

CHAPITRE 4

ANALYSE DES RÉSULTATS ET DISCUSSION

4.1. APERÇU DES SIGNAUX

En nous basant sur la méthodologie proposée au chapitre 3 et les différentes techniques d'extraction de caractéristiques connue dans la littérature, nous avons procédé par analyser notre signal de vibration recueilli du capteur qui a fourni des données défailtantes. Dans un premier temps pour réaliser notre système de surveillance de l'état de la machine nous avons extrait de ce signal la moyenne des fréquences crêtes. Étant donné que la fréquence de crête moyenne distingue avec succès les signaux nominaux des signaux défectueux, nous avons extrait la fréquence de crête moyenne de chaque segment de données de notre signal.

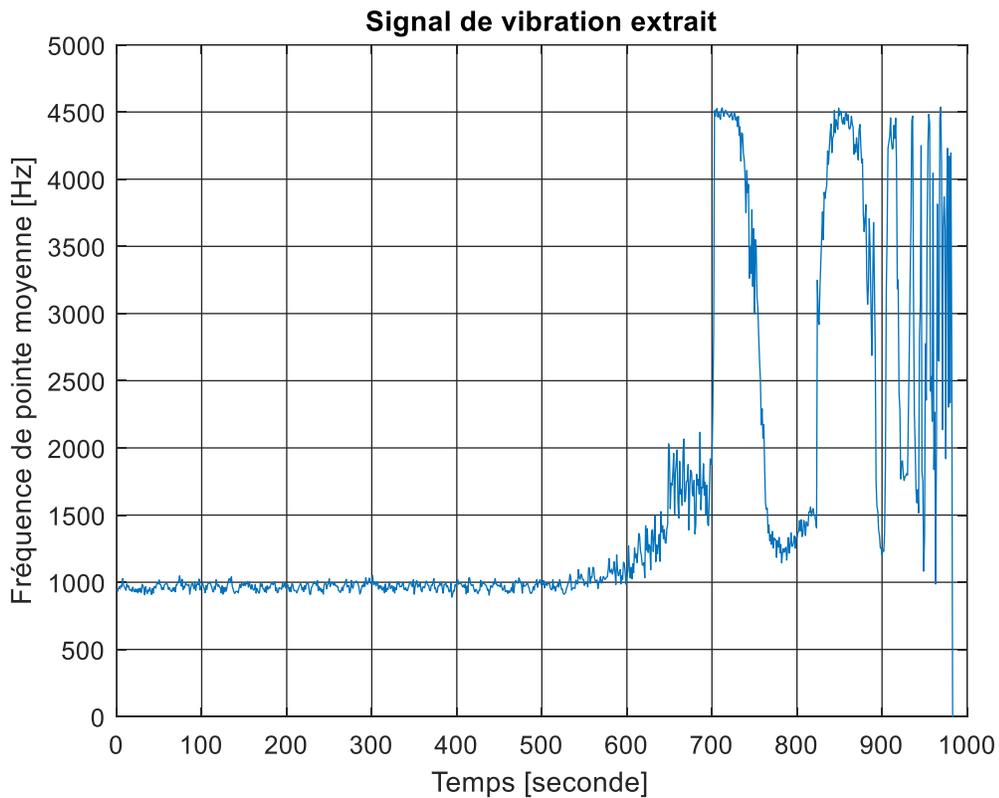


Figure 18 : Fréquences maximales moyennes extraites en fonction du temps.

4.2. ANALYSE DES CARACTÉRISTIQUES SÉLECTIONNÉES

Pour réaliser la classification des signaux grâce à l'apprentissage automatique en fonction des différentes classes que nous avons définies à savoir les trois niveaux de dégradation du signal de vibration (signal nominal, signal dégradé, signal défaillant) ; il nous faut un meilleur choix dans la sélection des caractéristiques que nous allons utiliser pour entraîner nos algorithmes de classification. Pour cela après plusieurs analyses des différentes caractéristiques nous présentons dans cette section, celle que nous avons retenue (Kurtosis et RMS) pour la réalisation de notre projet de recherche dû au fait qu'elles distinguent le milieu les niveaux de dégradation par rapport aux autres caractéristiques que nous avons extraites.

À la suite de la visualisation des signaux, nous sommes passés à l'étape de sélection des caractéristiques. Tout d'abord, nous allons présenter la méthode temporelle, ensuite nous présenterons la méthode statistique.

4.2.1. MÉTHODE TEMPS-FREQUENCE

Dans cette section, des caractéristiques représentatives sont extraites de chaque segment de données. Mais avant de sélectionner les fonctions à utiliser, nous allons tracer le spectrogramme des signaux de vibration. La visualisation des signaux dans le domaine temporel et fréquentiel peut aider à découvrir des modèles de signal qui indiquent une dégradation ou une défaillance.

Les représentations temps-fréquence décrivent comment le contenu spectral d'un signal évolue en fonction du temps. Ils peuvent être utilisés pour former les algorithmes qui peuvent identifier et extraire des modèles à partir des représentations temps-fréquence et également choisir parmi une variété de techniques qui peuvent générer des représentations temps-fréquence pour les signaux, notamment le spectrogramme.

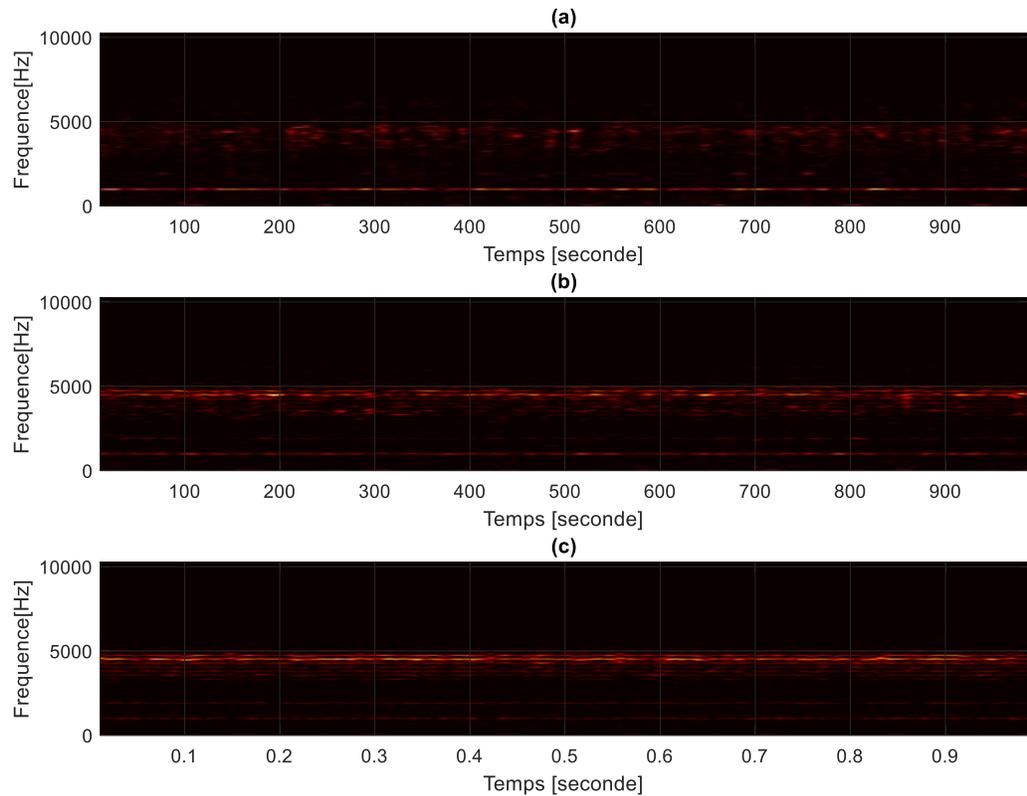


Figure 19 : Tracer du spectrogramme des signaux de vibration. (a)Signal nominal ; (b)Signal dégradé ; (c)Signal défaillant.

Avec le spectrogramme nous constatons que les trois niveaux de dégradation sont bien distingués sur la figure 15. Pour le signal nominal nous observons une concentration de la puissance du signal dans les fréquences situées au tour de 1000Hz et pour les autres types de signaux une concentration du niveau de signal dans des fréquences beaucoup plus élevées notamment pour le signal défaillant ou la majeure partie de la puissance du signal se concentre à des fréquences situées autour de 4000Hz et 5000Hz. Cela démontre l'utilité d'utiliser les méthodes temps-fréquence pour extraire les caractéristiques des signaux, car elle offre de meilleurs résultats pour entraîner nos algorithmes de classification.

4.2.2. MÉTHODE STATISTIQUE

Avec la boîte à outils Diagnostic Feature Designer du logiciel MATLAB, nous avons importé et visualisé l'ensemble des données de vibration dans le but d'extraire les caractéristiques essentielles pour réaliser la classification de nos signaux avec les algorithmes d'apprentissage automatique supervisée. Nous avons obtenu les résultats représentés sur la figure 16 qui montre que les caractéristiques obtenues avec le RMS sont meilleures, car elle distingue mieux les trois groupes de données en fonction du niveau de défaillance par rapport aux caractéristiques de Kurtosis.

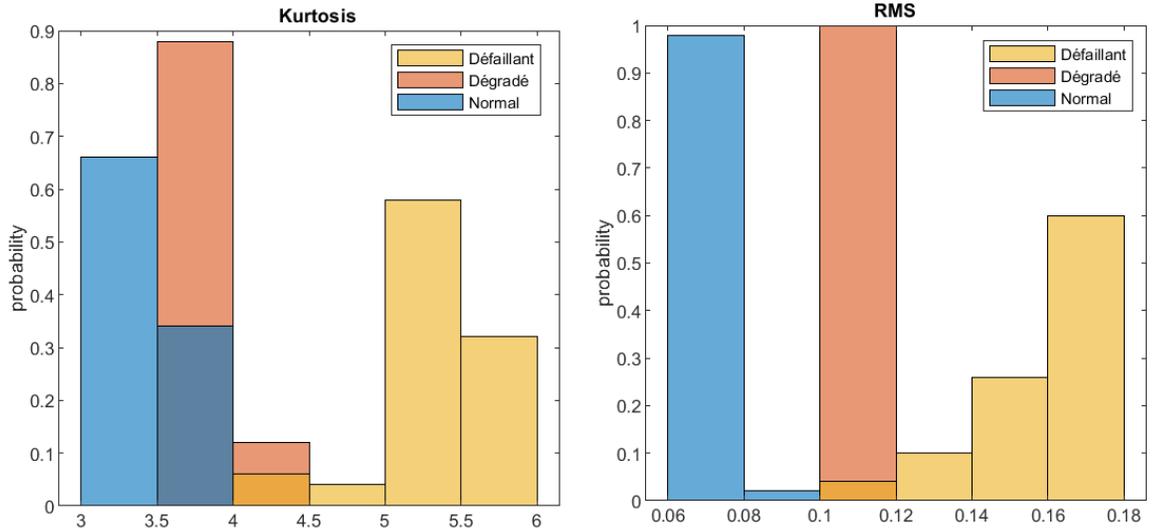


Figure 20 : Histogramme des caractéristiques des signaux de vibration.

La figure 16 nous présente l'histogramme des caractéristiques des signaux obtenues grâce à la méthode de Kurtosis et de RMS. Elle nous permet par la même occasion de comparer les deux méthodes et de constater que RMS dans le cas de notre projet offre une meilleure distinction des différents niveaux de dégradation de notre signal. Avec Kurtosis nous constatons que la probabilité pour qu'un signal nominal se situe sur une plage de valeur comprise entre [3 et 3.5] est de 65% et avec RMS cette probabilité est de 99% pour des valeurs comprises entre [0.06 et 0.08].

4.3. RÉSULTATS DES TESTS

Les résultats de notre projet de recherche montrent comment extraire des caractéristiques des données mesurées pour effectuer une surveillance de l'état d'une machine avec des capteurs installés pour recueillir les données de vibration. Sur la base des fonctionnalités extraites, des modèles dynamiques sont générés, validés et utilisés pour prévoir les pannes afin que des actions puissent être prises avant que les pannes réelles ne se produisent.

Dans l'objectif de mettre en place un système de prédiction de défaut des machines tournante capable de prédire les pannes avant que celle-ci ne produise et dans l'optique de fournir un système robuste et fiable, nous avons décidé de développer ce projet avec deux méthodes d'analyse de la machine. La première méthode est intuitive, car elle est basée sur une interprétation graphique de l'évolution la moyenne des fréquences crête dans le temps. Pour cela nous avons mis en place un système appelé système de surveillance de l'état de la machine qui va au fur et à mesure que la machine est en marche recevoir des données recueillies par les capteurs et mettre à jour le graphique qui représente les résultats du calcul de la moyenne des fréquences crête. Grâce aux seuils de bon fonctionnement qui sont fixés de base sur le graphique les opérateurs pourront vérifier s'il a une augmentation des valeurs de la moyenne et prendre des décisions d'intervenir avant que la panne ne survienne.

Cette première méthode a été développée dans le but de soutenir la seconde, qui est l'application de l'apprentissage automatique sur les signaux de nos capteurs plus précisément les algorithmes de classification des signaux détaillés dans la section de la méthodologie. Avec la combinaison de ces deux méthodes, nous serons plus aptes à interpréter et à valider un signal de dégradation ou de panne de la machine comme une vraie défaillance en limitant au maximum les risques de se tromper.

4.3.1. RÉSULTATS DU SYSTÈME DE SURVEILLANCE DE L'ÉTAT DE LA MACHINE

Dans cette section, la surveillance de l'état de la machine est effectuée à l'aide des seuils prédéfinis et des modèles dynamiques. Pour surveiller les conditions de fonctionnement de la machine, nous avons créé une alarme qui se déclenche si la fréquence de crête moyenne dépasse le seuil prédéfini. La prévision nous aide à mieux nous préparer à une panne potentielle ou même à arrêter la machine avant une panne. La figure 17 montre que les valeurs prévues de la fréquence de crête moyenne sont bien inférieures au seuil et au fur et à mesure que de nouvelles données arrivent, le système met à jour les paramètres du modèle et estime les valeurs en fonction des seuils prédéfinis. Une fois, le seuil atteint une alarme est automatiquement générée pour arrêter la machine ou informer les opérateurs d'un potentiel défaut.

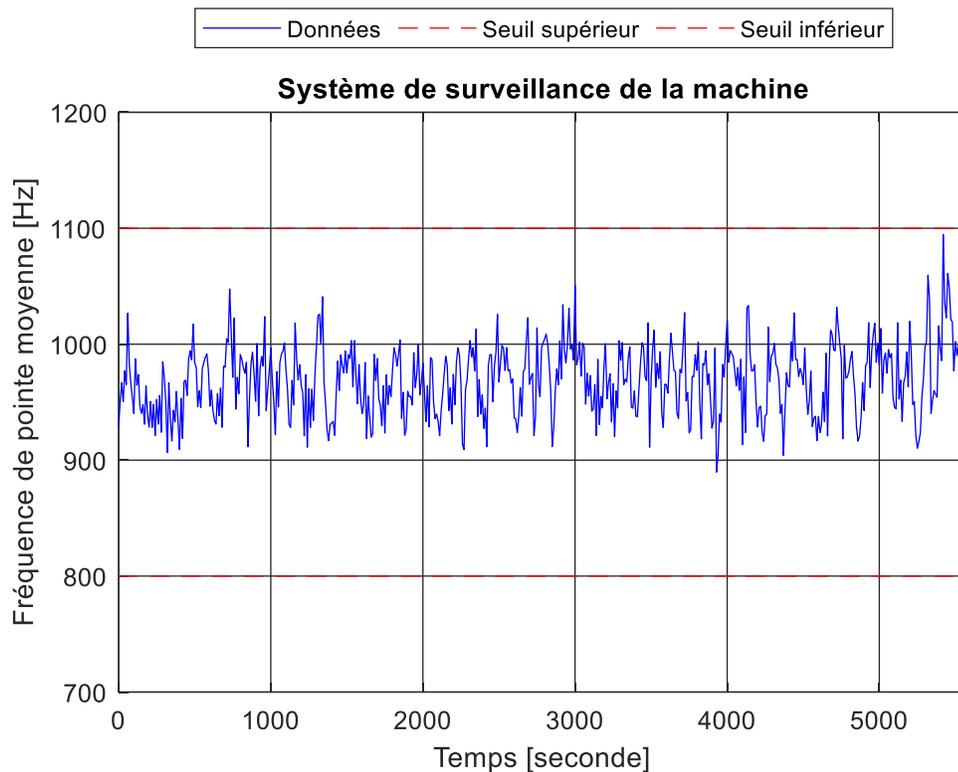


Figure 21 : Mise en œuvre du système de surveillance avec détection de panne.

4.3.2. RÉSULTATS DES TESTS DE CLASSIFICATION DES SIGNAUX

Une étape importante lors de l'utilisation de l'apprentissage automatique consiste à vérifier les performances des modèles que nous avons utilisés. L'évaluation d'un système de classification se base généralement sur un partage de la base de données en trois ensembles aussi appelés sous-ensemble :

- Un sous-ensemble de donnée d'apprentissage pour la conception du système de classification, c'est un jeu de données d'exemples utilisés pour l'apprentissage.
- Un sous-ensemble de données de test qui est indépendant de l'ensemble de données de formation, mais qui suit la même distribution de probabilité pour évaluer l'architecture du système de classification.
- Un sous-ensemble de données de validation qui est un ensemble de données exemples utilisés pour ajuster les hyperparamètres du classificateur et pour valider la pertinence de l'architecture choisie.

La distinction entre les deux premiers ensembles est nécessaire pour évaluer les performances du système de classification en ne testant pas le système avec des échantillons ayant servi pour sa conception. Le troisième sous-ensemble permet d'obtenir une validation indépendante des critères de sélection. Dans certains cas, la taille de la base de données ne permet pas d'avoir un partage en trois sous-ensembles ; dans ce cas on utilise uniquement un partage en deux : un sous-ensemble d'apprentissage et un sous-ensemble de tests. Il existe différents types de méthode pour réaliser le partage de la base de données en ensemble d'apprentissages et en ensemble de tests. Parmi ces méthodes, on trouve la validation croisée (k-fold) et la validation par holdout.

4.3.2.1. MÉTHODE DE VALIDATION CROISÉE (K-FOLD)

Une méthode d'évaluation des performances d'un algorithme d'apprentissage automatique est la validation croisée. Cette technique permet à l'algorithme de faire des prédictions en utilisant des données non utilisées pendant la phase de formation. La validation croisée partitionne un ensemble de données et utilise un sous-ensemble pour entraîner l'algorithme et les données restantes à tester. Étant donné que la validation croisée n'utilise pas toutes les données pour créer un modèle, il s'agit d'une méthode couramment utilisée pour éviter le surapprentissage pendant la formation.

Le principe de cette méthode de validation consiste à partitionner les données en k sous-ensembles (ou plis) choisis au hasard de taille à peu près égale. Un sous-ensemble est utilisé pour valider le modèle formé à l'aide des sous-ensembles restants. Ce processus est répété k fois, de sorte que chaque sous-ensemble soit utilisé exactement une fois pour la validation. L'ensemble de formation est ensuite utilisé pour former un algorithme d'apprentissage supervisé, et l'ensemble de tests est utilisé pour évaluer ses performances. Ce processus est répété plusieurs fois et l'erreur de validation croisée moyenne est utilisée comme indicateur de performance.

➤ RÉSULTATS DES TESTS AVEC LA VALIDATION CROISÉE

Avec cette approche, nous avons obtenu un taux de classification de 99.3% avec l'algorithme de Naïve Bayes qui est le meilleur parmi l'ensemble des algorithmes testés. La figure 18 nous montre les groupes de classification créée avec ce classificateur.

Pour réaliser les tests de validation avec la méthode de validation croisée, nous avons choisi de configurer nos paramètres de validation de la façon suivante : pour un ensemble de (403*20480 lignes) données, nous avons utilisé un paramètre de K égal à 5 pour avoir les meilleures performances de nos algorithmes de classification.

Avec ces paramètres, l'ensemble de données a été divisé en (403 diviser par 5) sous-ensemble pour servir de données d'apprentissage et de donnée de test. Et le résultat nous permet de constater sur la figure 18 un groupement en trois classes de nos données toujours en respectant les valeurs de Kurtosis et de RMS que nous avons obtenues lors de l'extraction de nos caractéristiques. Et avec le taux de réussite de 99.3%, nous pouvons conclure que le choix des caractéristiques utilisées est le meilleur choix. Mais toujours dans l'optique d'améliorée se taux de réussite nous avons opté de tester la deuxième méthode de validation Holdout avec des paramètres qui nous fourniront un meilleur taux de réussite.

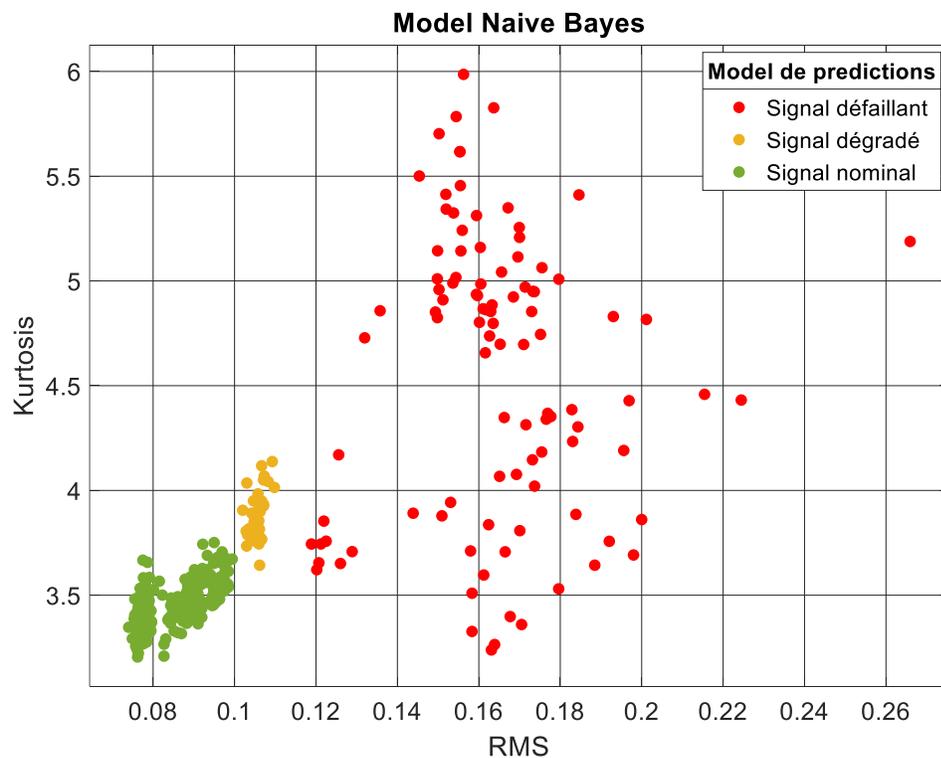


Figure 22 : Classification des signaux avec la méthode de validation croisée.

4.3.2.2. MÉTHODE DE VALIDATION HOLDOUT

Le holdout, c'est quand vous divisez votre ensemble de données en un ensemble « train » et « test ». L'ensemble d'apprentissages est ceux sur quoi le modèle est formé, et l'ensemble de tests est utilisé pour voir dans quelle mesure ce modèle fonctionne sur les données. Une division commune lors de l'utilisation de la méthode de holdout consiste à utiliser 80% des données pour la formation et les 20% restants des données pour les tests. Le principe de cette méthode de validation consiste à partitionner les données en exactement deux sous-ensembles d'un rapport spécifié pour la formation et la validation.

➤ RÉSULTATS DES TESTS AVEC LA VALIDATION HOLDOUT

La performance globale de cette approche proposée est la meilleure. Dans cette approche, nous avons obtenu un taux de classification de 100% avec tous les algorithmes testés. La figure 19 nous montre les groupes de classification créée avec cette méthode.

Pour réaliser les tests de validation avec la méthode de validation holdout, nous avons choisi de configurer nos paramètres de validation de la façon suivante : Pour un ensemble de (403*20480 lignes) données nous avons partitionné les données en deux groupes avec un pourcentage de 70% pour les données d'entraînement et un pourcentage de 30% qui seront utiliser comme données de test pour nos algorithmes de classification.

Avec ces paramètres l'ensemble de données a été divisé en (70% de 403 égales à 283) sous-ensemble pour servir de données d'apprentissage et en (30% de 403 égales à 120) de donnée de test. Et le résultat nous permet de constater sur la figure 19 un groupement en trois classes de nos données toujours en respectant les valeurs de Kurtosis et de RMS que nous avons obtenues lors de l'extraction de nos caractéristiques. Et avec un taux de réussite cette fois de 100%, nous pouvons conclure que le choix des caractéristiques utiliser est le meilleur

choix et que la validation Holdout avec des paramètres que nous avons utilisés offre un meilleur taux de réussite.

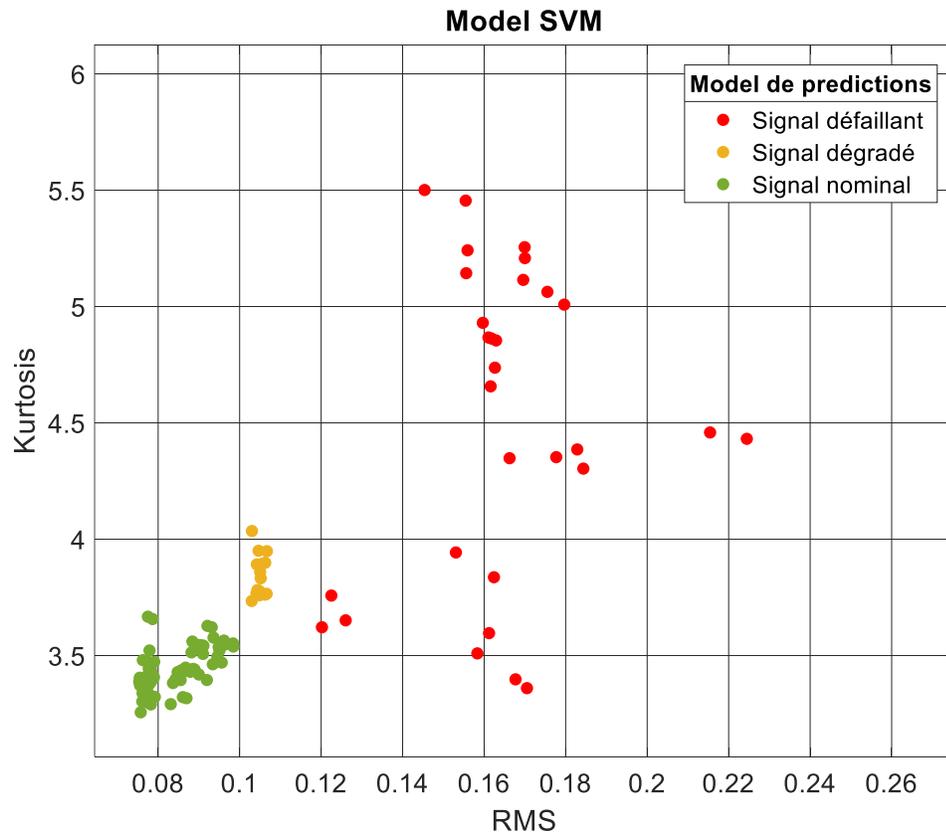


Figure 23 : Classification des signaux avec la méthode de la validation holdout.

4.3.2.3. MATRICE DE CONFUSION

L'évaluation d'un système de classification se base généralement sur la mesure du taux de classification. Pour avoir une meilleure interprétation des résultats, nous utilisons la matrice de confusion. Il s'agit d'une matrice carrée de taille $[M \times M]$ où M représente le nombre de classes. Chaque colonne de la matrice représente le nombre d'occurrences d'une classe prédite, alors que chaque ligne représente le nombre d'occurrences d'une classe réelle (ou de référence).

La matrice de confusion offre la possibilité d'analyse du taux de bonne classification (ou d'erreur) par classe. Les valeurs de la diagonale de la matrice représentent le nombre de formes correctement classifiées. La somme des valeurs par ligne correspond au nombre d'échantillons de test par classe. Le taux de classification par classe est donné par la valeur à la diagonale divisée par la somme des valeurs par ligne.

➤ **MATRICE DE CONFUSION AVEC LA VALIDATION CROISÉE**

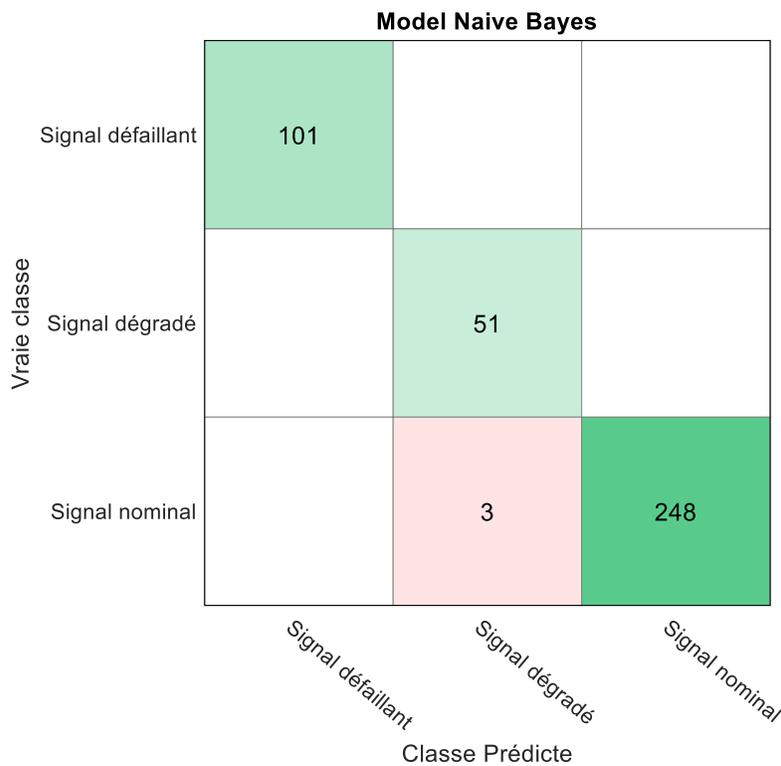


Figure 24 : Matrice de confusion avec la méthode de validation croisée.

La figure 20 nous présente la matrice de confusion obtenue avec la méthode de validation croisée. Nous pouvons valider le résultat du taux de réussite de 99.3% obtenue avec cette méthode, car avec un nombre d'échantillons de 251 du signal nominal utilisé, nous avons obtenu une reconnaissance de 248 échantillons et 3 échantillons du signal nominal ont été confondus comme échantillons de signal dégradé. Pour le reste, des échantillons ils ont très bien été repartis par l'algorithme de classification Naïve Bayes utiliser avec la validation croisée.

➤ **MATRICE DE CONFUSION AVEC LA VALIDATION HOLDOUT**

Model SVM

	Signal défaillant		
Vraie classe	Signal défaillant	30	
	Signal dégradé		15
	Signal nominal		75
		Signal défaillant	Signal dégradé
		Classe prédicte	

Figure 25 : Matrice de confusion avec la méthode de la validation holdout.

La figure 21 nous présente la matrice de confusion obtenue avec la méthode de validation holdout. Nous pouvons valider le résultat du taux de réussite de 100% obtenue avec cette méthode, car avec un nombre de 75 échantillons du signal nominal utilisé nous avons une reconnaissance des 75 échantillons comme échantillons de signal nominal. Et ces pareils avec les deux autres classes de signaux, nous observons une réussite de classification de 100% aussi. Donc l'ensemble des échantillons ont très bien été répartis par l'algorithme de classification SVM utiliser avec la validation holdout.

La méthode de validation holdout est généralement utilisée lorsqu'on dispose d'un très grand ensemble de données pour construire le modèle initial du projet de classification de signaux. Et la validation croisée utilise plusieurs fractionnements d'entraînement et de test, il faut donc plus de puissance de calcul et de temps pour l'exécution lors de l'utilisation de la méthode de validation croisée.

Les tableaux 4 et 5 ci-dessous montrent la comparaison des taux de réussites entre les différents types d'algorithmes de classification que nous avons entraînés et testés. Nous pouvons donc constater qu'avec la méthode de la validation croisée sur le tableau 4, nous avons obtenu un taux de 98.7% pour l'algorithme arbre de décision et un taux de 99.3% pour les autres types d'algorithmes à savoir le SVM, KNN et Naïve Bayes. Et pour la méthode de validation holdout sur le tableau 5, nous avons obtenu un taux de 100% pour les différents types d'algorithmes utilisés.

Tableau 4 : Résultats des tests avec la validation croisée

Algorithmes	Résultats
Arbre de décision	98.7%
Naïve Bayes	99.3%
SVM	99.3%
KNN	99.3%

Tableau 5 : Résultats des tests avec la validation holdout

Algorithmes	Résultats
Arbre de décision	100%
Naïve Bayes	100%
SVM	100%
KNN	100%

CHAPITRE 5

CONCLUSION ET PERSPECTIVES

5.1. CONCLUSION

Le présent travail a porté sur la conception d'un outil de surveillance de l'état d'une machine tournante et d'un outil de classification de signaux grâce à l'apprentissage automatique. Une étude de la littérature a été menée dans un premier temps. Et cela nous a permis d'identifier les technologies et outils existants pour surveiller les machines et les algorithmes d'apprentissage machine disponible pour analyser les signaux vibrations.

Le projet de recherche présenté dans cette étude propose, une solution permettant d'exploiter le potentiel des modèles d'intelligence artificielle pour l'amélioration du processus de maintenance dans l'industrie par la reconnaissance des défauts des machines. Le contexte décrit dans le chapitre 1 montre l'importance de posséder un programme de maintenance prédictive efficace. En effet, les arrêts non planifiés des chaînes de production pour cause de bris sont des sources de perte financière et de perte de temps affectant directement la santé budgétaire et la compétitivité des entreprises.

En conclusion, les informations qui sont obtenues grâce aux systèmes de surveillance et de l'apprentissage automatiques peuvent aider les responsables de la maintenance à préparer les ressources matérielles et humaines nécessaires avant que survienne une défaillance. Ainsi, les stratégies de maintenance traditionnelles impliquant une maintenance corrective et préventive peuvent être remplacées par une maintenance prédictive. Avec un excellent taux de reconnaissances à 100 %, cela démontre que l'utilisation de la reconnaissance des signaux de vibration peut permettre une prédiction plus précise des différents éléments en défaillances sur les machines. Par la suite de ce chapitre, nous allons mettre en lumière les perspectives d'amélioration et les apports d'un point de vue personnel de notre projet de recherche.

5.2. PERSPECTIVE

Malgré la qualité des recherches et le succès de l'approche proposée dans ce mémoire, il subsiste plusieurs limitations. Toutefois, pour améliorer notre système qui présente une insuffisance due au fait que nous avons travaillé sur des données expérimentales et non sur des données obtenues sur des machines en pleine production, il faudra procéder par une phase pilote de test en situation réelle avant de se lancer dans l'implémentation du système dans un milieu industriel.

De plus, durant ce projet de recherche, nous avons acquis de nouvelles connaissances en intelligence artificielle avec les algorithmes de classifications. Cette expérience nous a également donné la possibilité d'améliorer nos compétences en méthodologie de recherche, en communication et en rédaction. Nous avons aussi travaillé sur un article scientifique qui a été accepté et sera présenté lors d'une conférence internationale traitant des mêmes thématiques et sera publié dans un journal pour évaluation de notre projet de recherche.

Pour finir, cette maîtrise nous a donné l'envie de poursuivre dans le domaine de la recherche et de l'enseignement universitaire avec pour objectif de contribuer aux avancées scientifiques dans le monde et d'apprendre toujours plus.