**Project**

Les inondations peuvent-elles être prédites au Bangladesh ?

**1.Desicription du projet :**

Description du projet :

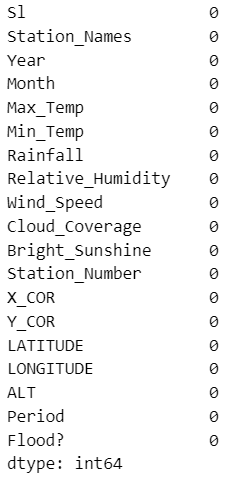
Peut-on prédire les inondations ? Si c'est le cas, des centaines de milliers de vies pourraient être sauvées.

L'objectif de ce projet est de tester plusieurs modèles de deep learning afin de développer et comparer les approches existantes, dans le but de sélectionner le ou les modèles les plus performants pour la prédiction. Nous nous fixons l'objectif que le modèle ait une précision (accuracy) d'au moins 0,94 pour le test et qu'il ait un rappel (recall) pour la classe 'inondation' plus élevé que pour la classe 'non inondation'.

**2.modèle utilisées :**

* KNN
* SVM avec tous les kernels existants
* Kmeans
* CNN\_RNN
* CNN\_GRU

**3.Traitement de données :**Le code Python data.isnull().sum() permet de vérifier l’existence de valeurs manquantes ou nulles.

****

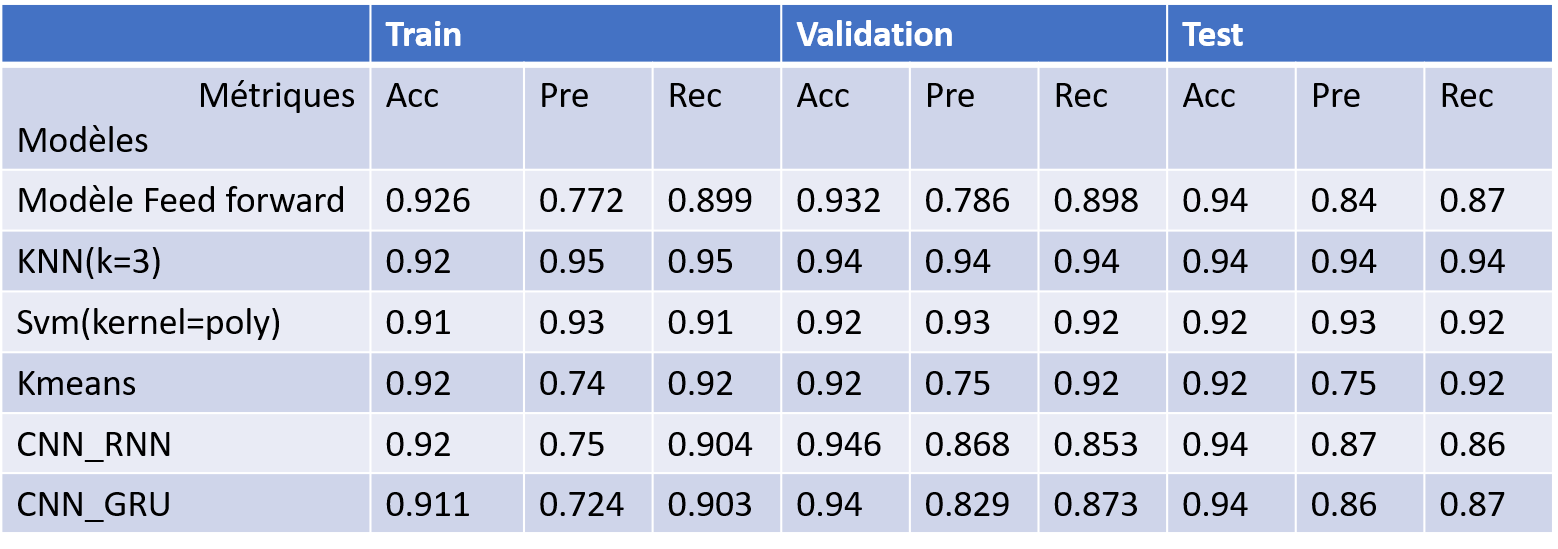
**4. Composition du dataset :**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Inondation | | Total |
| Oui | non |
| Train | 3307 | 13128 | 19742 |
| Validation | 401 | 1653 | 2054 |
| Test | 424 | 1631 | 2055 |

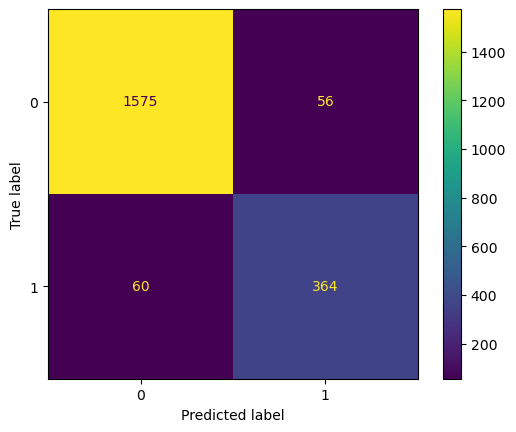
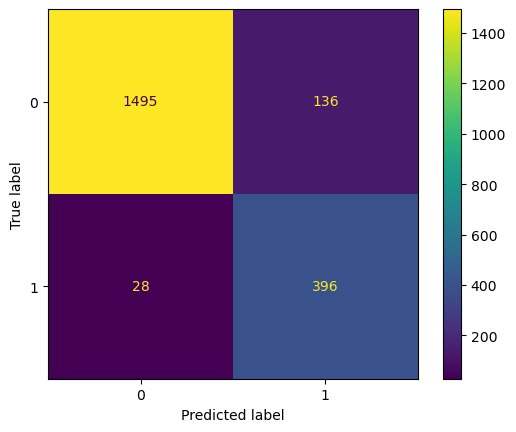
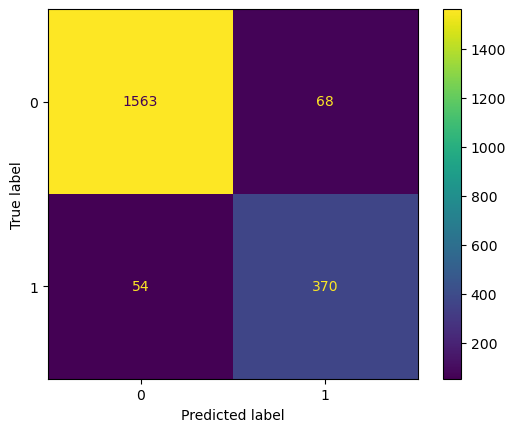
La base de données est non équilibrée, deux approches sont utilisées pour résoudre ce problème :

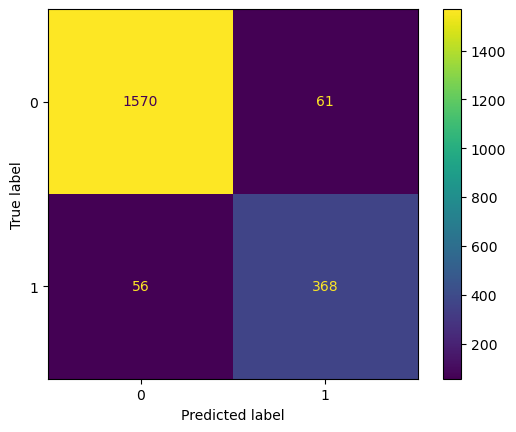
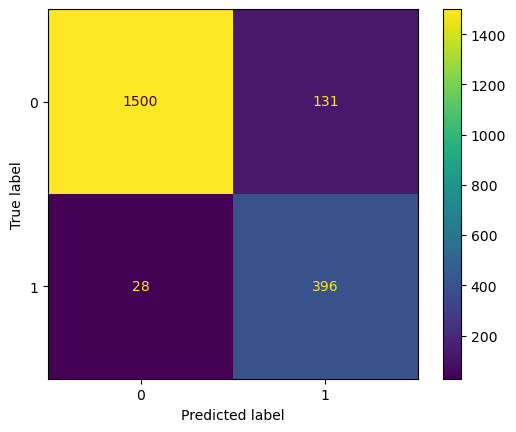
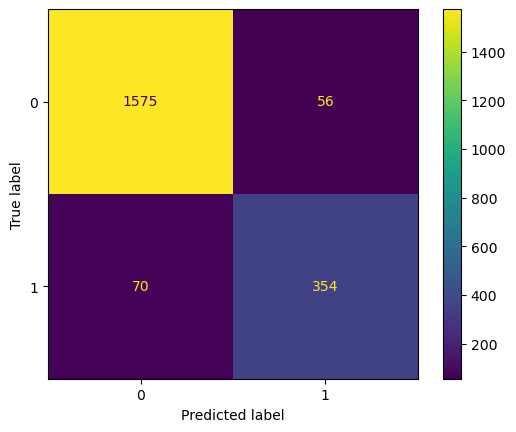
* Affecter des poids aux classes avec un poids plus élève à la classe minoritaire
* Augmentation des données

**5.** **Approche de classes avec poids :**

****

Il est possible d’observer que tous les modèles ont donnés une performance supérieure à 90 %





CNN\_RNN

SVM(kernel=Poly)

Feed forward

CNN\_GRU

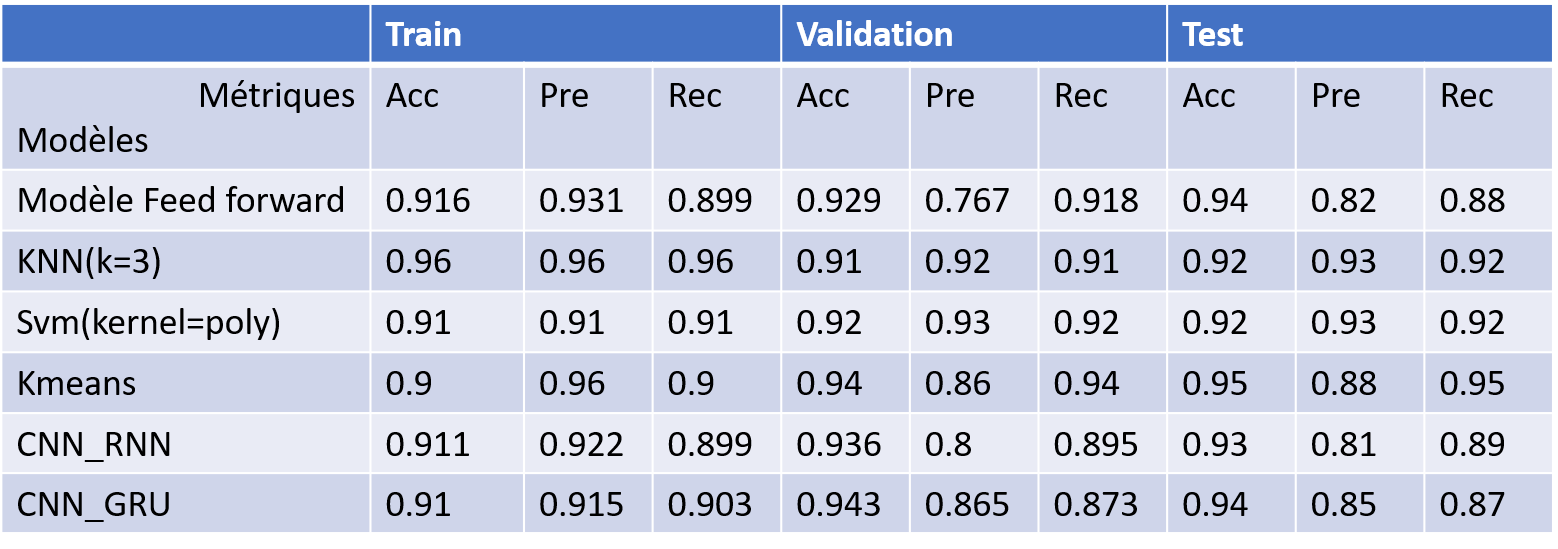
Kmeans

KNN(k=3)

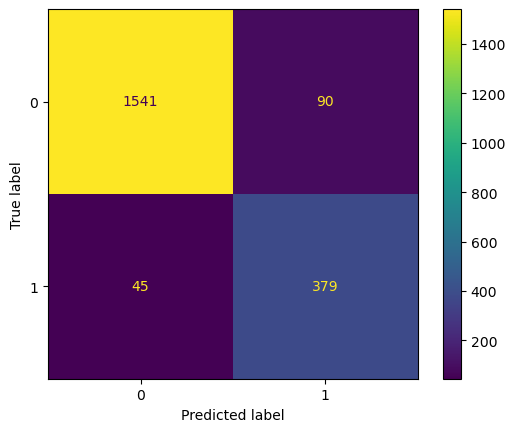
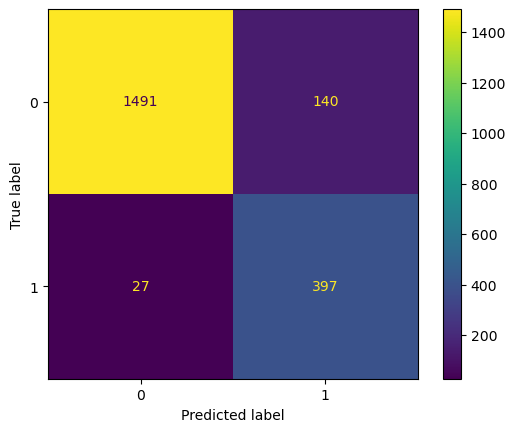
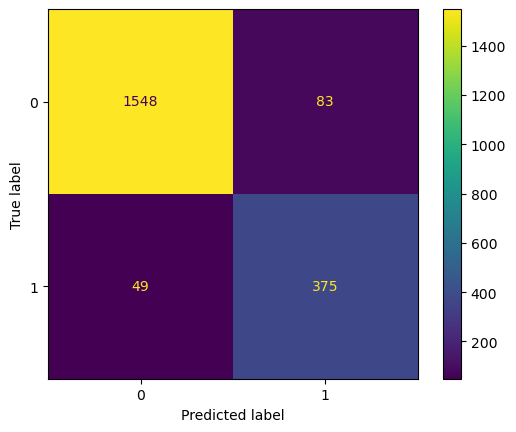
Il est possible de voir à partir des matrices de confusion obtenu pour chacun des modèles que le modèle feed\_forward ,svm avec kernel poly, le modèle kmeans et cnn\_GRU ont un recall plus élévé pour la classe inondation

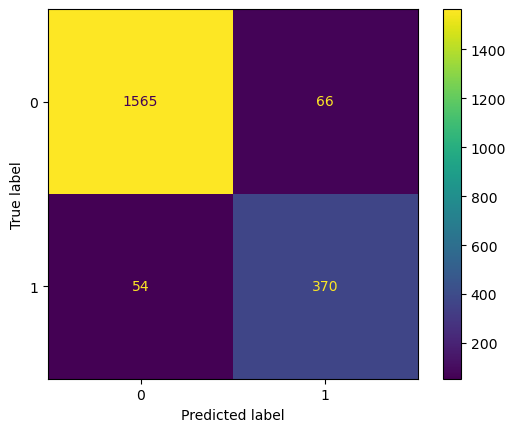
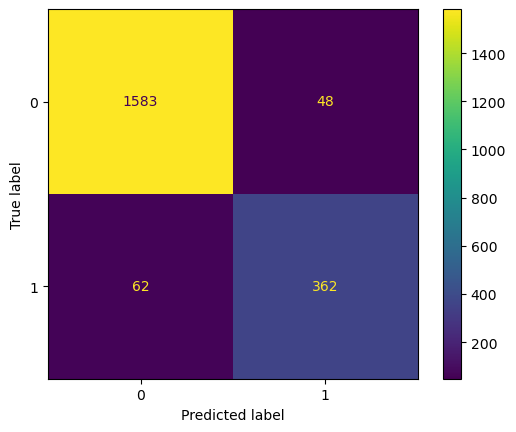
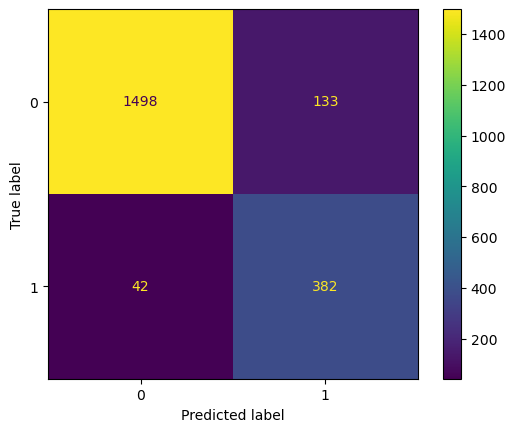
**6.** **Approche de classes avec Augmentation des données :**

Dans cette partie, nous allons utiliser la data augmentation et voir si elle a un impact sur l'amélioration des résultats des modèles. Pour cela, SMOTE sera utilisé.



Dans ce tableau il est possible de voir que tous les modèles ont donnés une performance supérieure à 90 %, cependant ces résultats sont presque les mêmes que ceux obtenu dans l’approche avec poids, donc l’approche avec augmentation de données n’a pas donné de meilleurs résultats.





CNN\_RNN

SVM(kernel=Poly)

Feed forward

CNN\_GRU

Kmeans

KNN(k=3)

De même pour l’approche avec augmentation des données il est possible de voir que le modèle feed\_forward ,svm avec kernel poly, le modèle kmeans et cnn\_GRU on un recall plus élévé pour la classe inondation

**7.Conclusion générale :**

* Les modèles sélectionnés pour la prédiction des inondations sont CNN\_DENSE, SVM avec kernel polynomial, K-Means, et CNN\_GRU
* Ces modéles ont atteint une accuracy supérieure à 0,94 et présentent un recall pour la classe 'inondation' plus élevé que pour la classe 'non-inondation’.
* Ces résultats permettent de conclure que les phénomènes d'inondation au Bangladesh peuvent être prédits avec une grande précision.