

Annexe 6. Informations descriptives spécifiques à fournir pour les fonctionnalités du dispositif médical s'appuyant sur des procédés d'apprentissage automatique (technologies relevant du champ de l'intelligence artificielle)

Remarques préliminaires

Dès lors que votre DM s'appuie sur au moins un procédé d'apprentissage automatique, vous devrez remplir cette grille pour apporter aux membres de la commission les informations qui leur sont nécessaires sur ce volet de votre DM. Intégrée au guide de dépôt en septembre 2020, elle sera modifiée autant que de besoin pour s'adapter aux évolutions technologiques.

Selon les cas, vous construirez une ou plusieurs grilles, le principe étant que vous renseigniez une grille pour chaque fonctionnalité « intelligente » du dispositif :

- dans le cas où il n'y a qu'une fonctionnalité s'appuyant sur des procédés d'apprentissage automatique : vous remplirez une seule grille. C'est notamment le cas lorsque l'intrication, ou la succession, de plusieurs procédés peut justifier leur regroupement au sein d'une même grille lorsqu'ils concourent à une même fonctionnalité « intelligente ».
- dans le cas d'un DM regroupant plusieurs fonctionnalités de ce type, vous renseignerez une grille par fonctionnalité.

Selon le type de technologie, certains items peuvent ne pas être adaptés. Dans ces cas, vous l'indiquerez, en apportant une justification. A l'inverse, vous pouvez également compléter les informations descriptives listées par toute information que vous jugeriez utile.

Grille descriptive

		Aide au remplissage
Finalité d'usage		
1	Rappeler l'usage revendiqué et le domaine d'application prévu du dispositif médical (DM) intégrant un ou plusieurs algorithmes d'apprentissage automatique	S'agit-il, par exemple : <ul style="list-style-type: none">– d'aider le patient à adapter la posologie de son traitement ?– de prédire ou détecter précocement la survenue d'un événement clinique ? Vous préciserez les pathologies ou situations cliniques visées, ou le caractère multidisciplinaire du DM, le cas échéant. Vous préciserez systématiquement l'utilisateur (patient ou professionnel).
2	Préciser l'intérêt des informations fournies ou des décisions prises par des procédés d'apprentissage automatique	On précise ici la fonctionnalité « intelligente » à laquelle l'apprentissage automatique a directement contribué. Par exemple : <ul style="list-style-type: none">– Détermination d'un score de gravité ?– Calcul d'une dose pour une adaptation de traitement ?

3	Rappeler les caractéristiques de la population visée et, s'il y a lieu, les caractéristiques pour lesquelles l'utilisation du DM n'est pas appropriée, du fait d'une non-indication, d'une contre-indication ou de facteurs influençant le résultat produit	Celles-ci peuvent être : <ul style="list-style-type: none"> – Démographiques (tranches d'âges, sexe...) – Physio-pathologiques (grossesse, personnes diabétiques ou asthmatiques, etc.) ou morphologiques (personnes amputées du membre inférieur, etc.) – Cliniques ou biologiques (stade de la maladie, etc.)
4	Décrire l'environnement de fonctionnement du système intelligent	Préciser notamment les conditions environnementales (météorologiques, luminosité, température, nature du terrain, etc.) permettant de caractériser le domaine de fonctionnement.

Données		
Description des échantillons utilisés pour l'apprentissage initial ou le réapprentissage du modèle		
5	Préciser les caractéristiques de la population dont les données d'apprentissage initial ou de réapprentissage du modèle sont extraites	Celles-ci peuvent être : <ul style="list-style-type: none"> Démographiques (tranches d'âges, sexe...) Physio-pathologiques (grossesse, personnes diabétiques ou asthmatiques, etc.) ou morphologiques (personnes amputées du membre inférieur, etc.) Cliniques ou biologiques (stade de la maladie, etc.) Distinguer la population à partir de laquelle les données d'apprentissage initial sont produites (entraînement, validation et test) de celle utilisée lors de la phase de réapprentissage (réentraînement, validation et test du système mis à jour), le cas échéant.
6	Préciser les caractéristiques de chaque échantillon utilisé pour l'apprentissage initial ou le réapprentissage du modèle	Sont attendues : leur fonction, leur taille et leur composition. Les variables incluses doivent être citées. La manière dont sont pris en compte les événements rares doit être décrite. Distinguer les bases de données des phases d'apprentissage initial (entraînement, validation et test) et en phase de réapprentissage (réentraînement, validation et test du système mis à jour), le cas échéant
7	Préciser la méthodologie de séparation ou de segmentation des échantillons	Par exemple, préciser les modalités de séparation (méthodes utilisées et proportions) et de segmentation (aléatoire, par date, par individu, etc.) des jeux de données d'entraînement, de validation et de test Distinguer les bases de données en phases d'apprentissage et de réapprentissage, le cas échéant.

Description des données d'entrée impliquées dans l'apprentissage initial ou le réapprentissage du modèle		
8	Préciser les caractéristiques des variables (type de variable, distribution...)	Distinguer les corpus d'entraînement, de validation et de test le cas échéant.
9	Indiquer le mode d'acquisition des variables et leur origine lors du processus d'apprentissage	Par exemple, une variable a-t-elle été saisie par un patient ? Provient-elle d'un capteur ? A-t-elle été générée à partir de modèles de patients virtuels ? Préciser si les variables ont été extraites de corpus de données ouverts ou achetés et indiquer lesquels, le cas échéant, ainsi que leur caractère pérenne ou non. Préciser les types de capteurs utilisés lors de l'acquisition des variables, le cas échéant.
10	Décrire les prétraitements appliqués aux données	Par exemple, les actions de nettoyage des données, de transformation, de réduction, d'augmentation (ajouts de bruits artificiels, de perturbations artificielles simulant des variations météorologiques ou des défauts capteurs, etc.) Préciser les données concernées et la proportion des données modifiées par ces prétraitements
11	Indiquer la proportion des données manquantes au sein des données brutes et décrire leur gestion.	Préciser les types de données manquantes (aléatoires ou prévisibles).
12	Expliquer les procédures mises en place pour détecter et gérer les données aberrantes, le cas échéant	En particulier, préciser la manière dont sont distinguées les données aberrantes (ex : données physiologiquement impossibles) des valeurs atypiques (ex : événements rares)
13	Justifier de la représentativité des échantillons utilisés pour l'apprentissage initial (entraînement, validation et test) de l'algorithme par rapport aux données auxquelles cet algorithme sera exposé une fois déployé	Une justification des critères de représentativité est attendue. Préciser notamment les outils et méthodes utilisés pour vérifier la représentativité des échantillons et détecter les biais potentiels. En cas d'apprentissage incrémental ou continu, indiquer l'impact potentiel des mises à jour sur l'ensemble des données d'apprentissage.
Description des données d'entrée impliquées dans la décision (une fois le dispositif médical déployé)		
14	Préciser les caractéristiques des variables (type, distribution...)	Indiquer les sources principales de différence entre les données d'entraînement, de validation, de test, et les données impliquées dans la prise de décision, une fois le système déployé (capteurs différents, conditions environnementales différentes, etc.).

15	Indiquer le mode d'acquisition des variables et leur origine	Par exemple, une variable a-t-elle été saisie par un patient ? Provient-elle d'un capteur ? Indiquer les réglages de plage de mesure et de sensibilité des instruments de mesure, le cas échéant.
16	Décrire les prétraitements appliqués aux données utilisées pour la prise de décision	Par exemple, les actions de nettoyage des données, de transformation, de réduction, etc.
17	Lister les variables de sortie (objets de la prédiction du modèle) et leurs caractéristiques (type, unité...)	Préciser les variables qui seront exploitées au regard de l'objectif. Préciser si elles sont exploitées par un autre composant du DM ou si elles sont communiquées à l'utilisateur (si tel est le cas de quelle manière)

Modèle : description de l'entraînement, de la validation et du test, avant et après le déploiement du DM

18	Décrire le type d'apprentissage utilisé	<p>S'agit-il d'un apprentissage automatique :</p> <ul style="list-style-type: none"> – supervisé – semi-supervisé, – non supervisé, – par renforcement, – fédéré, – centralisé, – autre ? <p>Ces propositions ne sont pas exclusives les unes des autres.</p>
19	Décrire le type de tâche automatisée par l'algorithme	<p>classification supervisée (déterminer des critères de classement),</p> <p>classification non supervisée (définir des classes),</p> <p>classement (ranger dans des classes),</p> <p>régression (prévision quantitative),</p> <p>segmentation,</p> <p>autre ?</p>
20	Préciser la fréquence de mise à jour	<p>L'apprentissage est-il :</p> <ul style="list-style-type: none"> – continu (système apprenant de manière autonome après son déploiement) ? – initial (algorithme conçu par apprentissage puis figé après le déploiement du DM) ? – ou incrémental (algorithme dont la mise à jour de la structure et/ou des paramètres après le

		déploiement du DM est supervisée par un humain et implique une validation a priori et/ou a posteriori) ?
21	Décrire les critères de sélection du modèle	<p>Par exemple, le taux d'erreur, le temps de calcul, le nombre et la nature des données disponibles, caractère explicable ou embarquable, etc.</p> <p>Ne pas rentrer dans le détail des données d'entrée du système (elles font l'objet des questions 5 à 17), ni des méthodes de test utilisées (elles font l'objet des questions 26 à 32)</p>
22	Décrire les différentes phases d'entraînement, de validation et de test, avant déploiement du DM	<p>Indiquer les différentes phases d'entraînement, de validation et de test en précisant notamment si elles s'appuient sur des données individuelles ou collectives.</p> <p>Ne pas rentrer dans le détail des méthodes de test mises en place (elles font l'objet des questions 26 à 32).</p>
23	Décrire les stratégies d'entraînement, de validation et de test des mises à jour, le cas échéant	<p>Indiquer les différentes phases de réentraînement, de validation et de test mises en œuvre une fois le DM déployé, en précisant notamment si elles s'appuient sur des données individuelles ou collectives.</p> <p>Préciser en particulier la fréquence de réentraînement, les variables impliquées et la période de prise en compte des données, le lieu de réalisation des calculs de réentraînement (en local sur le DM ou sur serveur).</p> <p>Ne pas rentrer dans le détail de la supervision et/ou intervention humaine dans ces phases (elles font déjà l'objet des questions 24 et 25) ni des méthodes de test des mises à jour (elles font déjà l'objet des questions 26 à 32).</p>
24	Décrire la manière dont les personnes intervenant dans le développement du système sont référencées	Préciser si les responsables humains ou raisons sociales impliqués à chaque étape du cycle de vie du DM intelligent (recueil des données, développement, qualification, utilisation et rétrocontrôle des DM dotés d'IA) peuvent être identifiés.
25	Indiquer, le cas échéant, dans quels cas un humain intervient dans le processus de réapprentissage	Par exemple, en cas d'apprentissage actif, préciser la fréquence et la qualification de l'intervenant. En cas d'annotation réalisée par un opérateur, préciser la qualification et le rôle de ce dernier.

Caractéristiques fonctionnelles		
Performance et qualification		
26	Décrire et justifier le choix des métriques utilisées pour la mesure de performance ...	Par exemple : Root-mean-square deviation, Area Under Curve, F1-score, ZoneMap, Jaccard
27	Décrire les traitements opérés qui ont eu un impact substantiel sur la performance	Par exemple, en cas de déséquilibre des classes dans le cadre d'une classification supervisée, indiquer si un rééquilibrage de classes a été réalisé, ainsi que la méthode utilisée.
28	Décrire les risques identifiés de sur et sous apprentissage et les méthodes mises en place pour y remédier	Un lien pourra notamment être établi avec les réponses apportées à la question 7 sur la séparation/segmentation des données.
29	Préciser si le système renvoie un niveau de confiance pour chacune de ses décisions	Il pourrait par exemple s'agir d'indiquer, pour un classifieur d'images, s'il renvoie les probabilités pour l'image d'entrée d'appartenir à chacune des classes
30	Décrire les méthodes de qualification du système d'apprentissage automatique	Préciser notamment le protocole de test mis en place et les procédures utilisées pour s'assurer de la répétabilité des mesures de performance et de la reproductibilité des expérimentations. En cas d'utilisation de méthodes formelles pour qualifier le système d'apprentissage automatique, justifier le choix des méthodes utilisées et la manière de définir les domaines sur lesquels les méthodes formelles ont été appliqués.
31	Indiquer les résultats des mesures de performance sur les différents jeux de données	Par exemple, les taux d'erreurs fournis par les métriques sur les bases d'entraînement, de validation et de test, selon la répartition réalisée Préciser si une base de données indépendante de celles d'entraînement, de validation et de test a été utilisée pour qualifier le modèle. Préciser, dans le cas d'analyse par preuves formelles, les résultats obtenus et le domaine de validité de ces résultats.
32	Préciser les seuils de performance choisis (valeurs limites, taux d'erreur maximal...) et expliquer le choix de ces seuils	

Robustesse du système		
33	Préciser les outils mis en place pour générer des exemples antagonistes en phase d'évaluation des performances et de qualification	
34	Préciser les outils mis en place pour surveiller les performances du système intelligent après son déploiement	Préciser notamment les mécanismes mis en place pour mesurer la dégradation et/ou la dérive du modèle (campagnes d'évaluation régulières, etc.), ainsi que les moyens et protocoles de traçage, d'archivage et d'analyse des dégradations de performance
35	Préciser les seuils choisis (valeurs limites, taux d'erreur maximal...) pour le suivi de la dégradation et/ou de la dérive du modèle et expliquer le choix de ces seuils	
36	Préciser les mesures mises en place en cas de détection automatique ou par l'utilisateur d'une dégradation ou d'une dérive du modèle	Par exemple : informations transmises à l'utilisateur, substitution de l'algorithme apprenant par un système expert, réentraînement...
Résilience du système		
37	Décrire le système mis en place pour la détection d'anomalie des données d'entrée en utilisation opérationnelle	Cela pourra par exemple concerner la détection des données en dehors du domaine de fonctionnement nominal du système intelligent
38	Décrire les impacts cliniques et techniques potentiels induits par des anomalies sur les données d'entrée du système d'apprentissage automatique	Par exemple, que se passera-t-il : En cas de non-correction des valeurs aberrantes ? En cas d'anomalie de saisie d'une valeur déclarative par le patient ? Du fait du niveau d'incertitude associé aux données d'entrée (données physiologiques, environnementales...) ? En cas d'indisponibilité d'une donnée ? En cas de perte d'intégrité de la donnée ?
39	Préciser les mesures mises en place en cas de détection automatique ou par l'utilisateur d'une anomalie (par exemple d'un dysfonctionnement endommageant les données d'entrée)	Par exemple : informations transmises à l'utilisateur, mode dégradé, substitution de l'algorithme apprenant par un système expert, intervention du clinicien ou du technicien, etc.

Explicabilité et interprétabilité		
40	Indiquer les éléments d'explicabilité mis à disposition par le dispositif intelligent	Préciser, le cas échéant, la ou les techniques d'explicabilité mises en place afin de permettre de comprendre les principaux facteurs ayant conduit à la décision prise ou proposée par l'algorithme d'apprentissage automatique. Préciser le destinataire de ces explications : utilisateur (soignant ou patient), développeur, etc. Indiquer également si les explications sont enregistrées pour analyse a posteriori par des experts (utilisateurs et/ou développeurs).
41	Indiquer les éléments d'interprétabilité, c'est-à-dire les paramètres (variables d'entrée, pondérations, etc.) influant sur la prise de décision, ainsi que la méthode utilisée pour les identifier	Pour les algorithmes ayant eu un apprentissage initial ou incrémental, ces paramètres sont-ils identifiés (par exemple au moyen de fonctions d'influence) ?
42	Préciser si les décisions et actions du dispositif intelligent sont confrontées aux recommandations professionnelles	Indiquer notamment si une confrontation des sorties de l'algorithme d'apprentissage automatique avec les recommandations professionnelles est réalisée en temps réel ou a posteriori. Préciser si ces comparaisons sont rendues accessibles aux utilisateurs. Par exemple, les sorties de l'algorithme d'apprentissage automatique sont-elles confrontées à celles d'un système expert modélisant des recommandations de prise en charge ?

Glossaire

Ce glossaire est uniquement destiné à accompagner cette grille descriptive des algorithmes d'apprentissage automatique dans le contexte de l'évaluation de dispositifs médicaux par la CNEDiMTS.

Terme	Définition	Source
Apprentissage automatique	Processus par lequel un algorithme évalue et améliore ses performances sans l'intervention d'un programmeur, en répétant son exécution sur des jeux de données jusqu'à obtenir, de manière régulière, des résultats pertinents.	15
Apprentissage non supervisé	Apprentissage automatique dans lequel l'algorithme utilise un jeu de données brutes et obtient un résultat en se fondant sur la détection de similarités entre certaines de ces données.	15
Apprentissage supervisé	Apprentissage automatique dans lequel l'algorithme s'entraîne à une tâche déterminée en utilisant un jeu de données assorties chacune d'une annotation indiquant le résultat attendu	15
Classement	Action de classer des objets, des personnes selon un certain ordre.	16
Classification supervisée	Technique qui consiste à regrouper des données en fonction de leur proximité permettant ainsi de différencier deux classes discrètes ou plus.	17
Dérive du modèle	Un algorithme d'IA dont les paramètres sont figés devient inconsistant avec son environnement si celui-ci a évolué.	18
Domaine d'emploi	Description de l'environnement et de la population visée, pour lesquels l'algorithme ou le programme est conçu.	-
Donnée	Représentation de l'observation d'une variable sur un élément, un individu ou une instance d'une population, destinée à faciliter son traitement.	-
Donnée brute	Donnée n'ayant subi aucune transformation depuis son observation initiale.	-
Donnée d'entrée	Donnée utilisée pour l'apprentissage ou la prise de décision du modèle.	-
Donnée de sortie	Valeur représentant tout ou partie de la décision prise par l'algorithme à partir des données d'entrée.	-

¹⁵ Journal officiel du 09/12/2018

¹⁶ <https://www.larousse.fr/dictionnaires/francais/classement/16405>

¹⁷ D'après définition ISO (en cours d'élaboration)

¹⁸ Tsymbal, A. (2004). The problem of concept drift: definitions and related work. Computer Science Department, Trinity College Dublin, 106(2), 58.

Terme	Définition	Source
Échantillon	Fraction représentative d'une population ou d'un univers statistique	19
Entraînement	Processus de l'apprentissage automatique pendant lequel le système d'intelligence artificielle construit un modèle à partir de données.	17, 20
Exemple antagoniste	Cas limite mettant en difficulté le système évalué.	-
Explicabilité	Capacité de mettre en relation et de rendre compréhensible les éléments pris en compte par l'algorithme, par exemple les variables d'entrée, et leurs conséquences, par exemple, sur la prévision d'un score, et ainsi sur la décision. Les explications doivent être adaptées au niveau de compréhension de la personne auxquelles elles sont destinées.	-
Hyperparamètre	Paramètre qui est réglé pendant les exécutions successives de l'entraînement du modèle afin notamment de contrôler le sous ou sur-apprentissage.	21
Information	Élément de connaissance traduit par un ensemble de données selon un code déterminé, en vue d'être conservé, traité ou communiqué. Une information est issue de l'interprétation d'une ou plusieurs données mises en commun.	22
Interprétabilité	Capacité de rendre compréhensible le fonctionnement d'un système d'intelligence artificielle. Un algorithme est « interprétable » lorsqu'on comprend précisément son fonctionnement, par exemple, lorsqu'un système expert modélise un arbre décisionnel.	17
Jeu de données	Groupe de données	-
Modèle	Construction mathématique générant une inférence ou une prédiction à partir de données d'entrée.	17
Paramètre	Coefficient d'un modèle que le système d'apprentissage automatique estime ou entraîne tout seul et qui a un impact sur les données de sortie.	21
Résilience	Capacité du système à maintenir sa conformité à des exigences de performance et/ou de sécurité en présence de données d'entrée extérieures à son domaine d'emploi (p. ex. en raison d'un défaut capteur).	-

¹⁹ Centre National de Ressources Textuelles et Lexicales www.cntrl.fr

²⁰ D'après la Déclaration de Montréal pour un développement responsable de l'intelligence artificielle

²¹ <https://developers.google.com/machine-learning/glossary>

²² <https://www.dictionnaire-academie.fr/article/A9I1218>

Terme	Définition	Source
Robustesse	Capacité d'un système à maintenir son niveau de performance quelles que soient les circonstances.	17
Segmentation « Segmentation des données » « Tâche de segmentation automatique »	Segmentation des données : division d'un corpus de données en plusieurs bases (par exemple d'entraînement, de validation et de test), soit à partir de critères objectifs (date, âge, etc.) soit de manière aléatoire. Tâche de segmentation automatique : extraction et reconnaissance automatique de zones d'intérêt à partir de données d'entrée (p. ex. une image).	23
Test	Processus consistant à rechercher des erreurs liées à l'exécution d'un algorithme ou d'un programme en s'appuyant sur des jeux de données d'entrée n'ayant pas été utilisés lors de la phase d'entraînement.	-
Validation	Processus consistant à expérimenter, observer et optimiser (hyperparamètres) le comportement du système lors de son exécution afin de s'assurer, dans le domaine d'emploi, de l'adéquation des données de sortie avec les résultats attendus.	17
Variable	Caractéristique (qualitative ou quantitative) observable d'un élément.	-

²³ Rakoto–Ravalontsalama, M. (1990). Méthodes de segmentation automatique d'image. Analyse quantitative des formes, Télé-détection, pp251-260.