

Prédiction conformelle basée sur les copules pour la régression multi-cibles

Soundouss Messoudi*, Sébastien Destercke*, Sylvain Rousseau*

*HEUDIASYC - UMR CNRS 7253, Université de Technologie de Compiègne
57 avenue de Landshut, 60203 COMPIEGNE CEDEX - FRANCE
<https://www.hds.utc.fr/>, prenom.nom@hds.utc.fr

Résumé. Peu de travaux traitent de la prédiction conformelle pour la régression multi-cibles. Cet article se concentre sur l'obtention de prédictions multivariées valides (c'est-à-dire calibrées en fréquence). Pour ce faire, nous proposons d'utiliser des copules pour la prédiction conformelle inductive, et nous montrons empiriquement que la méthode proposée garantit l'efficacité et la validité pour les problèmes de régression multi-cibles sur divers ensembles de données.

1 Introduction

En apprentissage automatique supervisé, il existe de nombreuses situations où il y a de multiples cibles, éventuellement corrélées. De plus, certaines applications sont critiques, comme en médecine. Un niveau de confiance ajouté aux prédictions d'un tel problème avec la prédiction conformelle permet de fournir des prédictions sous forme d'ensemble (en classification) ou d'intervalle (en régression) avec une garantie statistique selon la probabilité d'erreur choisie. Cependant, cette méthode est principalement utilisée en apprentissage à cible unique.

Dans ce papier¹, nous proposons une méthode de prédiction conformelle adaptée à la régression multi-cibles basée sur les copules, fonctions modélisant la dépendance entre des variables multidimensionnelles, afin d'exploiter les corrélations entre les cibles. Les sections 2 et 3 fournissent une brève introduction à la prédiction conformelle et à la régression multi-cibles. Nous présenterons ensuite notre approche dans la section 4 avec un rappel des principes et théorèmes de base des copules. Les expériences et leurs résultats sont décrits dans la section 5.

2 Régression conformelle inductive

La régression conformelle inductive [Papadopoulos et al. (2002)] fournit une prédiction d'intervalle au lieu d'une prédiction ponctuelle avec une garantie statistique choisie.

Soit $\{z_1 = (x_1, y_1), z_2 = (x_2, y_2), \dots, z_n = (x_n, y_n)\}$ un sac contenant les paires successives d'un objet $x_i \in X$ et de son étiquette à valeur réelle $y_i \in \mathbb{R}$. En supposant que ces données sont échangeables (une condition plus faible que l'i.i.d.), il est possible de prédire $y_{n+1} \in \mathbb{R}$ pour tout nouvel objet $x_{n+1} \in X$ en suivant le cadre conformel inductif.

1. Article original publié au journal Pattern Recognition [Messoudi et al. (2021)]

La première étape consiste à diviser les données d'entraînement original Z en un *ensemble d'apprentissage* $Z^{tr} = \{z_1, \dots, z_l\}$ et un *ensemble de calibration* $Z^{cal} = \{z_{l+1}, \dots, z_n\}$, avec $|Z^{cal}| = n - l$. Ensuite, un algorithme d'apprentissage automatique appelé *algorithme sous-jacent* est entraîné sur Z^{tr} pour obtenir une *mesure de non-conformité* pour évaluer l'étrangeté d'un exemple z_k en calculant le score de non-conformité α_k . Une mesure de non-conformité normalisée est donnée par $\alpha_i = |y_i - \hat{y}_i|/\sigma_i$ où σ_i est une valeur propre à chaque exemple x_i permettant de mesurer la difficulté de prédire sa cible y_i . Par exemple, σ_i peut être estimée avec $\sigma_i = \exp(\mu_i) + \beta$ où $\mu_i = \ln(|y_i - \hat{y}_i|)$ prédite par un perceptron multi-couches (MLP).

Pour un nouvel exemple x_{n+1} , l'algorithme sous-jacent associe à toute prédiction possible \hat{y} son score de non-conformité $\alpha_{n+1}^{\hat{y}}$ et calcule sa *p-value* avec :

$$p_v(\hat{y}_{n+1}) = \frac{|\{i = l + 1, \dots, n, n + 1 : \alpha_i \geq \alpha_{n+1}^{\hat{y}}\}|}{n - l + 1}. \quad (1)$$

La dernière étape consiste à choisir le *niveau de signifiante* $\epsilon \in (0, 1)$ pour obtenir une prédiction avec un *niveau de confiance* $1 - \epsilon$, étant la garantie statistique de couverture de la vraie valeur y_{n+1} par la prédiction d'intervalle $\hat{\mathbf{y}}_{n+1}$ telle que $\hat{\mathbf{y}}_{n+1} = \{\hat{y}_{n+1} \in \mathbb{R} : p_v(\hat{y}_{n+1}) > \epsilon\}$.

Le score de non-conformité normalisé est calculé avec la mesure de non-conformité standard (différence absolue entre la valeur réelle y_i et la valeur prédite \hat{y}_i) mise à l'échelle avec σ_i , un terme estimant la difficulté de prédire y_i . Il est calculé comme $\alpha_i = |y_i - \hat{y}_i|/\sigma_i$.

La séquence des scores de non-conformité $\alpha_{l+1}, \dots, \alpha_n$ pour tous les exemples dans Z^{cal} est obtenue et triée par ordre décroissant. Ensuite, l'indice du score de non-conformité α_s est calculé au fractile $(1 - \epsilon)$ en fonction du niveau de signifiante choisi ϵ , tel que :

$$\mathbb{P}\left(\frac{|y_i - \hat{y}_i|}{\sigma_i} \leq \alpha_s\right) \geq 1 - \epsilon, \quad (2)$$

qui devient une égalité si la méthode est parfaitement calibrée. Enfin, pour un nouvel exemple x_{n+1} , la prédiction d'intervalle couvrant la vraie cible y_{n+1} avec une probabilité $1 - \epsilon$ est :

$$\hat{\mathbf{y}}_{n+1} = [\hat{y}_{n+1} - \alpha_s \sigma_{n+1}, \hat{y}_{n+1} + \alpha_s \sigma_{n+1}]. \quad (3)$$

Avant d'introduire notre problème, notons qu'en supposant que notre méthode est bien calibrée et que $|y_i - \hat{y}_i|/\sigma_i$ est associé à une variable aléatoire Q , (2) peut être réécrite comme :

$$\mathbb{P}(Q \leq \alpha_s) = 1 - \epsilon := F_Q(\alpha_s), \quad (4)$$

ce qui sera utile pour traiter les copules et les sorties multi-cibles par la suite. Notons également que cela signifie que la spécification de ϵ définit de manière unique une valeur α_s .

3 Prédiction conformelle pour la régression multi-cibles

Dans la régression multi-cibles (RMC), l'espace X est le même que dans la régression standard, mais l'espace $Y \subset \mathbb{R}^m$ est constitué de m cibles à valeurs réelles. Cela signifie que les observations sont des paires (x_i, y_i) où chaque objet $x_i \in X$ est associé à un vecteur de dimension m à valeurs réelles $y_i = (y_i^1, \dots, y_i^m) \in Y$ (nous utiliserons les exposants pour

désigner les dimensions et les indices pour désigner les observations). Notre objectif est de fournir une méthode de RMC conformelle permettant d'exploiter les relations entre les cibles.

Soient $\hat{y}_{n+1}^j, \bar{y}_{n+1}^j$ respectivement les limites inférieure et supérieure des prédictions d'intervalle données par la mesure de non-conformité pour chaque cible Y^j étant donné une nouvelle instance x_{n+1} . Nous définissons l'hyper-rectangle $[\hat{y}_{n+1}]$ comme le produit cartésien suivant :

$$[\hat{y}_{n+1}] = \times_{j=1}^m [\hat{y}_{n+1}^j, \bar{y}_{n+1}^j]. \quad (5)$$

Cet hyper-rectangle forme le volume $\prod_{j=1}^m (\bar{y}_{n+1}^j - \hat{y}_{n+1}^j)$ auquel une prédiction globale y_{n+1} d'un nouvel exemple x_{n+1} doit appartenir pour être valide. L'objectif de la prédiction conformelle dans ce cadre est de satisfaire un niveau de signifiante global ϵ_g requis tel que $\mathbb{P}(y_{n+1} \in [\hat{y}_{n+1}]) \geq 1 - \epsilon_g$. Cette probabilité peut également s'écrire comme suit :

$$\begin{aligned} & \mathbb{P}(y_{n+1}^1 \in [\hat{y}_{n+1}^1, \bar{y}_{n+1}^1], \dots, y_{n+1}^m \in [\hat{y}_{n+1}^m, \bar{y}_{n+1}^m]) \\ &= \mathbb{P}\left(\frac{|y_{n+1}^1 - \hat{y}_{n+1}^1|}{\sigma_{n+1}^1} \leq \alpha_s^1, \dots, \frac{|y_{n+1}^m - \hat{y}_{n+1}^m|}{\sigma_{n+1}^m} \leq \alpha_s^m\right) \geq 1 - \epsilon_g. \end{aligned} \quad (6)$$

Ainsi, il faut trouver pour chaque cible j le score de non-conformité individuel α_s^j défini par un ϵ^j par cible pour qu'un niveau de confiance global $1 - \epsilon_g$ soit garanti. En étendant (4) et en considérant les variables aléatoires $Q^j = |y^j - \hat{y}^j|/\sigma^j, j \in \{1, \dots, m\}$, on obtient :

$$\mathbb{P}(Q^1 \leq \alpha_s^1, \dots, Q^m \leq \alpha_s^m) \geq 1 - \epsilon_g. \quad (7)$$

La distribution jointe dans (7) étant inconnue, nous exploitons la connexion entre cette équation et les copules. Avant cela, notons à nouveau que sous l'hypothèse que nous sommes bien calibrés, nous pouvons transformer (7) en

$$F(\alpha_s^1, \dots, \alpha_s^m) = 1 - \epsilon_g, \quad (8)$$

où F désigne ici la fonction de répartition jointe induite par \mathbb{P} .

4 Régression multi-cibles conformelle à base de copules

Cette section rappelle le principes des copules [Nelsen (1999)] avant de détailler comment nous les appliquons aux approches conformelles pour la régression multi-cibles.

4.1 Aperçu sur les copules

Soit $\mathbf{Q} = (Q^1, \dots, Q^m)$ un vecteur aléatoire de m dimensions composé des variables aléatoires Q^1, \dots, Q^m . Soit sa fonction de répartition jointe $F = F_Q : \mathbb{R}^m \rightarrow [0, 1]$. Cette f.r. contient deux informations importantes :

- La f.r. de chaque variable aléatoire Q^j tel que $F_j(q^j) = \mathbb{P}(Q^j \leq q^j) \forall j \in \{1, \dots, m\}$.
- La structure de dépendance entre elles.

Une copule à m dimensions $C : [0, 1]^m \rightarrow [0, 1]$ permet d'isoler cette structure de dépendance des marges. Elle est caractérisée par les propriétés suivantes :

Prédiction conformelle basée sur les copules pour la régression multi-cibles

1. $C(u^1, \dots, u^m) = 0$ si $u^j = 0$ pour au moins un $j \in \{1, \dots, m\}$.
2. $C(1, \dots, 1, u^j, 1, \dots, 1) = u^j$.
3. C est m -croissante.

Selon le théorème de [Sklar (1959)], il est possible de produire une fonction de distribution multivariée avec un mélange de distributions marginales univariées et une copule appropriée :

Théorème 4.1 (Théorème de Sklar). *Pour toute f.r. à m dimensions F avec des distributions marginales F_1, \dots, F_m , il existe une copule $C : [0, 1]^m \rightarrow [0, 1]$ telle que :*

$$F(\mathbf{q}) = F(q^1, \dots, q^m) = C(F_1(q^1), \dots, F_m(q^m)), \quad \mathbf{q} \in \mathbb{R}^m.$$

Si F_j est continue pour tout $j \in \{1, \dots, m\}$, alors C est unique.

4.2 Régression multi-cible conformelle à base de copules

Reprenons notre problème précédent consistant à trouver ϵ^j pour chaque cible j pour obtenir une confiance globale de $1 - \epsilon_g$. Considérons (8). Selon le théorème de Sklar, nous avons :

$$\begin{aligned} F(\alpha_s^1, \dots, \alpha_s^m) &= C(F_1(\alpha_s^1), \dots, F_m(\alpha_s^m)) \\ &= C(1 - \epsilon^1, \dots, 1 - \epsilon^m) \\ &= 1 - \epsilon_g \end{aligned}$$

où la deuxième ligne est obtenue à partir de (2). Il est clair que si nous connaissions la copule C , alors nous pourrions rechercher les valeurs ϵ^j fournissant la confiance globale désirée.

Un problème majeur est alors d'obtenir la copule modélisant la structure de dépendance entre les cibles et leurs niveaux de confiance. Comme les copules sont classiquement estimées à partir d'observations multi-variées, nous utiliserons ici les scores de non-conformité générés à partir de Z^{cal} . Plus précisément, si α_i^j est le score de non-conformité correspondant à la j^{eme} cible de z_i de Z^{cal} pour $i \in \{l+1, \dots, n\}$, nous estimons une copule C à partir de la matrice

$$A = \begin{bmatrix} \alpha_{l+1}^1 & \alpha_{l+1}^2 & \dots \\ \vdots & \ddots & \\ \alpha_n^1 & & \alpha_n^m \end{bmatrix}. \quad (9)$$

4.2.1 La copule indépendante

Cette copule considère les m cibles comme étant indépendantes. (7) devient alors :

$$\Pi(F_1(\alpha_s^1), \dots, F_m(\alpha_s^m)) = \prod_{j=1}^m F_j(\alpha_s^j) = \prod_{j=1}^m \mathbb{P}(Q^j \leq \alpha_s^j) \geq \prod_{j=1}^m (1 - \epsilon^j) = 1 - \epsilon_g.$$

Si on suppose que tous les niveaux de signifiante individuels $\epsilon^1, \dots, \epsilon^m$ sont égaux à la même valeur ϵ_t , alors :

$$\prod_{j=1}^m (1 - \epsilon^j) = (1 - \epsilon_t)^m = 1 - \epsilon_g.$$

Ainsi, le niveau de signifiante individuel ϵ_t est calculé avec $\epsilon_t = 1 - \sqrt[m]{1 - \epsilon_g}$, et ensuite utilisé pour obtenir les différents α_s^j pour chaque cible de la régression multi-cibles.

4.2.2 La copule de Gumbel

La copule de Gumbel est un membre de la famille des copules archimédiennes qui ne dépend que d'un seul paramètre θ , et donc est un bon représentant des copules paramétriques. Le paramètre θ est estimé à partir de la matrice A en utilisant l'estimateur de pseudo-vraisemblance maximale [Hofert et al. (2019)] avec une optimisation numérique. Une fois ceci obtenu, et pour un choix particulier de ϵ^j , on a l'expression de la copule de Gumbel :

$$C_G^{\hat{\theta}} = \exp - \left(\sum_{j=1}^m (-\ln(1 - \epsilon^j))^{\hat{\theta}} \right)^{1/\hat{\theta}} = \exp - \left(\sum_{j=1}^m (-\ln F_j(\alpha_s^j))^{\hat{\theta}} \right)^{1/\hat{\theta}} \quad (10)$$

Les estimations \hat{F}_j permettent de rechercher les valeurs ϵ^j qui rendront cette équation égale à $1 - \epsilon_g$. La solution est facile à obtenir analytiquement si $\epsilon^1 = \dots = \epsilon^m = \epsilon_t$, car on a alors que $\epsilon_t = 1 - (1 - \epsilon_g)^{1/\sqrt[m]{m}}$, et on peut donc obtenir $\alpha_s^1, \dots, \alpha_s^m$ en remplaçant F_j par \hat{F}_j .

Cette copule a été choisie car sa borne inférieure est la copule indépendante (avec $\hat{\theta} = 1$). Ainsi, on peut capturer l'indépendance si elle est vérifiée, et sinon chercher dans la direction de la dépendance positive. Une des raisons de ce choix est que les expériences précédentes [Messoudi et al. (2020)] indiquent que la copule indépendante donne des résultats trop conservatifs.

4.2.3 La copule empirique

Les copules paramétriques ont l'avantage de nécessiter moins de données pour être bien estimées, tout en ayant l'inconvénient d'induire un biais dans l'estimation susceptible de croître à mesure que le nombre de cibles augmente. La copule empirique présente une manière non paramétrique d'estimer les marginales directement à partir des observations. Ainsi, en posant \mathbf{u}_i les pseudo-observations qui remplacent les distributions marginales inconnues, nous avons alors que :

$$C_E(F_1(\alpha_s^1), \dots, F_m(\alpha_s^m)) = \frac{1}{n-l} \sum_{i=l+1}^n \prod_{j=1}^m \mathbb{1}_{u_i^j \leq F_j(\alpha_s^j)}. \quad (11)$$

En utilisant que $F_j(\alpha_s^j) = 1 - \epsilon^j$, nous pouvons alors rechercher les valeurs de ϵ^j , $j = 1, \dots, m$ qui rendront (11) égale à $1 - \epsilon_g$. Notons que dans ce cas, même en supposant que $\epsilon^1 = \dots = \epsilon^m = \epsilon_t$, une recherche algorithmique est nécessaire, qui est cependant facile puisque C_E est une fonction croissante (une simple recherche dichotomique suffit alors).

5 Expériences et résultats

Nous travaillons avec un réseau de neurones profond (NN) et une forêt aléatoire (RF) comme algorithmes sous-jacents, et nous comparons les trois copules pour montrer que leur ajout aux mesures de non-conformité fonctionne avec n'importe quel algorithme sous-jacent. Nous appliquons les étapes de régression conformelle décrites dans la section 2 comme suit :

1. Entraîner l'algorithme sous-jacent (NN ou RF) sur les données d'entraînement appropriées (X^{tr}, Y^{tr}), et prédire \hat{Y}^{cal} et \hat{Y}^{ts} pour les données de calibration et de test.

Noms	Exemples	Variabes	Cibles
music origin [Zhou et al. (2014)]	1059	68	2
sgemm [Nugteren et Codreanu (2015)]	241600	14	4

TAB. 1: Informations sur les ensembles de données de régression multi-cibles utilisés.

2. Entraîner le MLP de normalisation sur les données d'entraînement appropriées (X^{tr} , μ^{tr}) avec m sorties, et prédire μ^{cal} et μ^{ts} pour les données de calibration et de test.
3. Si nécessaire, obtenir une estimation de la copule C à partir de la matrice A des scores de non-conformité de calibration. Puis pour chaque niveau de signifiante global ϵ_g :
 - Obtenir $\epsilon^j = \epsilon_t$ pour $j \in \{1, \dots, m\}$ et calculer $\alpha_s = \{\alpha_s^1, \dots, \alpha_s^m\}$ pour toutes les cibles en utilisant Z^{cal} , selon les méthodes mentionnées dans la section 4.2.
 - Obtenir les prédictions d'intervalle pour les données de test avec (3).

L'implémentation a été faite en utilisant Python et est disponible sur Github². Nous utilisons huit jeux de données avec différents nombres de cibles et de tailles, mais nous présenterons ici les résultats pour deux jeux de données (les autres résultats peuvent être consultés sur notre article de journal [Messoudi et al. (2021)]). Ils sont résumés dans le tableau 1.

Pour évaluer la validité en régression multi-cibles, la précision est calculée en fonction de l'appartenance ou non de l'observation y à l'hyper-rectangle $[\hat{y}]$ selon ϵ_g . En le faisant pour plusieurs valeurs de ϵ_g et pour chaque copule, on obtient une courbe de précision qui doit être aussi proche que possible de la ligne de calibration, ligne où le taux d'erreur est exactement égal à ϵ_g pour un niveau de confiance $1 - \epsilon_g$. Les figures 1 et 2 détaillent ces résultats.

Les résultats montrent clairement que les meilleures performances sont obtenues en utilisant la copule Empirique, avec une courbe presque parfaitement alignée avec la ligne de calibration. Cela est dû au fait que la copule empirique utilise une estimation non paramétrique des marginales basée sur les observations, permettant au modèle de mieux s'adapter à la structure de dépendance de chaque jeu de données. Celle-ci est négligée lors de l'utilisation de la copule indépendante, puisque le lien entre les m cibles n'est pas exploité lors du calcul du ϵ_t . Par exemple, nous pouvons déduire que les cibles sont fortement liées pour "sgemm" par le grand écart entre les courbes de précision indépendante et empirique (Fig. 2a et 2b). Pour la copule de Gumbel, la courbe de précision est généralement plus proche de la ligne de calibration par rapport à la courbe de la copule indépendante. Cela confirme l'existence d'une structure de dépendance entre les cibles, puisque la borne inférieure de la copule de Gumbel est la copule indépendante, c.-à-d. si les cibles étaient indépendantes, les deux courbes correspondraient parfaitement. C'est le cas pour "music origin" (Fig. 1a et 1b).

Pour évaluer l'efficacité en régression multi-cibles, nous calculons le volume des prédictions obtenues avec (5), puis la valeur médiane de ces volumes d'hyper-rectangles pour chaque expérience (pour que l'estimation soit robuste face à de très grands hyper-rectangles).

Les résultats d'efficacité sont présentés dans les Figures 1c et 2c pour $\epsilon_g = 0.1$. Ils montrent que la copule indépendante présente un volume d'hyper-rectangle médian plus important que les copules de Gumbel et empirique, notamment dans les cas où l'existence d'une structure

2. <https://github.com/M-Soundouss/CopulaConformalMTR>

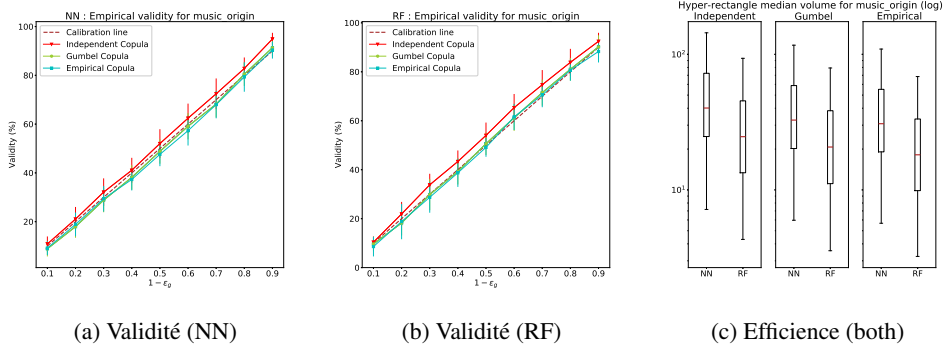


FIG. 1: Results for music origin.

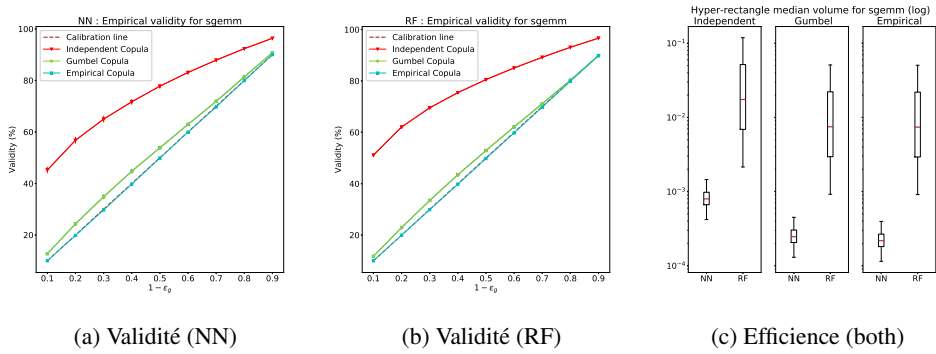


FIG. 2: Résultats pour sgemm.

de dépendance est confirmée par les courbes de calibration. Cela est dû au fait que la copule indépendante ignore la dépendance entre les scores de non-conformité, ce qui conduit à une surestimation de l'erreur globale de l'hyper-rectangle. Cet impact est évité pour la copule empirique car elle tire parti de cette dépendance pour construire de meilleures prédictions d'intervalle. Nous remarquons aussi que les diagrammes en boîte pour la copule empirique sont plus serrés que les deux autres, ce qui montre que les valeurs sont homogènes sur tous les plis par rapport à la copule indépendante par exemple, où la variation est beaucoup plus visible.

6 Conclusions et perspectives

Dans cet article, nous avons fourni un moyen assez facile d'obtenir des prédictions conformelles valides pour la régression multi-cibles. Pour ce faire, nous avons exploité un lien entre les scores de non-conformité et les copules, un outil couramment utilisé pour modéliser une distribution multivariée. Des expériences sur divers ensembles de données pour des copules représentatives montrent que la méthode permet d'améliorer l'hypothèse d'indépendance naïve.

Comme futures perspectives, nous aimerions explorer la flexibilité de notre cadre, par exemple en l’adaptant aux distributions prédictives conformelles plus riches [Vovk et al. (2017)], ou en proposant des protocoles permettant d’obtenir ϵ_g à partir de différents degrés de confiance individuels choisis. Nous aimerions aussi apprendre directement une fonction de coût prenant en compte la validité et l’efficacité [Colombo et Vovk (2020)] pour la régression multi-cibles, éventuellement en utilisant le volume de l’hyper-rectangle comme paramètre pour définir les valeurs de ϵ_t qui nous donnent le plus petit volume pour une même validité.

Références

- Colombo, N. et V. Vovk (2020). Training conformal predictors. In *Conformal and Probabilistic Prediction and Applications*, pp. 55–64. PMLR.
- Hofert, M., I. Kojadinovic, M. Mächler, et J. Yan (2019). *Elements of copula modeling with R*. Springer.
- Messoudi, S., S. Destercke, et S. Rousseau (2020). Conformal multi-target regression using neural networks. In *Conformal and Probabilistic Prediction and Applications*, pp. 65–83. PMLR.
- Messoudi, S., S. Destercke, et S. Rousseau (2021). Copula-based conformal prediction for multi-target regression. *Pattern Recognition*, 108101.
- Nelsen, R. B. (1999). An introduction to copulas, volume 139 of. *Lecture Notes in Statistics*.
- Nugteren, C. et V. Codreanu (2015). Cltune : A generic auto-tuner for opencl kernels. In *2015 IEEE 9th International Symposium on Embedded Multicore/Many-core Systems-on-Chip*, pp. 195–202. IEEE.
- Papadopoulos, H., K. Proedrou, V. Vovk, et A. Gammerman (2002). Inductive confidence machines for regression. In *European Conference on Machine Learning*, pp. 345–356. Springer.
- Sklar, M. (1959). Fonctions de repartition an dimensions et leurs marges. *Publ. inst. statist. univ. Paris* 8, 229–231.
- Vovk, V., J. Shen, V. Manokhin, et M.-g. Xie (2017). Nonparametric predictive distributions based on conformal prediction. In *Conformal and Probabilistic Prediction and Applications*, pp. 82–102. PMLR.
- Zhou, F., Q. Claire, et R. D. King (2014). Predicting the geographical origin of music. In *2014 IEEE International Conference on Data Mining*, pp. 1115–1120. IEEE.

Summary

Few works deal with conformal prediction for multi-target regression. This paper focuses on providing valid (i.e., frequency calibrated) multi-variate predictions. To do so, we propose to use copula functions for inductive conformal prediction, and empirically show that the proposed method ensures efficiency and validity for multi-target regression problems on various data sets.