

Questions pour le rapport

Partie 1

1. Quel a été le rôle des callbacks ModelCheckpoint et EarlyStopping ? Ont-ils été utiles pendant votre entraînement ?

Le rôle des callbacks ModelCheckpoint servait à garder automatiquement le meilleur modèle pendant l'entraînement et l'EarlyStopping arrêta l'entraînement quand le modèle ne progressait plus.

Oui les 2 modèles restent très utiles car cela évite de perdre du temps et de ne pas trop l'entraîner pour rien.

2. Analysez les courbes d'apprentissage. Le modèle est-il en sur-apprentissage ou sous-apprentissage ? Et la matrice de confusion, ça aide ?

En observant les courbes on peut voir que la précision atteint presque 99.9% en entraînement et environ 98,5% en validation. On peut voir un léger sur-apprentissage.

La matrice de confusion est utile car elle montre où le modèle se plante et visuellement cela permet de savoir où améliorer.

3. Quels ont été les principaux défis rencontrés ? Comment pourriez-vous améliorer les performances de ce premier modèle ?

Le premier défi est que le CNN de base surapprend très vite, c'est pourquoi nous avons utilisé le modèle pré-entraîné MobileNetV2. Le défi suivant est le redimensionnement vers 224x224 pour MobileNetV2, c'est pourquoi il est important de mettre les callbacks intelligents earlystopping et Modelcheckpoint.

4. Compare les performances entre le modèle simple et le modèle avec transfert d'apprentissage. Pourquoi y a une telle différence ?

Le modèle simple CNN fonctionne bien mais on voit qu'il atteint vite ses limites surtout quand les images sont difficiles à reconnaître. Il est très rapide à entraîner parce que les images sont petites.

Le modèle avec MobileNetV2 donne de meilleurs résultats car il reconnaît les images plus facilement et correctement du fait qu'il a déjà été entraîné sur un très grand ensemble d'images. Donc il a déjà appris à détecter pleins de formes.

Partie 2

1. Expliquez le rôle du paramètre `class_weight='balanced'`. Quelles autres techniques auriez-vous pu utiliser pour gérer le déséquilibre des classes ?

Le rôle du paramètre `class_weight='balanced'` est d'aider le modèle à mieux garder les données déséquilibrées. Il donne plus de poids aux erreurs sur la classe minoritaire. Il y a plusieurs autres techniques comme SMOTE ou undersampling.

2. Quel type d'apprentissage et quel type de problème est-ce ? Argumentez.

C'est clairement du machine learning supervisé car on a des exemples avec leurs étiquettes et le problème est une classification binaire.

3. Donnez un exemple de régression lié à ce type de données, mettez un exemple via argumentation et image

Voici un exemple de régression, si une banque veut prédire combien de client va dépenser demain ou dans les prochains jours, c'est une estimation d'un montant. C'est donc une régression.

