

PROJET : Système d'Aide à la Décision (SAD) pour le Diagnostic de Tumeurs Cérébrales

Module : Machine Learning (ALIF83)
2025 – 2026
Mohamed HAMIDI

1 Contexte Médical

Vous développez un **système d'aide à la décision (SAD)** pour un service de radiologie. Contrairement à un simple classificateur, votre système ne se contente pas de donner une étiquette, il doit agir comme un partenaire pour le médecin :

- **Évaluer le niveau de confiance** de chaque prédiction.
- **Recommander des actions** adaptées au degré de certitude.
- **Prioriser les cas urgents** nécessitant une expertise humaine immédiate.
- **Minimiser les faux négatifs** (risque vital critique).

Note : Ce projet simule un outil réel d'aide au radiologue et non un remplacement automatique.

2 Objectifs Pédagogiques

- Comprendre la différence fondamentale entre classification brute et aide à la décision.
- Implémenter un système à seuils de confiance multiples.
- Utiliser des métriques orientées "métier" (décision clinique).
- Créer un workflow de triage automatisé.

3 Données et Définitions

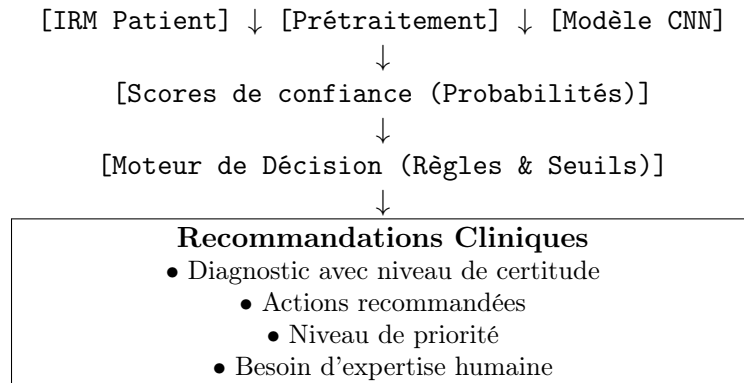
Dataset : Brain Tumor MRI Dataset (Kaggle).

Classes et Implications Cliniques :

1. **Gliome** (Tumeur maligne) → **URGENT**
2. **Méningiome** (Tumeur bénigne) → Surveillance
3. **Tumeur pituitaire** → Traitement spécialisé
4. **Pas de tumeur** → Rassurer le patient

4 Architecture du Système

Le flux de données du SAD est le suivant :



5 Cahier des Charges

Séance 1 : Exploration et Modèles de Base

Tâche 1 : Exploration et Prétraitement Identique à la version standard (chargement, redimensionnement, augmentation de données).

Tâche 2 : Régression Logistique avec Calibration

- Entraîner une régression logistique multinomiale.
- **Calibration** : Utiliser *Platt Scaling* ou *Isotonic Regression* pour obtenir des probabilités fiables.
- Analyser la distribution des scores : identifier les prédictions incertaines ($\text{max_prob} < 0.7$).

Tâche 3 : MLP avec Gestion de l'Incertainité

- Architecture MLP avec sortie probabiliste.
- **Objectif** : Détecter les limitations du modèle.

Séance 2 : CNN et Système de Décision

Tâche 4 : CNN Optimisé pour la Décision Architecture CNN classique, mais avec ajout de **Temperature Scaling** en sortie pour calibrer les probabilités finales. Sauvegarder les activations pour analyse future.

Tâche 5 : Moteur de Décision Clinique Implémenter les règles métiers suivantes :

```
1 # Seuils de confiance
2 SEUIL_HAUTE_CONFIANCE = 0.85
3 SEUIL_CONFIANCE_MOYENNE = 0.65
4 SEUIL_CONFIANCE_FAIBLE = 0.50
5
6 # Logique de decision
7 if max_probabilite >= SEUIL_HAUTE_CONFIANCE:
8     decision = "Diagnostic automatique valide"
9     action = "Rapport envoye au medecin traitant"
10    priorite = determiner_urgence(classe_predite)
11 elif max_probabilite >= SEUIL_CONFIANCE_MOYENNE:
12     decision = "Diagnostic probable - Revision recommandee"
13     action = "Validation par radiologue junior"
```

```

14     priorite = "Normale (48h)"
15 elif max_probabilite >= SEUIL_CONFIANCE_FAIBLE:
16     decision = "Cas incertain"
17     action = "Revision par radiologue senior"
18     priorite = "Elevee (24h)"
19 else:
20     decision = "Incertitude elevee"
21     action = "Double lecture obligatoire + IRM complementaire"
22     priorite = "Urgente (12h)"

```

Listing 1 – Logique de Décision et Seuils

```

1 # Tolerance asymetrique pour minimiser faux negatifs
2 if classe_predite == "Pas de tumeur":
3     if max_probabilite < 0.95: # Seuil tres eleve exige
4         action = "Verification obligatoire (risque faux negatif)"

```

Listing 2 – Gestion des Faux Négatifs (Sécurité)

Tâche 6 : Tableau de Bord de Décision Générer un rapport textuel automatisé pour chaque patient :

Rapport Automatisé

```

=====
RAPPORT D'AIDE A LA DECISION
=====
Patient ID: P_12345          Date: 01/02/2026

PREDICTION PRINCIPALE
-----
Classe: Gliome
Confiance: 91.3%
Niveau de certitude: ELEVE [OK]

SCORES PAR CLASSE
-----
• Gliome: 91.3%
• Meningiome: 5.2%
• Tumeur pituitaire: 2.1%
• Pas de tumeur: 1.4%

RECOMMANDATIONS CLINIQUES
-----
Diagnostic: Gliome detecte (haute confiance)
Action: Referer immediatement en oncologie
Priorite: [!] URGENT - Prise en charge sous 12h
Revision humaine: Optionnelle (validation finale)

ELEMENTS D'ATTENTION
-----
• Tumeur maligne suspectee
• IRM de controle recommandee
=====

```

Tâche 7 : Analyse de Performance du SAD Calculer les métriques orientées "métier" :

- **Taux de couverture automatique** : % de cas gérés sans intervention humaine.
- **Accuracy par tranche de confiance** : L'accuracy doit être $> 95\%$ quand la confiance est > 0.85 .
- **Analyse Coût-Bénéfice** :

$$Cot_{total} = (FN \times 1000) + (FP \times 100) + (Revision \times 50)$$

6 Livrables Attendus

6.1 Notebook Jupyter

Le notebook doit contenir :

1. Introduction au SAD (Classification vs Décision).
2. Modèles calibrés (RegLog, MLP, CNN).
3. **Moteur de décision** : Implémentation des règles.
4. **Simulation** : 20 exemples de rapports générés.
5. Analyse critique et éthique.

6.2 Fonctions Python

Vous devez implémenter les fonctions suivantes :

```
1 def predire_avec_confiance(image, model):
2     """Retourne prediction + scores de confiance"""
3     //code
4
5 def generer_recommandation(probabilites, seuils):
6     """Applique les regles de decision"""
7     //code
8
9 def calculer_incertitude_mc_dropout(image, model, n_iter=20):
10    """Estime l'incertitude via Monte Carlo Dropout"""
11    //code
12
13 def creer_rapport_decision(patient_id, prediction, confiance):
14    """Genere le rapport formate"""
15    //code
```

Bonne Chance !!