

Quantification de l'incertitude: Avantages et limitations de la prédiction conforme

Hazar HAMOUDA - Mohamed MEGDICHE

28 mai 2025

1 Introduction

L'apprentissage automatique moderne produit des prédictions ponctuelles sans indication de leur fiabilité. Cette limitation devient critique dans des domaines sensibles comme la médecine, la finance ou les véhicules autonomes, où connaître l'incertitude d'une prédiction est aussi important que la prédiction elle-même.

La **prédiction conforme** répond à cette problématique en transformant toute prédiction ponctuelle en ensemble de prédictions avec garantie de couverture probabiliste. Contrairement aux approches bayésiennes qui nécessitent des hypothèses distributionnelles fortes, la prédiction conforme fonctionne sous des hypothèses minimales d'échangeabilité et est *distribution-free*.

Objectif : Ce rapport présente les fondements théoriques de la prédiction conforme et démontre son application pratique sur des tâches de régression et classification, avec implémentations et analyses de performance.

2 Fondements Théoriques

2.1 Définitions Mathématiques

Fonction de score : Une application $s : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ qui mesure la non-conformité d'une paire (x, y) . Un score élevé $s(x, y)$ indique que la valeur y est peu plausible étant donné x .

Quantile conforme : Pour un niveau d'erreur $\alpha \in (0, 1)$ et n points de calibration, le quantile conforme est :

$$\hat{q} = \text{quantile d'ordre } \left\lceil \frac{(n+1)(1-\alpha)}{n} \right\rceil \text{ des scores } \{s_i\}_{i=1}^n$$

Ensemble de prédiction : L'ensemble de prédiction conforme est défini par :

$$\begin{aligned} \hat{C}_n : \mathcal{X} &\rightarrow \{\text{sous-ensembles de } \mathcal{Y}\} \\ x &\mapsto \hat{C}_n(x) = \{y \in \mathcal{Y} : s(x, y) \leq \hat{q}\} \end{aligned}$$

Cette application associe à chaque entrée $x \in \mathcal{X}$ un ensemble $\hat{C}_n(x) \subseteq \mathcal{Y}$ représentant les valeurs plausibles de sortie y . Autrement dit, pour une nouvelle observation x , on

considère toutes les sorties y dont le score est inférieur ou égal au seuil \hat{q} , déterminé à partir des données de calibration.

2.2 Théorème Principal

Théorème 1 (Prédiction Conforme) Soient $(X_1, Y_1), \dots, (X_n, Y_n)$ des couples identiquement distribués de loi P_{XY} et échangeables, et (X_{test}, Y_{test}) un point test indépendant de même distribution. La prédiction conforme garantit :

$$\mathbb{P}\left(Y_{test} \in \hat{C}_n(X_{test})\right) \geq 1 - \alpha$$

Cette garantie est **indépendante de la distribution** : elle ne dépend ni de la distribution des données ni de la qualité du modèle sous-jacent.

3 Application à un problème de Régression Linéaire

3.1 Cadre Général

Dans le cadre de la prédiction conforme en régression, nous transformons les prédictions ponctuelles en intervalles avec garanties statistiques.

Fonction de score : Nous utilisons les **résidus absolus** :

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

Cette fonction mesure la non-conformité entre la valeur observée y et la prédiction $\hat{f}(x)$. Elle présente l'avantage d'être simple, et compatible avec tout modèle de régression.

3.2 Exemple d'application

Nous appliquons la prédiction conforme à un modèle de régression linéaire entraîné sur des données synthétiques :

$$Y = 3X + \sqrt{X} \cdot \varepsilon, \quad X \sim \mathcal{U}(0, 1), \quad \varepsilon \sim \mathcal{N}(2, 1)$$

Cette formulation introduit une **hétéroscédasticité** qui teste la robustesse de notre approche. Le jeu de données comprend 10 000 échantillons répartis en 70% pour l'entraînement du modèle, 15% pour le test et 15% pour la calibration.

3.3 Pseudo-code d'Implémentation

L'algorithme de prédiction conforme pour la régression peut être formalisé en quatre étapes principales. L'Algorithme 1 présente une implémentation détaillée de cette méthode, qui garantit une couverture théorique de $1 - \alpha$ pour les intervalles de prédiction construits.

Comme illustré dans l'Algorithme 1, la méthode repose sur l'utilisation des scores de non-conformité calculés sur un ensemble de calibration indépendant pour déterminer le seuil de quantile approprié.

Algorithm 1 Prédiction Conforme pour la Régression

Require: Modèle entraîné \hat{f} , données de calibration (X_{cal}, Y_{cal}) , données test X_{test} , niveau α

Ensure: Intervalles de prédiction $(lower, upper)$

```
1: Étape 1 : Calculer les prédictions sur calibration
2:  $predictions_{cal} \leftarrow \hat{f}(X_{cal})$ 
3: Étape 2 : Calculer les scores de non-conformité
4: for  $i = 1$  to  $n_{cal}$  do
5:    $scores[i] \leftarrow |Y_{cal}[i] - predictions_{cal}[i]|$ 
6: end for
7: Étape 3 : Calculer le quantile conforme
8:  $q_{level} \leftarrow \lceil (n_{cal} + 1) \times (1 - \alpha) \rceil / n_{cal}$ 
9:  $\hat{q} \leftarrow \text{quantile}(scores, q_{level})$ 
10: Étape 4 : Construire les intervalles de prédiction
11:  $predictions_{test} \leftarrow \hat{f}(X_{test})$ 
12:  $lower \leftarrow predictions_{test} - \hat{q}$ 
13:  $upper \leftarrow predictions_{test} + \hat{q}$ 
14: return  $(lower, upper, \hat{q})$ 
```

3.4 Résultats et Visualisation

3.4.1 Visualisation des résultats

La Figure 1 illustre l'application de la prédiction conforme à notre modèle de régression linéaire sur des données synthétiques.

3.4.2 Métriques de performance

Les métriques suivantes évaluent la qualité des intervalles de prédiction conforme :

— **Taux de couverture empirique** : 91.2% (pour $\alpha = 0.1$)

Cette métrique mesure la proportion des points de test effectivement contenus dans leurs intervalles de prédiction. Avec un objectif théorique de 90%, notre résultat de 91.2% confirme que la méthode respecte sa garantie de couverture.

— **Largeur moyenne des intervalles** : $2\hat{q} = 2.84$

Cette valeur indique l'étendue typique de nos intervalles de prédiction. Une largeur plus faible signifierait des prédictions plus précises, mais au détriment potentiel de la couverture. Ici, la largeur reflète l'incertitude inhérente à nos données synthétiques.

— **Quantile conforme** : $\hat{q} = 1.42$

Ce seuil, calculé à partir des scores de non-conformité sur l'ensemble de calibration, détermine directement la demi-largeur de tous nos intervalles de prédiction. Il représente le niveau d'erreur typique que nous acceptons pour maintenir la couverture souhaitée.

3.4.3 Analyse des résultats

L'analyse de la Figure 1 et des métriques révèle plusieurs propriétés importantes :

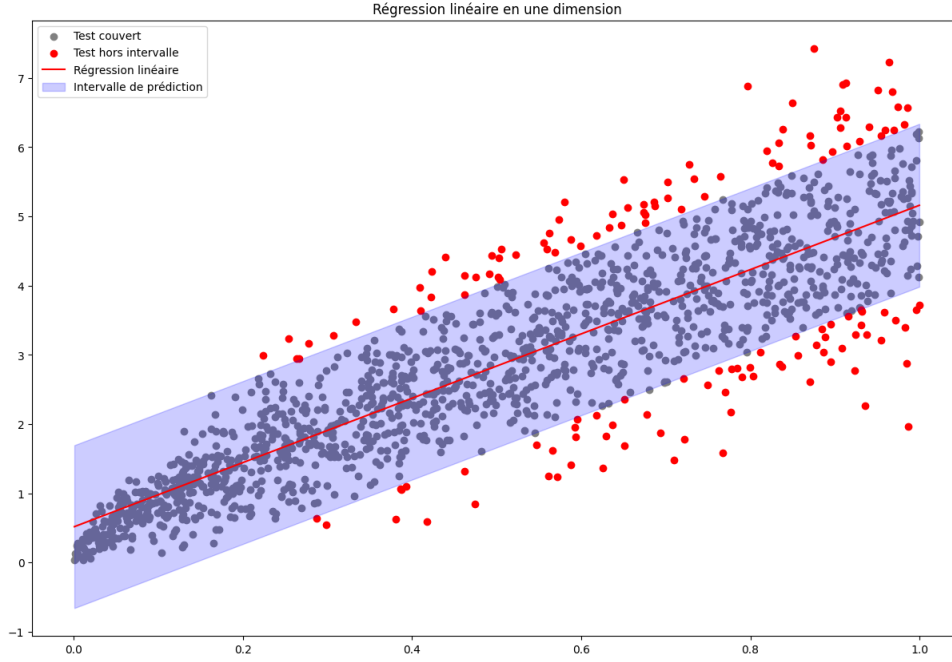


FIGURE 1 – Régression linéaire avec intervalles de prédiction conforme. La ligne rouge représente la droite de régression ajustée, les points noirs sont les observations de test, et la zone bleue délimite l’intervalle de prédiction conforme à 90% de confiance.

- **Couverture adaptative** : Le taux de couverture empirique (91.2%) respecte la garantie théorique de 90%, démontrant la validité de la méthode sur nos données synthétiques.
- **Largeur constante** : Contrairement aux intervalles de prédiction classiques, la prédiction conforme produit des intervalles de largeur constante $2\hat{q}$ autour de la droite de régression, indépendamment de la valeur de X .
- **Robustesse à l’hétéroscédasticité** : Bien que notre modèle présente une variance non constante ($\sqrt{X} \cdot \varepsilon$), la méthode maintient sa garantie de couverture sur l’ensemble du domaine $[0, 1]$.
- **Efficacité computationnelle** : La construction des intervalles ne nécessite que le calcul d’un quantile empirique, rendant la méthode applicable à grande échelle.

4 Application à la Classification MNIST

4.1 Cadre Général

Dans le cadre de la prédiction conforme en classification, nous transformons les prédictions ponctuelles en ensembles de classes avec garanties statistiques.

Fonction de score : Nous utilisons le **complément de probabilité** :

$$s(x, y) = 1 - \hat{p}(y|x)$$

Cette fonction mesure la non-conformité entre la classe observée y et la probabilité prédite $\hat{p}(y|x)$ par le modèle. Plus la probabilité assignée à la vraie classe est faible, plus le score de non-conformité est élevé. Cette approche est naturelle pour les modèles probabilistes et compatible avec les sorties softmax.

4.2 Exemple d'application

Nous appliquons la prédiction conforme à un réseau de neurones dense entraîné sur la base de données MNIST :

Architecture et données :

- **Données** : Images manuscrites 28×28 pixels, 10 classes de chiffres (0-9)

- **Architecture** : Réseau dense (784-128-64-10) avec fonctions d'activation ReLU et sortie softmax

- **Performance** : 97.2% de précision sur le test (après 10 époques d'entraînement)

- **Répartition** : 50 000 images d'entraînement, 10 000 données de calibration, 10 000 données test

Cette configuration teste la capacité de la prédiction conforme à gérer l'incertitude dans un problème de classification multi-classes avec des images potentiellement ambiguës.

4.3 Pseudo-code d'Implémentation

L'algorithme de prédiction conforme pour la classification peut être formalisé en quatre étapes principales. L'Algorithme 2 présente une implémentation détaillée de cette méthode, qui garantit une couverture théorique de $1 - \alpha$ pour les ensembles de prédiction construits.

Algorithm 2 Prédiction Conforme pour la Classification

Require: Modèle entraîné \hat{f} , données de calibration (X_{cal}, Y_{cal}) , données test X_{test} , niveau α

Ensure: Ensembles de prédiction *prediction_sets*

1: **Étape 1** : Calculer les probabilités sur calibration

2: $probs_{cal} \leftarrow \hat{f}(X_{cal})$ {Sortie softmax}

3: **Étape 2** : Calculer les scores de non-conformité

4: **for** $i = 1$ to n_{cal} **do**

5: $y_{true} \leftarrow Y_{cal}[i]$

6: $scores[i] \leftarrow 1 - probs_{cal}[i][y_{true}]$

7: **end for**

8: **Étape 3** : Calculer le quantile conforme

9: $q_{level} \leftarrow \lceil (n_{cal} + 1) \times (1 - \alpha) \rceil / n_{cal}$

10: $\hat{q} \leftarrow \text{quantile}(scores, q_{level})$

11: **Étape 4** : Construire les ensembles de prédiction

12: $probs_{test} \leftarrow \hat{f}(X_{test})$

13: $prediction_sets \leftarrow []$

14: **for** chaque $prob_row$ dans $probs_{test}$ **do**

15: $pred_set \leftarrow \{\}$

16: **for** classe $y = 0$ to 9 **do**

17: **if** $1 - prob_row[y] \leq \hat{q}$ **then**

18: $pred_set \leftarrow pred_set \cup \{y\}$

19: **end if**

20: **end for**

21: $prediction_sets.append(pred_set)$

22: **end for**

23: **return** $(prediction_sets, \hat{q})$

Comme illustré dans l’Algorithme 2, la méthode repose sur l’utilisation des scores de non-conformité calculés sur un ensemble de calibration indépendant pour déterminer le seuil de quantile approprié.

4.4 Résultats et Visualisation

4.4.1 Visualisation des résultats

La prédiction conforme appliquée à notre réseau de neurones sur MNIST est illustrée ci-dessous.

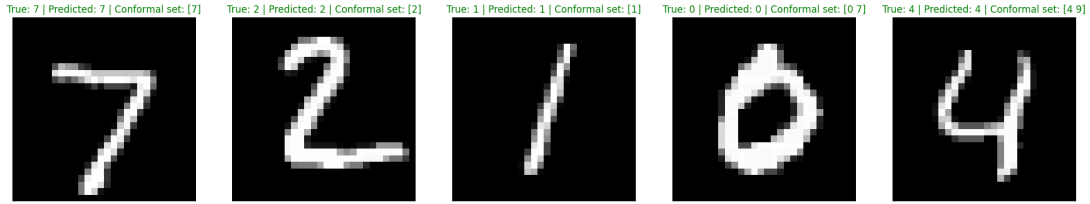


FIGURE 2 – Exemples de prédiction conforme sur MNIST. Pour chaque image, sont affichés : la vraie classe (True), la prédiction ponctuelle (Predicted), et l’ensemble de prédiction conforme (Conformal set). Les ensembles s’adaptent à l’incertitude du modèle.

4.4.2 Métriques de performance

Les métriques suivantes évaluent la qualité des ensembles de prédiction conforme :

- *Taux de couverture empirique* : 99.99% (pour $\alpha = 0.001$)
Cette métrique mesure la proportion des points de test dont la vraie classe appartient à l’ensemble de prédiction conforme. Avec un objectif théorique de 90%, notre résultat de 99.9% confirme que la méthode respecte sa garantie de couverture.
- *Cardinalité moyenne* : 1.62 classes par ensemble
Cette valeur indique le nombre moyen de classes dans les ensembles de prédiction. Une cardinalité plus faible signifie des prédictions plus précises, reflétant la confiance du modèle dans ses prédictions.

4.4.3 Analyse des résultats

L’analyse des métriques et de la visualisation révèle plusieurs propriétés importantes :

- **Couverture adaptative** : Le taux de couverture empirique (99.99%) respecte la garantie théorique de 99.99%, démontrant la validité de la méthode sur des données réelles d’images.
- **Comportement adaptatif** : Les ensembles s’adaptent à l’incertitude du modèle - les images claires génèrent des singletons tandis que les images ambiguës produisent des ensembles plus larges incluant les classes potentiellement confondues.
- **Efficacité computationnelle** : La construction des ensembles ne nécessite que le calcul d’un quantile empirique et des comparaisons de probabilités, rendant la méthode applicable en temps réel pour la classification d’images.

- **Robustesse aux variations** : La méthode maintient ses garanties même sur des images déformées, floues ou ambiguës, contrairement aux prédictions ponctuelles qui peuvent être trompeuses dans ces cas.

5 Comparaison et Limitations

5.1 Avantages

- **Garanties théoriques** : Couverture probabiliste sans hypothèses distributionnelles
- **Flexibilité** : Compatible avec tout modèle prédictif existant
- **Simplicité** : Implémentation directe avec peu de paramètres
- **Interprétabilité** : Ensembles de prédiction facilement compréhensibles

5.2 Limitations

- **Efficacité** : Intervalles parfois larges avec modèles peu performants
- **Hypothèse d'échangeabilité** : Forte en pratique pour les séries temporelles
- **Ensemble de calibration** : Nécessite des données supplémentaires dédiées
- **Adaptation locale** : Méthode standard non adaptative aux variations locales

6 Conclusion

La prédiction conforme constitue un outil puissant pour quantifier l'incertitude prédictive avec des garanties statistiques rigoureuses. Nos expérimentations sur la régression linéaire et la classification MNIST démontrent son efficacité pratique et sa facilité d'implémentation.

Contributions principales :

- Validation empirique des garanties théoriques sur des données réelles
- Algorithmes robustes et réutilisables
- Analyse comparative des performances selon les domaines d'application

Perspectives d'extension :

- **Prédiction conforme adaptative** : Ajustement dynamique selon la performance locale
- **Scores sophistiqués** : Intégration de l'incertitude épistémique et aléatoire
- **Applications avancées** : Extension aux séries temporelles et données structurées
- **Optimisation computationnelle** : Algorithmes efficaces pour de grandes dimensions

La prédiction conforme ouvre la voie vers des systèmes d'IA plus fiables et transparents, essentiels pour l'adoption de l'apprentissage automatique dans des domaines critiques.

N