의해야 할까?

X : 구간 별 교통상황



도착 예정 소요 시간 : ?

10 days of New Data

Coverage(Truck): 90%

Coverage(Bus): 90%

Coverage(Car): 90%

Coverage(Motorcycle): 90%

Joint Coverage = 90%?

Joint Coverage = 0.9^4=65.6%?

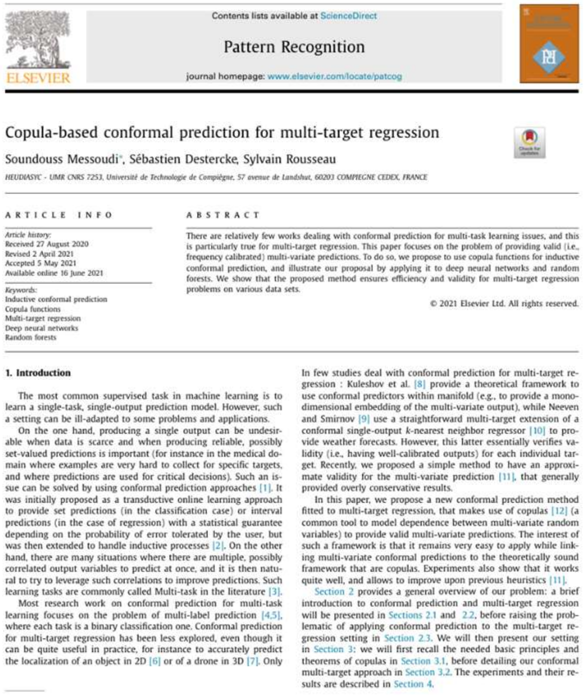
PAPER 1 : Copula 기반 Conformal Prediction

Copula-based Conformal Prediction for Multi-Target Regression

❖ 출력간 의존관계를 반영해 더 효율적인 예측영역을 만들자

- **Problem**
  - 출력별 CP만으로는 joint validity를 직접 확보할 수 없음
  - 의존관계를 무시한 calibration은 과도하게 보수적일 수 있음
- **Approach**
  - 각 출력별 Score의 의존관계를 copula로 추정
  - 목표 Joint coverage로부터 output-wise coverage level 결정
- **Contribution**
  - 출력의존관계를 반영한 joint calibration
  - Validity와 prediction-region efficiency의 동시 개선

2021 Pattern Recognition



Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/patcog

Copula-based conformal prediction for multi-target regression

Soundouss Messoudi*, Sébastien Destercke, Sylvain Rousseau

HEUDIASYC - UMR CNRS 7253, Université de Technologie de Compiègne, 57 avenue de Landshut, 60203 COMPIEGNE CEDEX, FRANCE

ARTICLE INFO

Article history:
Received 27 August 2020
Revised 2 April 2021
Accepted 5 May 2021
Available online 16 June 2021

Keywords:
Inductive conformal prediction
Copula functions
Multi-target regression
Deep neural networks
Random forests

ABSTRACT

There are relatively few works dealing with conformal prediction for multi-task learning issues, and this is particularly true for multi-target regression. This paper focuses on the problem of providing valid (i.e., frequency calibrated) multi-variate predictions. To do so, we propose to use copula functions for inductive conformal prediction, and illustrate our proposal by applying it to deep neural networks and random forests. We show that the proposed method ensures efficiency and validity for multi-target regression problems on various data sets.

1. Introduction

The most common supervised task in machine learning is to learn a single-task, single-output prediction model. However, such a setting can be ill-adapted to some problems and applications.

On the one hand, producing a single output can be undesirable when data is scarce and when producing reliable, possibly set-valued predictions is important (for instance in the medical domain where examples are very hard to collect for specific targets, and where predictions are used for critical decisions). Such an issue can be solved by using conformal prediction approaches [1]. It was initially proposed as a transductive online learning approach to provide set predictions (in the classification case) or interval predictions (in the case of regression) with a statistical guarantee depending on the probability of error tolerated by the user, but was then extended to handle inductive processes [2]. On the other hand, there are many situations where there are multiple, possibly correlated output variables to predict at once, and it is then natural to try to leverage such correlations to improve predictions. Such learning tasks are commonly called Multi-task in the literature [3].

Most research work on conformal prediction for multi-task learning focuses on the problem of multi-label prediction [4,5], where each task is a binary classification one. Conformal prediction for multi-target regression has been less explored, even though it can be quite useful in practice, for instance to accurately predict the localization of an object in 2D [6] or of a drone in 3D [7]. Only In few studies deal with conformal prediction for multi-target regression : Kuleshov et al. [8] provide a theoretical framework to use conformal predictors within manifold (e.g., to provide a mono-dimensional embedding of the multi-variate output), while Neeven and Smirnov [9] use a straightforward multi-target extension of a conformal single-output k-nearest neighbor regressor [10] to provide weather forecasts. However, this latter essentially verifies validity (i.e., having well-calibrated outputs) for each individual target. Recently, we proposed a simple method to have an approximate validity for the multi-variate prediction [11], that generally provided overly conservative results.

In this paper, we propose a new conformal prediction method fitted to multi-target regression, that makes use of copulas [12] (a common tool to model dependence between multi-variate random variables) to provide valid multi-variate predictions. The interest of such a framework is that it remains very easy to apply while linking multi-variate conformal predictions to the theoretically sound framework that are copulas. Experiments also show that it works quite well, and allows to improve upon previous heuristics [11].

Section 2 provides a general overview of our problem: a brief introduction to conformal prediction and multi-target regression will be presented in Sections 2.1 and 2.2, before raising the problematic of applying conformal prediction to the multi-target regression setting in Section 2.3. We will then present our setting in Section 3: we will first recall the needed basic principles and theorems of copulas in Section 3.1, before detailing our conformal multi-target approach in Section 3.2. The experiments and their results are described in Section 4.

Messoudi, Soundouss, Sébastien Destercke, and Sylvain Rousseau. "Copula-based conformal prediction for multi-target regression." Pattern Recognition 120 (2021): 108101.

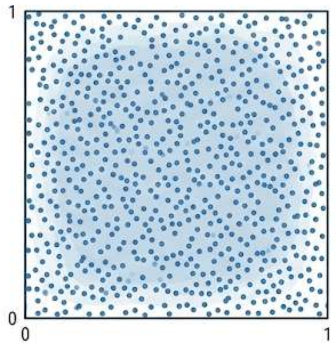
PAPER 1 : Copula 기반 Conformal Prediction

Copula-based Conformal Prediction for Multi-Target Regression

❖ Copula란?

- 문제의식: 결합분포를 통째로 다루면 복잡하다 → "각 변수의 개별 분포"와 "변수 간 순수 의존 구조"로 분리
- Sklar 정리: $F\left(s^{(1)}, \ldots, s^{(m)}\right) = C\left(F_1\left(s^{(1)}\right), \cdots, F_{\mathrm{m}}\left(s^{(m)}\right)\right) = C\left(u^{(1)}, \ldots, u^{(m)}\right)$
  - $F_1\left(s^{(1)}\right), \cdots, F_{\mathrm{m}}\left(s^{(m)}\right)$: 각 출력 변수 별 non-conformity score의 주변누적분포(marginal cdf)(=백분위수 u)
  - C : copula함수, 정의역은 $[0,1]^{\mathrm{m}}$
- 각 변수의 개별 백분위수 값들을 변수 간 의존 관계에 따라 합성하여 전체 결합누적확률을 산출하는 함수

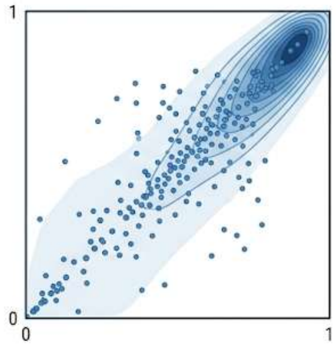
변수 간 완전 독립 가정



Independence Copula

$$C_{\Pi}\left(u^{(1)}, \ldots, u^{(m)}\right) = \prod_{j=1}^{m} u^{(j)}$$

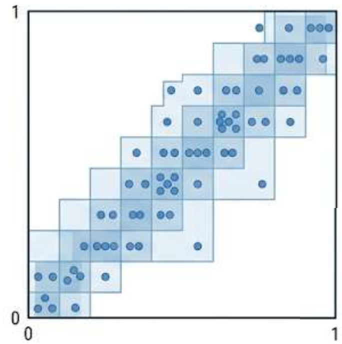
양의 상관관계 및 상위 극단값 의존 가정



Gumbel Copula

$$C_G^{\theta}\left(u^{(1)}, \ldots, u^{(m)}\right) = \exp\left(-\left[\sum_{j=1}^{m}\left(-\ln u^{(j)}\right)^{\theta}\right]^{1/\theta}\right)$$

사전 가정 없음 (비모수적 데이터 반영)



Empirical Copula

$$C_E\left(u^{(1)}, \ldots, u^{(m)}\right) = \frac{1}{n_{\text{cal}}} \sum_{i=1}^{n_{\text{cal}}} \mathbb{I}\left(u_i^{(1)} \le u^{(1)}, \ \ldots, \ u_i^{(m)} \le u^{(m)}\right)$$

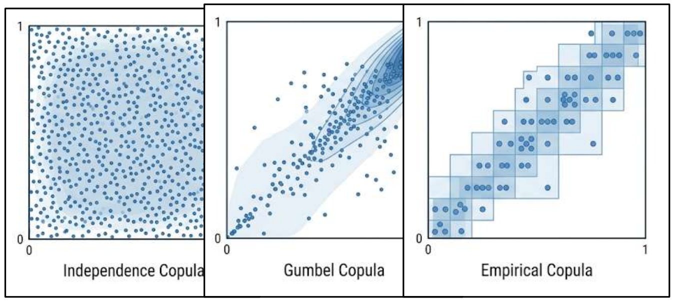
Data Mining Quality Analytics

PAPER 1 : Copula 기반 Conformal Prediction

Copula-based Conformal Prediction for Multi-Target Regression

❖ Copula를 통해 어떻게 Joint coverage를 보장하는가?

- Copula를 통해 “모든 출력차원이 들어갈 확률(joint coverage)”을 만족하기 위한 **각 output별 coverage**를 구할 수 있다.
- Joint coverage 제약을 각 출력별 calibration level로 분해



$$C\left(1-\alpha^{(1)}, \ldots, 1-\alpha^{(m)}\right) = 1-\alpha_{\text{joint}}$$

Joint Coverage를 만족하기 위해

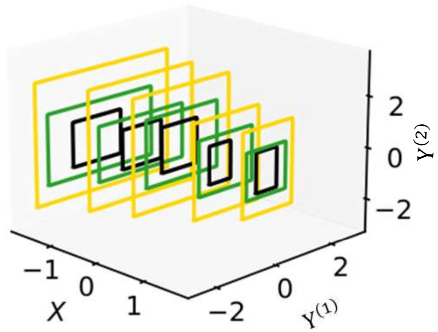
각 output에서 몇 % coverage를 요구할 것인가?

각 output $\hat{y}^{(j)}$별
Conformal Prediction

요구 coverage : $1-\alpha^{(\mathrm{i})}$

Prediction region

$$\hat{y}^{(j)}(x_{n+1}) \pm \hat{q}^{(j)}\hat{\sigma}^{(j)}(x_{n+1}) \quad (j = 1, \ldots, m)$$



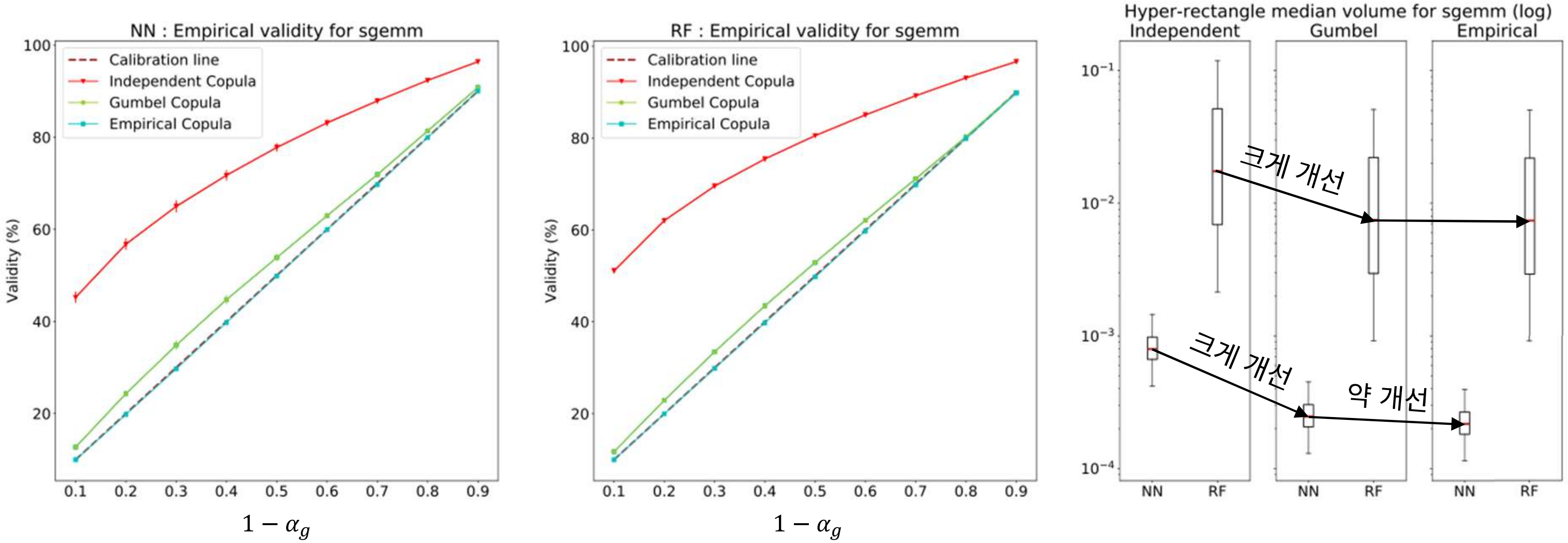
Hyperrectangle형태의 예측영역

PAPER 1 : Copula 기반 Conformal Prediction

Copula-based Conformal Prediction for Multi-Target Regression

❖ 실험 결과

- Validity: Independent의 과도한 보수성(Over-coverage)을 해소하고, Empirical Copula가 목표 Coverage에 가장 정밀하게 수렴
- Efficiency: 변수 간 의존관계를 반영(Gumbel, Empirical)함으로써 과도하게 보수적인 output-wise coverage를 완화하고 부피를 대폭 축소



(a) Validity (NN) (b) Validity (RF) (c) Efficiency (both)

Messoudi, Soundouss, Sébastien Destercke, and Sylvain Rousseau. "Copula-based conformal prediction for multi-target regression." Pattern Recognition 120 (2021): 108101.

PAPER 2 : ST-DQR

Calibrated Multiple-Output Quantile Regression with Representation Learning

❖ 예측영역의 형태를 고정하지 말고, 데이터가 정하게 하자

- **Problem**
  - 기존 prediction region은 제한된 geometry에 의존 (예 : hyperrectangle)
  - 실제 $Y \mid X$의 구조에 따라 불필요한 영역이 포함될 수 있음
- **Approach**
  - Conditional VAE 기반 representation learning
  - Latent space에서 directional quantile regression
  - Nonlinear decoding으로 flexible region 복원
  - Calibration으로 finite-sample marginal coverage 확보
- **Contribution**
  - Flexible Prediction Region
  - 동일 target coverage에서 efficiency 개선

2023 JMLR

Journal of Machine Learning Research 24 (2023) 1-48 Submitted 10/21; Revised 9/22; Published 1/23

Calibrated Multiple-Output Quantile Regression with Representation Learning

Shai Feldman SHAI.FELDMAN@CS.TECHNION.AC.IL
Department of Computer Science
Technion—Israel Institute of Technology
Technion City, Haifa 32000, Israel

Stephen Bates STEPHENBATES@CS.BERKELEY.EDU
Departments of Electrical Engineering and Computer Science and of Statistics
University of California, Berkeley
Berkeley, CA 94720, USA

Yaniv Romano YROMANO@TECHNION.AC.IL
Departments of Electrical and Computer Engineering and of Computer Science
Technion—Israel Institute of Technology
Technion City, Haifa 32000, Israel

Editor: Samuel Kaski

Abstract

We develop a method to generate predictive regions that cover a multivariate response variable with a user-specified probability. Our work is composed of two components. First, we use a deep generative model to learn a representation of the response that has a unimodal distribution. Existing multiple-output quantile regression approaches are effective in such cases, so we apply them on the learned representation, and then transform the solution to the original space of the response. This process results in a flexible and informative region that can have an arbitrary shape, a property that existing methods lack. Second, we propose an extension of conformal prediction to the multivariate response setting that modifies any method to return sets with a pre-specified coverage level. The desired coverage is theoretically guaranteed in the finite-sample case for any distribution. Experiments conducted on both real and synthetic data show that our method constructs regions that are significantly smaller compared to existing techniques.

Keywords: conformal prediction, uncertainty quantification, quantile regression, multiple regression, variational auto-encoder

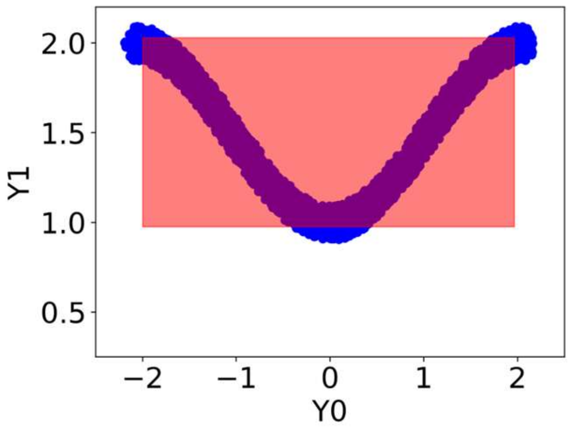
Feldman, Shai, Stephen Bates, and Yaniv Romano. "Calibrated multiple-output quantile regression with representation learning." Journal of Machine Learning Research 24.24 (2023): 1-48.

PAPER 2 : ST-DQR

Calibrated Multiple-Output Quantile Regression with Representation Learning

❖ 고정된 Prediction region의 형태를 사전에 제한하면 어떤 비효율이 발생하는가?

- 많은 multivariate UQ 방법은 prediction region을 사전에 정한 geometry 안에서 구성
- 고정된 형태 제약이 low-density 영역까지 포함하여 불필요하게 큰 Region Volume을 형성



실제 Y 분포에 맞는 형태로 예측 영역을 유연하게 만들 수 없을까?

Data Mining Quality Analytics

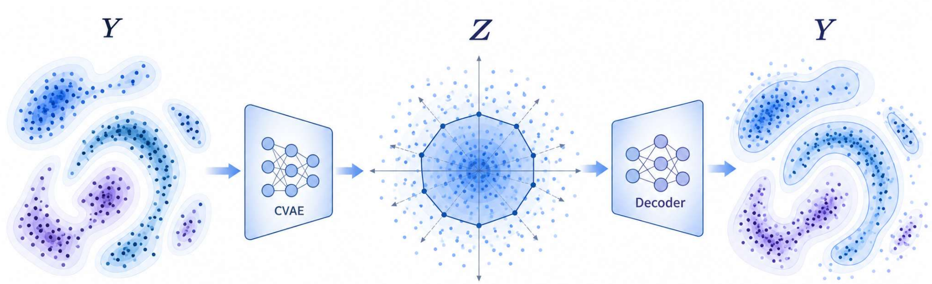
PAPER 2 : ST-DQR

Calibrated Multiple-Output Quantile Regression with Representation Learning

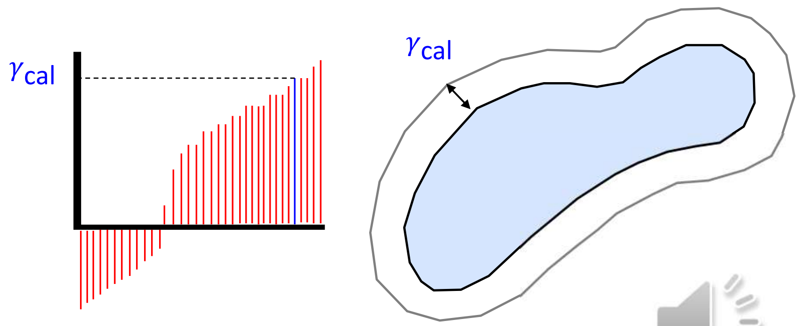
❖ Y 분포에 맞는 예측 영역 형태를 어떻게 만들 수 있을까?

- Stage 1 (Representation Learning): CVAE 기반의 비선형 디코딩으로 복잡한 geometry를 반영하여 효율적인 Region을 생성
- Stage 2 (Conformal Calibration): Calibration 데이터를 바탕으로 Region 크기를 보정하여 finite-sample coverage를 보증

Stage 1 : Representation Learning



Stage 2 : Conformal Calibration



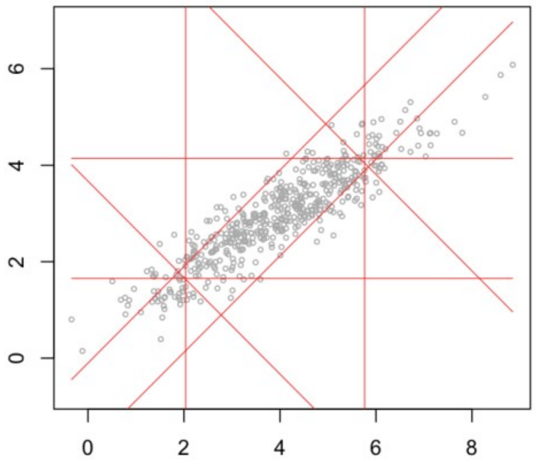
PAPER 2 : ST-DQR

Calibrated Multiple-Output Quantile Regression with Representation Learning

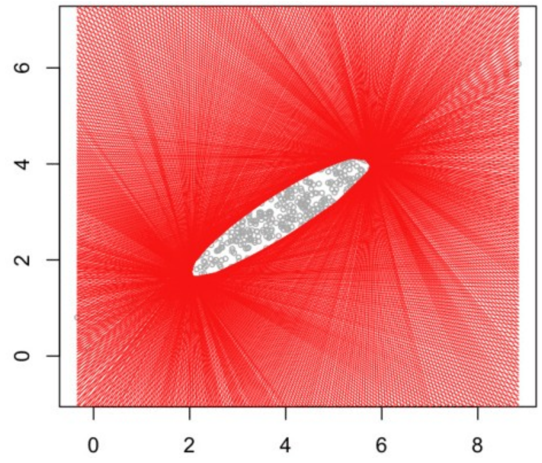
❖ DQR(Directional Quantile Regression)이란?

- 다차원 반응 변수(Y) 공간에서 여러 방향 벡터(u)로 데이터를 사영(projection)한 뒤, 각 방향마다 단변량 분위수 회귀를 풀어 반공간(half-space)의 경계면을 추정
- 모든 방향의 반공간을 교집합하여 영역을 깎아내므로 **반드시 볼록(Convex)한 형태만 가능**하고, 방향이 늘어날 수록 내부 포함 확률이 감소해 목표치보다 **Coverage가 떨어지는(Under-coverage)** 구조적 한계를 지님

M=8방향 (α=0.1, 목표 90%)



M=512방향 (α=0.1, 목표 90%)



Data Mining Quality Analytics

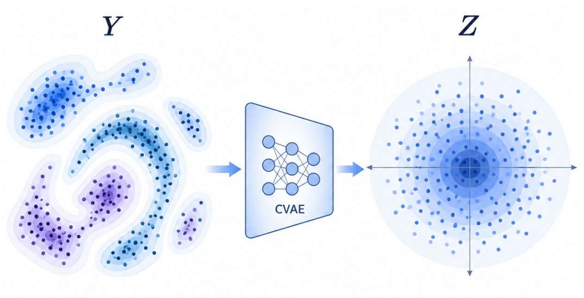
Feldman, Shai, Stephen Bates, and Yaniv Romano. "Calibrated multiple-output quantile regression with representation learning." Journal of Machine Learning Research 24.24 (2023): 1-48.

PAPER 2 : ST-DQR

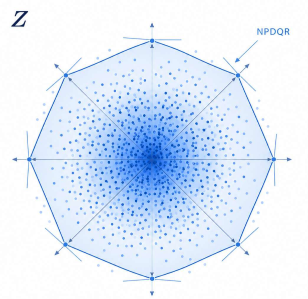
Calibrated Multiple-Output Quantile Regression with Representation Learning

❖ 1단계: ST-DQR (Spherically Transformed DQR)

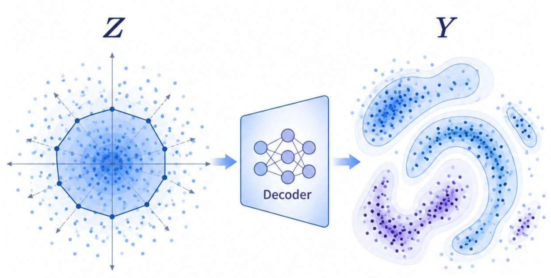
- Step 1 (Encoding): CVAE를 통해 복잡한 출력 공간(Y)을 다루기 쉬운 구형(Spherical) 정규분포 잠재 공간(Z)으로 매핑
- Step 2 (DQR): 등고선이 볼록한 잠재 공간에서 방향 분위수 회귀(NPDQR)를 적용하여 안정적인 **Convex Region 구축**
- Step 3 (Decoding): 도출된 잠재 영역을 비선형 디코더로 역변환하여, **실제 데이터 분포 형태를 따르는 예측영역** 완성



Step 1 : CVAE mapping



Step 2 : DQR



Step 3 : Decoding

DQR의 볼록(convex)이라는 제약을 단순한 잠재공간으로 Mapping하여 해결

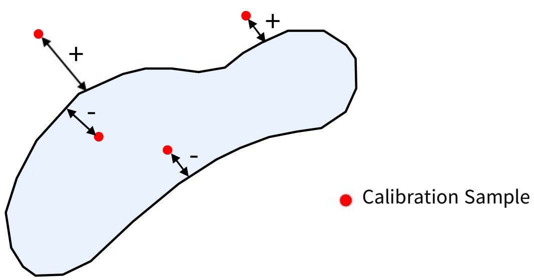
PAPER 2 : ST-DQR

Calibrated Multiple-Output Quantile Regression with Representation Learning

❖ 2단계: 다변량 적합 예측 보정

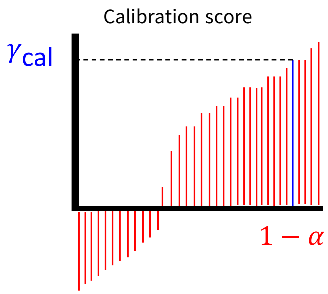
- DQR의 본질적 한계(무한 반공간의 교집합)로 인해 영역은 목표 Coverage에 미달하는 Under-coverage가 발생함
- 별도의 Calibration 데이터로 경계면과의 거리를 측정하고 영역을 보정(팽창/수축)하여 유한 표본하에 목표 Coverage($\geq 1-\alpha$)를 엄밀히 보증함

최소거리 기반 Non-conformity Score 정의



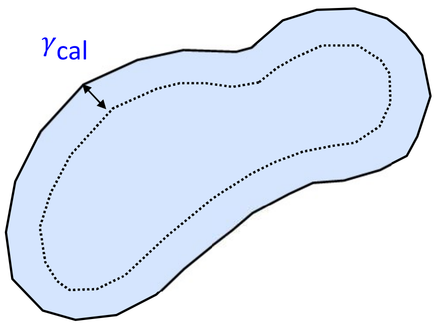
$$E_i = \begin{cases} \min_{a \in R(X_i)} \|Y_i - a\|_2 & \text{if } Y_i \notin R(X_i) \quad (\text{영역 밖: 양수}) \\ -\min_{b \in R(X_i)^c} \|Y_i - b\|_2 & \text{if } Y_i \in R(X_i) \quad (\text{영역 안: 음수}) \end{cases}$$

보정 임계값 산출



$$\gamma_{\text{cal}} = (1-\alpha)(1+1/n_2) \text{ 번째 분위수 값}$$

최종 예측 영역 생성



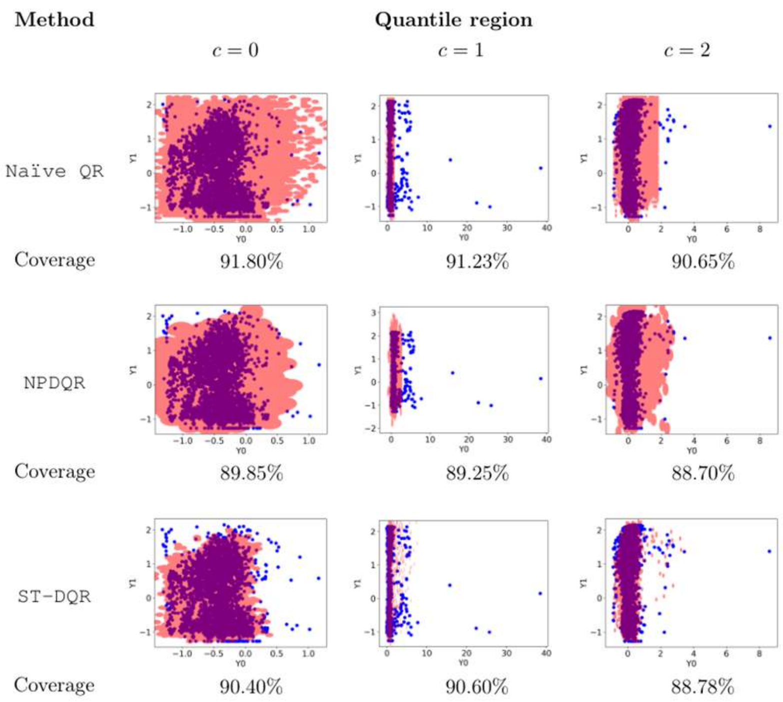
$$S^{\gamma_{\text{cal}}}(X_{n+1}) = \left\{ y \in \mathbb{R}^d : \min_{a \in R(X_{n+1})} \|y - a\|_2 \leq \gamma_{\text{cal}} \right\}$$

PAPER 2 : ST-DQR

Calibrated Multiple-Output Quantile Regression with Representation Learning

❖ 실험결과

- Validity: 보정 단계 덕분에 목표 coverage를 유지
- Efficiency: 같은 coverage에서 영역 크기가 감소



Naïve QR: 차원별 분위수 독립 추정

NPDQR: 신경망 기반 볼록 다면체 영역

ST-DQR: 제안 방법

Feldman, Shai, Stephen Bates, and Yaniv Romano. "Calibrated multiple-output quantile regression with representation learning." Journal of Machine Learning Research 24.24 (2023): 1-48.

PAPER 2 : ST-DQR

Calibrated Multiple-Output Quantile Regression with Representation Learning

❖ 실험결과

- Validity: 보정 단계 덕분에 목표 coverage를 유지
- Efficiency: 같은 coverage에서 영역 크기가 감소

	Coverage rate			Relative area of quantile regions		
Data Set name	ST-DQR	Naïve QR	NPDQR	ST-DQR	Naïve QR	NPDQR
bio	90.0	90.002	89.892	1	1.223	1.222
house	90.157	89.978	89.876	1	1.168	1.149
blog_data	90.145	90.064	90.016	1	1.551	1.821
meps_19	90.144	89.892	89.865	1	2.215	2.169
meps_20	89.997	90.0	89.913	1	2.27	2.209
meps_21	89.899	89.879	89.676	1	2.191	2.212

Naïve QR: 차원별 분위수 독립 추정

NPDQR: 신경망 기반 볼록 다면체 영역

ST-DQR: 제안 방법

Feldman, Shai, Stephen Bates, and Yaniv Romano. "Calibrated multiple-output quantile regression with representation learning." Journal of Machine Learning Research 24.24 (2023): 1-48.

PAPER 2 : ST-DQR

Calibrated Multiple-Output Quantile Regression with Representation Learning

❖ Prediction Region은 모든 입력(x)에서 동일한 Coverage를 보장하는가?

- Conformal Calibration은 샘플 전체 평균(Marginal)만 보장하므로, X에 따른 이분산성(Heteroscedasticity) 및 국소적 데이터 밀도 차이로 인해 국소 구간에서 Coverage 변동이 발생함

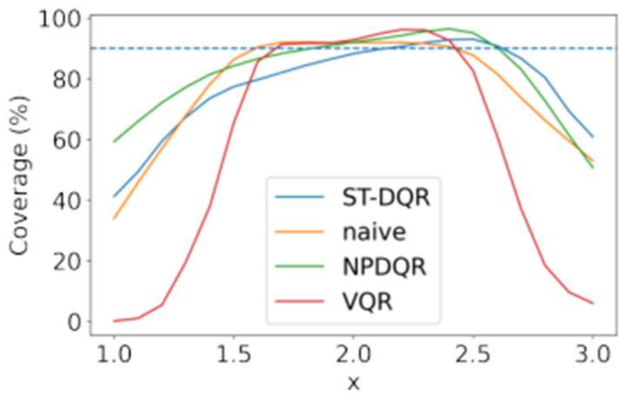


Figure 12: Conditional coverage achieved over the non-linear synthetic data with 10000 samples.

각 Input Condition에서 Coverage를 더 안정적으로 만들 수 없을까?

Feldman, Shai, Stephen Bates, and Yaniv Romano. "Calibrated multiple-output quantile regression with representation learning." Journal of Machine Learning Research 24.24 (2023): 1-48.

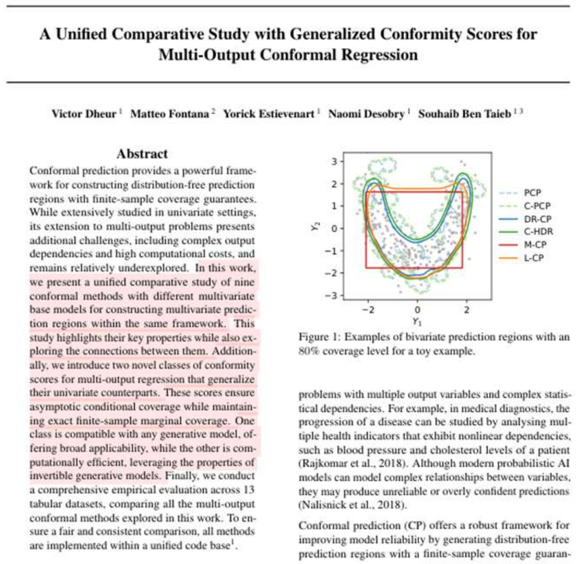
PAPER 3 : Generalized Conformity Scores

A Unified Comparative Study with Generalized Conformity Scores for Multi-Output Conformal Regression

❖ 입력마다 달라지는 Conformity Score의 분포를 정규화하자

- **Problem**
  - 기존 conformity score의 distribution이 X에 따라 달라질 수 있음
  - 하나의 global threshold는 input별로 서로 다른 coverage를 유발할 수 있음
- **Approach**
  - 두 종류의 generalized conformity score
    - CDF-based conformity scores (C-PCP 방법)
    - Latent-based conformity score (L-CP 방법)
- **Contribution**
  - Exact finite-sample marginal coverage 유지
  - **Asymptotic Conditional Coverage**
  - 9개 multi-output CP 방법의 unified comparison

2025 ICML



A Unified Comparative Study with Generalized Conformity Scores for Multi-Output Conformal Regression

Victor Dheur[1] Matteo Fontana[2] Yorick Estievenart[1] Naomi Desobry[1] Souhaib Ben Taieb[1 3]

Abstract

Conformal prediction provides a powerful framework for constructing distribution-free prediction regions with finite-sample coverage guarantees. While extensively studied in univariate settings, its extension to multi-output problems presents additional challenges, including complex output dependencies and high computational costs, and remains relatively underexplored. In this work, we present a unified comparative study of nine conformal methods with different multivariate base models for constructing multivariate prediction regions within the same framework. This study highlights their key properties while also exploring the connections between them. Additionally, we introduce two novel classes of conformity scores for multi-output regression that generalize their univariate counterparts. These scores ensure asymptotic conditional coverage while maintaining exact finite-sample marginal coverage. One class is compatible with any generative model, offering broad applicability, while the other is computationally efficient, leveraging the properties of invertible generative models. Finally, we conduct a comprehensive empirical evaluation across 13 tabular datasets, comparing all the multi-output conformal methods explored in this work. To ensure a fair and consistent comparison, all methods are implemented within a unified code base[1].

Figure 1: Examples of bivariate prediction regions with an 80% coverage level for a toy example.

problems with multiple output variables and complex statistical dependencies. For example, in medical diagnostics, the progression of a disease can be studied by analysing multiple health indicators that exhibit nonlinear dependencies, such as blood pressure and cholesterol levels of a patient (Rajkomar et al., 2018). Although modern probabilistic AI models can model complex relationships between variables, they may produce unreliable or overly confident predictions (Nalisnick et al., 2018).

Conformal prediction (CP) offers a robust framework for improving model reliability by generating distribution-free prediction regions with a finite-sample coverage guaran-

Dheur, V., et al(2025). "A Unified Comparative Study with Generalized Conformity Scores for Multi-Output Conformal Regression." Proceedings of the 42nd International Conference on Machine Learning, in Proceedings of Machine Learning Research267:13444-13485

PAPER 3 : Generalized Conformity Scores

A Unified Comparative Study with Generalized Conformity Scores for Multi-Output Conformal Regression

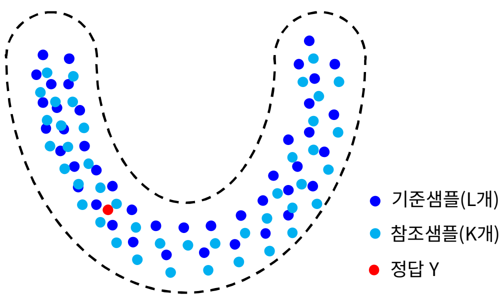
❖ Score scale을 입력(X)에 관계없도록 어떻게 정규화 할 수 있을까?

1) CDF Based Conformity Score (C-PCP)

- 입력 X=x조건부 누적분포함수(CDF)를 통해 균등분포($\mathcal{U}(0,1)$)로 정규화함으로써, 점수 분포의 x의존성을 제거

$$\tilde{Y}^{(1)}, \ldots, \tilde{Y}^{(L)} \overset{i.i.d.}{\sim} \hat{F}_{Y|X=x}$$

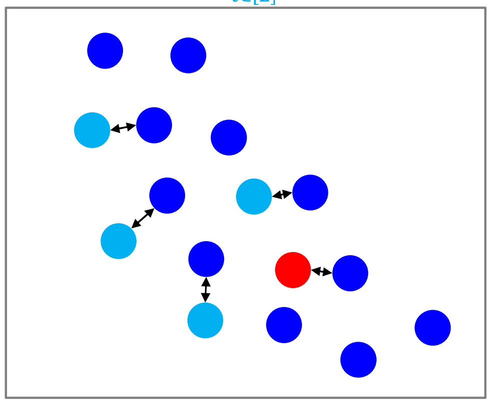
$$\hat{Y}^{(1)}, \ldots, \hat{Y}^{(K)} \overset{i.i.d.}{\sim} \hat{F}_{Y|X=x}$$



조건부 생성 모델 통한 예측 분포$\hat{F}_{Y|X=x}$에서 Montecarlo 샘플링 L+K개

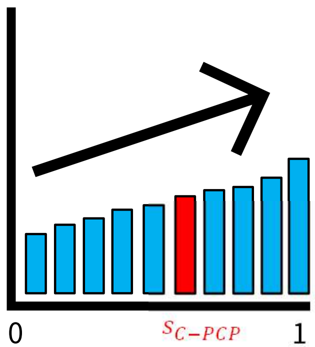
$$d(x, y) = \min_{l \in [L]} \| y - \tilde{Y}^{(l)} \|_2$$

$$d\left(x, \hat{Y}^{(k)}\right) = \min_{l \in [L]} \| \hat{Y}^{(k)} - \tilde{Y}^{(l)} \|_2$$



Calibration set 각 샘플 x별로 **참조샘플 $\hat{Y}^{(k)}$**과 **정답 Y**의 **기준샘플 $\tilde{Y}^{(l)}$**에 대한 **최단거리d**를 구함

$$s_{C-PCP}(x, y) = \frac{1}{K} \sum_{k=1}^{K} \mathbb{I}\left(d\left(x, \hat{Y}^{(k)}\right) \leq d(x, y)\right)$$



정답 Y의 최단거리d 의 Empirical CDF(=백분위수)를 구함 **(Score의 X 의존성 제거)**

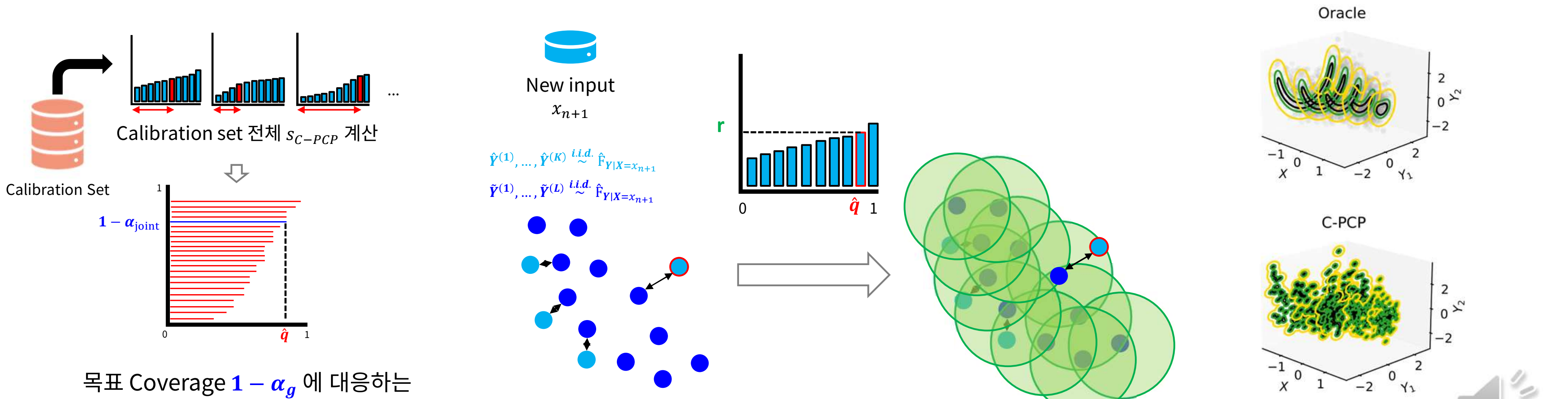
PAPER 3 : Generalized Conformity Scores

A Unified Comparative Study with Generalized Conformity Scores for Multi-Output Conformal Regression

❖ Score scale을 입력(X)에 관계없도록 어떻게 정규화 할 수 있을까?

1) CDF Based Conformity Score (C-PCP 방법)

- 새로운 입력(x)마다 거리d의 조건부 분위수를 통해 구의 반지름(r)을 동적으로 적응시킴으로써, 표본 샘플링만으로 조건부 불확실성을 반영한 다변량 예측 영역을 생성



목표 Coverage $\mathbf{1-\alpha_g}$ 에 대응하는 s_{C-PC} Score 임계치 $\hat{q}$ 추출

임계치 $\hat{q}$에 해당하는 **참조샘플**의 최단거리를 **반지름(r)**으로 하여 **기준샘플 L개** 주위에 **구를 씌워 합집합 예측영역**을 만듦

$s_{\text{C-PCP}} = F_{D|X}(d \mid X) \mid X \sim \text{Uniform}(0,1)$

점근적 가정($|\mathcal{D}_{cal}| \to \infty, K \to \infty$)하에 조건부 Coverage를 보장

PAPER 3 : Generalized Conformity Scores

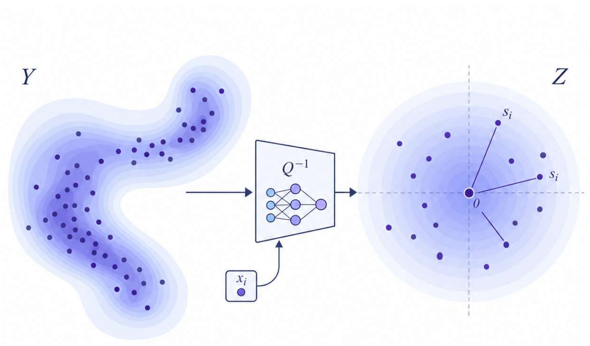
A Unified Comparative Study with Generalized Conformity Scores for Multi-Output Conformal Regression

❖ Score scale을 입력(X)에 관계없도록 어떻게 정규화 할 수 있을까?

2) Latent Conformity Scores (L-CP)

- 가역 생성 모델(Q^{-1})을 통해 복잡한 출력 공간Y를 표준정규 잠재공간Z으로 매핑하여 단순한 원점 거리$\|z\|$)로 임계 초구($\hat{q}$)를 구한 뒤, 이를 다시 출력 공간Y로 복원
- 입력 X에 따른 모든 비선형성과 이분산성을 가역 함수가 흡수하여 잠재변수 Z를 X와 독립인 표준정규분포로 정규화하므로, 모든 x 지점마다 점근적 가정하에 1-α의 Conditional Coverage를 유지

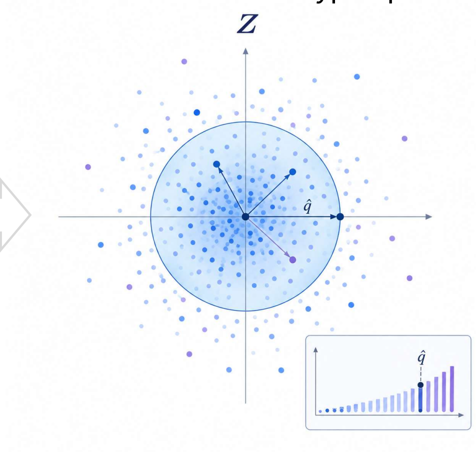
역매핑 및 점수화



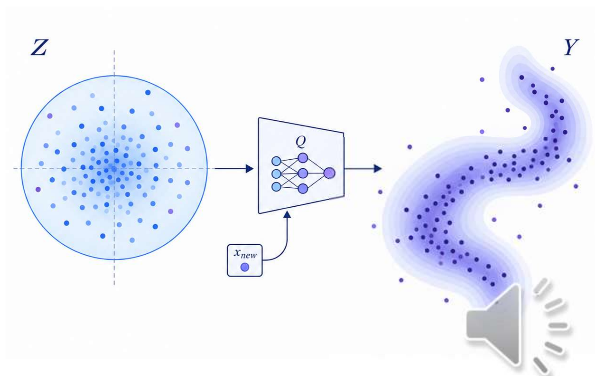
$$S_{\text{L-CP}} = \|Q^{-1}(Y;X)\|_2 \mid X = x \sim \chi_d \quad (\forall x)$$

점근적 가정($\hat{Q} \to Q^*$, $|\mathcal{D}_{\text{cal}}| \to \infty$)하에
조건부 Coverage를 보장

Calibration in Latent hypersphere



출력 공간 복원

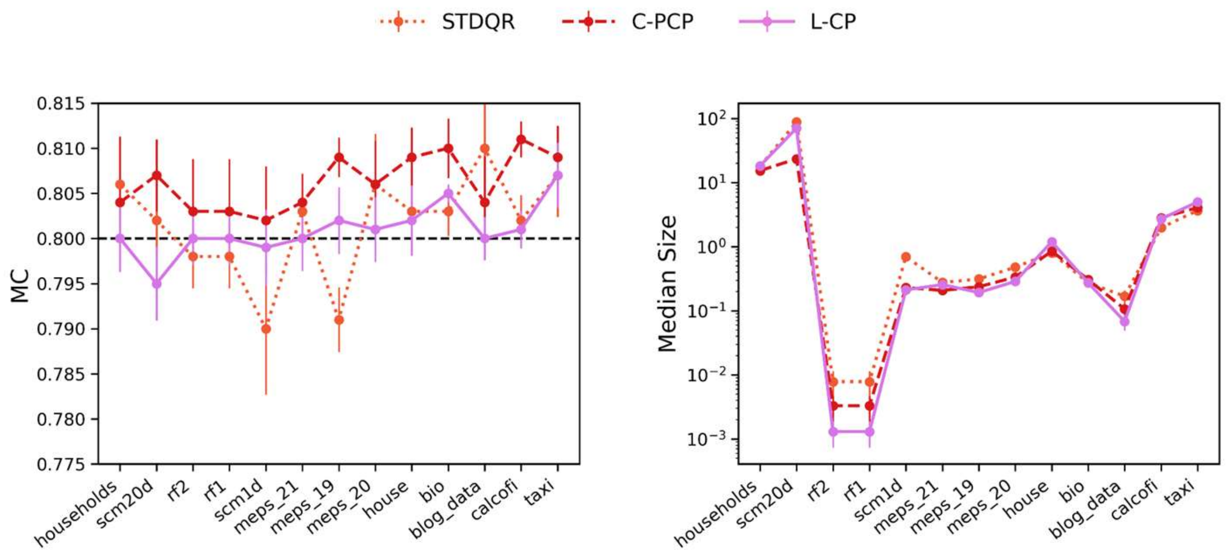


PAPER 3 : Generalized Conformity Scores

A Unified Comparative Study with Generalized Conformity Scores for Multi-Output Conformal Regression

❖ 실험결과 : ST-DQR방법론 대비 Validity와 Efficiency를 유지하였는가?

- Maginal Coverage(Validity): L-CP와 C-PCP는 전 데이터셋에서 목표를 충족하며, ST-DQR은 일부 데이터셋에서 약간 미달
- Median Size(Efficiency) : 거의 모든 DataSet에서 L-CP, C-PCP가 우수.



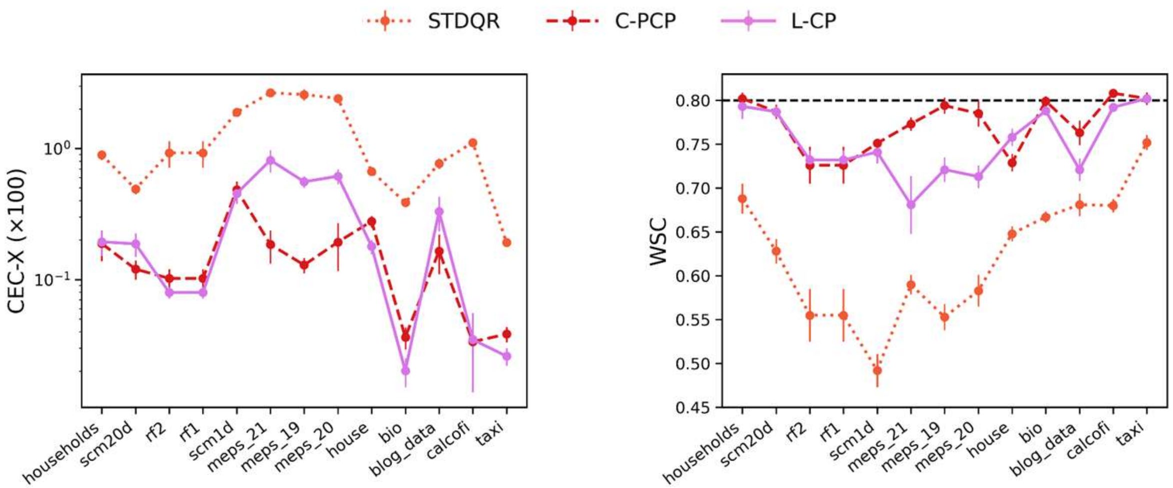
Dheur, V., et al(2025). “A Unified Comparative Study with Generalized Conformity Scores for Multi-Output Conformal Regression.” Proceedings of the 42nd International Conference on Machine Learning, in Proceedings of Machine Learning Research267:13444-13485

PAPER 3 : Generalized Conformity Scores

A Unified Comparative Study with Generalized Conformity Scores for Multi-Output Conformal Regression

❖ 실험결과 : ST-DQR 방법론 대비 조건부 Coverage를 개선하였는가?

- 국소 조건부 오차(CEC)를 기존 대비 최대 10~100배 대폭 낮추고 Worst Slab Coverage(WSC)를 65~80% 선으로 방어하여, 우수한 Conditional Validity를 입증함.



- CEC-X: 입력 X구간별 Coverage의 목표치 대비 평균 오차 (낮을수록 우수).
- WSC: 가장 성능이 붕괴된 최악 구간의 Coverage 수치 (높을수록 우수).

Dheur, V., et al(2025). “A Unified Comparative Study with Generalized Conformity Scores for Multi-Output Conformal Regression.” Proceedings of the 42nd International Conference on Machine Learning, in Proceedings of Machine Learning Research267:13444-13485

Conclusion

Conclusions

Conclusion: Toward Prediction Regions Aligned with Uncertainty Structure

- PAPER1 : Messoudi 2021
  - 문제정의 : Output-wise CP로 joint calibration을 효율적으로 확보하기 어려움
  - 방법 : Copula-based calibration
  - 결과 : Dependence-aware joint calibration & efficiency
  - 한계 : Fixed-form region geometry
- PAPER2 : Feldman 2023
  - 문제정의 : 제한된 region geometry가 불필요한 volume을 포함
  - 방법 : Representation learning + ST-DQR
  - 결과 : Flexible geometry & statistical efficiency
  - 한계 : Conditional coverage 미보장
- PAPER3 : Dheur 2025
  - 문제정의 : Conformity score의 조건부분포가 입력 x에 따라 달라진다.
  - 방법 : CDF-based / latent-based generalized conformity scores
  - 결과 : asymptotic conditional validity

Multi-output CP는 Efficiency와 Conditional Validity를 개선하기 위해,
예측영역을 실제 조건부 결합 불확실성 구조에 맞게 바꿔옴

THANK YOU