

Documentation
T-DAT-902

NOSTRADAMOVIES

POSTERS STANDARDISATION YOU SAY?



L'impact d'une affiche de film ne doit pas être sous-estimé, car elle peut donner une première impression du film qu'elle représente. Les agents de marketing et les designers le savent bien, et ils mettent souvent en œuvre des stratégies pour transmettre des messages précis à travers ces affiches. Cependant, grâce aux avancées technologiques et à l'existence de bases de données bien étiquetées ainsi que de nombreuses librairies Python. Il est maintenant possible de prédire le genre, l'année et le pays de production d'un film rien qu'en analysant son affiche.

Dans ce document, nous présenterons les méthodes que nous avons mises en place pour réaliser de telles prédictions. Un poster de film contient plusieurs informations cruciales que nous expliquerons en détail pourquoi nous avons tiré parti. Notre objectif était d'acquérir un ensemble de données complet et d'utiliser des techniques d'apprentissage automatique pour analyser les affiches et extraire des informations clés telles que le genre, l'année et le pays de production.

1. Première approche

Après avoir étudié précédemment un projet où nous avons réussi à déterminer si un poumon était malade ou non en utilisant une base d'images. Nous avons décidé d'adopter une approche similaire pour comprendre les enjeux du projet.

Nous avons opté pour l'utilisation d'un CNN (réseau de neurones convolutif), tout comme dans le projet précédent, en utilisant l'intégralité de la banque d'images disponible. Cette approche nous a permis d'obtenir nos premiers résultats.

Prenons l'exemple de l'affiche d'une série Netflix intitulée "Mercredi". Voici notre première tentative :



```
# Prédiction sur un poster donné
poster_filename = r"/Users/dimitriraymond/Documents/Epitech/Semestre 2/DAT/outer.jpg"
img = Image.open(posters_filename).convert("RGB")
img = img.resize(posters_size)
X_pred = np.array(img).reshape(1, posters_size[0], posters_size[1], 3)
y_pred = model.predict(X_pred)
predicted_genre = metadata["genre"].unique()[np.argmax(y_pred)]
print("Predicted genre:", predicted_genre)
```

```
1/1 [=====] - 0s 19ms/step
Predicted genre: Comedy, Music
```

En effet, nous reconnaissons que l'approche que nous avons utilisée jusqu'à présent est relativement simple, et bien que les résultats obtenus soient cohérents, nous n'avons pas encore exploité toutes les informations présentes dans l'image.

Dans notre étude initiale, nous nous sommes principalement concentrés sur l'utilisation du CNN pour extraire des caractéristiques visuelles générales de l'affiche, telles que les couleurs, les formes et les motifs. Cependant, nous sommes conscients que d'autres informations présentes dans l'image, telles que les éléments spécifiques du scénario, les personnages ou les symboles représentés, pourraient également être pertinentes pour prédire le genre.

Afin de prendre en compte ces informations supplémentaires, nous envisageons d'explorer des techniques plus avancées, telles que le traitement du langage naturel (NLP) pour analyser les titres, les descriptions associés aux affiches ou encore d'utiliser des modèles de vision pour détecter des objets spécifiques dans l'image. Cela nous permettra de tirer le meilleur parti de toutes les informations disponibles et d'améliorer nos prédictions.

2. Informations disponibles

En effet, pour améliorer les résultats de notre modèle, il est essentiel de disposer d'un dataset riche en informations. Ce dataset devrait inclure toutes les informations pertinentes qui peuvent être extraites à partir d'une image.

Prenons l'affiche de la série "Mercredi" comme exemple. Dans cette image, plusieurs éléments indiquent que c'est un film d'horreur, comme des couleurs sombres et saturées, des symboles effrayants, des images de personnages ou de scènes terrifiantes, ainsi que des éléments de typographie spécifiques.

Pour construire un dataset plus complet, nous pourrions collecter un ensemble d'images d'affiches de films d'horreur existants. En examinant et en étiquetant ces images de manière approfondie, nous pourrions extraire des caractéristiques spécifiques qui sont récurrentes dans le genre de l'horreur. Cela inclurait des informations telles que les couleurs dominantes, les motifs, les éléments iconographiques, les expressions faciales des personnages, etc.

En enrichissant notre dataset avec ces informations, notre modèle aurait une base plus solide pour apprendre les caractéristiques distinctives des films d'horreur et améliorer ainsi sa capacité à prédire avec précision le genre, l'année et le pays de production des films à partir de leurs affiches.

3. Deuxième approche

Les méthodes utilisées dans la première approche n'ont pas été fructueuses. Notre modèle n'était pas cohérent, comme on nous l'a fait remarquer. Afin de remédier à cela, nous devons récupérer un ensemble de données pertinentes sur nos affiches et images d'entraînement, que nous pourrions utiliser comme features lors de la création de notre modèle. Nous avons rassemblé ces données pour former notre ensemble de données (dataset).



A. La couleur prédominante

Effectivement, nous pouvons observer que l'affiche de la série "Mercredi" est plutôt sombre, ce qui peut être courant dans de nombreux films d'horreur. Cette observation nous conduit à réfléchir à la possibilité d'utiliser des bibliothèques Python pour extraire les pixels de l'image et calculer la couleur dominante présente.

Pour atteindre cet objectif, nous pouvons utiliser des bibliothèques populaires de traitement d'images en Python telles que OpenCV ou PIL (Python Imaging Library). Ces bibliothèques nous permettront de lire l'image et d'extraire les valeurs de couleur des pixels.

En parcourant les pixels de l'image, nous pourrions compter combien de pixels correspondent à chaque valeur de couleur (RVB). En identifiant la couleur la plus fréquente, nous pourrions ainsi déterminer la couleur dominante présente dans l'image.

Il est important de noter que cette approche se concentre spécifiquement sur l'analyse des couleurs et ne prend pas en compte les autres informations visuelles ou textuelles présentes dans l'affiche. Cependant, cela peut être une caractéristique intéressante à prendre en compte lors de la prédiction du genre d'un film, notamment pour des genres spécifiques comme l'horreur, où des schémas de couleurs sombres et oppressantes sont souvent utilisés.

```
# Chemin vers l'image
chemin_image = r"/Users/dimitriraymond/Documents/Epitech/Semestre 2/DAT/mercredi.jpeg"

# Appeler la fonction pour récupérer la couleur dominante
couleur_dominante = get_couleur_dominante(chemin_image)

# Afficher le résultat
print("Couleur dominante (R, G, B) :", couleur_dominante)

✓ 0.0s

Couleur dominante (R, G, B) : (1, 2, 22)

# Exemple d'utilisation
couleur_rgb = couleur_dominante # Remplacez par vos propres valeurs RVB
couleur_nominale = convertir_en_couleur_nominale(couleur_rgb)

# Afficher le résultat
print("Couleur nominale :", couleur_nominale)

✓ 0.0s

Couleur nominale : noir
```

B. Visage et expression du visage

Dans le cadre de notre projet, nous utilisons la détection des visages à l'aide de classificateurs en cascade Haar pour localiser les visages sur les affiches de films. En comptant le nombre de visages détectés, nous pouvons distinguer les films mettant en avant un seul personnage des films avec un groupe de personnages, ce qui est crucial pour prédire leur genre. De plus, nous analysons les émotions exprimées sur les visages en utilisant un modèle de reconnaissance des émotions pré-entraîné. Pour les affiches présentant deux visages, nous prenons en compte l'émotion du visage le plus mis en avant, c'est-à-dire celui qui est le plus grand et occupe une position centrale. Quant aux affiches avec plus de deux visages, nous identifions l'émotion dominante. Cette analyse nous fournit des indices précieux sur le ton et le contenu émotionnel du film, ce qui est essentiel pour affiner nos prédictions de genre.



C. Année de production

L'année de production d'un film peut également être utile pour déterminer le genre d'une affiche. En effet, les tendances cinématographiques évoluent en fonction des différentes époques. Les affiches étant déjà classées selon leur année de sortie, il est relativement facile d'intégrer cette information dans notre dataset. Par conséquent, nous avons pris la décision de l'inclure dans notre ensemble de données.

En tenant compte de l'année de production dans notre modèle, nous pouvons considérer les influences culturelles et artistiques propres à chaque période. Par exemple, les films d'horreur des années 1980 peuvent présenter des esthétiques et des thèmes distincts de ceux des années 2000. En incorporant cette dimension temporelle, nous espérons améliorer la précision de notre modèle dans la prédiction du genre des films à partir de leurs affiches.

D. Le Texte

Dans certaines affiches, le titre peut jouer un rôle important dans la détermination du thème du film. En examinant attentivement l'affiche en question, nous pouvons identifier plusieurs textes visibles. Il serait bénéfique de les extraire et d'établir le thème principal qui en ressort.

Le titre d'un film peut fournir des indications précieuses sur le genre, l'intrigue ou l'ambiance générale du film. Par exemple, si le titre contient des mots clés tels que "horreur", "mystère" ou "science-fiction", il y a de fortes chances que le film soit associé à ces genres respectifs. En analysant le texte présent sur l'affiche, nous pouvons donc obtenir des informations supplémentaires pour affiner notre prédiction du genre du film.

Pour relever et analyser le texte, nous pourrions utiliser des techniques de traitement du langage naturel (NLP) en utilisant des bibliothèques Python telles que NLTK (Natural Language Toolkit) ou spaCy. Ces bibliothèques nous permettraient de détecter les mots

clés, d'effectuer une classification thématique ou même d'extraire des concepts spécifiques liés au film.



```
[nltk_data] Error loading stopwords: <urlopen error [SSL:
[nltk_data] CERTIFICATE_VERIFY_FAILED] certificate verify failed:
[nltk_data] unable to get local issuer certificate (_ssl.c:992)>
[nltk_data] Error loading averaged_perceptron_tagger: <urlopen error
[nltk_data] [SSL: CERTIFICATE_VERIFY_FAILED] certificate verify
[nltk_data] failed: unable to get local issuer certificate
[nltk_data] (_ssl.c:992)>
Sujets du texte (noms) : ['fun', '[', 'peel', 'apie', 'peter', 'world', 'need', 'friends', 'wlan', 'hue', 'glen', 'case', 'kay', 'place', 'tamiluams']
Prénoms du texte : ['[']
Mots les plus fréquents : [('fun', 1), ('[', 1), ('peel', 1), ('apie', 1), ('peter', 1), ('inacold', 1), ('world', 1), ('need', 1), ('friends', 1), ('_', 1)]
Synonymes des mots les plus fréquents : ['fun', 'Peel', 'Peter', 'universe', 'need', 'friend']
```

Grâce à l'utilisation de plusieurs bibliothèques Python, nous sommes en mesure de récupérer une partie du texte présent dans l'image. Il est important de noter que nous n'avons pas pu récupérer l'intégralité du texte et qu'il peut y avoir des fautes d'orthographe dans les extraits relevés. Cependant, ce qui nous intéresse le plus, c'est le thème qui se dégage de ces extraits de texte. De plus, nous avons réussi à extraire les synonymes les plus fréquents, tels que "fun" et "friend", qui pourraient être utilisés dans notre dataset.

E. Titre

Nous avons mis en œuvre une solution pour détecter et regrouper les titres présents dans une image en utilisant des techniques de traitement d'image et de clustering. Le processus commence par charger l'image, puis la convertir en niveaux de gris. En appliquant un seuillage sur l'image, un masque binaire est obtenu, permettant de distinguer les zones correspondant aux titres. Les contours des titres sont ensuite extraits de l'image et le texte de chaque titre est extrait à l'aide de la reconnaissance optique de caractères (OCR).

Une fois que les titres ont été extraits et leurs positions identifiées, ils sont transformés en un ensemble de données structuré. Les étapes de prétraitement incluent l'encodage des variables catégorielles, la mise à l'échelle des coordonnées des titres et une réduction de dimension à l'aide de l'analyse en composantes principales (PCA). Ces étapes permettent de représenter les titres de manière numérique, prête pour le clustering.

Le modèle de clustering utilisé est K-means, qui a été préalablement entraîné. Les titres extraits sont assignés à des clusters prédéfinis en utilisant le modèle K-means. Les étiquettes prédites des clusters sont renvoyées, permettant ainsi de regrouper les titres similaires.

En résumé, ce code offre une méthode automatisée pour détecter et regrouper les titres présents dans une image, permettant une analyse rapide et structurée du contenu visuel.

4. Dataset

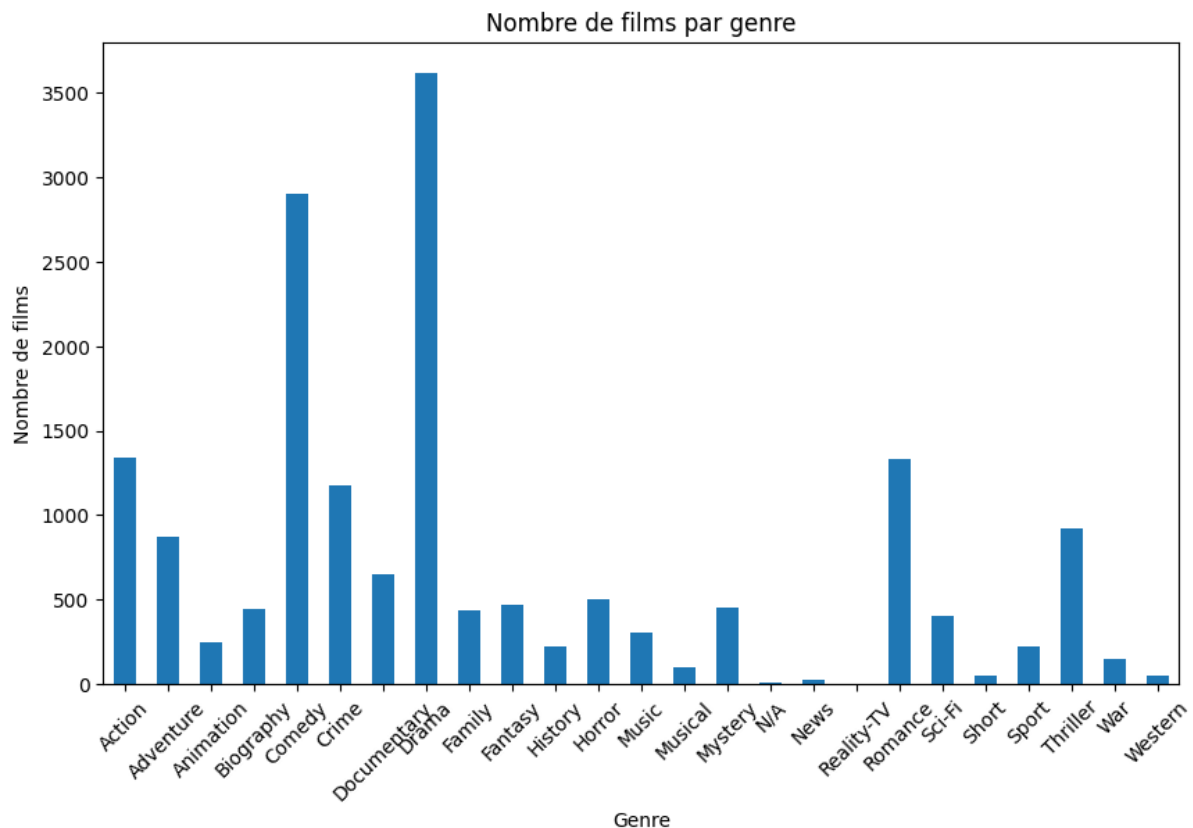
Une fois que toutes ces informations ont été relevées, nous pouvons les combiner à l'aide d'un processus de fusion pour chaque photo, afin d'obtenir une représentation complète. Cette fusion consolidera les caractéristiques visuelles, les couleurs dominantes, les années de production, les extraits de texte et les synonymes récupérés à partir des affiches.

Une méthode courante pour rassembler ces informations est de créer un fichier CSV (Comma-Separated Values). Cette approche facilite l'organisation et l'accès aux données pour une analyse ultérieure.

En assemblant toutes ces informations dans un fichier CSV, nous créons un ensemble de données complet et structuré qui servira de base pour l'entraînement et l'évaluation de notre modèle de prédiction du genre des films à partir des affiches.

5. Analyse des données

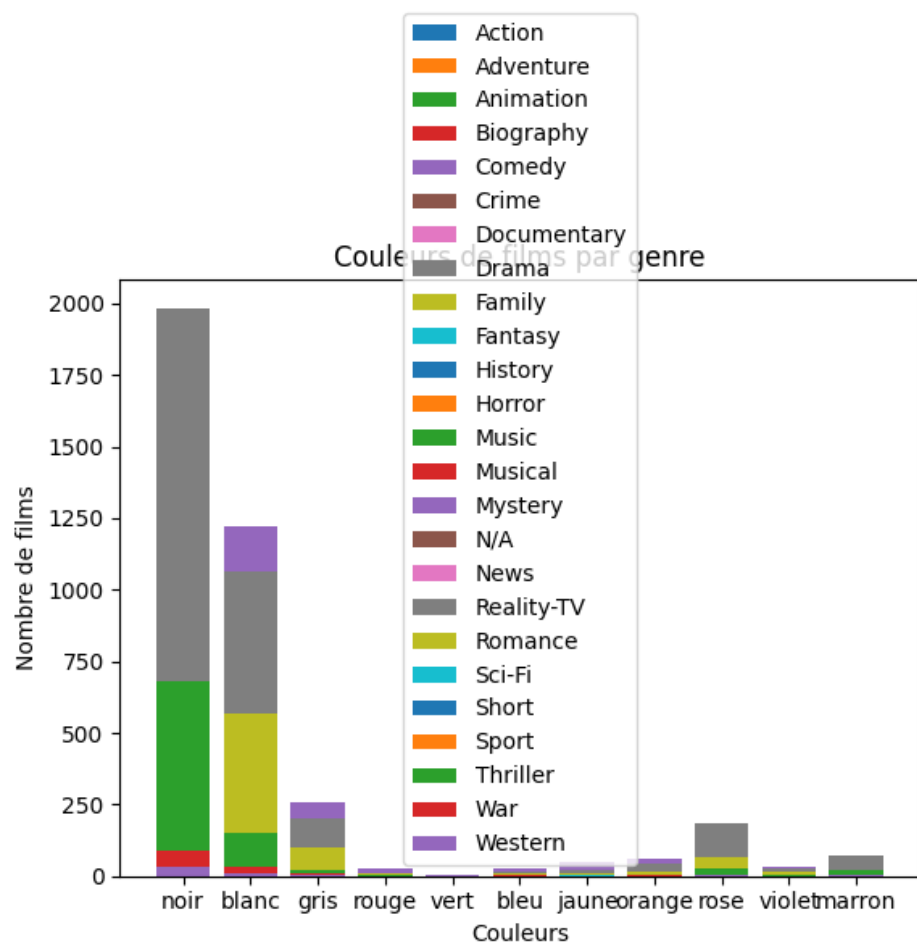
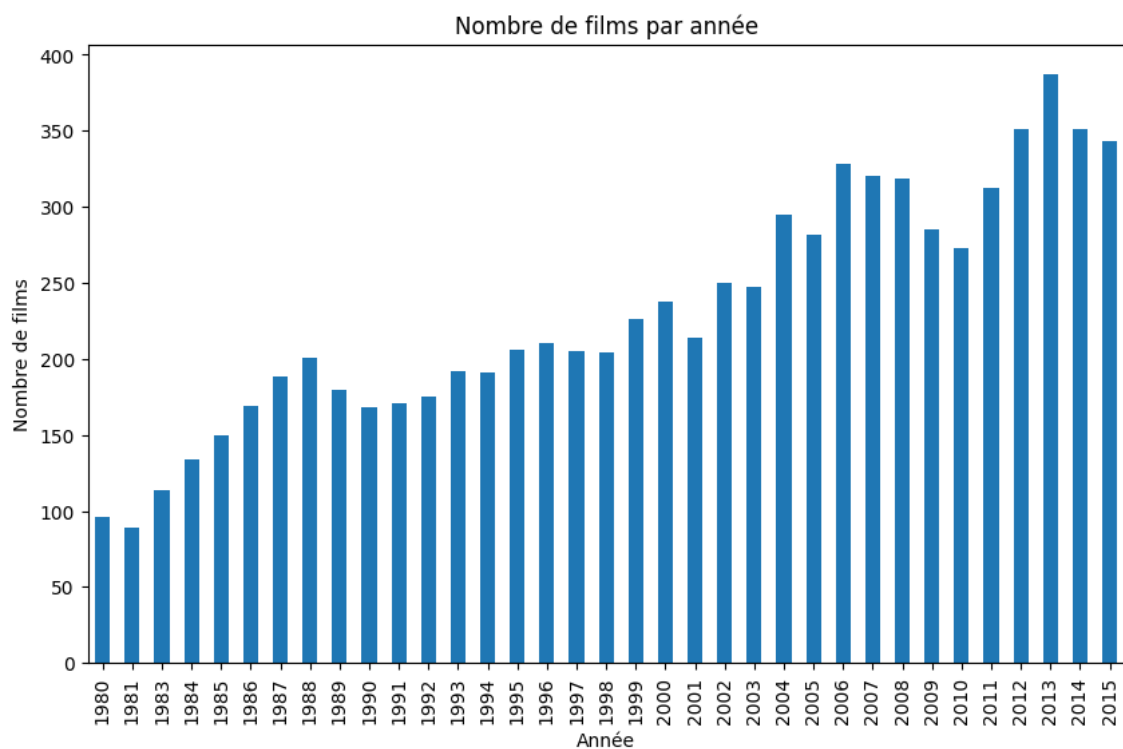
Nous avons effectué des statistiques avec les données qu'on a récupérées.



Une observation importante révèle que les genres "Documentary" et "Biography" contiennent le plus grand nombre d'images dans notre dataset. En revanche, les genres "New Reality-TV" et "Western" contiennent moins de 100 images chacun. Cette disparité dans la quantité de données pose la question suivante : devons-nous conserver ces genres dans notre dataset ?

La décision de garder ou non ces genres avec un nombre limité d'images dépend de plusieurs facteurs. Tout d'abord, il est important d'évaluer l'importance de ces genres dans notre analyse et dans le contexte de notre objectif. Si ces genres ont une signification particulière ou si leur inclusion est essentielle pour notre étude, nous pouvons envisager de les conserver malgré le nombre réduit d'images.

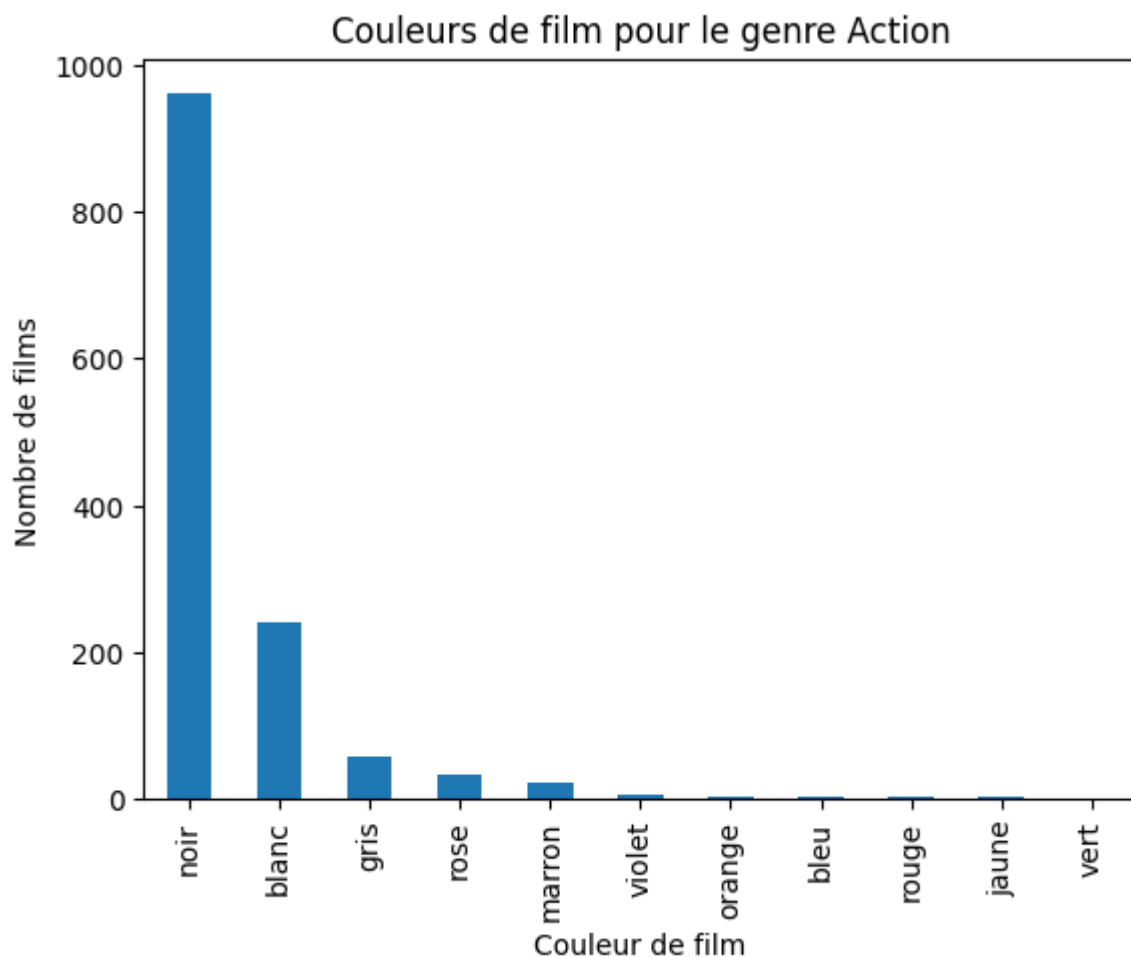
Cependant, il est également crucial de considérer la représentativité et la qualité des prédictions que notre modèle pourrait faire pour ces genres avec un nombre limité de données. Si la quantité insuffisante d'images entraîne une incertitude élevée ou une faible précision dans les prédictions pour ces genres, il serait alors préférable de les exclure temporairement de notre dataset.

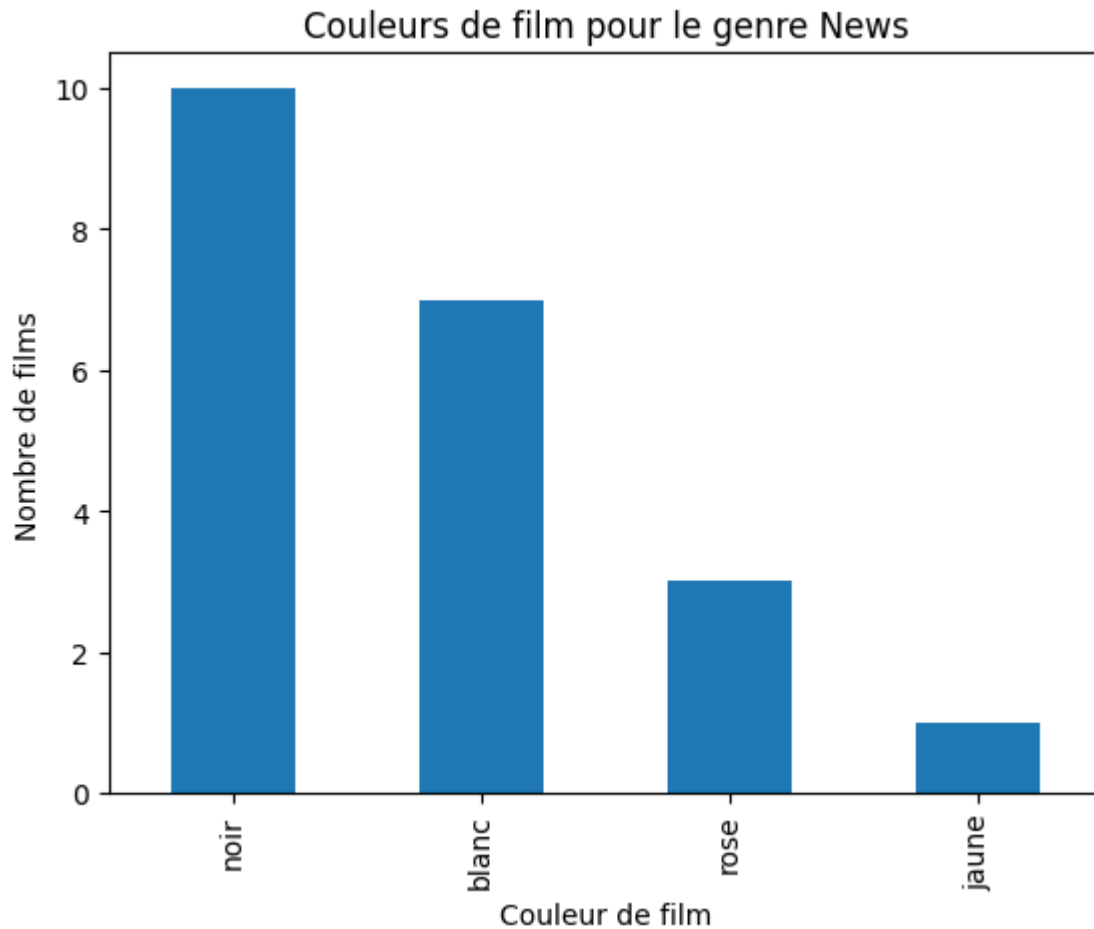


Dans ce graphique, nous pouvons observer en premier lieu que les images en noir et blanc sont les plus représentées. Cette tendance semble cohérente avec les genres de films qui ont traditionnellement utilisé ce type de palette visuelle. Par ailleurs, nous remarquons que les autres couleurs représentées ne dépassent pas la centaine d'images. Il est évident qu'il manque des données pour constituer un dataset cohérent dans ce contexte.

En poursuivant notre analyse, nous constatons que parmi les images au couleur noir, les genres les plus représentés sont le thriller et le drame. Cette observation est intéressante, car ces genres cinématographiques sont souvent associés à des tensions narratives.

De plus, nous remarquons que dans les images en couleur blanche, le genre de la romance est très présent. Cette constatation renforce la cohérence de nos analyses, car le choix d'une palette de couleurs plus claire et lumineuse peut évoquer des thèmes sentimentaux.





Face à ce nombre limité de résultats, nous avons pris la décision d'ajouter la deuxième couleur prédominante dans notre dataset afin d'enrichir les informations disponibles.

En effet, en considérant uniquement la couleur dominante, nous pourrions manquer des nuances et des variations significatives dans les affiches de films. En ajoutant la deuxième couleur prédominante, nous introduisons une dimension supplémentaire dans notre analyse, ce qui peut potentiellement améliorer la précision de nos prédictions.

6. Résultat

Dans un premier temps, nous avons collecté plusieurs données à partir des affiches de films, notamment la date de sortie, la couleur dominante, le nombre de visages, l'émotion principale (pour les images avec plusieurs visages, nous avons sélectionné l'émotion du visage le plus en avant) et le texte complet apparaissant sur l'image.

Concernant les labels (genres), nous avons une liste complète comprenant : Action, Adventure, Animation, Biography, Comedy, Crime, Documentary, Drama, Family, Fantasy, History, Horror, Music, Musical, Mystery, N/A, News, Reality-TV, Romance, Sci-Fi, Short, Sport, Thriller, War, Western.

Cependant, en utilisant ces données, nous avons obtenu une précision (accuracy) très basse de 0,13. Pour améliorer cela, nous avons apporté plusieurs modifications :

1. Suppression de la colonne "Date" car cette information n'est pas toujours présente sur les affiches et n'est donc pas pertinente pour la prédiction.
2. Ajout de nouvelles données : position du visage (pour les images avec plusieurs visages, nous avons utilisé la position du visage le plus en avant), position du titre, une deuxième émotion et une deuxième couleur (couleur secondaire).
3. Pour augmenter la précision, nous avons regroupé certains genres qui sont souvent associés. Ainsi, en sortie de notre modèle, nous avons les catégories suivantes : 'Comedy', 'Documentary', 'Horror', 'N/A', 'Short', 'Sport', 'Western', 'Action_Adventure_Sci-Fi', 'Animation_Family_Fantasy', 'Music_Musical', 'Crime_Mystery_Thriller', 'Drama_Romance', 'History_Biography_War', 'Reality_TV_News'.
4. Nous avons également testé différents types d'algorithmes de classification et avons finalement retenu RandomForestClassifier qui nous a donné la meilleure précision.

Grâce à ces ajustements, nous avons considérablement amélioré les performances de notre modèle, atteignant ainsi une précision (accuracy) de 0,33. Cette amélioration est significative par rapport à la précision initiale de 0,13 que nous avons obtenue.

Cette augmentation de la précision démontre l'efficacité des changements que nous avons apportés, tels que l'ajout de nouvelles caractéristiques et la sélection d'un meilleur algorithme de classification. En obtenant une précision de 0,33, notre modèle est capable de prédire avec plus de fiabilité le genre des films.

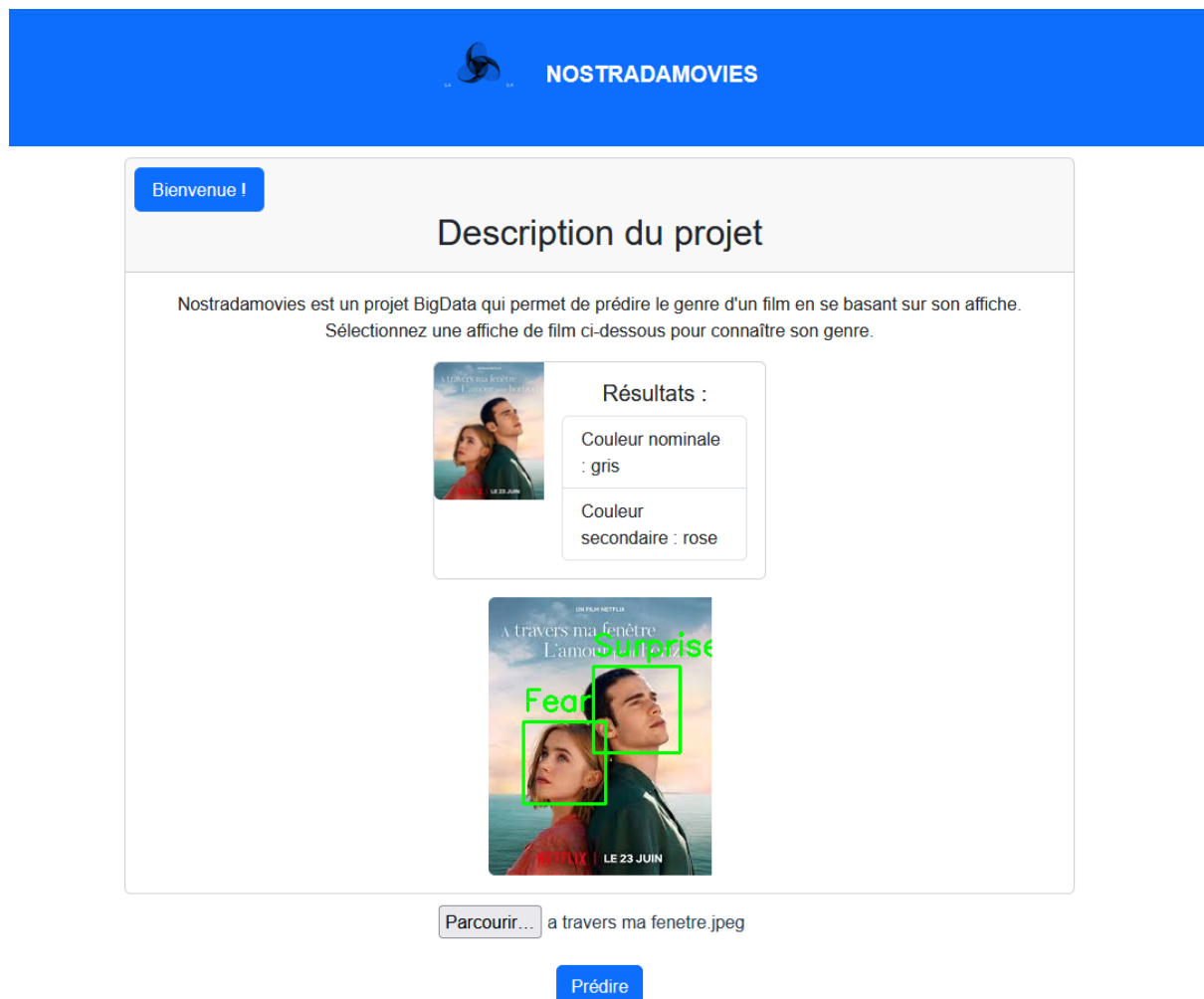
Ces résultats encourageants témoignent de l'importance de l'optimisation de notre approche, en tenant compte des informations pertinentes extraites des affiches et en

utilisant des techniques de classification appropriées. Nous pourrions continuer à travailler sur l'amélioration de notre modèle afin d'obtenir des résultats encore plus précis.

7. Interface graphique

Nous avons développé une interface graphique conviviale basée sur Vue.js. Cette interface permet à l'utilisateur de sélectionner une affiche de film à partir de son ordinateur et de demander une prédiction de genre. À l'écran, nous affichons des informations extraites de l'image, telles que le cadre entourant les visages (s'il y en a sur l'affiche) et les émotions détectées. Enfin, nous affichons le genre prédit par notre modèle.

Cette interface offre une expérience interactive où les utilisateurs peuvent facilement interagir avec notre système de prédiction de genre.



Conclusion :

En cette fin d'année, nous arrivons à la conclusion de notre projet de prédiction du genre des films à partir des affiches, et nous sommes fiers des accomplissements que nous avons réalisés. Nous avons parcouru un long chemin, en explorant différentes approches, en ajustant et en améliorant nos méthodes, et en finissant par obtenir des résultats prometteurs.

Ce projet nous a permis de plonger dans le monde fascinant de l'apprentissage automatique et de l'analyse d'images. Nous avons appris à exploiter les ressources disponibles, telles que les bibliothèques Python et les techniques de traitement d'image, pour extraire des informations précieuses des affiches de films. Cela nous a ouvert les yeux sur les multiples facettes que les affiches peuvent révéler, allant des couleurs et des visages aux émotions et au texte.

Au cours de notre parcours, nous avons également découvert l'importance d'un bon ensemble de données. Nous avons réalisé l'effort de collecter et d'étiqueter des affiches de films pour créer un dataset complet et représentatif. Cette étape cruciale nous a permis d'entraîner notre modèle et d'améliorer ses performances au fil du temps.

L'aboutissement de notre projet est la création d'une interface utilisateur conviviale, où les utilisateurs peuvent interagir avec notre modèle en téléchargeant une affiche et en recevant une prédiction de genre en retour. Cela ajoute une dimension pratique à notre travail, en le rendant accessible et utile pour les amateurs de cinéma.

En regardant en arrière, nous sommes conscients des défis auxquels nous avons été confrontés et des ajustements que nous avons dû apporter en cours de route. Cependant, ces défis ont également été des opportunités d'apprentissage et de croissance, nous permettant de mieux comprendre les subtilités de l'analyse d'images et de l'apprentissage automatique.

Alors que nous nous apprêtons à passer à une nouvelle année, nous gardons en tête les possibilités d'amélioration et de développement futur de notre projet. Nous envisageons d'explorer de nouvelles techniques, d'élargir notre ensemble de données et de perfectionner notre modèle pour obtenir des résultats encore plus précis et significatifs.

En fin de compte, ce projet nous a non seulement permis d'acquérir de nouvelles compétences techniques, mais il a également renforcé notre passion pour la créativité et l'innovation dans le domaine de l'intelligence artificielle. Nous sommes reconnaissants de cette expérience enrichissante et excitante, et nous sommes impatients de poursuivre notre exploration et notre apprentissage dans les projets futurs.

En cette période de transition, nous tenons à remercier tous ceux qui ont contribué à notre projet, nos collaborateurs, nos mentors et tous ceux qui nous ont soutenus tout au long de ce voyage. Votre soutien et votre engagement ont été inestimables.

Alors que nous nous tournons vers l'avenir, nous sommes animés par un sentiment d'optimisme et de détermination. Nous continuerons à explorer les possibilités offertes et à

repousser les limites de ce qui est possible. Car c'est grâce à de tels projets que nous contribuons à façonner le futur de la technologie et à créer un monde meilleur.

Au revoir à cette année passionnante, et bienvenue à de nouvelles aventures et de nouveaux défis passionnants dans les années à venir !