

# Analyse des sentiments : les critiques de films sont-elles bonnes ou mauvaise

## **Objectif :**

Vous entraînerez un classificateur binaire pour effectuer une analyse des sentiments sur un ensemble de données IMDB.

## **Contexte :**

Ce carnet forme un modèle d'analyse des sentiments pour classer les critiques de films comme positives ou négatives , en fonction du texte de la critique. Il s'agit d'un exemple de classification binaire ( ou à deux classes), un type de problème d'apprentissage automatique important et largement applicable.

Vous utiliserez l' ensemble de données de critiques de films volumineux qui contient le texte de 50 000 critiques de films de la base de données Internet Movie . Ceux-ci sont divisés en 25 000 avis pour la formation et 25 000 avis pour les tests. Les ensembles d'entraînement et de test sont équilibrés , ce qui signifie qu'ils contiennent un nombre égal d'avis positifs et négatifs.

1. Télécharger les données, explorer le répertoire et charger le jeu de données.

`https://ai.stanford.edu/~amaas/data/sentiment/aclImdb\_v1.tar.gz`

2. Notez que les critiques contiennent du texte brut (avec des signes de ponctuation et des balises HTML occasionnelles comme `<br/>` ).