



الجمهورية الجزائرية الديمقراطية الشعبية

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

وزارة التعليم العالي والبحث العلمي

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

المدرسة الوطنية العليا للتكنولوجيا والهندسة - عنابة

ECOLE NATIONALE SUPERIEURE DE TECHNOLOGIE ET D'INGENIERIE - ANNABA

Département d'Électronique, d'Électrotechnique et d'Automatique.

## MEMOIRE

En vue d'obtention du diplôme de MASTER

Domaine : Science et Technologie

Filière : Electrotechnique

Spécialité : Réseaux électrique

Présenté par

**SEMARI DOUHA**

**GHEDJATI NOUHA**

Évaluation des méthodes de détection d'événements et de  
Classification pour l'identification des appareils dans les systèmes NILM

Encadrant :

**Dr. DEBBAH.A**

Co-Encadrant:

**Pr. DOGHMANE. N**

Membres du jury :

- |                       |           |              |
|-----------------------|-----------|--------------|
| ▪ Dr. Faouzi BOUCHARB | Président | ENSTI Annaba |
| ▪ Dr. Adel. DJELLAL   | Examineur | ENSTI Annaba |
| ▪ Dr. Hichem .MAYACHE | Examineur | ENSTI Annaba |

Année 2025

## Dédicace :

---

إلى أمي الغالية، التي غرست في نفسي حب العلم وقيم الصبر والاجتهاد، إلى من ساندتني بدعائها واحتوائها وتشجيعها دون كلل، أهديك هذا العمل عربون محبة وامتنان، وتقديرًا لِعطاء لا يُقاس.

إلى أسرتي الكريمة، التي وقّرت لي كل سبل الدعم والاستقرار، وكانت السند الحقيقي في كل المراحل، لكم أرفع أسمى عبارات الشكر والتقدير.

إلى زميلتي الفاضلة، شريكة هذا المشروع، التي تميزت بروح التعاون والانضباط والمسؤولية، أُعبر لك عن عميق امتناني لتفانيك.

وإلى صديقتي العزيزات، اللاتي كنّ دائمًا مصدرًا للتشجيع والدعم النفسي والمعنوي، لكنّ كل التقدير والمحبة.

Douha

## Dédicace :

---

الحمد لله الذي وفقنا لإنجاز هذا العمل، وله الفضل أولاً وأخيراً.  
إلى والديّ الكريمين، اللذين مهّدا لي طريق العلم بتضحياتهما، ومنحاني الثقة لأصل إلى هذا اليوم.  
إلى إخوتي وأختي، سند العائلة ورفاق الدرب، لمساندتهم الدائمة دعمهم المعنوي الذي كان له أبلغ الأثر.  
إلى رفيقتي في هذه الرحلة التي كانت دعماً في لحظات الشك واليأس.  
إلى صديقتي وزميلاتي، اللاتي شاركنني رحلة العلم بكل ما فيها من تعبٍ وفرحٍ وأمل.  
أهديكم ثمرة هذا الجهد، عربون وفاء وامتنان.

Nouha

## Remerciement :

---

*Par la grâce d'Allah, qui nous a donné la force, la patience et la persévérance nécessaires, nous arrivons aujourd'hui au terme de ce travail.*

*Au moment de conclure ce mémoire, nous souhaitons adresser nos plus chaleureux remerciements à celles et ceux sans qui ce travail n'aurait pu voir le jour.*

*Nos remerciements les plus sincères s'adressent à notre encadrant, Dr. DEBBAH, pour sa rigueur intellectuelle, sa patience remarquable et la pertinence de ses conseils tout au long de ce travail.*

*Nous tenons également à remercier notre co-encadrant, Pr. DOGHMANE, pour son accompagnement bienveillant et son soutien précieux durant la réalisation de ce mémoire.*

*Notre reconnaissance s'étend également à l'ensemble de l'équipe pédagogique de l'ENSTI, ainsi qu'aux intervenants professionnels, pour la qualité de l'enseignement théorique et pratique dispensé tout au long de notre formation.*

*Enfin, nous ne saurions conclure sans exprimer notre profonde gratitude envers nos familles, et tout particulièrement nos parents. Leur amour, leur soutien inconditionnel et leurs encouragements constants ont été notre force motrice face aux défis de ce parcours.*

*À toutes et à tous, nous exprimons notre plus sincère reconnaissance et notre gratitude la plus profonde.*

## الملخص

يستكشف مشروع التخرج هذا سُبل تحسين أنظمة المراقبة غير الاقتحامية لاستهلاك الكهرباء (NILM) بهدف تجزئة استهلاك الطاقة داخل المنازل، وذلك لتشجيع الحفاظ على الطاقة في مواجهة التحديات البيئية والاقتصادية. يركز المشروع على التقييم المقارن لمجموعات مختلفة من خوارزميات كشف الأحداث (العتبة، TFED CUSUM)، وطرق التصنيف (CNN، SVM، XGBoost، RF) لتحديد الأجهزة الكهربائية انطلاقاً من إشارة استهلاك إجمالية. كشفت الدراسة، التي أجريت على مجموعة بيانات ENERTALK، أن الخوارزميات القائمة على أشجار القرار، وخاصة XGBoost و RF، عند دمجها مع طرق الكشف CUSUM و TFED، تتفوق باستمرار على المنهجيات الأخرى من حيث مقاييس الدقة (Precision)، والاستدعاء (Recall)، ومقياس F1، والصحة (Accuracy)، مما يثبت مانتها وإمكاناتها العالية لتحقيق إدارة طاقة ذكية وفعالة.

الكلمات المفتاحية: المراقبة غير الاقتحامية، كشف الأحداث، NILM، تعلم الآلة.

## Résumé :

Ce projet de fin d'études explore l'optimisation des systèmes de surveillance non intrusive de la consommation électrique (NILM) pour la désagrégation de l'énergie au sein des foyers, visant à encourager la conservation énergétique face aux défis environnementaux et économiques. Il se concentre sur l'évaluation comparative de différentes combinaisons d'algorithmes de détection d'événements (Seuil, CUSUM et TFED) et de méthodes de classification (RF, XGBoost, SVM, CNN) pour identifier les appareils électriques à partir d'un signal de consommation agrégé. L'étude, menée sur le jeu de données ENERTALK, révèle que les algorithmes basés sur les arbres de décision, notamment XGBoost et RF, lorsqu'ils sont associés aux méthodes de détection CUSUM et TFED, surpassent constamment les autres approches en termes de précision, rappel, score F1 et exactitude, démontrant leur robustesse et leur potentiel pour une gestion énergétique intelligente et efficace.

**Mots clés :** Surveillance non intrusive, Détection d'événements, NILM, Apprentissage automatique

## Abstract:

This final-year project explores the optimization of Non-Intrusive Load Monitoring (NILM) systems for household energy disaggregation, aiming to promote energy conservation in response to environmental and economic challenges. It focuses on the comparative evaluation of various combinations of event detection algorithms (Threshold, CUSUM and TFED) and classification methods (Random Forest, XGBoost, SVM, CNN) to identify electrical appliances from an aggregated consumption signal. The study, conducted on the ENERTALK dataset, reveals that tree-based algorithms, particularly XGBoost and Random Forest, when coupled with CUSUM and TFED detection methods, consistently outperform other approaches in terms of precision, recall, F1-score, and accuracy, demonstrating their robustness and potential for smart and efficient energy management.

**Keywords:** Non-intrusive monitoring, Event detection, NILM, ENERTALK, Machine learning

Liste des abréviations :

---

<b>NILM</b>	Non-Intrusive Load Monitoring (Surveillance Non-Intrusive des Charges)
<b>FSM</b>	Finite State Machines (Machines à États Finis)
<b>CVD</b>	Continuous Variable Devices (Appareils à Variation Continue)
<b>EMI</b>	Electromagnetic Interference (Interférences Électromagnétiques)
<b>TFED</b>	Tukey's Fence Event Detection (Détection d'événements par les bornes de Tukey)
<b>CUSUM</b>	Cumulative Sum (Somme Cumulative)
<b>IQR</b>	InterQuartile Range (Écart Interquartile)
<b>RF</b>	Random Forest
<b>XGBoost</b>	eXtreme Gradient Boosting
<b>SVM</b>	Support Vector Machines (Machines à Vecteurs de Support)
<b>CNN</b>	Convolutional Neural Networks (Réseaux de Neurones Convolutifs)
<b>RNN</b>	Recurrent Neural Networks (Réseaux de Neurones Récurrents)
<b>kHz / MHz</b>	kilohertz / mégahertz

Liste des figures :

---

<b>Figure 0-1</b> : (a) Cadre général de l'approche NILM, (b) Données de charge agrégées obtenues à l'aide d'un seul point de mesure, (c) Différents types de charge en fonction de leur schéma de consommation d'énergie [4].....	6
<b>Figure 2-1</b> Diagramme de la somme cumulative (CUSUM) pour la détection des points hors contrôle.....	11
<b>Figure 2-2</b> : Visualisation de l'écart interquartile (IQR) et des bornes de détection d'outliers .....	12
<b>Figure 2-3</b> : Schéma du fonctionnement de forêt d'arbre de décision .....	13
<b>Figure 2-4</b> : les machines à vecteurs de support (SVM) .....	14
<b>Figure 2-5</b> : Le réseau de neurones convolutifs .....	15
<b>Figure 3-1</b> : Pipeline de traitement NILM.....	19
<b>Figure 4-1</b> Précision obtenu pour chaque combinaison de détection et classification.....	23
<b>Figure 4-2</b> Rappel (Recall) obtenu pour chaque combinaison de détection et classification..	24
<b>Figure 4-3</b> : F1-Score obtenu pour chaque combinaison de détection et classification .....	25

<b>Figure 4-4</b> Exactitude (Accuracy) obtenue pour chaque combinaison de détection et classification.....	26
<b>Figure 4-5</b> : Perte logarithmique obtenu pour chaque combinaison de détection et classification.....	27

*Liste des tableaux :*

---

<b>Tableau 2-1</b> : Comparaison des méthodes de détection d'événements en NILM.....	12
<b>Tableau 3-1</b> : Hyperparamètres Optimaux des Modèles.....	20

Table des matières :

---

<i>Introduction générale</i> .....	1
<i>Chapitre 01 : Fondements et Concepts des Systèmes NILM</i> .....	4
1.1 Définition :.....	5
1.2 Principe de fonctionnement :.....	5
1.2.1 Les différents types des appareils électriques :.....	5
1.2.2 Les étapes principales des systèmes NILM : .....	6
1.2.2.1 L'acquisition des données : .....	6
1.2.2.2 Détection d'événements et extraction de caractéristiques :.....	7
1.2.2.2.1 Catégories des caractéristiques principales.....	7
1.2.2.2.2 Catégories des méthodes de détection : .....	7
1.2.2.3 L'identification des charges : .....	8
1.3 Applications du NILM dans les environnements industriels et domestiques.....	8
1.4 Principaux Défis Actuels dans les systèmes NILM.....	8
<i>Chapitre 2 : Méthodes de Détection d'Événements et Modèles de Classification en NILM</i> .....	9
2.1 Techniques de Détection d'Événements .....	10
2.1.1 Détection Basée sur le Seuil :.....	10
2.1.2 Méthode de somme cumulative (CUSUM) :.....	10
2.1.3 Détection d'événements basée sur les bornes de Tukey (TFED):.....	12
2.1.4 Comparaison des méthodes de détection d'événements en NILM : .....	12
2.2 L'apprentissage automatique.....	13
2.2.1 Forêt d'arbre de décision (RF) : .....	13
2.2.2 Contraction d'EXtreme Gradient Boosting (XGBoost) :.....	14
2.2.3 Les Machines à Vecteurs de Support (SVM) :.....	14
2.2.4 Les Réseaux de Neurones Convolutifs (CNN) : .....	15
2.3 Métriques d'Évaluation pour la Classification.....	15
2.3.1 L'Accuracy .....	16
2.3.2 La Précision.....	16
2.3.3 Le Rappel .....	16
2.3.4 F1-Score .....	16
2.3.5 Perte Logarithmique.....	16
<i>Chapitre 3 : Cadre Expérimental et Mise en Œuvre des Méthodes NILM</i> .....	17
3.1 Contexte Expérimental : .....	18
3.1.1 Jeu de donnée : .....	18

3.1.2	Environnement de Travail :.....	18
3.2	Protocole d'Évaluation des Pipelines.....	18
3.2.1	Composants des Pipelines Évalués .....	18
3.2.2	Déroulement de l'Expérimentation.....	19
<i>Chapitre 4 : Analyse Comparative des performances des pipelines NILM.....</i>		<i>22</i>
4.1	Analyse des Métriques de Performance .....	23
4.1.1	Précision :.....	23
4.1.2	Rappel : .....	24
4.1.3	F1-Score : .....	25
4.1.4	Exactitude:.....	25
4.1.5	Perte Logarithmique (Log Loss) : .....	26
4.2	Discussion des Performances des Différentes Combinaisons Détection-Classification 27	
<i>Conclusion.....</i>		<i>28</i>
<i>Bibliographie.....</i>		<i>30</i>

*Introduction générale*

---

Aujourd'hui, le monde est confronté à un sentiment d'urgence croissant concernant les efforts de conservation de l'énergie. Les enjeux environnementaux, économiques et sociaux poussent à repenser en profondeur notre manière de gérer l'énergie. La consommation d'énergie des foyers représente, selon les estimations, environ 33% de la consommation totale d'énergie dans le monde. Cette proportion met en lumière le potentiel considérable des ménages dans la transition énergétique globale. Il a été démontré que la sensibilisation des usagers, via un retour d'information sur la consommation de leurs appareils électriques domestiques, constitue un levier efficace, permettant d'atteindre une réduction de la consommation de l'ordre de 20% [1]. Cette statistique est fondamentale, car elle souligne que l'efficacité énergétique n'est pas uniquement une question d'optimisation technique, mais également une question de comportement humain et d'information. La capacité à fournir aux consommateurs des données granulaires et compréhensibles sur leur consommation est donc un catalyseur essentiel pour des changements comportementaux significatifs.

La gestion énergétique s'impose ainsi comme une priorité stratégique, bien au-delà de sa seule dimension technique. Historiquement, la surveillance de la consommation des appareils électriques s'est appuyée sur des méthodes intrusives [2], consistant à installer des compteurs d'énergie sur chaque appareil pour surveiller et contrôler leur consommation électrique de manière individuelle. Cependant, cette approche présente plusieurs inconvénients majeurs : le coût d'installation augmente linéairement avec le nombre d'appareils à surveiller, certains appareils peuvent être difficilement accessibles pour l'installation de capteurs, et le système devient contraignant à reconfigurer à chaque ajout, déplacement ou retrait d'un appareil. Ces limitations techniques et économiques ont freiné l'adoption généralisée de solutions de surveillance détaillée des appareils. Pour surmonter ces obstacles et fournir efficacement les informations granulaires nécessaires pour répondre aux enjeux énergétiques globaux, le développement d'une approche alternative est devenu impératif.

C'est dans ce contexte que la méthodologie du Non-Intrusive Load Monitoring (NILM), que l'on peut traduire par "surveillance non intrusive de la consommation", a émergé comme une solution aux contraintes des approches intrusives. Le concept du NILM, introduit par George W. Hart en 1992 [3], repose sur un principe fondamentalement différent : surveiller l'ensemble des charges électriques d'une installation en analysant la consommation agrégée, mesurée en un seul point du réseau. Un algorithme de désagrégation est ensuite appliqué pour estimer la puissance individuelle consommée par chaque appareil. Cette approche unique, qui permet de déduire des informations détaillées à partir d'une source unique et non invasive, a propulsé le NILM du statut de concept de recherche à celui de priorité stratégique dans la gestion énergétique moderne. Cette transition est le reflet d'une prise de conscience que les défis énergétiques actuels exigent des solutions innovantes, capables de s'intégrer de manière transparente dans les infrastructures existantes sans engendrer de coûts prohibitifs ou de complexités d'installation.

La performance d'un système de surveillance non intrusive de la consommation (NILM) repose sur l'efficacité de deux étapes importantes et successives : la détection d'événements, qui identifie les instants de changement de consommation, et la classification, qui attribue ces changements à un appareil spécifique. Pour chacune de ces étapes, il existe une grande diversité de méthodes algorithmiques dont le choix impacte directement la précision globale du système. La sélection et l'optimisation de ces algorithmes sont donc essentielles pour la viabilité pratique du NILM.

La présente recherche se concentre sur l'évaluation comparative de différents algorithmes de détection d'événements et de classification utilisés dans les systèmes NILM pour la désagrégation énergétique. L'étude vise à évaluer ces algorithmes et à les comparer en fonction de métriques de performance clés telles que l'exactitude (accuracy), la précision (precision), le rappel (recall), le F1-score et la perte logarithmique (Log Loss). Ce rapport est structuré en quatre chapitres. Le premier chapitre introduit les fondements et concepts des systèmes NILM. Le deuxième chapitre détaille les méthodes de détection d'événements et les modèles d'apprentissage automatique employés pour la classification. Le troisième chapitre décrit le protocole expérimental et le processus des pipelines NILM. Enfin, le quatrième chapitre présente et analyse les résultats obtenus, offrant une discussion approfondie des performances des différentes combinaisons algorithmiques.

*Chapitre 01 : Fondements et Concepts des Systèmes NILM*

---

Ce chapitre est dédié à l'établissement des bases théoriques et conceptuelles nécessaires à la compréhension des systèmes de Surveillance Non-Intrusive des Charges.

### **1.1 Définition :**

Le NILM, parfois appelée désagrégation de l'énergie, est une approche permet d'identifier les appareils électriques en fonctionnement sans avoir besoin d'installer des capteurs sur chaque appareil individuel. En observant uniquement la consommation électrique globale au niveau du tableau électrique principal [3]

Mathématiquement, le problème de désagrégation du NILM peut être exprimé par l'équation suivante [2] :

$$P_{totale}(t) = \sum_{n=1}^N P_n(t) + e(t) \quad (1.1)$$

Où :

- $P_{totale}(t)$  est la puissance totale mesurée à l'instant  $t$ ,
- $P_n(t)$  est la puissance estimée de l'appareil  $n$ ,
- $e(t)$  est l'erreur résiduelle de modélisation.

### **1.2 Principe de fonctionnement :**

Les approches NILM identifient les appareils en repérant ces signatures électriques au sein du signal de consommation global d'une installation électrique.

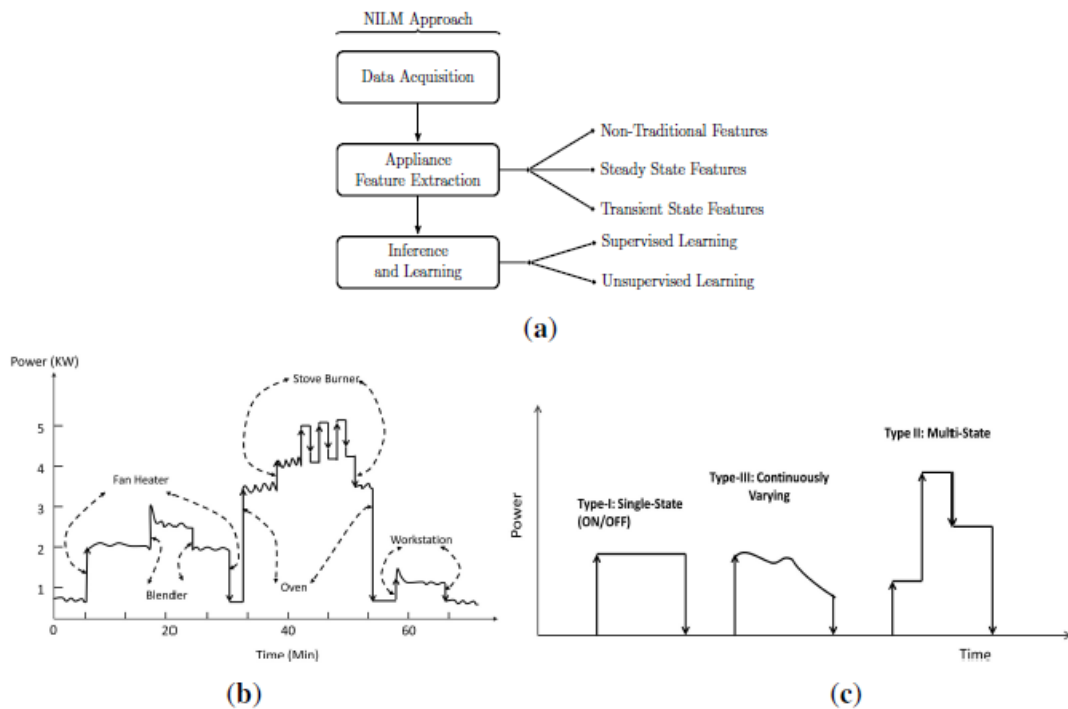
Chaque appareil électrique à sa propre empreinte énergétique, qu'on appelle sa signature de charge ou signature d'appareil, L'identification des appareils dépend fortement de ces signatures, qui sont elles-mêmes influencées par la catégorie de l'appareil, Dans le cadre des systèmes NILM, les appareils électriques sont communément catégorisés en fonction de leurs modèles de consommation énergétique, ce qui facilite leur identification et la modélisation de leur comportement [3].

#### **1.2.1 Les différents types des appareils électriques :**

- Les appareils de Type I, caractérisés par un état binaire (marche/arrêt). Ces dispositifs basculent simplement entre un état de consommation nulle et un état de consommation active stable.
- Les appareils de Type II sont des machines à états finis (FSM). Ils possèdent un nombre défini d'états de fonctionnement distincts, chacun associé à un niveau de puissance spécifique, comme les différentes phases d'un cycle de lavage.
- Les appareils de Type III sont désignés comme des dispositifs à variation continue (CVD). Leur particularité réside dans une consommation de puissance qui n'est pas fixe

mais variable en permanence, sans états discrets. Cela inclut par exemple les variateurs de lumière ou les appareils à vitesse modulée.

- Les appareils de Type IV sont les consommateurs permanents. Ces dispositifs restent actifs sur de longues périodes, souvent des jours ou des semaines, et maintiennent une consommation énergétique relativement constante. Un réfrigérateur ou un routeur sont des exemples éloquents de cette catégorie. [3]



**Figure 0-1 :** (a) Cadre général de l’approche NILM, (b) Données de charge agrégées obtenues à l’aide d’un seul point de mesure, (c) Différents types de charge en fonction de leur schéma de consommation d’énergie [4]

## 1.2.2 Les étapes principales des systèmes NILM :

Comme le montre la figure 1-1 (a), Une approche NILM se compose de trois étapes : l’acquisition des données, l’extraction des caractéristiques et l’identification des charges.

### 1.2.2.1 L’acquisition des données :

Les méthodologies NILM reposent sur l’analyse de données électriques acquises via un unique capteur positionné sur le tableau électrique principal. La nature des données collectées varie cependant selon les études et les technologies de mesure employées.

Une première approche consiste à mesurer directement les grandeurs fondamentales du courant alternatif, telles que la tension efficace (V), le courant efficace (A) et, par leur produit, la puissance apparente (S) [5] [6]

En complément de ces mesures directes, d'autres études exploitent des caractéristiques dérivées, calculées à partir des grandeurs primaires. Celles-ci incluent notamment la puissance réactive (Q) et le facteur de puissance

Au-delà des grandeurs fondamentales, des études plus avancées exploitent des signatures électriques de plus haut niveau, telles que les interférences électromagnétiques (EMI), la distorsion harmonique [7] [8] ou encore l'analyse des phénomènes transitoires à l'allumage ou à l'extinction des appareils [9] [10]

L'extraction de ces différentes caractéristiques est directement conditionnée par la fréquence d'échantillonnage du signal électrique. La littérature scientifique classe généralement les approches NILM en deux grandes catégories :

- **NILM à Haute Fréquence** : Cette approche utilise des fréquences d'échantillonnage élevées, typiquement de l'ordre du kilohertz (kHz) au mégahertz (MHz). Elle permet de capturer des détails fins du signal comme les transitoires et les harmoniques, offrant ainsi des signatures très riches.
- **NILM à Basse Fréquence**: Cette approche se contente d'une fréquence d'échantillonnage faible, généralement de 1 Hz jusqu'à une mesure toutes les quelques minutes.

### ***1.2.2.2 Détection d'événements et extraction de caractéristiques :***

Après la collecte des données, il est impératif d'extraire des caractéristiques pertinentes des séries temporelles électriques afin de détecter les transitions d'état des appareils.

#### **1.2.2.2.1 Catégories des caractéristiques principales**

Ces caractéristiques se regroupent en trois catégories principales [4]:

1. Les caractéristiques en régime permanent sont dérivées du fonctionnement stable de l'appareil, incluant notamment la puissance active (P) et la puissance réactive (Q) [11]
2. Les caractéristiques en régime transitoire sont, quant à elles, issues des phases de changement d'état (par exemple, les pics de courant, les temps de réponse, ou les enveloppes spectrales) [12]

3. les caractéristiques non traditionnelles englobent des informations contextuelles au-delà des mesures électriques directes. Elles comprennent des éléments tels que l'heure de la journée, la fréquence d'usage, la distribution ON/OFF et les corrélations entre le fonctionnement de différents appareils [13].

#### **1.2.2.2.2 Catégories des méthodes de détection :**

On peut diviser les méthodes de détection d'événements en trois grandes catégories [14]:

- Les méthodes basées sur des règles simples : Elles utilisent des connaissances déjà connues pour fixer des seuils ou des paramètres précis afin de détecter les événements [3] [15]

- Les modèles probabilistes adoptent un cadre stochastique, où la détection d'un changement est formalisée comme un problème de test d'hypothèse basé sur des métriques statistiques [16] [17]
- Les approches par filtre apparié, relevant du traitement du signal, recherchent la correspondance entre un prototype de signature d'événement et le signal mesuré via une opération de corrélation [18] [19]

### ***1.2.2.3 L'identification des charges :***

La phase d'identification de charge utilise des caractéristiques de charge agrégées pour identifier les appareils individuels en appliquant des algorithmes d'apprentissage automatique ou des techniques de traitement du signal. Les modèles d'apprentissage automatique à ce stade peuvent être classés en apprentissage supervisé et non supervisé. Le modèle peut être sélectionné en fonction des exigences de la solution NILM, telles que le coût, la précision et le temps.

## **1.3 Applications du NILM dans les environnements industriels et domestiques**

- Surveillance énergétique et sensibilisation des consommateurs [20]
- Optimisation de la consommation énergétique des machines
- Détection d'anomalies et maintenance prédictive

## **1.4 Principaux Défis Actuels dans les systèmes NILM**

- **Diversité des appareils :** Il est très difficile de créer un modèle unique capable de reconnaître tous les appareils, en raison de la grande variabilité des signatures électriques entre différentes marques et modèles au sein d'une même catégorie [4] : Il ne suffit pas d'apprendre à reconnaître "un réfrigérateur" ; un modèle de 2010 n'aura pas la même signature électrique qu'un modèle à technologie inverter de 2023
- **Absence de datasets de référence**
- **Appareils à faible consommation :** Les appareils consommant peu d'énergie ont des signatures très similaires
- **Dépendance à l'apprentissage supervisé :** La plupart des solutions NILM nécessitent une phase d'entraînement avec des données étiquetées. Cela pose deux problèmes :
  - Une simple modification du réseau par le fournisseur d'énergie (ex: correction du facteur de puissance) peut rendre les signatures apprises obsolètes [15].
  - Les systèmes ne peuvent pas identifier un nouvel appareil qui n'était pas dans la base de données d'entraînement initiale.

*Chapitre 2 : Méthodes de Détection d'Événements et Modèles de Classification en  
NILM*

---

Ce chapitre présente les approches adoptées pour détecter les événements électriques et classifier les charges dans un système NILM. Il expose les principes des méthodes retenues, les raisons de leur sélection, ainsi que les indicateurs d'évaluation utilisés pour mesurer leur performance.

## 2.1 Techniques de Détection d'Événements

La première étape fondamentale de l'approche événementielle en NILM consiste à détecter avec précision les changements de consommation, appelés événements, au sein du signal agrégé, où un événement correspond à une transition du signal d'un régime stationnaire à un nouveau régime stationnaire distinct [21]. La littérature scientifique propose une vaste gamme de méthodes pour cette tâche, allant de techniques statistiques classiques à des approches plus complexes. Dans le cadre de ce mémoire, afin de mener une analyse comparative pertinente, nous avons sélectionné un panel d'algorithmes représentatifs et validés.

Nous présenterons ainsi : la détection par seuil simple, qui constitue la référence historique et la plus intuitive ; la méthode de Tukey (TFED), reconnue pour sa robustesse statistique face au bruit ; et enfin, l'algorithme CUSUM, une technique de contrôle de processus efficace pour identifier les dérives du signal. Les sections qui suivent détailleront le principe théorique de chacune de ces approches.

### 2.1.1 Détection Basée sur le Seuil :

La méthode la plus simple et la plus couramment utilisée en NILM pour détecter les événements est la détection par seuil. Elle consiste à surveiller la variation instantanée de la puissance (ou du courant) et à signaler un événement dès que cette variation dépasse un seuil prédéfini [22] :

$$|\Delta P_t| = |P_t - P_{t-1}| \geq S \quad (2.1)$$

- $|\Delta P_t|$ : variation de puissance entre deux instants.
- $S$ : seuil de détection défini selon les caractéristiques du système.

### 2.1.2 Méthode de somme cumulative (CUSUM) :

L'algorithme de somme cumulative est une technique utilisée pour suivre les changements dans une série de données, appliquée au domaine du NILM pour détecter les événements de mise en marche ou d'arrêt des appareils électriques en surveillant les variations de la puissance consommée [23] .

Le principe repose sur l'analyse continue d'une série temporelle de puissance, notée  $\{x_n\}_{n=1}^{\infty}$  à travers une fenêtre glissante de taille fixe  $m$ .

### Etape 01 : calcule la moyenne locale

Dans un premier temps, une moyenne de référence locale notée  $k_m$  est calculée sur une fenêtre glissante de taille  $m$ . Cette moyenne correspond à la moyenne des valeurs de puissance sur l'intervalle courant et se calcule selon l'équation suivante :

$$k_m = \frac{1}{m} \sum_{i=n}^{n+m-1} x_i \quad (2.2)$$

Où :

- $x_i$  représente la consommation instantanée à l'instant  $i$
- $m$  est la taille de la fenêtre glissante

### Etape 02 : Statistique de test CUSUM

De manière récursive une statistique de test cumulée notée  $Z_n$  est mise à jour. Cette statistique permet d'accumuler les écarts positifs entre la consommation instantanée mesurée  $x_n$  et la moyenne locale  $k_m$ , Sa formule est la suivante:

$$Z_n = \max(0, Z_{n-1} + x_n - k_m), \text{ avec } Z_0 = 0 \quad (2.3)$$

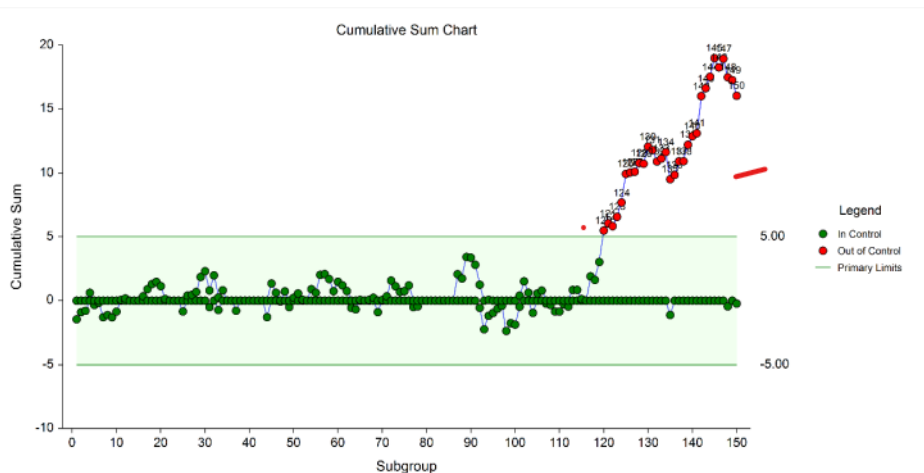
#### Critère de détection d'événements :

Un événement est détecté lorsque la valeur de  $Z_n$  dépasse un niveau prédéfini  $H$ .

Lorsque  $Z_n > H$ , cela signifie que la somme des variations positives s'est accumulée de manière significative, indiquant un changement d'état durable et non de simples fluctuations aléatoires.

La période où  $Z_n$  est positif mais reste inférieur à  $H$  peut être considérée comme une phase de "suspicion" ou de "pré-alerte" où un événement pourrait être en train de se former.

Une fois l'événement détecté, l'instant précis de l'événement peut être estimé en analysant le début de la croissance de  $Z_n$



**Figure 2-1** Diagramme de la somme cumulée (CUSUM) pour la détection des points hors contrôle [41]

### 2.1.3 Détection d'événements basée sur les bornes de Tukey (TFED):

La méthode de l'écart interquartile (IQR), introduite par Tukey en 1977 [24] est couramment utilisée pour détecter les valeurs aberrantes dans un ensemble de données, son principe d'application repose sur la distinction statistique entre le bruit et les événements

**Etape 01** : calcul de l'IQR

$$IQR = Q_3 - Q_1 \quad (2.4)$$

Où :

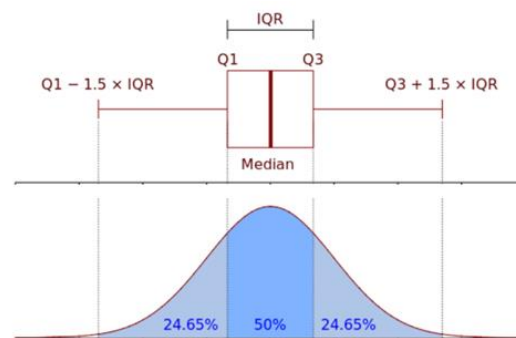
- $Q_1$  est le premier quartile 25%
- $Q_3$  est le troisième quartile 75%

**Etape 02** : détection des valeurs aberrantes

Une valeur  $x$  est considérée comme aberrante si :

$$x < Q_1 - k \times IQR \text{ ou } x > Q_3 + k \times IQR$$

Avec :  $k$  facteur de tukey  $1.5 \leq k \leq 3$



**Figure 2-2** : Visualisation de l'écart interquartile (IQR) et des bornes de détection d'outliers [40]

### 2.1.4 Comparaison des méthodes de détection d'événements en NILM :

**Tableau 2-1** : Comparaison des méthodes de détection d'événements en NILM [25] [26]

Méthode	Avantages	Limites
Seuil simple	- la simplicité d'implémentation	- sensible au bruit - limite en précision
Cusum	- adapté au temps réel et systèmes embarqués - robuste au bruit - réduisant les faux positifs.	- nécessite un paramétrage précis des seuils - une performance réduite face aux charges à consommation variable
TFED	- meilleure localisation temporelle - un taux de précision (99 %) sur des jeux de données standards - faible taux de faux positifs	- Méthode récente - Nécessite un échantillonnage adapté

## 2.2 L'apprentissage automatique

Une fois les événements détectés et leurs caractéristiques extraites, la seconde étape cruciale du processus NILM est leur classification, qui vise à attribuer chaque événement à l'appareil l'ayant provoqué. Pour cette fin, de nombreux algorithmes d'apprentissage automatique ont été appliqués avec succès. Notre étude s'est portée sur un éventail de modèles performants et représentatifs de différentes familles d'algorithmes, tous largement cités dans la littérature NILM. Nous aborderons ainsi : les Machines à Vecteurs de Support (SVM), pour leur efficacité à trouver des frontières de décision optimales ; la Forêt Aléatoire (Random Forest) et XGBoost, comme exemples puissants d'approches ensemblistes ; et enfin, les Réseaux de Neurones Récurrents (RNN), spécifiquement conçus pour modéliser la nature séquentielle des données de consommation. La suite de cette partie sera consacrée à l'exposé des fondements théoriques de ces quatre modèles.

### 2.2.1 Forêt d'arbre de décision (RF) :

L'algorithme de la Forêt Aléatoire est une méthode d'apprentissage ensembliste conçue pour pallier une faiblesse majeure de l'arbre de décision unique : son instabilité. Un arbre seul est en effet très sensible aux variations des données d'apprentissage, où de petites perturbations peuvent altérer radicalement sa structure et ses prédictions [27]. Pour surmonter ce problème, la Forêt Aléatoire construit une multitude d'arbres de décision décorrélés les uns des autres en introduisant une double source d'aléa : le bagging, qui consiste à entraîner chaque arbre sur un sous-échantillon aléatoire des données, et la sélection aléatoire d'un sous-ensemble de caractéristiques à chaque nœud. Alors qu'un arbre isolé peut être affecté par le bruit, l'agrégation des prédictions de nombreux arbres diversifiés permet de neutraliser les erreurs individuelles et de converger vers un modèle plus stable et plus précis [27] [28]. La prédiction finale est ensuite déterminée par un vote majoritaire en classification ou par la moyenne des prédictions en régression.

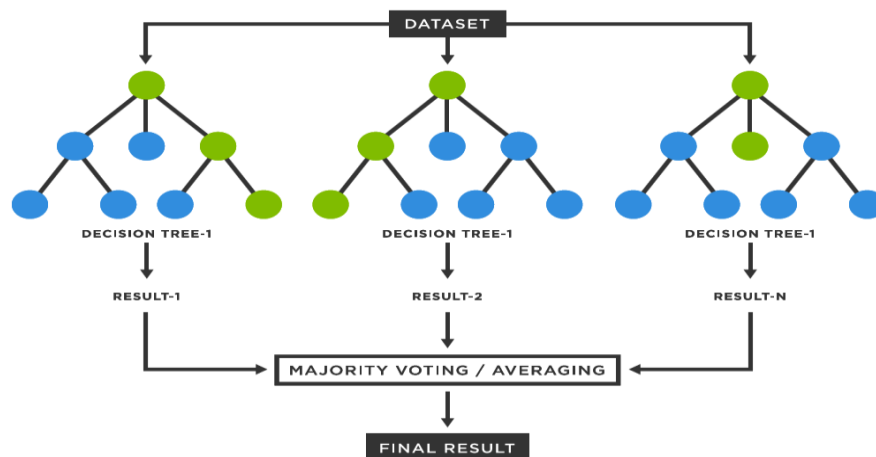


Figure 2-3 : Schéma du fonctionnement de forêt d'arbre de décision [40]

### 2.2.2 Contraction d'EXtreme Gradient Boosting (XGBoost):

Est un algorithme d'apprentissage supervisé de type ensembliste, réputé pour sa haute performance en classification et en régression. Son principe de fonctionnement repose sur le Gradient Boosting, une technique qui construit des modèles de manière séquentielle, typiquement des arbres de décision. Contrairement à des méthodes comme la Forêt Aléatoire qui construisent des arbres indépendamment, XGBoost entraîne chaque nouvel arbre pour corriger spécifiquement les erreurs résiduelles commises par les arbres précédents. Ce processus itératif permet d'affiner progressivement la prédiction globale. La force de XGBoost réside dans ses optimisations poussées : il intègre nativement des mécanismes de régularisation pour prévenir le sur apprentissage, une gestion intelligente des valeurs manquantes, ainsi que des optimisations système comme la parallélisations des calculs, ce qui en fait une solution à la fois rapide, robuste et extrêmement précise [29].

### 2.2.3 Les Machines à Vecteurs de Support (SVM) :

Un algorithme d'apprentissage supervisé puissant, utilisé pour des tâches de classification et de régression. Le principe fondamental du SVM en classification consiste à déterminer un hyperplan optimal qui sépare les données en différentes classes. Cet hyperplan est celui qui maximise la marge, c'est-à-dire la distance entre lui-même et les points de données les plus proches de chaque classe (nommés les "vecteurs de support"). Pour les problèmes où une séparation linéaire n'est pas possible, le SVM emploie l'astuce du noyau (*kernel trick*), qui projette les données dans un espace de plus grande dimension où une séparation linéaire devient réalisable. La performance du modèle dépend alors du choix d'un noyau adapté (linéaire, polynomial, RBF) et du réglage des hyperparamètres qui contrôlent la complexité du modèle et la tolérance aux erreurs. Cette capacité à gérer des frontières de décision complexes rend le SVM particulièrement robuste [30].

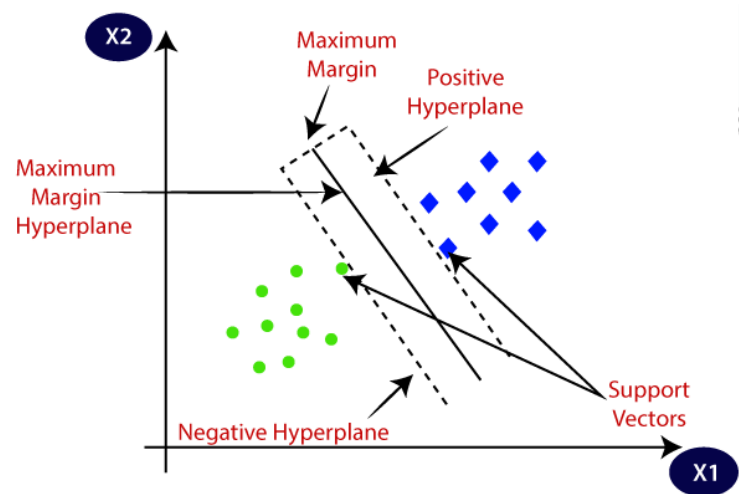


Figure 2-4 : les machines à vecteurs de support (SVM) [30]

## 2.2.4 Les Réseaux de Neurones Convolutifs (CNN) :

Les Réseaux de Neurones Convolutifs sont une famille d'architectures d'apprentissage dont la conception s'inspire directement du fonctionnement du cortex visuel biologique [31] .

Leur principale force est de pouvoir traiter efficacement des données qui possèdent une structure de grille, comme des séries temporelles ou des spectrogrammes.

L'efficacité des CNN repose sur deux principes ingénieux qui exploitent la structure même des données :

1. **La Localité des Connexions** : Au lieu de connecter tous les neurones entre eux, un neurone d'une couche de convolution ne "regarde" qu'une petite zone de l'entrée, son "champ récepteur". Il se spécialise ainsi dans l'analyse de motifs locaux.
2. **Le Partage de Poids** : Le même "détecteur de motif", que l'on appelle un filtre, est réutilisé et glissé sur toute l'étendue des données. Le réseau apprend donc une seule fois à reconnaître une caractéristique (comme un contour ou une texture) et peut ensuite la trouver n'importe où dans les données.

L'architecture d'un CNN, dont les bases ont été posées par les travaux sur le modèle LeNet [32], est une succession de couches qui transforment progressivement l'information :

- La couche de convolution applique des filtres pour créer des cartes de caractéristiques.
- La fonction d'activation introduit une non-linéarité essentielle pour que le modèle puisse apprendre des relations complexes.
- La couche de pooling (ou sous-échantillonnage) simplifie l'information en réduisant sa dimension, ce qui rend le modèle plus léger et plus robuste.
- La couche entièrement connectée, à la fin du processus, agit comme un classificateur final pour prendre la décision.

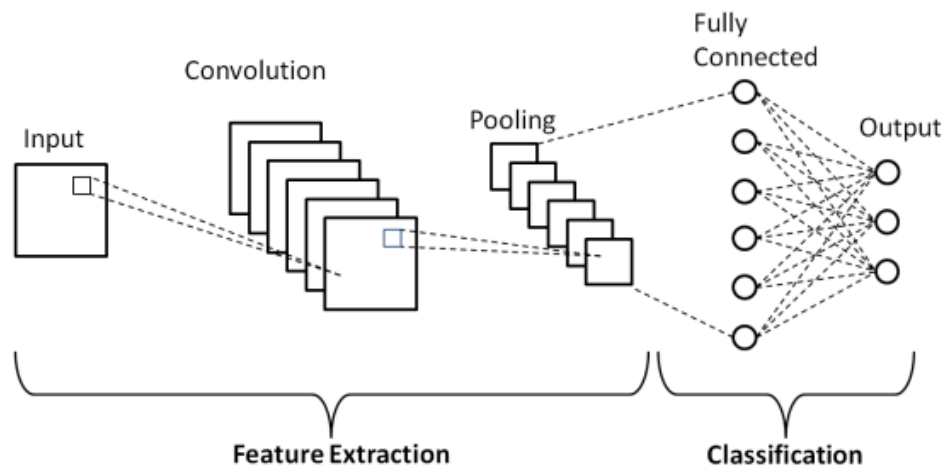


Figure 2-5 : Le réseau de neurones convolutifs [39]

## 2.3 Métriques d'Évaluation pour la Classification

Pour évaluer et comparer objectivement les performances de nos modèles de classification, nous avons retenu plusieurs métriques standards issues de la matrice de confusion [30] chacune offre une perspective différente sur l'efficacité du modèle

### 2.3.1 L'exactitude

La mesure la plus intuitive de la performance. Elle représente simplement la proportion de prédictions correctes sur l'ensemble des prédictions effectuées. Bien qu'utile pour une vue d'ensemble, elle peut être trompeuse si les classes sont déséquilibrées.

$$Accuracy = \frac{Vrais\ Positifs + Vrais\ Négatifs}{Total\ des\ prédictions} \quad (2.5)$$

### 2.3.2 La Précision

La mesure de pertinence des prédictions positives. Elle répond à la question : "Lorsque le modèle prédit une classe, à quel point est-il fiable ?". C'est le rapport entre les vrais positifs et l'ensemble des prédictions positives.

$$Précision = \frac{Vrais\ Positifs}{Vrais\ Positifs + Faux\ Positifs} \quad (2.6)$$

### 2.3.3 Le Rappel

Mesure l'exhaustivité du modèle. Il répond à la question : "Le modèle a-t-il réussi à identifier toutes les instances positives qui existent réellement ?". C'est le rapport entre les vrais positifs et l'ensemble des instances qui étaient réellement positives.

$$Rappel = \frac{Vrais\ Positifs}{Vrais\ Positifs + Faux\ Négatifs} \quad (2.7)$$

### 2.3.4 F1-Score

La Précision et le Rappel sont souvent en opposition. Le F1-Score a été introduit pour offrir une mesure unique qui équilibre ces deux aspects [31]. C'est la moyenne harmonique de la Précision et du Rappel, ce qui le rend particulièrement pertinent pour évaluer un modèle de manière équilibrée

$$F1 - Score = 2 \times \frac{Précision \times Rappel}{Précision + Rappel} \quad (2.8)$$

### 2.3.5 Perte Logarithmique

La Perte Logarithmique évaluant la confiance du modèle dans sa prédiction. Elle mesure à quel point les probabilités prédites sont proches des probabilités réelles (qui sont 1 pour la vraie classe et 0 pour les autres) [2]. Un score de perte logarithmique faible indique un modèle bien calibré. À l'inverse, la métrique pénalise lourdement un modèle qui est "sûr de lui" (prédit une haute probabilité) mais qui se trompe.

$$Log\ Loss = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^M y_{ij} \log(p_{ij}) \quad (2.9)$$

Où :

$N$  est le nombre d'échantillons  $M$  le nombre de classes,  $y_{ij}$  est 1 si l'échantillon  $i$  appartient à la classe  $j$  (et 0 sinon),  $p_{ij}$  est la probabilité prédite par le modèle que l'échantillon  $i$  appartienne à la classe  $j$ .

*Chapitre 3 : Cadre Expérimental et Mise en Œuvre des Méthodes NILM*

---

Ce chapitre détaille la mise en œuvre pratique de l'étude, décrivant le jeu de données utilisé, l'environnement de travail, et le protocole rigoureux d'évaluation des pipelines NILM. L'accent est mis sur les étapes de prétraitement, de détection d'événements, et d'optimisation des hyperparamètres, qui sont fondamentales pour la fiabilité et la reproductibilité des résultats.

### **3.1 Contexte Expérimental :**

#### **3.1.1 Jeu de donnée :**

L'ensemble de nos expérimentations repose sur le jeu de données public ENERTALK. Ce dernier est constitué de mesures de puissance active et réactive, collectées à une fréquence de 15 Hz sur une période de 29 à 122 jours au sein de 22 foyers coréens. [32]

Ce jeu de données a été sélectionné pour la haute qualité de ses enregistrements et la disponibilité simultanée des signaux de puissance agrégée et individuelle, ce qui est essentiel pour la validation de nos approches. Pour les besoins de cette étude, nous avons spécifiquement concentré notre analyse sur les données de la maison identifiée par le code house\_00 ce choix s'est imposé car elle est la seule habitation où le signal agrégé correspond à la somme des consommations de tous les appareils surveillés.

#### **3.1.2 Environnement de Travail :**

Toutes les implémentations et analyses ont été réalisées en utilisant le langage de programmation Python (version 3.9). L'écosystème logiciel s'appuie sur plusieurs bibliothèques scientifiques reconnues, notamment Pandas, NumPy, Scikit-learn, XGBoost, TensorFlow/Keras, Matplotlib et Seaborn.

### **3.2 Protocole d'Évaluation des Pipelines**

Notre approche expérimentale consiste à évaluer plusieurs chaînes de traitement complètes. Une telle chaîne, souvent appelée "pipeline" en apprentissage automatique [33], représente le processus complet depuis la détection d'un événement jusqu'à sa classification finale. Notre démarche vise donc à tester systématiquement chaque modèle de classification sur les différents jeux de données produits par les méthodes de détection, afin d'identifier la combinaison la plus performante.

#### **3.2.1 Composants des Pipelines Évalués**

Chaque pipeline évalué est constitué de deux composants successifs : une méthode de détection et un modèle de classification

Trois méthodes de détection, dont les principes ont été détaillés dans le chapitre 2, ont été implémentées pour générer les jeux de données d'événements :

- CUSUM (Cumulative Sum)
- TFED (tukey fancy Event Detection)

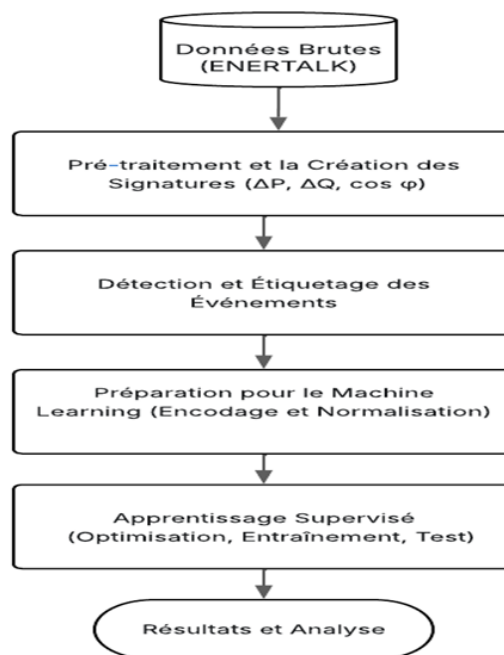
- Seuillage Simple

Quatre modèles de classification, présentés en détail dans le chapitre 2 :

- Random Forest (RF)
- XGBoost
- Support Vector Machine (SVM)
- Réseau de Neurones Convolutif (CNN)

### 3.2.2 Déroulement de l'Expérimentation

Pour résumer notre travail, voici le chemin complet que nous avons suivi, illustré par le logigramme de la Figure 3-1



**Figure 3-1** : Pipeline de traitement NILM

Notre point de départ est constitué des enregistrements bruts de la maison house\_00. Ces données ont d'abord subi une phase de prétraitement des données et extraction des signatures. Cette étape a consisté à synchroniser et ré-échantillonner les signaux, puis à calculer pour chaque instant une "signature" numérique composée des caractéristiques clés : la variation de puissance active ( $\Delta P$ ), la variation de puissance réactive ( $\Delta Q$ ) et le facteur de puissance ( $\cos \varphi$ ).

Une fois les données préparées, la phase de détection d'événements a été initiée, nous avons appliqué nos trois approches – CUSUM, TFED, et Seuillage Simple – de manière indépendante, générant ainsi trois jeux de données distincts, chacun possédant sa propre population d'événements.

C'est sur ces trois jeux de données que s'est déroulée la phase d'apprentissage supervisé. Pour chaque jeu de données, le processus a été le suivant :

1. Étiquetage des événements pour identifier l'appareil source.
2. Normalisation : La dernière étape du prétraitement a été la mise à l'échelle des caractéristiques. Il est en effet essentiel d'éviter que des attributs aux ordres de grandeur très différents ne biaisent l'apprentissage. Pour ce faire, nous avons employé le StandardScaler de la bibliothèque Scikit-learn [33] . Concrètement, cette technique standardise chaque caractéristique en lui soustrayant sa moyenne et en la divisant par son écart-type. Le résultat est que toutes les caractéristiques sont recentrées autour de zéro avec une variance unitaire, ce qui assure une base de comparaison équitable pour l'algorithme
3. Optimisation des Hyperparamètres par Recherche en Grille (Grid Search) :

Il est largement reconnu dans la communauté de l'apprentissage automatique que la performance d'un modèle, même très puissant, peut être considérablement dégradée s'il n'est pas correctement réglé. En effet, un algorithme performant doté d'hyperparamètres inadaptés donnera souvent des résultats médiocres [34].

C'est pourquoi nous ne pouvions pas nous contenter des paramètres par défaut et avons mis en place une phase d'optimisation

Notre choix s'est porté sur la méthode de recherche en grille avec validation croisée (GridSearchCV). L'idée derrière cette technique [33] est assez intuitive : plutôt que de tester des réglages au hasard, on définit à l'avance une "grille" de valeurs qui nous semblent prometteuses pour chaque hyperparamètre clé. L'algorithme se charge alors d'un travail méthodique : il teste toutes les combinaisons possibles de ces réglages, une par une. Pour éviter que notre choix ne soit un coup de chance sur une seule découpe des données, chaque combinaison est évaluée plusieurs fois grâce à la validation croisée (k-fold). C'est seulement la configuration qui obtient le meilleur score en moyenne sur ces différentes évaluations qui est gardée, ce qui nous garantit un choix plus robuste.

Nous avons appliqué ce processus de recherche systématique à tous nos modèles (RF, XGBoost, SVM et CNN) et pour chacun des trois jeux de données.

Pour illustrer concrètement ce processus, le tableau 3-1 récapitule les réglages optimaux que nous avons obtenus pour les modèles lorsqu'ils ont été testés sur les données de CUSUM.

**Tableau 3-1** : Hyperparamètres Optimaux des Modèles (k-fold=5)

Modèle	Hyperparamètre	Valeur Optimale
Random Forest	N estimators	200
	Max features	Sqrt
	Min samples leaf	2
	Min samples split	2

Xgboost	Colsample bytree	0,7
	gamma	0
	Learning rate	0,01
	Max depth	3
	N estimators	100
	subsample	0,7
SVM	C	100
	gamma	1
	kernel	RBF
CNN	Num_filters	32
	kernel_size	3
	dropout_rate	0,3
	learning_rate	0,001

#### 4. Entraînement et évaluation

Chaque jeu de données issu des méthodes de détection (CUSUM, TFED (k=1,5), Seuillage Simple) a été divisé en 75 % pour l'entraînement et 25 % pour le test, en conservant une distribution équilibrée des classes. L'étude a porté sur 7 appareils électroménagers : Machine à laver (WM), Purificateur d'eau (WP), Frigo kimchi (FK), Micro-ondes (M), Cuisinier à riz (RC), Réfrigérateur (F), et Téléviseur (TV). L'évaluation des modèles s'est faite à l'aide de plusieurs métriques : exactitude, F1-score, Précision, Rappel, ainsi que la perte logarithmique pour mesurer la qualité des probabilités prédites.

*Chapitre 4 : Analyse Comparative des performances des pipelines NILM*

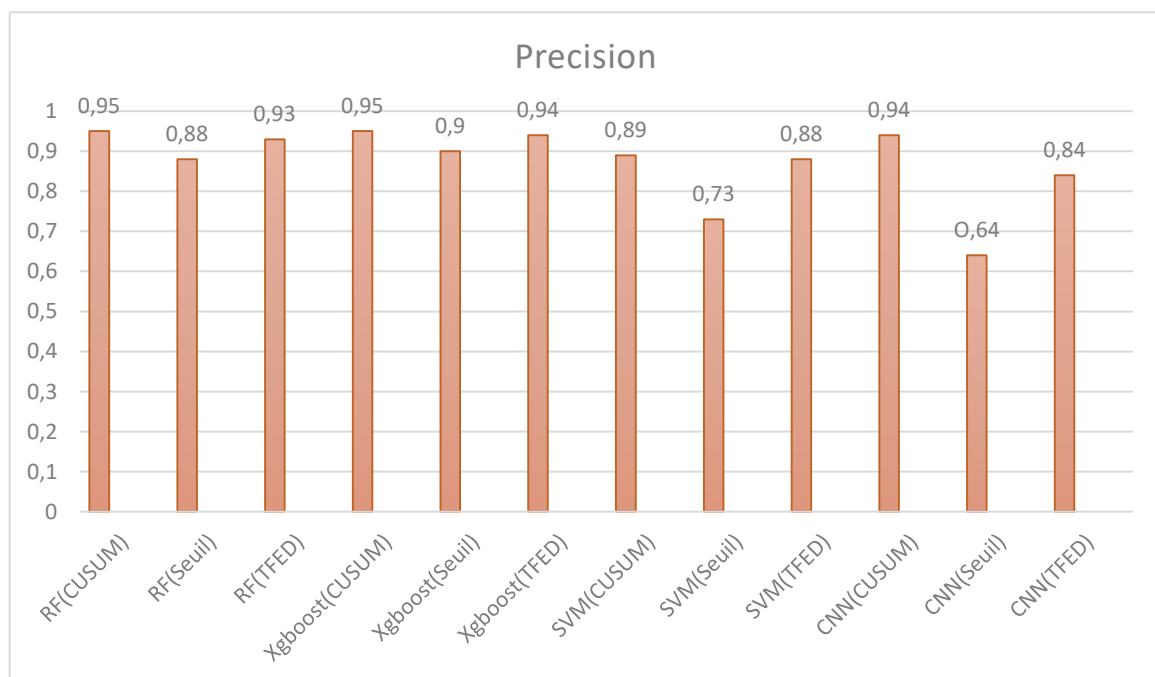
---

Ce chapitre est dédié à la présentation et à l'analyse détaillée des résultats obtenus lors des expérimentations menées sur les différentes configurations des pipelines de détection et de classification pour les systèmes NILM. L'objectif principal est d'évaluer la performance de chaque méthode à travers un ensemble de métriques d'évaluation clés, afin d'identifier les approches les plus efficaces et robustes pour la désagrégation des consommations électriques. Nous examinerons en profondeur la Précision, le Rappel, le Score F1, l'Exactitude et la Perte Logarithmique (Log Loss) de chaque modèle, en comparant leurs performances respectives dans diverses conditions. Cette analyse comparative permettra de dégager des conclusions significatives quant à l'aptitude de chaque pipeline à répondre aux défis posés par l'identification non-intrusive des appareils électriques.

#### 4.1 Analyse des Métriques de Performance

Les résultats des différentes combinaisons de méthodes de détection et de classification sont présentés et analysés ci-dessous pour chaque métrique.

##### 4.1.1 Précision :



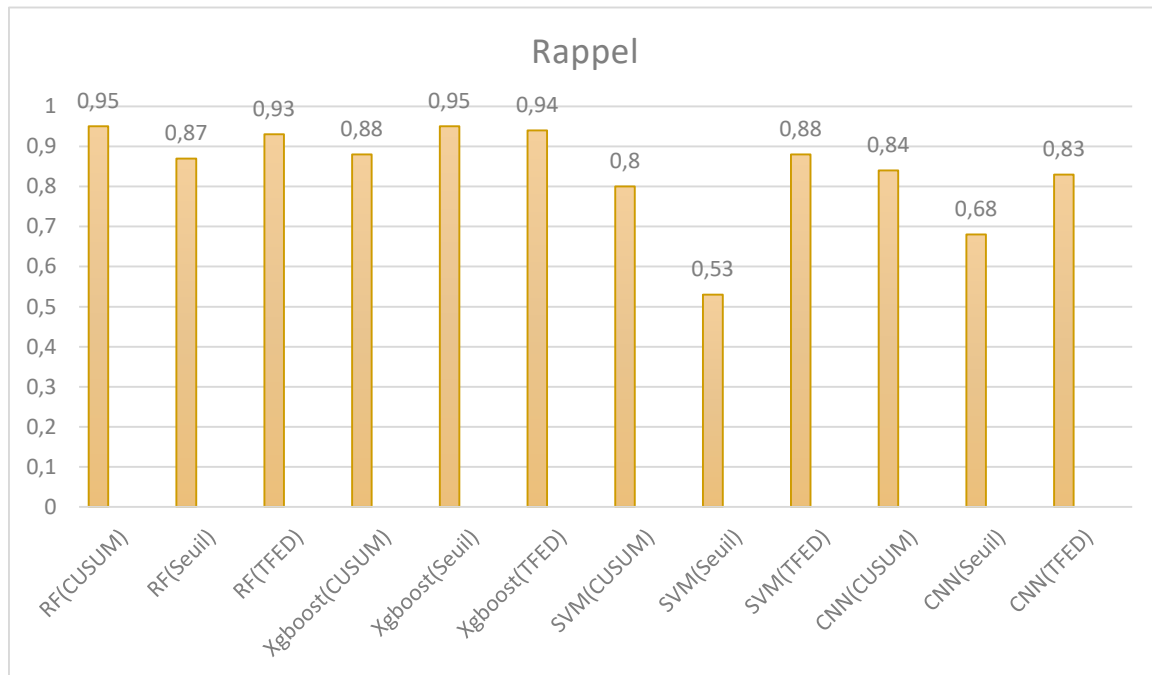
**Figure 4-1** Précision obtenu pour chaque combinaison de détection et classification

La métrique de Précision, illustrée dans la figure 4.1, indique la proportion de prédictions positives vraies parmi toutes les prédictions positives pour chaque méthode de pipeline.

- RF(CUSUM), Xgboost (CUSUM) et Xgboost (TFED) ont atteint la plus haute précision à 0,95, 0,95 et 0,94 respectivement, démontrant leur forte capacité à minimiser les faux positifs.
- RF(TFED) et Xgboost (Seuil) ont également montré des valeurs de précision élevées de 0,93 et 0,90.

- SVM(CUSUM), SVM(TFED) et CNN(CUSUM) ont maintenu une bonne précision à 0,89, 0,88 et 0,94.
- Inversement, SVM(Seuil) et CNN(Seuil) ont présenté une précision plus faible à 0,73 et 0,64, suggérant un taux plus élevé de classifications de faux positifs pour ces méthodes.

#### 4.1.2 Rappel :

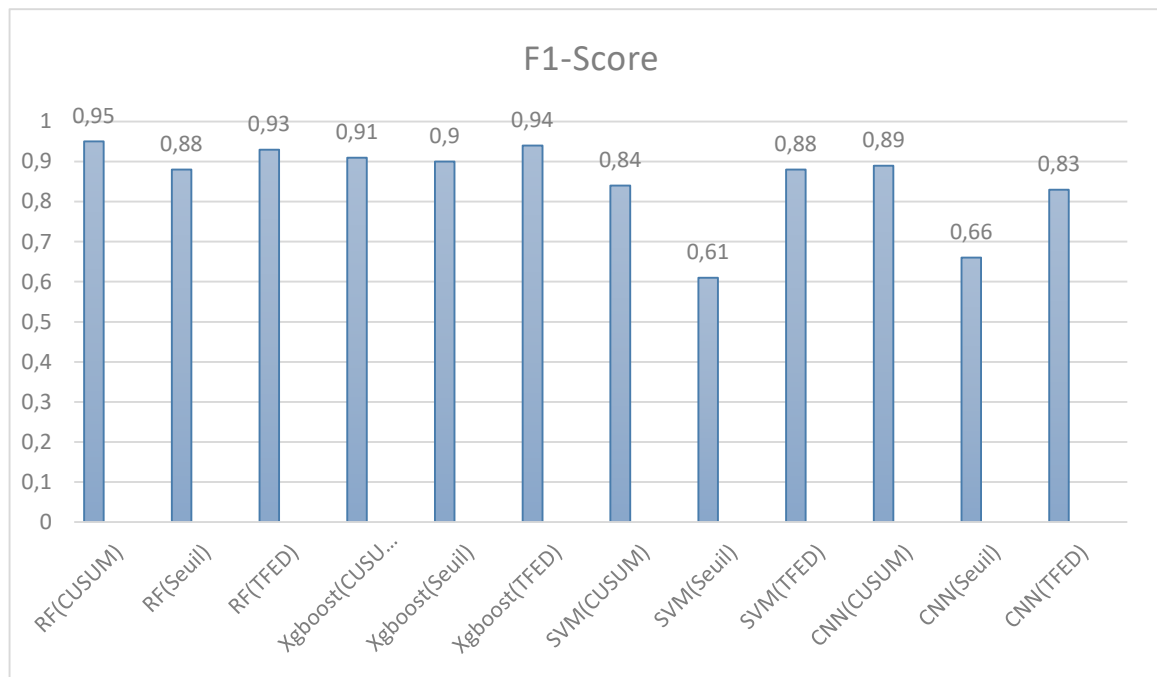


**Figure 4-2** Rappel (Recall) obtenu pour chaque combinaison de détection et classification

La métrique de Rappel, représentée dans la figure 4.2, mesure la proportion de prédictions positives vraies parmi toutes les instances positives réelles.

- RF(CUSUM) et Xgboost(CUSUM) ont atteint le rappel le plus élevé à 0,95, indiquant leur efficacité à identifier un grand pourcentage de cas positifs réels.
- Xgboost(TFED) et RF(TFED) ont également montré de fortes valeurs de rappel de 0,94 et 0,93.
- SVM(CUSUM), SVM(TFED), CNN(CUSUM) et CNN(TFED) ont bien performé avec des scores de rappel de 0,80, 0,88, 0,84 et 0,83, respectivement.
- Similaire à la précision, SVM(Seuil) et CNN(Seuil) ont montré les valeurs de rappel les plus basses à 0,53 et 0,68, suggérant que ces méthodes pourraient manquer un nombre significatif d'instances positives réelles

### 4.1.3 F1-Score :



**Figure 4-3 :** F1-Score obtenu pour chaque combinaison de détection et classification

Le Score F1, présenté dans la figure 4.3 est la moyenne harmonique de la Précision et du Rappel, offrant une mesure équilibrée de l'exactitude d'un modèle.

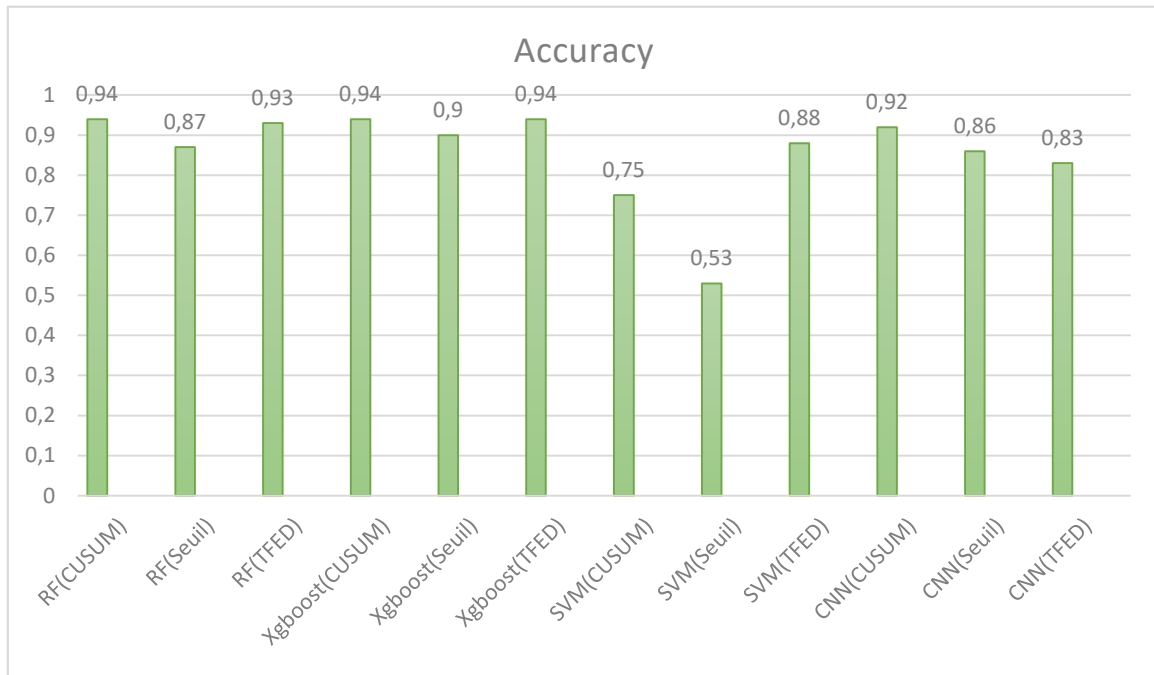
- RF(CUSUM) et Xgboost(TFED) ont enregistré les Scores F1 les plus élevés de 0,95 et 0,94, respectivement, indiquant leur performance globale supérieure en équilibrant la précision et le rappel.
- RF(TFED) et Xgboost(CUSUM) ont également atteint des Scores F1 élevés de 0,93 et 0,91.
- Xgboost(Seuil), SVM(CUSUM), SVM(TFED) et CNN(CUSUM) ont également démontré de bonnes performances avec des Scores F1 de 0,90, 0,84, 0,88 et 0,89.
- SVM(Seuil) et CNN(Seuil) ont de nouveau montré les Scores F1 les plus bas à 0,61 et 0,66, soulignant leurs difficultés à atteindre un équilibre entre la précision et le rappel.

### 4.1.4 Exactitude:

La métrique d'Exactitude, montrée dans la figure 4.4, représente la proportion de prédictions correctes (à la fois les vrais positifs et les vrais négatifs) parmi le nombre total de cas.

- Xgboost (CUSUM), Xgboost (TFED) et RF (TFED) ont atteint les valeurs d'exactitude les plus élevées de 0,94, 0,94 et 0,93, respectivement, démontrant leur forte justesse globale dans les prédictions.

- RF (CUSUM) et Xgboost (Seuil) ont également très bien performé avec des valeurs d'exactitude de 0,94 et 0,90.
- SVM (TFED) et CNN (CUSUM) ont également montré une exactitude robuste à 0,88 et 0,92.
- Conformément aux métriques précédentes, SVM (Seuil) et CNN (Seuil) ont eu les scores d'exactitude les plus bas, à 0,53 et 0,86.



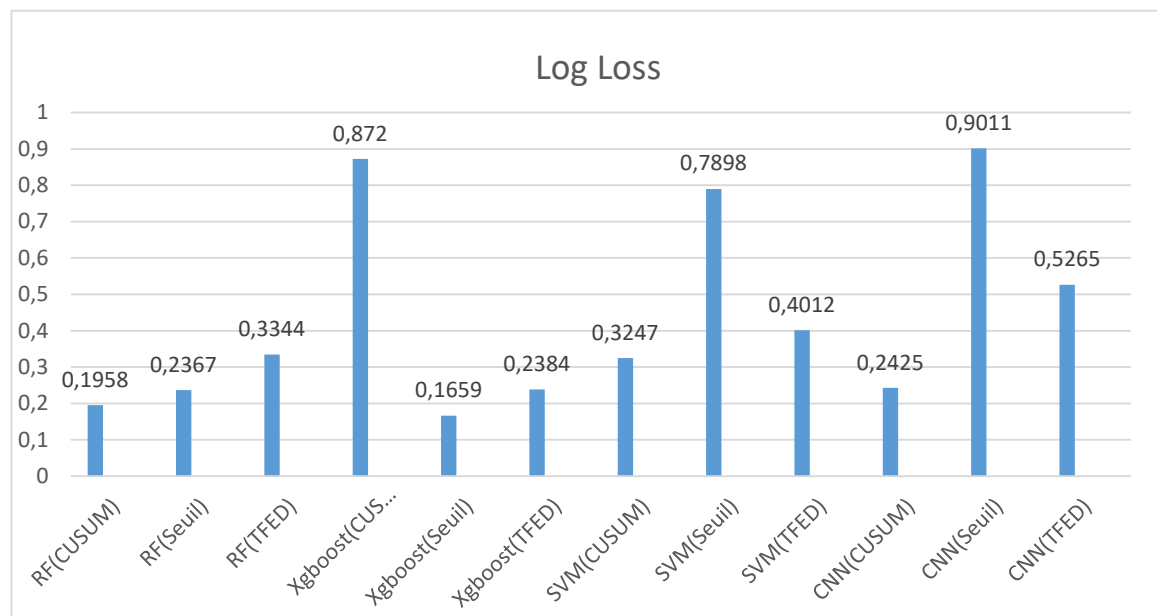
**Figure 4-4** Exactitude (Accuracy) obtenue pour chaque combinaison de détection et classification

#### 4.1.5 Perte Logarithmique (Log Loss) :

La métrique de Perte Logarithmique, présentée dans la figure 4.5, quantifie l'incertitude des prédictions en pénalisant les classifications erronées, avec des valeurs plus faibles indiquant une meilleure performance.

- Xgboost (Seuil) a présenté la plus faible Perte Logarithmique à 0,1659, indiquant des prédictions très précises et confiantes.
- RF (CUSUM), RF (Seuil), Xgboost (TFED) et CNN (CUSUM) ont également montré des valeurs de Perte Logarithmique relativement faibles, à 0,1958, 0,23, 0,2384 et 0,2425 respectivement, reflétant leurs bonnes performances.
- Inversement, Xgboost (CUSUM) et SVM (Seuil) ont eu des valeurs de Perte Logarithmique significativement plus élevées de 0,872 et 0,7898, suggérant une incertitude plus élevée et des performances prédictives moindres pour ces configurations spécifiques.

- SVM (CUSUM), SVM (TFED) et CNN (TFED) ont enregistré des valeurs de Perte Logarithmique modérées de 0,3247, 0,4012 et 0,5265.



**Figure 4-5** : Perte logarithmique obtenu pour chaque combinaison de détection et classification

#### 4.2 Discussion des Performances des Différentes Combinaisons Détection-Classification

À travers toutes les métriques évaluées (Précision, Rappel, Score F1, Exactitude et Perte Logarithmique), les méthodes de classification Xgboost et Random Forest (RF), particulièrement lorsqu'elles sont combinées avec les méthodes de détection CUSUM et TFED, ont constamment démontré des performances supérieures pour les systèmes NILM. Ces combinaisons ont généralement atteint des scores élevés en précision, rappel, score F1 et exactitude, tout en maintenant des valeurs de perte logarithmique compétitives (ou, dans le cas de Xgboost (Seuil), exceptionnellement basses). Les méthodes SVM et CNN ont montré plus de variabilité dans leurs performances, SVM (Seuil) et CNN (Seuil) étant généralement moins efficaces sur la plupart des métriques. Ces résultats suggèrent que les méthodes d'ensemble comme Xgboost et Random Forest sont très appropriées pour une détection et une classification NILM robustes en raison de leur capacité à équilibrer efficacement divers aspects de performance.

*Conclusion*

---

Cette étude a mené une évaluation comparative des algorithmes de détection d'événements et de classification au sein des systèmes de Surveillance Non-Intrusive des Charges (NILM), une technologie essentielle pour la gestion énergétique moderne. Face à l'urgence croissante de la conservation de l'énergie et aux limites des méthodes de surveillance intrusives, le NILM offre une voie prometteuse pour sensibiliser les consommateurs et optimiser la consommation domestique.

Les résultats expérimentaux, obtenus sur le jeu de données ENERTALK, ont clairement démontré que la performance globale d'un pipeline NILM est fortement influencée par la combinaison synergique de ses composants. Il a été établi que la qualité de la détection d'événements est primordiale ; les méthodes de classification les plus avancées ne peuvent compenser les lacunes d'une détection d'événements de faible qualité, comme en témoignent les performances médiocres des pipelines utilisant la détection par seuil simple.

En revanche, les méthodes de classification basées sur l'apprentissage ensembliste, à savoir XGBoost et Random Forest, se sont avérées constamment supérieures. Leur capacité inhérente à réduire la variance et le biais, à gérer des données bruyantes et à intégrer des interactions complexes entre les caractéristiques les rend particulièrement adaptées aux défis de la désagrégation des charges. Ces modèles ont atteint des niveaux élevés de précision, de rappel, de F1-score et d'exactitude, en particulier lorsqu'ils étaient associés aux méthodes de détection CUSUM et TFED, qui fournissent des événements plus robustes et temporellement mieux localisés. La performance exceptionnelle de Xgboost (Seuil) en termes de Log Loss, malgré une détection plus simple, suggère une calibration de modèle remarquable pour cette configuration spécifique. Cependant, la Perte Logarithmique élevée de Xgboost (CUSUM) malgré d'excellentes métriques d'exactitude met en lumière l'importance d'évaluer la confiance du modèle, au-delà de la simple justesse des prédictions.

En synthèse, les combinaisons Xgboost-CUSUM, Xgboost-TFED et Random Forest-CUSUM/TFED émergent comme les approches les plus robustes et efficaces pour la détection et la classification des charges dans les systèmes NILM. Ces résultats confirment la pertinence des méthodes d'ensemble pour les applications NILM et fournissent des orientations claires pour le développement de solutions de gestion énergétique plus fiables et précises.

Pour les recherches futures, il serait pertinent d'explorer des approches d'apprentissage non supervisé ou semi-supervisé afin de surmonter la dépendance aux données étiquetées, un défi majeur pour la scalabilité du NILM. L'intégration de techniques d'apprentissage en ligne pourrait également permettre aux systèmes de s'adapter dynamiquement aux nouveaux appareils ou aux changements dans les habitudes de consommation sans nécessiter de réentraînement complet. Enfin, la validation de ces pipelines sur un éventail plus large de jeux de données réels, provenant de diverses régions et types de foyers, serait essentielle pour confirmer leur généralisable et leur applicabilité à l'échelle mondiale

## ***Bibliographie***

---

- [1] K. C. A. e. al, «Is disaggregation the holy grail of energy efficiency? The case of electricity,» *Energy Policy*, vol. 52, p. 213–234, 2013.
- [2] P. A. S. e. al, «Non-intrusive load monitoring: A review,» *IEEE Transactions on Smart Grid*, pp. 1-1, 2022.
- [3] G. Hart, «Nonintrusive appliance load monitoring,» *Proceedings of the IEEE*, vol. 80, n° 112, p. 1870–1891, 1992.
- [4] A. Z. e. al, «Non-intrusive load monitoring approaches for disaggregated energy sensing: a survey,» *Sensors (Switzerland)*, vol. 12, n° 112, p. 16838–16866.
- [5] S. Makonin, *Real-time embedded low-frequency load disaggregation*, 2014.
- [6] V. Mehta, *Principles of Electronics*, Upper Saddle River, 1980.
- [7] C. Y. e. al, «Comparative study of event detection methods for non-intrusive appliance load monitoring,» *Energy Procedia*, p. 1840–1843, 2014.
- [8] W. C. e. al, «Harmonics Load Signature Recognition by Wavelets Transforms,» chez *International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT)*, 2000.
- [9] K. Lee, «Estimation of variable-speed-drive power consumption from harmonic content,» *IEEE Transactions on Energy Conversion (abrégé : IEEE Trans. Energy Convers.)*, p. 566–574, 2005.
- [10] W. Wichakool, «Modeling and estimating current harmonics of variable electronic loads,» *IEEE Transactions on Power Electronics*, p. 2803–2811, 2009.
- [11] M. e. al., «Circuit-level load monitoring for household energy management,» *IEEE Pervasive Computing (abrégé : IEEE Pervas. Comput.)*, vol. 10, p. 40–48, 2011.
- [12] F. e. al., *An experimental study on electrical signature identification of non-intrusive load monitoring (NILM) systems*, Heidelberg, 2011.
- [13] H. K. e. al., «Unsupervised disaggregation of low frequency power measurements,» chez *Proceedings of the 2011 SIAM International Conference on Data Mining (SDM'11)*, 2011.
- [14] K. D. A. e. al., «Event detection for Non Intrusive Load Monitoring,» chez *Proceedings of IECON 2012 – 38th Annual Conference of the IEEE Industrial Electronics Society*, 2012.
- [15] M. B. e. J. Voss, «Detecting patterns of appliances from total load data using a dynamic programming approach,» chez *Proceedings of the 4th IEEE International Conference on Data Mining (ICDM 2004)*, 2004.

- [16] M. B. e. al., «User-Centered Nonintrusive Electricity Load Monitoring for Residential Buildings,» *Journal of Computing in Civil Engineering*, vol. 25, n° %16, p. 471–480, 2011.
- [17] Y. J. e. al., «A Time-Frequency Approach for Event Detection in Non-Intrusive Load Monitoring,» chez *Proceedings of SPIE*, vol. 8050, 2013.
- [18] S. B. L. e. al., «Transient Event Detection in Spectral Envelope Estimates for Nonintrusive Load Monitoring,» *IEEE Transactions on Power Delivery*, vol. 10, n° %13, p. 1200–1210, 1995.
- [19] S. R. S. e. al., «Nonintrusive Load Monitoring and Diagnostics in Power Systems,» *IEEE Transactions on Instrumentation and Measurement*, vol. 57, n° %17, p. 1445–1454, 2008.
- [20] «Conv-NILM-Net, a causal and multi-appliance model for energy source separation,» chez *ECML PKDD 2022 (European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases)*, 2022.
- [21] R. G. e. al., «Energy management using non-intrusive load monitoring techniques – State-of-the-art and future research directions,» *Sustainable Cities and Society*, vol. 62, 2020.
- [22] M. F. e. al., «Thresholding methods in NILM: A comprehensive review,» *The Journal of Supercomputing*, 2023.
- [23] Z. Z. e. al., «A novel CUSUM-based approach for event detection in smart metering,» chez *IOP Conference Series: Materials Science and Engineering*, 2018.
- [24] J. W. Tukey, *Exploratory Data Analysis*, Addison-Wesley, 1977.
- [25] «Mesure non intrusive : La technologie NILM de Smart Impulse,» *Smart Impulse* . [En ligne].
- [26] Hugo Michel, «Projet NILM – Détection d’activité non intrusive à partir de la courbe de charge,» [En ligne]. Available: [https://www.hugomichel.io/project/nilm\\_project/](https://www.hugomichel.io/project/nilm_project/). [Accès le 14 juin 2025].
- [27] C. Strobl, «An introduction to recursive partitioning: Rationale, application and characteristics of classification and regression trees, bagging and random forests,» *Psychological Methods*, vol. 14, p. 323–348, 2009.
- [28] T. G. Dietterich, «An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization,» *Machine Learning*, vol. 40, p. 139–157, 2000.
- [29] A. P. e. al., «XGBoost Algorithm and Its Comparative Analysis,» *International Journal of Novel Research and Development (IJNRD)*, vol. 7, n° %112, 2022.
- [30] V. Deshpande, «Medium,» 3 avril 2023. [En ligne]. Available: <https://medium.com/@vdeshpande551/understanding-the-inner-workings-of-svm-from-data-preprocessing-to-model-deployment-cdc9d72a2d34>. [Accès le 2025 juin 10].

- [31] D. H. e. a. Hubel, «Receptive fields and functional architecture of monkey striate cortex,» *The Journal of Physiology*, vol. 195, n° %11, p. 215–243, 1968.
- [32] Y. LeCun, «Backpropagation Applied to Handwritten Zip Code Recognition,» *Neural Computation*, vol. 1, n° %14, p. 541–551, 1989.
- [33] T. Hastie, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, New York, NY, USA: Springer, 2009.
- [34] C. J. v. Rijsbergen, *Information Retrieval*, UK: Butterworth-Heinemann, 1979.
- [35] C. e. a. Shin, «The ENERTALK dataset, 15 Hz electricity consumption data from 22 houses in Korea,» *Scientific Data*, vol. 6, n° %1193, 2019.
- [36] F. P. e. al, «Scikit-learn: Machine Learning in Python,» *Journal of Machine Learning Research*, vol. 12, p. 2825–2830, 2011.
- [37] I. G. e. al, *Deep Learning*, MIT Press, 2016.
- [38] Jorgecardete, «Convolutional Neural Networks: A Comprehensive Guide,» Février 2024. [En ligne]. Available: <https://medium.com/thedeephub/convolutional-neural-networks-a-comprehensive-guide-5cc0b5eae175>. [Accès le juin 2025].
- [39] Z. Haytam, «Outliers Detection in PySpark #2 – Interquartile Range,» 2019.
- [40] Y. B. e. al, «Hybrid model based on TF-IDF, chi-square, and random forest for Arabic sentiment analysis,» *ResearchGate*, 2024.
- [41] N. S. Software, «CUSUM Charts,» [En ligne]. Available: [https://www.ncss.com/wp-content/themes/ncss/pdf/Procedures/NCSS/CUSUM\\_Charts.pdf](https://www.ncss.com/wp-content/themes/ncss/pdf/Procedures/NCSS/CUSUM_Charts.pdf).