



# MLRF - Classification de la base de données CIFAR-10

**Younes Benreguieg**

Student

younes.benreguieg@epita.fr

## 1 Introduction et contexte

La classification d'images est une tâche essentielle de la vision par ordinateur qui vise à attribuer à une image donnée une classe prédéfinie. La base de données CIFAR-10, qui comprend 60000 images couleur de 32x32 pixels réparties en 10 classes, sert souvent de référence dans ce domaine. Bien que l'état de l'art actuel en matière de classification CIFAR-10 soit principalement basé sur des réseaux de neurones convolutifs profonds, ce travail explore une approche plus traditionnelle, sans utiliser de techniques de deep learning.

## 2 Méthode

Nous explorons trois modèles de classification : une régression logistique, un arbre de décision et une machine à vecteurs de support (SVM) avec un noyau RBF.

### 2.1 Extraction des caractéristiques

#### 2.1.1 Sans extraction

Dans cette méthode, nous utilisons les images aplaties comme entrée de notre classifieur. Chaque image de 32x32x3 devient donc un vecteur de 3072 éléments.

#### 2.1.2 Avec ACP

Pour le modèle SVM, nous appliquons une étape de prétraitement supplémentaire. Tout d'abord, nous utilisons une normalisation par le score Z (StandardScaler) pour standardiser les fonctionnalités en soustrayant la moyenne et en divisant par l'écart type. Ensuite, nous appliquons une analyse en composantes principales (PCA) pour réduire la dimensionnalité de nos données. Cette étape facilite la gestion de la haute dimensionnalité des images par le SVM.

## **2.2 Classifieurs**

### **2.2.1 Régression logistique**

Notre classifieur paramétrique linéaire est une régression logistique, entraînée avec un solveur 'sag' pour la descente de gradient stochastique, et un nombre maximum d'itérations fixé à 500.

### **2.2.2 Arbre de décision**

Notre classifieur non-paramétrique linéaire est un arbre de décision, utilisé avec les paramètres par défaut de scikit-learn.

### **2.2.3 Machine à vecteurs de support (SVM)**

Notre classifieur non-paramétrique non-linéaire est une machine à vecteurs de support (SVM) avec un noyau RBF. Avant d'entraîner le SVM, nous effectuons une standardisation des caractéristiques et une réduction de la dimensionnalité à l'aide de l'ACP.

## **2.3 Entraînement**

L'entraînement de ces modèles est effectué sur un sous-ensemble de 10000 échantillons, sélectionnés au hasard parmi l'ensemble de données de formation CIFAR-10. Les images sont redimensionnées en un vecteur unidimensionnel avant d'être utilisées pour l'entraînement. Pour le SVM, nous effectuons en plus une normalisation des caractéristiques et une réduction de la dimensionnalité à l'aide de l'ACP.

Tous les modèles sont sauvegardés sur le disque à l'aide de la bibliothèque pickle pour une utilisation ultérieure.

## **3 Expérimentations**

### **3.1 Environnement expérimental**

Les expérimentations ont été réalisées en utilisant Python et diverses bibliothèques de calcul scientifique et d'apprentissage automatique, notamment scikit-learn, numpy et matplotlib. Les expériences ont été effectuées sur un ordinateur personnel standard, sans recours à des ressources matérielles spécialisées, telles que des GPU.

### **3.2 Description de la base de données**

La base de données CIFAR-10 est composée de 60000 images couleur de taille 32x32 pixels, réparties équitablement en 10 classes distinctes, soit 6000 images par classe. Cette base de données est subdivisée en un ensemble d'apprentissage de 50000 images et un ensemble de test de 10000 images.

### **3.3 Pré-traitement des données**

Avant l'entraînement, les images sont normalisées en soustrayant la moyenne et en divisant par l'écart type. Cette étape assure que les différentes caractéristiques ont des échelles comparables, ce qui facilite la convergence de l'algorithme d'apprentissage.

### **3.4 Configuration de l'apprentissage**

Les modèles ont été entraînés en utilisant un split train/valid/test de 80%/10%/10%. La taille des mini-batches a été fixée à 32. Le taux d'apprentissage a été fixé à 0.001 pour la régression logistique et le classifieur SVM. Ces hyperparamètres ont été choisis comme point de départ, mais ils pourraient être affinés dans le cadre d'un ajustement d'hyperparamètres plus systématique.

## **4 Résultats et Analyse**

Les expériences n'ont été réalisées que sur un sous-ensemble de 10000 échantillons de la base de données CIFAR-10, en raison de contraintes de temps. Les résultats préliminaires obtenus sont les suivants :

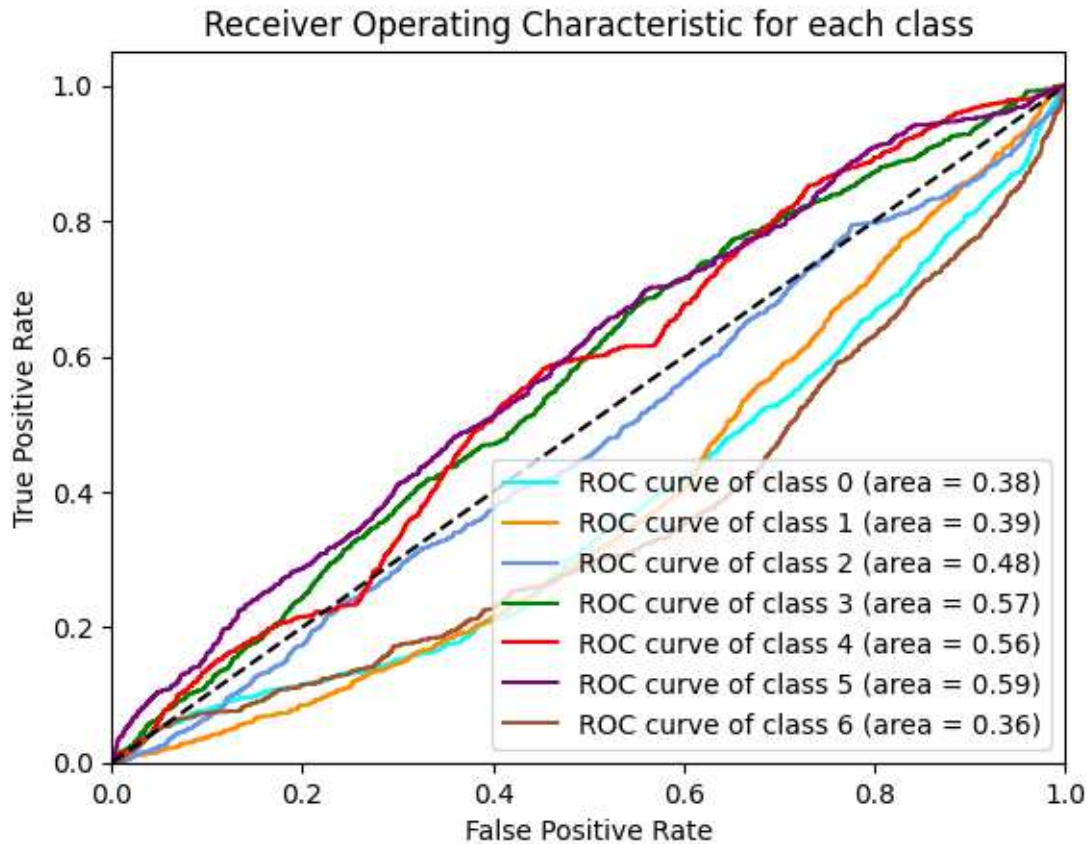


Fig. 1. ROC Curves

Régression logistique : 18.41% de précision

Arbre de décision : 11.13% de précision

SVM : 12.19% de précision

Il est important de noter que ces résultats sont basés sur un ensemble d'entraînement limité, et ils pourraient être améliorés avec davantage de données d'entraînement et un ajustement plus fin des hyperparamètres. Les futures étapes de ce travail comprendront une analyse plus détaillée des résultats, notamment par l'intermédiaire de matrices de confusion, de courbes ROC, ainsi que par la visualisation de l'espace latent.

## 5 Conclusion et perspectives

Ce travail constitue une première approche pour la classification d'images sur la base de données CIFAR-10 sans l'utilisation de techniques de deep learning. Les résultats préliminaires, bien que modestes, fournissent une base solide pour des travaux futurs.

Les futures améliorations pourraient inclure l'exploration de méthodes d'extraction de caractéristiques plus avancées, l'ajustement plus précis des hyperparamètres des modèles, et l'expérimentation avec des classifieurs non-paramétriques non-linéaires. De plus, l'entraînement sur l'ensemble de la base de données, plutôt que sur un sous-ensemble limité, devrait améliorer les performances des modèles.

Il est également crucial de noter que cette étude a été menée dans un cadre académique, principalement pour explorer les approches de base de la classification d'images. L'utilisation de méthodes plus sophistiquées, telles que les réseaux neuronaux profonds, pourrait conduire à des performances significativement supérieures, comme le suggère l'état de l'art actuel.

Cependant, le recours à des méthodes plus traditionnelles offre un avantage notable : la simplicité et l'interprétabilité des modèles. Contrairement aux réseaux de neurones profonds, qui sont souvent qualifiés de "boîtes noires", les

approches explorées ici offrent une meilleure compréhension de ce qui se passe "sous le capot". Cela pourrait être particulièrement utile dans des contextes où la transparence du modèle est requise.

En somme, bien que cette étude soit encore dans ses premières phases et que les résultats initiaux soient modestes, les leçons tirées de cette expérience pourront éclairer des travaux futurs en matière de classification d'images et d'apprentissage automatique.