

"La vraie sagesse est moins présomptueuse que la folie. Le sage doute souvent et change d'avis ;
l'insensé est obstiné et ne doute pas ; il sait tout sauf sa propre ignorance."

— Akhenaton, pharaon de la XVIIIe dynastie —

Table des matières

Résumé	3
Abstract	4
Remerciements	5
Introduction	10
Chapitre 1: Contexte générale, définitions et état de l'art	12
1.1 Contexte général	13
1.2. Intelligence Artificielle et modèles prédictifs	15
1.2.1 Machine learning	16
1.2.2 Deep learning	16
1.2.3 Évaluation des modèles prédictifs	17
1.2.4 Algorithmes d'optimisation et fonctions d'activation	19
1.3. Etude de cas dans la littérature	21
1.4.1 Etude n°1: Prediction of Blast-Induced Ground Vibrations: A Comparison Between Empirical and Artificial-Neural-Network Approaches	23
1.3.2 Etude n°2: A Novel Artificial Intelligence Approach to Predict BlastInduced Ground Vibration in Open-Pit Mines Based on the Firefly Algorithm and Artificial Neural Network.	26
1.3.3 Etude n°3: A new combination of artificial neural network and K-nearest neighbors models to predict blast-induced ground vibration and air-overpressure.	29
1.3.4 Etude n°4:Predicting blasting propagation velocity and vibration frequency using artificial neural networks.	31
1.3.5 Etude n°5:Prediction of blast-induced ground vibration using artificial neural networks.	35
Chapitre 2 : Collecte, analyse et préparation des données	38
2.1 Processus de collecte des données dans les sites miniers	39
2.2 Forme et prétraitement des données	41
2.2.1 La forme des données	41
2.2.2 Prétraitement des données	41
2.3 Analyse statistique des données par régression linéaire et détermination de l'intervalle de confiance	43
2.3.1 Loi d'atténuation	43
2.3.2 Détermination de L'Intervalle de confiance	45
Chapitre 3: Modèles de prédiction des vibrations induites par les tirs à l'explosif à l'aide de l'Intelligence Artificielle	46
3.1 Développement et optimisation d'un Réseau de Neurones Artificiels (DL)	47
3.1.1 Évaluation et validation du modèle de Réseau de Neurones (Train set)	50
3.1.2 Évaluation et validation du modèle de Réseau de Neurones (Test set)	51

3.1.3 Test du modèle sur de nouvelles données	52
3.2 Interprétation des résultats	52
Chapitre 4: Augmentation du jeu de données, développement et déploiement web des modèles prédictifs	53
4.1 Techniques de l'augmentation de données utilisées	54
4.1.1 Interpolation	54
4.1.2 Jittering	54
4.1.3 Combinaison des deux méthodes	55
4.2 Développement et optimisation d'un Réseau de Neurones Artificiels (DL)	56
4.2.1 Évaluation et validation du modèle de Réseau de Neurones (Train set)	56
4.2.2 Évaluation et Validation du Modèle de Réseau de Neurones (Test set)	57
4.2.3 Test du modèle sur de nouvelles données	58
4.2.4 Comparaison entre les résultats du modèle avant et après augmentation	59
4.3 Déploiement web des modèles prédictifs	60
4.3.1 Framework: Flask	60
4.3.2 Développement de l'application web	61
Conclusion et perspectives	62
Références	65

Résumé

Cette étude porte sur la prédiction des vibrations induites par les tirs à l'explosif en utilisant des techniques d'intelligence artificielle, notamment le développement de réseaux de neurones artificiels (ANN). Le travail consiste à concevoir, entraîner, et évaluer ces modèles prédictifs afin d'améliorer leur précision. Une attention particulière a été portée à l'augmentation des données, en utilisant des techniques comme l'interpolation et le jittering, pour renforcer la capacité du modèle à généraliser et à mieux s'adapter aux variations des données réelles.

En complément, une interface web a été développée pour permettre l'utilisation pratique des modèles prédictifs. Cette interface, construite avec le framework Flask, permet aux utilisateurs de tester les modèles avec de nouvelles données et d'obtenir des prédictions de manière intuitive.

L'intégration de ces éléments assure non seulement une amélioration des performances des modèles mais aussi une accessibilité aisée à ces outils d'intelligence artificielle.

Abstract

This study focuses on the prediction of vibrations induced by blasting using artificial intelligence techniques, in particular the development of artificial neural networks (ANNs). The work involved designing, training and evaluating these predictive models in order to improve their accuracy. Particular attention was paid to data augmentation, using techniques such as interpolation and jittering, to enhance the model's ability to generalize and better adapt to variations in real data.

In addition, a web interface has been developed to enable practical use of the predictive models. This interface, built with the Flask framework, enables users to test the models with new data and obtain predictions intuitively. The integration of these elements ensures not only improved model performance, but also greater accessibility to these artificial intelligence tools.

Remerciements

Je tiens à exprimer ma profonde gratitude à **Pr. Moulay Ali Bekri** pour m'avoir offert l'opportunité de réaliser ce projet, une expérience inestimable qui a enrichi mes compétences et mes connaissances.

Je tiens également à exprimer ma sincère gratitude au **Pr. Imad Kadiri** pour son encadrement exceptionnel et son aide précieuse tout au long de la réalisation de ce travail. Son expertise approfondie et ses conseils avisés ont été essentiels pour orienter et faire progresser ce projet de manière significative.

Je tiens à exprimer ma reconnaissance à **Pr. Mrani Nabil** et **Dr. Ouazzani Nassima** pour leurs remarques pertinentes et leur encadrement attentif, qui ont grandement contribué à la qualité de ce travail.

Enfin, je souhaite exprimer ma profonde gratitude à mes parents, ma sœur et mes amis pour leur soutien indéfectible et leur encouragement constant. Leur présence et leur motivation ont été des sources inestimables de force et d'inspiration, me permettant ainsi de mener à bien ce projet avec détermination et succès.

Liste des figures, tableaux et abréviations

Liste des Figures

Fig(1.1) : John McCarthy (1927-2011) <i>et</i> Alan Turing (1912-1954)
Fig(1.2) : L'intelligence artificielle et ses sous-domaines
Fig(1.3) : Vue du ciel de la zone d'étude, Pérou
Fig(1.4) : Le modèle MiniMate
Fig(1.5) : Résumé du modèle
Fig(1.6) : Vue du ciel de la zone d'étude, Vietnam
Fig(1.7) : Vue du ciel de la zone d'étude, Iran
Fig(1.8) : La structure du modèle
Fig(1.9) : Vue du ciel de la zone d'étude, Espagne
Fig(1.10) : La structure du modèle
Fig(1.11) : Résultat Test-set de PPV(mm/s)
Fig(1.12) : Vue du ciel de la zone d'étude, Iran
Fig(1.13) : Résultat Test-set de PPV(mm/s)
Fig(2.1) : Processus des tirs à l'explosif
Fig(2.2) : Interval de confiance des données
Fig(3.1) : Architecture du modèle
Fig(3.2.1) : Les Points prédits par rapport aux points mesurés (Train)
Fig(3.2.2) : Les Points prédits dans l'intervall de confiance (Train)
Fig(3.3.1) : Les Points prédits par rapport aux points mesurés (Test)

Fig(3.3.2) : Les Points prédits dans l'intervale de confiance (Test)
Fig(4.1.1) : Les Points prédits par rapport aux points mesurés (Train)
Fig(4.1.2) : Les Points prédits dans l'intervale de confiance (Train)
Fig(4.2.1) : Les Points prédits par rapport aux points mesurés (Test)
Fig(4.2.2) : Les Points prédits dans l'intervale de confiance (Test)
Fig(4.3) : Interface utilisateur web

Liste des tableaux

Tab(1.1) : Les études sur l'utilisation de l'intelligence artificielle et de l'apprentissage automatique et profond pour la prédiction des vibrations induites par les tirs à l'explosif
Tab(1.2) : Paramètres du réseau de neurones
Tab(1.3) : Les valeurs moyennes de RMSE , R2
Tab(1.4) : Paramètres du réseau de neurones
Tab(1.5) : Paramètres du réseau de neurones
Tab(3.1) : Paramètres du réseau de neurones
Tab(3.2) : Test du modèle
Tab(4.1) : Test du modèle après augmentation
Tab(4.2) : Comparaison entre les deux modèle

Liste des abréviations

IA	Intelligence Artificielle
ML	Machine Learning
DL	Deep Learning
Adam	Adaptive Moment Estimation
ReLU	Rectified Linear Unit
ANN	Artificial Neural Network
CNN	Convolutional Neural Networks
RNN	Recurrent Neural Networks
SVM	Support Vector Machine
SVR	Support Vector Regressor
KNN	K-Nearest Neighbors
FFA-ANN	Firefly Algorithm -Artificial Neural Network
CART	Classification and Regression Trees
RMSE	Root-Mean-Square Deviation
MSE	Mean Squared Error
MAE	Mean Absolute Error
R²	R-Squared_Coefficient of determination
BIGV	Blast Induced Ground Vibrations
PPV	Peak Particle Velocity
AOp	Aerial Overpressure

Introduction

L'intelligence artificielle (IA) et l'apprentissage automatique (Machine Learning, ML) connaissent une croissance rapide dans divers domaines, y compris le Génie Civil. Ce domaine, crucial pour le développement des infrastructures et des bâtiments, bénéficie de manière significative des avancées en IA et ML, en améliorant la précision, l'efficacité et la sécurité des projets.

L'IA permet de simuler des processus cognitifs humains, tels que la reconnaissance de motifs et la prise de décision, tandis que l'apprentissage automatique permet aux systèmes de s'améliorer automatiquement à partir de l'expérience. L'apprentissage profond (Deep Learning, DL), une branche avancée de l'IA, utilise des réseaux de neurones pour modéliser des structures de données complexes et extraire des caractéristiques pertinentes.

Dans le Génie Civil, ces technologies sont utilisées pour la maintenance prédictive des infrastructures, la gestion des ressources, la modélisation et la simulation de scénarios complexes, ainsi que la sécurité et la surveillance des chantiers. Cependant, une application spécifique de ces technologies est particulièrement pertinente pour notre étude : la prédiction des vibrations induites par les tirs à l'explosif.

Les tirs à l'explosif sont couramment utilisés dans le Génie Civil pour des opérations comme le creusement de tunnels, l'extraction minière et la démolition de structures. Cependant, ces explosions génèrent des vibrations qui peuvent impacter les structures environnantes et présenter des risques pour la sécurité. La prédiction précise de ces vibrations, mesurées en termes de vitesse particulaire maximale (Peak Particle Velocity, PPV) en mm/s, est essentielle pour minimiser ces impacts.

L'intégration de l'IA, du ML et du DL permet de développer des modèles prédictifs sophistiqués capables d'estimer avec précision l'intensité et la propagation des vibrations. Ces modèles utilisent des données historiques sur les tirs à l'explosif, des caractéristiques géotechniques du site et des paramètres spécifiques des tirs pour fournir des prévisions fiables. En s'appuyant sur ces prédictions, les ingénieurs peuvent planifier et mettre en œuvre des dispositions efficaces pour protéger les infrastructures et assurer la sécurité.

Ainsi, l'utilisation de l'intelligence artificielle pour la prédiction des vibrations induites par les tirs à l'explosif ouvre de nouvelles perspectives pour améliorer la précision des prévisions, augmenter la sécurité et optimiser les coûts. Les avancées en IA et ML offrent des outils puissants pour relever ce défi complexe du Génie Civil, en rendant les projets plus résilients et durables.

Chapitre 1 : Contexte générale, définitions et état de l’art

1.1 Contexte général

Techniquement, le principal objectif du tir à l'explosif est de fracturer le massif rocheux de manière à obtenir des blocs compatibles avec les engins de reprise ou de chargement et les installations de concassage-broyage. Une fragmentation optimale améliore l'efficacité de broyage et permet d'économiser des milliers de kilowatts-heures d'énergie. Il est également souhaitable que le tas abattu ait une forme et une compacité adaptées à la mobilité des engins de chargement.

Parmi les objectifs à atteindre, on cite une régularité et un faible endommagement des fronts d'exploitation, facilitant le positionnement et la réalisation des volées suivantes. Enfin, un objectif de plus en plus prioritaire est de réduire les nuisances au voisinage du site d'exploitation. Les explosifs sont utilisés dans la plupart des projets de développement à grande échelle tels que les lignes à grande vitesse, les projets routiers et les centrales hydroélectriques pour détruire les roches et les bétons. Plusieurs techniques peuvent être employées pour améliorer la productivité du site, comme l'augmentation des diamètres du trou de forage, l'ajustement de la charge totale ou le contrôle de la charge par délai. Cependant, contrôler l'énergie explosive impliquée dans le dynamitage des mines est difficile, car elle est une source majeure de perturbations telles que les vibrations, les bruits et les projections de fragments de roche.

Les vibrations induites par les tirs à l'explosif (Blast Induced Ground Vibrations, BIGV) peuvent poser divers problèmes en fonction de l'origine des vibrations, de la nature des constructions soumises à ces vibrations, et de la densité et sensibilité des populations environnantes. Avec l'urbanisation croissante autour des sites d'exploitation, l'établissement de seuils de confort et de sécurité est devenu nécessaire pour réduire les dommages tels que les bruits, les projections de fragments rocheux et les vibrations. Les vibrations sont quantifiées en termes de vitesse de pointe particulaire (Peak Particle Velocity, PPV) et de sa fréquence associée. L'intensité de ces vibrations dépend de divers paramètres contrôlables (plan de tir, caractéristiques de l'explosif) et incontrôlables (caractéristiques mécaniques de la roche, géologie, présence de l'eau).

Historiquement, l'effet de ces paramètres sur la PPV a été évalué statistiquement par des méthodes simples basées sur la régression linéaire. Cependant, ces dernières années, l'introduction de l'intelligence artificielle (IA) et de l'apprentissage automatique (Machine Learning, ML) a révolutionné la prédiction des vibrations induites par les tirs à l'explosif. Les réseaux de neurones

artificiels (Artificial Neural Networks, ANN) et d'autres techniques de ML, comme la régression par vecteurs de support (Support Vector Regression, SVR) et les arbres de décision (Decision Trees), permettent de modéliser des relations complexes entre les paramètres de tir et la réponse vibratoire.

De plus, l'apprentissage profond (Deep Learning, DL) utilise des architectures avancées comme les réseaux de neurones convolutifs (Convolutional Neural Networks, CNNs) et les réseaux de neurones récurrents (Recurrent Neural Networks, RNNs) pour capturer des niveaux d'abstraction élevés dans les données. Ces techniques prennent en compte plusieurs variables d'entrée, telles que le nombre de trous, nombre de rangés, l'espacement et la longueur des trous de forage, le poids de la charge explosive par délai, le bourrage, la profondeur maximale, (ND) et la distance entre le point de tir et le point de prédiction des vibrations en termes de PPV et de fréquence. Bien que les modèles ANN offrent des résultats satisfaisants avec un grand nombre de données, ils fonctionnent souvent comme des boîtes noires, faisant abstraction des aspects physiques des phénomènes mis en jeu.

Les différentes méthodes et approches mentionnées servent à évaluer et à prédire les vibrations induites dans la roche à divers points du site d'exploitation. La question suivante concerne les dommages potentiels que ces vibrations peuvent causer aux structures environnantes, nécessitant l'établissement de seuils de vibrations en fonction des enjeux à protéger. Sur la base des phénomènes de résonance, la réponse des structures aux vibrations du sol dépend de la fréquence des vibrations par rapport à la fréquence fondamentale de la structure. Les normes, comme la USBMRI 8507, la norme britannique BS 7385 et la norme allemande DIN 4150, fixent les seuils de vibrations autorisés en fonction de la PPV et de sa fréquence associée.

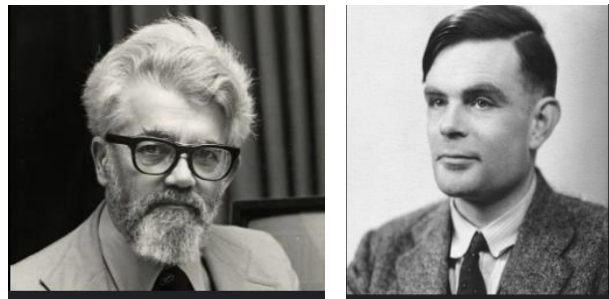
C'est dans ce contexte riche et complexe que notre étude a été entreprise. Des modèles intelligents ont été développés afin de prédire la PPV des vibrations induites par les tirs à l'explosif. Des mesures de vibrations ont été réalisées sur plusieurs sites avec des caractéristiques géologiques, hydrologiques, topographiques, et géomécaniques bien documentées. Ces mesures ont été exploitées rigoureusement pour entraîner et valider des modèles basés sur des techniques d'apprentissage automatique et d'apprentissage profond. Malgré les défis techniques, les résultats

obtenus offrent des perspectives enrichissantes pour la prédiction des vibrations et l'évaluation des impacts sur les structures et les populations environnantes.

1.2. Intelligence Artificielle et modèles prédictifs

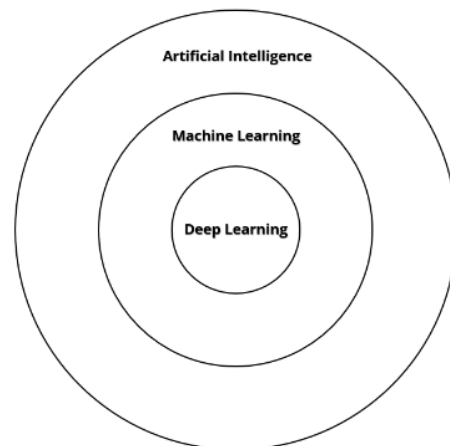
L'intelligence artificielle (IA) est un domaine de l'informatique qui vise à créer des machines capables de simuler l'intelligence humaine. L'histoire de l'IA remonte aux années 1950, lorsque des pionniers figurent dans la **figure 1.1** comme *Alan Turing* et *John McCarthy* ont commencé à explorer la possibilité de machines intelligentes. Turing a proposé le célèbre, “test de *Turing*” pour évaluer l'intelligence d'une machine, tandis que McCarthy a organisé la première conférence sur l'IA en 1956, marquant la naissance officielle du domaine.

Figure 1.1: John McCarthy (1927-2011) et Alan Turing (1912-1954)



Au cours des décennies suivantes, l'IA a connu des hauts et des bas, souvent appelés "hivers de l'IA", où l'enthousiasme et le financement ont fluctué. Cependant, les avancées technologiques et théoriques ont permis des progrès significatifs. Les années 1980 ont vu l'émergence des systèmes experts, capables de reproduire le raisonnement humain dans des domaines spécifiques. Les années 2000 et 2010 ont été marquées par l'essor de l'apprentissage automatique (Machine Learning, ML) et l'apprentissage profond (Deep Learning, DL)

Figure 1.2 L'intelligence artificielle et ses sous-domaines



(**figure 1.2**), grâce à la disponibilité de grandes quantités de données et à la puissance de calcul accrue.

1.2.1 Machine learning

L'apprentissage automatique (Machine Learning) est une branche de l'intelligence artificielle qui permet aux systèmes de s'améliorer automatiquement à partir de l'expérience (Bishop, 2006). Dans le contexte de la régression, utilisée pour prédire des valeurs continues, plusieurs techniques sont couramment employées. Parmi celles-ci, la régression par vecteurs de support (Support Vector Regression, SVR) est notable pour sa capacité à gérer des données non linéaires grâce à l'utilisation de noyaux (Drucker et al., 1997). Les arbres de décision (Decision Trees) sont également populaires en raison de leur simplicité et de leur capacité à capturer des relations complexes entre les variables (Quinlan, 1986). D'autres techniques incluent la régression linéaire, qui modélise la relation entre les variables par une ligne droite (Seber & Lee, 2012), la régression des forêts aléatoires (Random Forest Régression), qui utilise un ensemble d'arbres de décision pour améliorer la précision, et les réseaux de neurones artificiels, qui sont particulièrement puissants pour modéliser des relations complexes grâce à leurs architectures multicouches (LeCun et al., 2015). Ces techniques sont essentielles pour construire des modèles prédictifs robustes dans divers domaines, y compris le génie civil (Zeng et al., 2021).

1.2.2 Deep learning

L'apprentissage profond (Deep Learning) est une sous-branche avancée de l'apprentissage automatique qui utilise des réseaux de neurones artificiels composés de multiples couches pour modéliser des relations complexes dans les données. Dans le contexte de la régression, le Deep Learning est particulièrement important en raison de sa capacité à gérer des ensembles de données volumineux et hétérogènes tout en capturant des patterns complexes et non linéaires qui seraient difficiles à modéliser avec des techniques de régression traditionnelles.

Les réseaux de neurones profonds, sont capables de transformer les données brutes en représentations de haut niveau grâce à leur architecture multicouche. Cette capacité est cruciale pour des applications où les relations entre les variables sont très complexes, comme dans la prédiction des vibrations induites par les tirs à l'explosif. Le Deep Learning peut intégrer une multitude de caractéristiques géotechniques, environnementales et opérationnelles pour fournir des prévisions précises et fiables de la vitesse particulière maximale (PPV).

Ainsi, l'importance du Deep Learning dans la régression réside dans sa robustesse et sa flexibilité, permettant de construire des modèles prédictifs hautement performants. Ces modèles sont essentiels pour optimiser les décisions et les stratégies dans divers domaines, incluant le génie civil, en augmentant la précision des prédictions et en réduisant les risques associés.

1.2.3 Évaluation des modèles prédictifs

Pour évaluer la performance d'un modèle de régression, plusieurs métriques sont couramment utilisées. Ces métriques permettent de quantifier l'écart entre les valeurs prédites par le modèle et les valeurs réelles observées. Parmi les métriques les plus populaires, on trouve l'erreur quadratique moyenne (Mean Squared Error, MSE), la racine carrée de l'erreur quadratique moyenne (Root Mean Squared Error, RMSE), l'erreur absolue moyenne (Mean Absolute Error, MAE) et le coefficient de détermination (R^2).

Ces métriques sont détaillés ci-dessous :

- **Erreur quadratique moyenne (Mean Squared Error, MSE):**

La MSE mesure la moyenne des carrés des erreurs, c'est-à-dire les différences entre les valeurs prédites et les valeurs réelles. La formule mathématique est :

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \text{ (éq. 1)}$$

où n est le nombre de points de données, y_i est la valeur réelle, et \hat{y}_i est la valeur prédite.

- **Racine carrée de l'erreur quadratique moyenne (Root Mean Squared Error, RMSE):**

La RMSE est la racine carrée de la MSE. Elle donne une indication de l'erreur moyenne en unités de la variable observée. La formule est :

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \text{ (éq. 2)}$$

- **Erreur absolue moyenne (Mean Absolute Error, MAE):**

La MAE mesure la moyenne des valeurs absolues des erreurs. Contrairement à la MSE et à la RMSE, elle ne pénalise pas les grandes erreurs de manière disproportionnée. La formule est :

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (\text{éq. 3})$$

- **Coefficient de détermination (R^2):**

Le coefficient de détermination, ou R^2 , mesure la proportion de la variance des données observées qui est expliquée par le modèle. Un R^2 proche de 1 indique un bon ajustement du modèle aux données. La formule est :

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (\text{éq. 4})$$

où \bar{y} est la moyenne des valeurs réelles.

Ces métriques fournissent des informations complètes sur la performance d'un modèle de régression. La MSE et la RMSE sont particulièrement sensibles aux grandes erreurs, tandis que la MAE offre une mesure plus robuste aux outliers. Le coefficient R^2 permet de comprendre la proportion de variance expliquée par le modèle, offrant une vision globale de sa précision. L'utilisation combinée de ces métriques permet d'obtenir une évaluation complète et équilibrée des performances d'un modèle de régression, essentiel pour des applications complexes comme la prédiction des vibrations induites par les tirs à l'explosif dans le génie civil.

1.2.4 Algorithmes d'optimisation et fonctions d'activation

- **Algorithmes d'optimisation : ADaptive Moment estimation**

L'algorithme Adam (Adaptive Moment Estimation) est une méthode d'optimisation largement utilisée en apprentissage profond pour ajuster les poids des réseaux de neurones. Il combine les avantages des algorithmes de descente de gradient stochastique (SGD) avec moment, en utilisant non seulement la moyenne des gradients mais aussi la variance des gradients. Cela permet à Adam de s'adapter dynamiquement au taux d'apprentissage pour chaque paramètre.

L'algorithme Adam est défini par les équations suivantes :

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \text{ (éq. 5)}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \text{ (éq. 6)}$$

$$\widehat{m}_t = \frac{m_t}{1 - \beta_1^t} \text{ (éq. 7) , } \widehat{v}_t = \frac{v_t}{1 - \beta_2^t} \text{ (éq. 8)}$$

$$\theta_t = \theta_{t-1} - \frac{\alpha m_t}{\sqrt{\widehat{v}_t} + \epsilon} \text{ (éq. 9)}$$

Où :

- g_t est le gradient de la fonction de coût à l'étape t .
- β_1 et β_2 sont les coefficients d'exponentiation pour le moment de premier et second ordre.
- ϵ est un petit terme ajouté pour éviter la division par zéro.
- α est le taux d'apprentissage.

- **Fonctions d'activation : Rectified Linear Unit et Linear**

Les fonctions d'activation jouent un rôle crucial dans le fonctionnement des réseaux de neurones, car elles introduisent des non-linéarités dans le modèle. Cette non-linéarité permet au réseau de neurones de capturer des relations complexes entre les variables d'entrée et de sortie, rendant le modèle capable d'apprendre et de représenter des phénomènes variés et difficiles à modéliser avec des techniques linéaires classiques.

- La fonction **ReLU** (Rectified Linear Unit) est l'une des fonctions d'activation les plus largement utilisées dans les réseaux de neurones profonds. Sa popularité provient de sa simplicité et de son efficacité pour entraîner des réseaux de neurones, en particulier dans les tâches de régression. La fonction ReLU est définie par l'équation suivante :

$$ReLU(x) = \max(0, x) \text{ (éq. 9)}$$

Cette équation signifie que la fonction retourne la valeur d'entrée x si celle-ci est positive, et zéro sinon. En d'autres termes, ReLU active une unité neuronale seulement si l'entrée est positive. Si l'entrée est négative ou nulle, la sortie de cette unité sera également nulle.

- La fonction d'activation **Linear** est beaucoup plus simple que ReLU, car elle ne transforme pas l'entrée de manière non linéaire. Elle est définie par l'équation suivante :

$$f(x) = x \text{ (eq. 10)}$$

Cette fonction est également appelée fonction d'identité car elle retourne simplement la valeur d'entrée telle quelle. Dans un réseau de neurones, une couche avec une fonction d'activation linéaire ne fait qu'appliquer une transformation linéaire aux données.

1.3. Etude de cas dans la littérature

Dans le domaine de la prédiction des vibrations induites par les tirs à l'explosif, de nombreuses études ont été menées pour développer des modèles précis et robustes. Ces recherches visent principalement à minimiser les impacts négatifs des explosions, tels que les dommages structurels et les nuisances pour les populations environnantes, tout en optimisant les opérations de dynamitage pour des industries comme l'exploitation minière, la construction de routes et la création d'infrastructures. Traditionnellement, des méthodes statistiques simples, telles que la régression linéaire, ont été utilisées pour modéliser la relation entre divers paramètres de tir et les vibrations résultantes. Cependant, ces méthodes présentent des limites en termes de précision et de capacité à capturer des relations complexes entre les variables.

Avec l'avènement des technologies de l'intelligence artificielle, en particulier l'apprentissage automatique et l'apprentissage profond, de nouvelles approches ont été explorées pour améliorer les prédictions de la vitesse particulaire maximale (PPV) et d'autres caractéristiques des vibrations. Les techniques de machine learning, comme les arbres de décision, les forêts aléatoires, et la régression par vecteurs de support (SVR), ont montré leur potentiel pour modéliser des données complexes et hétérogènes. De plus, l'apprentissage profond, avec ses réseaux de neurones multicouches, offre des capacités inégalées pour capturer des patterns non linéaires et complexes, rendant les prédictions encore plus précises.

Les études récentes se concentrent également sur l'intégration de multiples sources de données, telles que les propriétés géotechniques, les caractéristiques de l'explosif, et les conditions environnementales, pour construire des modèles plus complets et fiables. Ces modèles intelligents sont évalués à l'aide de métriques rigoureuses telles que la MSE, la RMSE, la MAE et le coefficient de détermination R^2 , assurant ainsi leur robustesse et leur précision. L'utilisation de ces nouvelles approches basées sur l'IA permet non seulement d'améliorer la sécurité et l'efficacité des opérations de dynamitage, mais aussi de réduire les impacts environnementaux et sociaux, répondant ainsi aux exigences croissantes en matière de durabilité et de responsabilité sociétale.

Dans cette section, nous passerons en revue les principales études et avancées dans ce domaine, en mettant en lumière les méthodologies employées, les résultats obtenus et les contributions significatives à la compréhension et à la prédiction des vibrations induites par les tirs à l'explosif.

Le tableau ci-dessus représente différentes études sur l'utilisation de l'intelligence artificielle et de l'apprentissage automatique et profond pour la prédiction des vibrations induites par les tirs à l'explosif :

Tab(1.1) : Etude de cas sur l'utilisation de l'intelligence artificielle et de l'apprentissage automatique et profond pour la prédiction des vibrations induites par les tirs à l'explosif

N°	Titre	Année
1	Prediction of Blast-Induced Ground Vibrations: A Comparison Between Empirical and Artificial-Neural-Network Approaches. (Velasquez, Luis F., 2023)	2023
2	A Novel Artificial Intelligence Approach to Predict BlastInduced Ground Vibration in Open-Pit Mines Based on the Firefly Algorithm and Artificial Neural Network. (Shang et al., 2019)	2020
3	A new combination of artificial neural network and K-nearest neighbors models to predict blast-induced ground vibration and air-overpressure.(Amiri et al.,2016)	2016
4	Predicting blasting propagation velocity and vibration frequency using artificial neural networks. (Álvarez-Vigil et al., 2012)	2012
5	Prediction of blast-induced ground vibration using artificial neural networks. (Monjezi M. et al.,2011)	2011

1.4.1 Etude n°1: Prediction of Blast-Induced Ground Vibrations: A Comparison Between Empirical and Artificial-Neural-Network Approaches. (Velasquez, Luis F, 2023)

1.4.1.1 Zone d'Étude

La mine d'or à ciel ouvert est située entre 3200 et 3500 mètres au-dessus du niveau de la mer, à 480 km au nord-ouest de Lima. au nord-ouest de Lima, dans le district de Huamachuco, région de La Libertad, Pérou. L'emplacement général de la propriété est indiqué dans la **figure 1.3** ci-contre. L'exploitation utilise la lixiviation en tas pour obtenir de l'or et le transforme en barres doré. Les réserves d'or actuelles sont de 443,2 koz (prouvées et probables), Le débit moyen de l'usine est de 36 000 tonnes par jour et la teneur en or est de 0,33 g/tonne. Des roches sédimentaires du bassin mésozoïque de l'ouest du Pérou sont présentes dans la zone du projet. Le gisement de sulfures contient des minéralisations de cuivre primaire et secondaire ainsi que de l'or. Le gisement d'oxyde consiste en une minéralisation d'oxyde d'or à faible teneur. Le gisement contient également une minéralisation porphyrique Cu-Au (Mo) qui s'étend sur une longueur de 2,2 km du sud au nord et sur une largeur d'environ 1,1 km.

Figure 1.3: Vue du ciel de la zone d'étude, Pérou



Il est également important de mentionner que dans la zone d'influence directe, plus de 1500 habitants résident dans trois communautés : La Arena, La Ramada et El Alisar. La zone la plus peuplée est La Arena. La communauté de La Arena compte environ 226 familles, dont 80 % des maisons ont été construites à l'aide d'un système de Fosse Déchets décharge 33 béton. La communauté de La Ramada compte environ 200 familles réparties dans 140 maisons, dont 90 % ont été construites en briques de terre et en adobe. ont été construites en briques de terre et en adobe. D'autres bâtiments importants situés à proximité du projet sont l'école secondaire de La Ramada, où plus de 90 élèves suivent régulièrement des cours. Ramada, où plus de 90 élèves suivent régulièrement les cours. Dans la communauté d'El Alisar, environ 65 familles vivent dans 40 maisons, principalement construites en briques de terre et en adobe

1.3.1.2 Méthodes Utilisées

- Collecte des données

Les enregistrements des vibrations ont été recueillis par *Orica Mining Services Peru S.A.*, qui fournit des services de dynamitage à la société minière *La Arena S.A.* Les données ont été enregistrées de janvier à octobre 2022 à l'aide de sismographes Instantel, en particulier le modèle MiniMate Plus illustré à la **figure 1.4**.

Figure 1.4: Le modèle MiniMate

Au cours du processus de collecte des données, les paramètres de dynamitage, y compris la charge, l'espacement, la longueur de la tige, la profondeur du trou, le diamètre du trou, le facteur de poudre, le poids maximal de la charge par délai, le système d'initiation, la densité de l'explosif et le sous-forage. d'initiation entre autres, ont également été enregistrés. Pour chaque tir, la PPV a été mesurée à l'aide de sismographes installés à des points stratégiques dans les communautés locales.



- ANN models :

Le premier ensemble de données de la communauté de *La Arena* a comporté 168 observations et le second ensemble de données de la communauté de *La Ramada* a comporté 70 observations. La version 2022a de MATLAB a été utilisée pour construire, valider et tester le modèle ANN.

Paramètres du réseau de neurones dans le **tableau 1.2** ci-dessous: .

Tab(1.2) : Paramètres du réseau de neurones.

Paramètres d'entrée (unité)	Paramètres de sortie
Système d'initiation (-) Temps de sommeil (jours) Total des explosifs (kg) Densité explosive (g/cm3) Charge (m) Espacement (m) Nombre de trous chargés (Production+Tampon) (-) Longueur de la tige (m) Facteur poudre (kg/m3) Distance de surveillance (m) Poids maximal de la charge par délai (kg)	Vitesse maximale des particules (PPV)

Une fois que l'algorithme de *Levenberg-Marquardt* a été sélectionné pour l'entraînement, ce processus se poursuit jusqu'à ce qu'un critère d'arrêt soit rempli. L'apprentissage se poursuit jusqu'à ce que l'erreur de validation augmente consécutivement pendant un certain nombre d'itérations.

Fig(15) : Résumé du modèle

Algorithm

Data division: Random

Training algorithm: Levenberg-Marquardt

Performance: Mean squared error

Training Results

Training start time: 19-Nov-2022 21:22:34

Layer size: 12

	Observations	MSE	R
Training	134	0.1106	0.7412
Validation	17	0.2279	0.6638
Test	17	0.1776	0.5115

Le résumé du modèle présenté dans la **figure 1.5** contient des informations sur l'algorithme d'apprentissage et les résultats de l'apprentissage. Les valeurs de *MSE* et du coefficient de corrélation de Pearson (*R*) sont indiquées. Pour trouver le *RMSE*, le *MSE* a été élevé au carré, et pour calculer le *R2*, le coefficient de corrélation de Pearson a été élevé au carré.

1.3.1.3 Résultats :

Il a été observé que la moyenne la plus élevée de R^2 dans l'ensemble de test était l'architecture 11-16-1 comme l'indique le **tableau 1.3**, ce qui signifie 11 entrées, 16 neurones dans la couche cachée et 1 sortie. Le réseau optimal a été sélectionné sur la base de l'ensemble de test car le modèle ne disposait pas de ces données lors de l'entraînement initial.

Tab(1.3) : Les valeurs moyennes de RMSE , R^2 pour La Ramada et La Arena comme ensemble de données

Model	Network architecture		Training		Testing	
	Hidden layer	Neurons in the hidden layer	Avg. RMSE	Avg. R^2	Avg. RMSE	Avg. R^2
1	1	12	0.250	0.681	0.301	0.538
2	1	13	0.277	0.613	0.378	0.464
3	1	14	0.248	0.680	0.356	0.542
4	1	15	0.293	0.582	0.336	0.602
5	1	16	0.246	0.678	0.317	0.630
6	1	17	0.260	0.647	0.321	0.587
7	1	18	0.264	0.650	0.320	0.625
8	1	19	0.280	0.649	0.319	0.586
9	1	20	0.233	0.729	0.404	0.506
10	1	21	0.261	0.687	0.389	0.529
11	1	22	0.304	0.593	0.388	0.534
12	1	23	0.220	0.746	0.359	0.576

1.3.2 Etude n°2: A Novel Artificial Intelligence Approach to Predict BlastInduced Ground Vibration in Open-Pit Mines Based on the Firefly Algorithm and Artificial Neural Network. (Shang et al., 2019)

1.3.2.1 Zone d'Étude

L'étude est réalisée dans une mine à ciel ouvert située dans la province de *Binh Duong*, au *Vietnam* (**figure 1.6**). Les caractéristiques spécifiques de la zone d'étude comprennent :

Fig(1.6) : Vue du ciel de la zone d'étude, Vietnam



- **Géologie** : La zone est composée principalement de roches dures telles que le granit, ce qui influence la propagation des vibrations.
- **Paramètres des Explosions** : Les données incluent les quantités d'explosifs utilisées, les distances entre les points de mesure et les points d'explosion, et la configuration des charges explosives.

- **Conditions Environnementales** : Les conditions climatiques et environnementales de la zone, telles que l'humidité et la température, sont également prises en compte car elles peuvent affecter la propagation des ondes de choc.

1.3.2.2 Méthodes Utilisées

- **Collecte des données**

Pour collecter les données nécessaires à cette étude, un plan spécifique a été mis en œuvre. Des instruments Micromate (Instantel, Canada) ont été utilisés pour mesurer les vibrations du sol causées par les opérations de dynamitage dans des sites sensibles. Un récepteur GPS de haute précision, le X91B, a été utilisé pour calculer la distance de surveillance. Les autres paramètres ont été recueillis et exportés à partir de 83 conceptions de dynamitage.

Cinq paramètres indépendants ont été considérés pour prédire les vibrations du sol causées par les dynamitages :

- W : Poids de l'explosif utilisé
- S : Distance entre les trous de dynamitage
- R : Distance de surveillance
- B : Diamètre des trous de dynamitage
- P : Pression des gaz de l'explosion

Ces paramètres ont été choisis en raison de leur influence significative sur l'intensité des vibrations du sol induites par les dynamitages.

La vitesse de particule maximale (PPV) (valeur à prédire) est la mesure standard la plus connue pour illustrer les vibrations du sol induites par les dynamitages et est largement appliquée dans la recherche.

Les instruments Micromate ont été placés à des sites sensibles pour mesurer les vibrations du sol. Pour calculer la distance de surveillance, un récepteur GPS de haute précision a été utilisé. Les données recueillies ont été utilisées pour développer un modèle hybride d'intelligence artificielle pour prédire les vibrations du sol.

- **ANN models :**

Un modèle ANN est développé pour apprendre les relations complexes entre les paramètres d'entrée et les vibrations du sol. Le réseau est composé de plusieurs couches de neurones avec des fonctions d'activation non linéaires.

Optimisation avec l'Algorithme des Lucioles , ce dernier est utilisé pour optimiser les poids et les biais du réseau ANN. L'algorithme des Lucioles est inspiré du comportement des lucioles et est efficace pour résoudre des problèmes d'optimisation globale. Le modèle FA-ANN est entraîné avec un sous-ensemble des données collectées et validé avec le reste des données pour évaluer sa performance.

1.3.2.3 Résultats

La section des résultats de l'étude présente une évaluation comparative des différents modèles utilisés pour prédire les vibrations induites par les explosions (PPV) dans les mines de la carrière de Tan Dong Hiep. Les modèles évalués incluent KNN, SVM, CART et FFA-ANN. Les performances des modèles ont été mesurées à l'aide de plusieurs indicateurs de performance : RMSE, MAE, et R^2 .

- **Indicateurs de Performance des Modèles**

- KNN (K-Nearest Neighbors)
 - RMSE : 1.021
 - MAE : 0.661
 - R^2 : 0.851
- SVM (Support Vector Machine)
 - RMSE : 0.822
 - MAE : 0.501
 - R^2 : 0.927
- CART (Classification and Regression Trees)
 - RMSE : 0.821

- MAE : 0.639
- R^2 : 0.899
- FFA-ANN (Firefly Algorithm - Artificial Neural Network)
 - RMSE : 0.464
 - MAE : 0.356
 - R^2 : 0.966

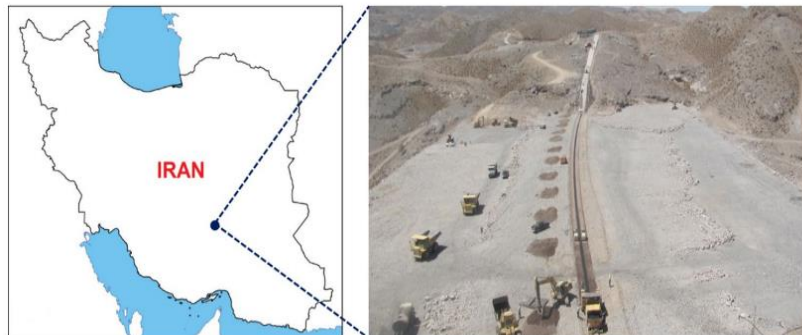
Le modèle FFA-ANN a démontré la performance la plus dominante avec les meilleurs scores sur tous les indicateurs de performance, suivi par les modèles SVM, CART, et enfin KNN. Cela est illustré par une gradation de couleur allant du blanc (moins performant) au rouge (plus performant), où le modèle FFA-ANN est représenté en rouge.

1.3.3 Etude n°3: A new combination of artificial neural network and K-nearest neighbors models to predict blast-induced ground vibration and air-overpressure.(Amiri et al.,2016)

1.3.3.1 Zone d'Étude

L'étude s'est concentrée sur la région du barrage de la rivière *Shur*, une zone (**figure 1.7**) où l'exploitation minière implique des opérations de dynamitage. Les paramètres clés de cette région incluent la granulométrie de la roche, la

Fig(1.7) : Vue du ciel de la zone d'étude,Iran



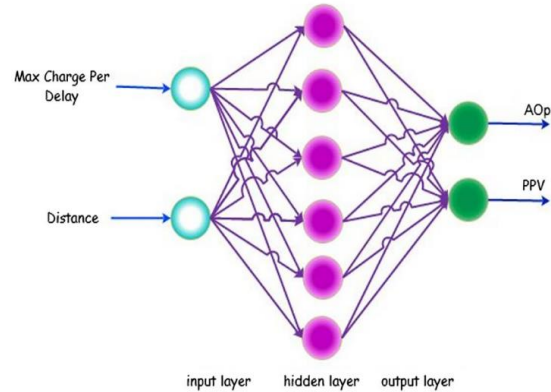
densité des matériaux, et les caractéristiques géologiques spécifiques. L'environnement de cette région a un impact direct sur les vibrations du sol et les ondes de choc générées par le dynamitage

1.3.3.2 Méthodes Utilisées

- **ANN models**

Dans cette recherche, un modèle de perceptron multicouche (MLP) a été utilisé pour prédire les vibrations du sol induites par des explosions (PPV) et la surpression aérienne (AOp). La structure de l'ANN (**figure 1.8**) comprend trois couches: entrée, cachée et sortie. Les variables indépendantes sont placées dans la couche d'entrée, tandis que les variables dépendantes (AOp et PPV) sont dans la couche de sortie. Le nombre de couches cachées et de neurones dans ces couches est déterminé par essais et erreurs.

Fig(1.8) : La structure du modèle



Le modèle ANN a été testé avec 61 ensembles de données pour l'entraînement et 20 pour les tests. Les performances du modèle ont été évaluées en utilisant le coefficient de corrélation (R^2) et l'erreur quadratique moyenne (RMSE). Un modèle est considéré excellent si R^2 est de 1 et RMSE est de 0.

Le modèle combiné ANN-KNN utilise à la fois ANN et K-nearest neighbors (KNN) pour améliorer les prédictions. Le processus comprend plusieurs étapes :

- **Normalisation des données :** Les données sont normalisées pour s'assurer que toutes les variables sont sur une échelle comparable.
- **Clustering K-means :** Les échantillons d'entraînement sont groupés en clusters similaires.
- **Prédicteur KNN :** Trouve les K plus proches voisins de l'échantillon de test et lui attribue la classe majoritaire parmi ces voisins.
- **Prédicteur ANN :** Pour chaque cluster, un ANN est entraîné avec les échantillons de ce cluster. L'ANN prédit les échantillons de test situés dans le même cluster.
- **Combinaison des résultats :** Les résultats prédits par ANN et KNN sont combinés en utilisant une pondération.

- **Padding** : Pour réduire les risques de sous-estimation, une petite constante est ajoutée aux résultats finaux.

1.3.3.3 Résultats :

Les résultats obtenus montrent que le modèle ANN-KNN (une combinaison d'ANN et de KNN) fournit des prédictions plus précises des vibrations du sol (PPV) et des surpressions d'air (AOp) par rapport aux modèles individuels. Les performances des modèles ont été évaluées à l'aide de trois indices : RMSE (Root Mean Square Error), et R^2 (coefficient de détermination). Les valeurs des indices pour différents modèles sont présentées ci-dessous :

- Prédiction de AOp :
 - ANN : $R^2 = 0.93$, RMSE = 3.34
 - ANN-KNN : $R^2 = 0.95$, RMSE = 1.7
- Prédiction de PPV :
 - ANN : $R^2 = 0.94$, RMSE = 0.28
 - ANN-KNN : $R^2 = 0.96$, RMSE = 0.24

Avec des valeurs plus élevées de R^2 , et des valeurs plus basses de RMSE, on conclut que le modèle ANN-KNN a donné une meilleure précision et fiabilité des prédictions .

1.3.4 Etude n°4: Predicting blasting propagation velocity and vibration frequency using artificial neural networks. (Álvarez-Vigil et al., 2012)

1.3.4.1 Zone d'Étude

La zone d'étude est située près de la ville de *Langreo*, à proximité de la ville d'*Oviedo*, dans le nord de l'*Espagne*, dans une carrière exploitée par la *Bahoto Mining Company*. L'accès à la mine se fait par le kilomètre 9 de la route AS-243 (*Oviedo-Frieres*). La **figure 1.9** ci-contre montre l'emplacement de la zone d'étude, où l'altitude la plus élevée est atteinte est (591 m). La carrière de *Bahoto* se trouve au-dessus d'une montagne calcaire dans la formation de *Valdeteja*.

Fig(1.9) : Vue du ciel de la zone d'étude, Espagne



1.3.4.2 Méthodes Utilisées

- **Collecte des données**

Les vibrations ont été surveillées à l'aide de sismographes conçus pour mesurer la vitesse et la fréquence des vibrations. Il n'existe pas de consensus sur le type d'équipement de mesure des vibrations le plus précis, surtout pour les vibrations dans les structures résultant de l'interaction entre le souffle et la structure. Pour cette recherche, un accélérographe Etna Strong Motion de Kinematics a été utilisé, équipé de stations triaxiales à haute plage dynamique et de mémoire interne pour stocker les données.

Pour garantir la précision des mesures, les directives de l'International Society of Explosives Engineers (ISEE) ont été suivies :

- Le capteur doit être placé sur ou dans le sol, du côté de la structure faisant face à l'explosion.
- L'emplacement du capteur doit refléter les niveaux de vibrations transmises par le sol à la structure.
- Le canal longitudinal doit être orienté vers le souffle et la direction enregistrée.

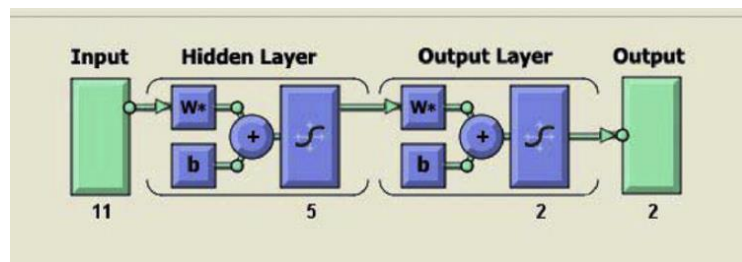
- Si l'accélération dépasse 0,2 fois la gravité, le capteur peut être découplé. Fixer le sismographe au sol peut être nécessaire selon les niveaux d'accélération.

Pour cette étude, les capteurs ont été boulonnés au sol sur des surfaces rocheuses et cimentées et installés uniquement pendant la période de test pour éviter les problèmes électriques et météorologiques. Les accélérographes étaient orientés avec les axes X, Y et Z respectivement vers l'est, le nord et verticalement pour chaque essai.

- **ANN models**

Un réseau de neurones à propagation avant (**figure 1.10**) comme modèle le plus approprié pour résoudre un problème d'identification dans l'analyse prédictive des vibrations dans la carrière de *Bahoto*. La solution

Fig(1.10) : La structure du modèle



est générée en établissant successivement des relations entrée/sortie jusqu'à atteindre un optimum.

Paramètres du réseau de neurones dans le **tableau 1.3** ci-dessous: .

Tab(1. 4) : Paramètres du réseau de neurones.

Paramètres d'entrée (unité)	Paramètres de sortie
Vitesse de détonation de l'explosif (km/s)	Vitesse maximale des particules (PPV)
Nombre de trous de mines (-)	Fréquence de vibration (F)
Charge totale (kg)	
Charge instantanée (kg)	
Charge (m)	
Espacement entre les trous de mines (m)	
Longueur du trou de mine (m)	
Diamètre du trou de mine (mm)	
Distance entre l'explosion et le point de contrôle (m)	
Disposition relative du point de contrôle de l'explosion (°)	
RMR (-) (rock mass rating) taux de masse rocheuse	

Configuration du réseau:

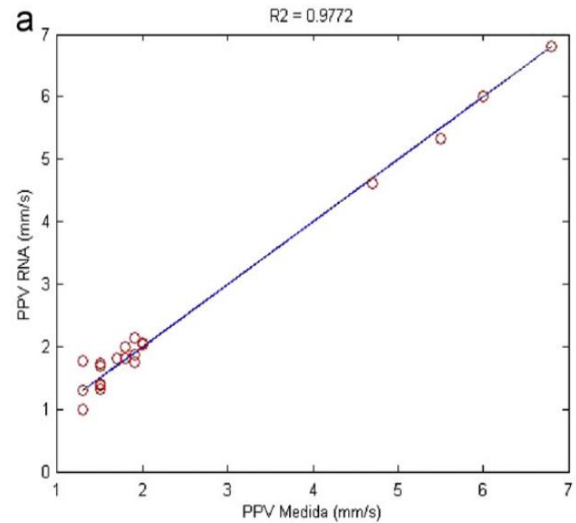
- Nombre de neurones d'entrée : 11
- Nombre de neurones de sortie : 2 (PPV et F)
- Nombre de couches cachées : 1
- Nombre de neurones dans la couche cachée : 5
- Nombre de cycles d'entraînement : 100
- Nombre de données d'entraînement : 40
- Nombre d'éléments de validation : 20

- Nombre d'éléments de test : 20
- Erreur tolérée : 0,05

1.4.4.3 Résultats

Les résultats montrent la qualité de la prédiction en utilisant la méthodologie des ANN. L'indicateur de qualité considéré était le coefficient de corrélation (R^2) entre les valeurs prédites et observées (**figure 1.10**). La prédiction était basée sur les données d'entrée. Les coefficients de corrélation entre les valeurs prédites par l'ANN et les valeurs mesurées étaient de 0,98 pour la vitesse de pointe des particules (PPV), des valeurs proches de 1 indiquant une bonne qualité de la prédiction.

Fig(1.11) : Résultat Test-set de PPV(mm/s)

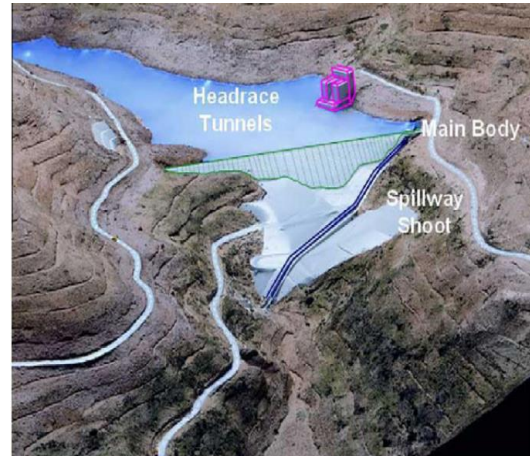


1.3.5 Etude n°5: Prediction of blast-induced ground vibration using artificial neural networks. (Monjezi M. et al., 2011)

1.3.5.1 Zone d'Étude

L'étude de cas se concentre sur le projet *Siahbisheh* en Iran (**figure 1.11**), où des mesures de vibrations de dynamitage ont été effectuées. Une base de données contenant 182 ensembles de données de vibrations a été collectée à différents emplacements stratégiques et vulnérables autour du projet. Ces données comprennent les paramètres d'entrée nécessaires pour le modèle ANN ainsi que les valeurs mesurées de PPV. L'étude de cas sert de fondement pour la validation et l'évaluation du modèle ANN développé, démontrant son applicabilité pratique dans un contexte réel.

Fig(1.12) : Vue du ciel de la zone d'étude, Iran



1.3.5.2 Méthodes Utilisées

- **Collecte des données**

La section des ensembles de données décrit les 182 enregistrements de vibrations de dynamitage utilisés pour la modélisation des réseaux de neurones. Les données collectées comprennent des informations sur la charge maximale par délai, la distance du point de surveillance au point de dynamitage, le bourrage, et la profondeur du trou. Ces paramètres d'entrée sont utilisés pour former le modèle ANN, tandis que les valeurs de PPV mesurées servent de référence pour évaluer les performances du modèle. L'exhaustivité et la précision des données sont cruciales pour le succès de la modélisation.

- ANN models

L'architecture du réseau de neurones feed-forward back-propagation utilisé pour la modélisation. Le réseau a été conçu avec une topologie 4-10-5-1, ce qui signifie qu'il comporte quatre neurones d'entrée, deux couches cachées avec respectivement dix et cinq neurones, et un neurone de sortie. Cette architecture a été choisie après des essais et des ajustements pour optimiser les performances du modèle. Les fonctions d'activation et les algorithmes d'apprentissage utilisés sont également discutés, soulignant l'importance de chaque composant dans la capacité du modèle à faire des prédictions précises.

Paramètres du réseau de neurones dans le **tableau 1.4**. ci-dessous: .

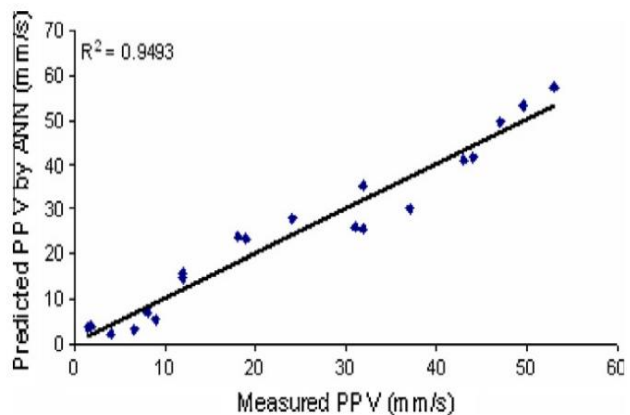
Tab(1.5) : Paramètres du réseau de neurones.

Paramètres d'entrée (unité)	Paramètres de sortie
Charge maximale par délai (kg) Distance du front de taille (m) Tige (cm) Profondeur du trou (m)	Vitesse maximale des particules (PPV)

1.3.5.3 Résultats

L'évaluation du modèle de ANN montre que la topologie optimale 4-10-5-1 est plus précise que les méthodes empiriques et statistiques conventionnelles pour prédire les PPV. Les performances du modèle sont mesurées en utilisant des métriques telles que le coefficient de détermination (R^2) et l'erreur absolue moyenne (MAE). Les résultats dans la **figure 1.12** montrent un coefficient de détermination

Fig(1.13) : Résultat Test-set de PPV(mm/s)



de 0,9493 et une erreur absolue moyenne de 0,002, indiquant une très bonne concordance entre les PPV mesurés et prédits. Cette évaluation confirme la supériorité du modèle ANN par rapport aux approches traditionnelles pour ce type de prédiction.

→ En conclusion, il est important de noter que les études précédentes dans ce domaine n'ont pas évalué leurs modèles dans un intervalle de confiance par rapport à leurs jeux de données. De plus, ces études ont souvent utilisé des jeux de données comprenant moins de 300 lignes. Sans l'application de techniques d'augmentation des données, les modèles développés sur de tels jeux de données peuvent manquer de robustesse et risquent de ne pas être performants lorsqu'ils sont confrontés à de nouvelles données.

Chapitre 2 : Collecte, analyse et préparation des données

2.1 Processus de collecte des données dans les sites miniers

La collecte de données dans les sites miniers est un processus crucial pour la prédiction des vibrations induites par les tirs à l'explosif. Les paramètres clés à mesurer incluent le nombre des trous (**NT**), le nombre des rangées de trous (**NR**), le diamètre des trous (**ND**), la charge d'explosif (**CUM**), le bourrage (**BF**), la profondeur maximale du trou (**PM**) et la distance entre le point de tir et l'instrument de mesure de la vibration (**Distance**).

Les tirs à l'explosif dans les sites miniers sont effectués en suivant un processus bien défini qui inclut les étapes suivantes:

➤ **Disposition des Trous de Forage:**

- Le site de tir est divisé en plusieurs rangées de trous de forage.
- Chaque rangée peut comporter un nombre variable de trous, selon le plan de tir.

➤ **Profondeur et Diamètre des Trous:**

- Chaque trou de forage est creusé à une profondeur spécifique, déterminée par la nature de la roche et les objectifs de fragmentation.
- Le diamètre des trous est également une variable critique qui influence la quantité d'explosif et l'efficacité de la fragmentation.

➤ **Chargement de l'Explosif:**

- Les trous sont remplis avec une charge d'explosif. La quantité d'explosif dans chaque trou est soigneusement calculée pour assurer une fragmentation optimale.
- La charge d'explosif est un paramètre essentiel pour prédire les vibrations induites par le tir.

➤ **Bourrage:**

- Après avoir chargé l'explosif, le trou est bourré. Le bourrage consiste à remplir le trou avec des matériaux inertes (comme de l'argile ou des débris rocheux) pour contenir l'explosion et diriger l'énergie vers la roche à fragmenter.
- Le bourrage aide à contrôler les projections et à maximiser l'efficacité de l'explosion.

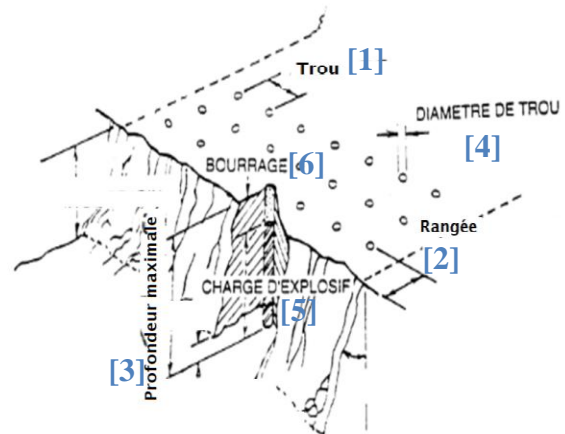
➤ **Distance de Mesure:**

- La distance entre le point de tir et l'instrument de mesure des vibrations est cruciale. Cette distance influence directement les lectures de PPV et doit être précisément mesurée.

Fig(2.1) : Processus des tirs à l'explosif

La **figure 2.1** illustre ce processus:

- Les trous. [1]
- Les rangées de trous de forage.[2]
- Les profondeurs [3] et diamètres des trous.[4]
- La charge d'explosif [5] et le bourrage dans chaque trou. [6]
- La distance entre le point de tir et l'instrument de mesure.



Voici l'explication détaillée de chaque paramètre dans le contexte de ce dessin:

➤ **Rangées de Trous:**

- Chaque rangée est clairement indiquée, montrant comment les trous sont alignés pour optimiser la fragmentation.

➤ **Profondeur et Diamètre des Trous:**

- Les variations de profondeur et de diamètre entre les trous sont marquées pour montrer leur impact sur la charge et le bourrage.

➤ **Charge d'Explosif et Bourrage:**

- La section de chaque trou illustre la distribution de l'explosif et le matériau de bourrage, soulignant l'importance de leur placement pour une explosion efficace.

➤ **Distance de Mesure:**

- La distance entre le point de tir (où l'explosif est détoné) et l'instrument de mesure (qui enregistre les vibrations) est représentée pour montrer son rôle dans la collecte des données.

2.2 Forme et prétraitement des données

2.2.1 La forme des données

Le jeu de données utilisé dans cette étude comprend 153 lignes et 8 colonnes. Les données d'entrée incluent les paramètres suivants :

- **NT** : le nombre de trous
- **NR** : le nombre de rangées de trous
- **PM** : la profondeur maximale des trous
- **BF** : le bourrage.
- **CUM** : la charge d'explosif
- **ND** : le diamètre du trou.
- **Distance** : la distance entre le point de tir et l'instrument de mesure

La donnée de sortie, **PPV** (vitesse particulaire maximale), est utilisée comme variable cible dans les modèles prédictifs. Ce jeu de données est essentiel pour la construction et la validation des modèles d'intelligence artificielle destinés à prédire les vibrations induites par les tirs à l'explosif.

2.2.2 Prétraitement des données

Le prétraitement des données est une étape essentielle dans la création de modèles prédictifs robustes. Il permet de nettoyer, transformer, et préparer les données brutes pour qu'elles puissent être exploitées efficacement par les algorithmes d'apprentissage automatique .

- **Détection des données aberrantes :**

La détection des données aberrantes est une étape cruciale du prétraitement, car les valeurs anormales peuvent significativement influencer les performances des modèles prédictifs. Une

méthode couramment utilisée pour identifier ces valeurs aberrantes est le calcul du *z-score*, qui mesure combien d'écart-types une donnée s'écarte de la moyenne de l'ensemble des données.

La formule mathématique du *z-score* est l'équation (11) :

$$z\text{-score} = \frac{x - \mu}{\sigma} \text{ (eq. 11)}$$

où :

- X est la valeur de la donnée,
- μ est la moyenne de l'ensemble des données,
- σ est l'écart-type de l'ensemble des données.

L'algorithme du calcul du *z-score* se déroule comme suit :

- **Définition du seuil** : Un seuil est fixé pour identifier les outliers, généralement défini à 3, ce qui correspond à 3 écart-types de la moyenne.
- **Calcul de la moyenne** : La moyenne des données est calculée pour servir de point de référence central.
- **Calcul de l'écart-type** : L'écart-type des données est déterminé pour mesurer la dispersion des valeurs autour de la moyenne.
- **Calcul du z-score pour chaque valeur** : Pour chaque donnée, le *z-score* est calculé en soustrayant la moyenne de cette valeur, puis en divisant le résultat par l'écart-type.
- **Vérification du z-score** : Le *z-score* calculé est comparé au seuil prédéfini. Si la valeur absolue du z-score dépasse ce seuil, la donnée est considérée comme une donnée aberrante.
- **Identification des outliers** : Les valeurs dont le *z-score* dépasse le seuil sont ajoutées à une liste d'outliers, permettant ainsi de les isoler pour un traitement ultérieur.

- **Remplacement des données aberrantes:**

Le processus de remplacement assure que les données aberrantes sont traitées de manière appropriée, minimisant ainsi leur impact sur la qualité et la précision des modèles prédictifs:

- **Remplacement des données de sortie aberrantes** : Les valeurs aberrantes dans les données de sortie, telles que la vitesse particulaire maximale (PPV), sont remplacées en utilisant la moyenne des valeurs non manquantes. Cette approche est effectuée à l'aide d'un imputer simple (*SimpleImputer* de *Sklearn*) qui calcule la moyenne de chaque colonne et remplace les valeurs aberrantes par ces moyennes.
- **Remplacement des données d'entrée aberrantes** : Les valeurs aberrantes dans les données d'entrée, telles que les paramètres de tir, sont traitées en utilisant la méthode des K plus proches voisins (*KNN*). Cette méthode remplace les valeurs aberrantes par la moyenne des valeurs des K plus proches voisins, où K est généralement fixé à 3. Cela permet d'estimer les valeurs manquantes ou aberrantes en se basant sur les informations des observations les plus proches.

2.3 Analyse statistique des données par régression linéaire et détermination de l'intervalle de confiance

2.3.1 Loi d'atténuation

La vitesse de pointe des particules (PPV) est l'un des paramètres les plus utilisés pour étudier les vibrations de tir, car il s'agit d'un paramètre bien connu du domaine, qui est facilement mesurable et qui a été relié à des dommages structuraux. La première corrélation empirique acceptée pour évaluer les vibrations d'explosion a été développée par (Koch, 1958). Elle est basée sur les mesures systématiques de PPV et sa dépendance à la distance entre les géophones et le point de dynamitage, comme indiqué dans l'équation (12) :

$$PPV = kQ^{\beta}R^{-\eta} \text{ (eq. 12)}$$

Où :

- PPV est la vitesse de pointe des particules.

- k , β et η sont des facteurs empiriques, qui peuvent varier selon les conditions géologiques et géomécaniques du site, en plus des techniques de forage et de dynamitages appliquées.
- Q est le poids de charge explosive par délai
- R est la distance entre le point de tir et le point de mesure.

L'équation (12), appelée loi d'atténuation, a été légèrement modifiée par USBM (US Mines Office) (Elevli et Arpaz, 2010), suivant l'équation (13). La distance scalaire (SD) est définie selon l'équation (14).

$$PPV = k(SD)^{-\beta} \text{ (éq. 13)}$$

$$SD = \frac{R}{Q^\alpha} \text{ (éq. 14)}$$

Où α dépend de la forme de la charge explosive ($\frac{1}{2}$ pour explosion sphérique et $\frac{1}{3}$ pour explosion cylindrique). De plus, k et β représentent des paramètres spécifiques du site. D'autres équations empiriques ont été suggérées par différents chercheurs pour décrire l'atténuation des vibrations, comme (Duvall et al., 1958), (Ghosh et al., 1985), (Gupta et al., 1988) et (Roy et al., 1993). Cependant, les équations (7) et (8) restent encore largement utilisées par les industriels. Récemment, de nombreux auteurs ont utilisé ces équations en les comparant aux PPV mesurées, afin d'évaluer le poids de charge par retard (Müller et al., 2007 ; Monjezi et al., 2012).

Dans le cadre de notre analyse, nous avons adopté une simplification de ces équations empiriques en utilisant la relation suivante pour la SD:

$$SD = \frac{R}{\sqrt{Q}} \text{ (éq. 15)}$$

Où :

- Q est le poids de charge explosive par délai
- R est la distance entre le point de tir et le point de mesure.

Cette formulation simplifiée permet de modéliser la relation entre les paramètres de tir de manière plus directe et est adaptée aux spécificités de notre étude.

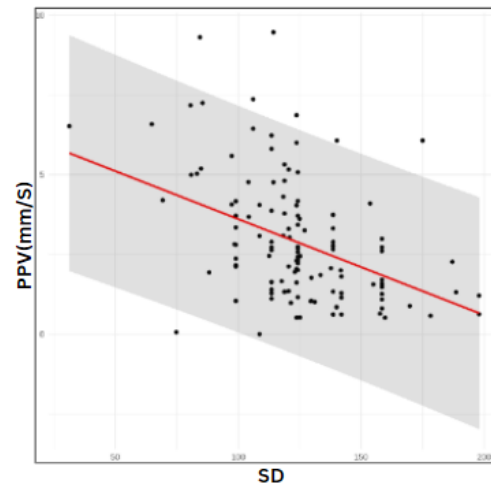
2.3.2 Détermination de L'Intervalle de confiance

Pour déterminer l'intervalle de confiance des prédictions de la vitesse de pointe des particules (PPV), nous avons utilisé l'équation (13) :

Un code en langage R a été employé pour réaliser cette analyse et déterminer les intervalles de confiance associés aux prédictions dans la **figure 2.2** ci-dessous:

- **Régression linéaire** : Une régression linéaire est réalisée pour modéliser la relation entre la vitesse de pointe des particules (PPV) et la distance scalaire (SD). Ce modèle permet d'étudier comment les valeurs de PPV peuvent être prédites en fonction des valeurs de SD.
- **Ajout des Intervalles à la Base de Données** : Les prévisions ainsi que les bornes inférieure et supérieure des intervalles de confiance sont ajoutées aux données originales. Cela permet de visualiser les prédictions du modèle avec leur variabilité.

Fig(2.2) : Intervall de confiance des données



Chapitre 3 : Modèles de prédiction des vibrations induites par les tirs à l'explosif à l'aide de l'Intelligence Artificielle

3.1 Développement et optimisation d'un Réseau de Neurones Artificiels (DL)

Le développement et l'optimisation d'un Réseau de Neurones Artificiels (DL) pour prédire les vibrations induites par les tirs à l'explosif impliquent plusieurs étapes clés, allant de la préparation des données à la conception de l'architecture du réseau, en passant par la régularisation pour éviter le surapprentissage (overfitting).

Paramètres du réseau de neurones dans le **tableau 3.1** ci-dessous: .

Tab(3.1) : Paramètres du réseau de neurones.

Paramètres d'entrée (unité)	Paramètres de sortie
Nombre de trous Nombre de rangées Profondeur maximale (m) Bourrage (m) Charge unitaire maximale (kg) Diamètre du trou Distance du tir (m)	Vitesse maximale des particules (mm/s)

Étapes du Développement

- **Division des Données** : Les données disponibles ont été divisées en deux ensembles : 80 % des données pour l'entraînement et 20 % pour les tests. Cette division assure que le modèle est évalué sur des données qu'il n'a pas vues pendant l'entraînement, ce qui permet de vérifier sa capacité de généralisation.
- **Normalisation des Données** : Avant de passer à l'entraînement, les caractéristiques d'entrée ont été normalisées en utilisant un *StandardScaler*. Cette étape est essentielle car

elle met toutes les variables sur une même échelle, ce qui permet au modèle de converger plus rapidement et de manière plus stable. La normalisation a été appliquée à l'ensemble d'entraînement, puis les mêmes paramètres ont été utilisés pour transformer l'ensemble de test.

- **Architecture du Réseau de Neurones (figure(3.1))**

- **Conception du Modèle** : L'architecture du réseau est composée de plusieurs couches :
 - Une première couche dense avec 128 neurones utilisant la fonction d'activation *ReLU*. Cette couche est régularisée avec une combinaison de régularisation *L1* et *L2* pour éviter l'overfitting.
 - Une deuxième couche dense avec 64 neurones, également avec activation *ReLU* et régularisée avec une régularisation *L1* et *L2* moins forte.
 - Une troisième couche dense avec 32 neurones, avec les mêmes paramètres de régularisation que la deuxième couche.
 - Une quatrième couche dense avec 8 neurones et activation *ReLU*.
 - Une couche de sortie dense avec 1 neurone et activation *linéaire*, adaptée pour une tâche de régression où l'objectif est de prédire une valeur continue.

- **Régularisation pour Éviter le surapprentissage**

- **Régularisation L1 et L2** : L1 et L2 sont utilisées pour pénaliser les poids du modèle, ce qui aide à réduire la complexité du modèle et donc à éviter le surapprentissage. La régularisation L1 encourage la sparsité (moins de poids non nuls), tandis que L2 empêche les poids de devenir trop grands.
- **Optimisation** : L'algorithme d'optimisation *Adam* a été choisi avec un taux d'apprentissage de 0.001. Cet optimiseur est adapté car il ajuste le taux d'apprentissage pour chaque paramètre, ce qui permet une convergence rapide et efficace.

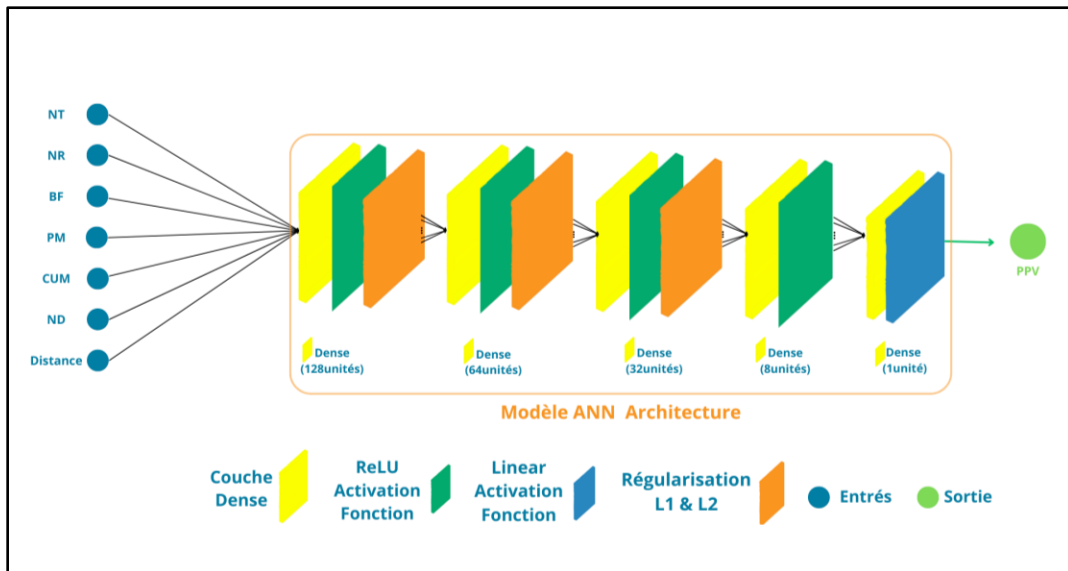
- **Techniques d'Optimisation et de Validation**

- **Early Stopping** : Un callback *EarlyStopping* a été utilisé pour surveiller la perte de validation et arrêter l'entraînement si la perte ne s'améliore plus après 20 époques. Cette

technique évite le surapprentissage en arrêtant le modèle au point où il commence à surapprendre.

- **Réduction du Taux d'Apprentissage** : Un callback *ReduceLROnPlateau* a été ajouté pour réduire le taux d'apprentissage lorsque la performance sur les données de validation stagne. Cette approche permet au modèle de converger plus finement vers un minimum local.

Fig(3.1) : Architecture du modèle

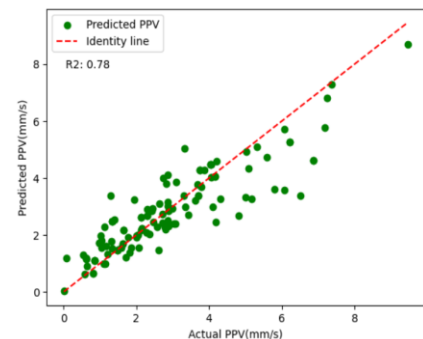


3.1.1 Évaluation et validation du modèle de Réseau de Neurones (Train set)

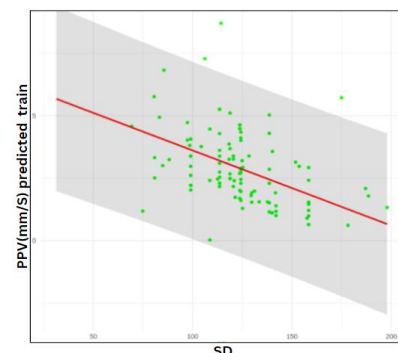
L'évaluation du modèle de réseau de neurones sur l'ensemble d'entraînement est une étape cruciale pour comprendre comment le modèle s'est ajusté aux données sur lesquelles il a été formé. Plusieurs métriques sont utilisées pour évaluer la performance du modèle, dont le Mean Squared Error (MSE), le Mean Absolute Error (MAE), et le R-squared (R^2).

- **Mean Squared Error (MSE)** : L'erreur quadratique moyenne sur l'ensemble d'entraînement est de **1.9817**. Cette métrique évalue la moyenne des carrés des erreurs, fournissant ainsi une indication sur l'écart type moyen entre les valeurs prédites et les valeurs réelles.
- **Mean Absolute Error (MAE)** : L'erreur absolue moyenne est de **0.6089**. Le MAE mesure la moyenne des erreurs absolues entre les prédictions et les observations réelles, offrant une mesure intuitive de l'exactitude du modèle.
- **R-squared (R^2)** : Le coefficient de détermination est de **0.7761** pour l'ensemble d'entraînement (**figure 3.2.1**). Ce score montre que **77,61 %** de la variance dans les données d'entraînement est expliquée par le modèle, ce qui indique un ajustement raisonnablement bon.
- En complément de ces métriques, une évaluation a été réalisée en utilisant un intervalle de confiance à **95%** (**figure 2.2**) pour vérifier la robustesse du modèle. Il a été observé que seulement **2.5%** des points prédits se trouvent en dehors de cet intervalle de confiance (**figure 3.2.2**), ce qui est inférieur au seuil de **5%**, rendant le modèle validé pour le train set.

Fig(2.2.1) : Les Points prédits par rapport aux points mesurés (Train)



Fig(2.2.2) : Les Points prédits dans l'intervalle de confiance (Train)



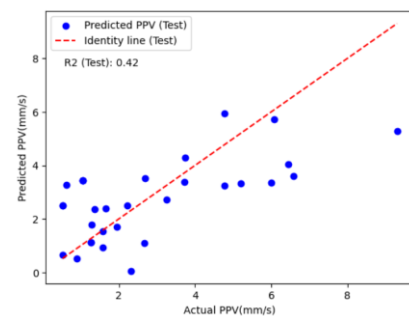
3.1.2 Évaluation et validation du modèle de Réseau de Neurones (Test set)

L'évaluation sur l'ensemble de test est essentielle pour vérifier la capacité de généralisation du modèle, c'est-à-dire sa performance sur des données qu'il n'a pas vues lors de l'entraînement.

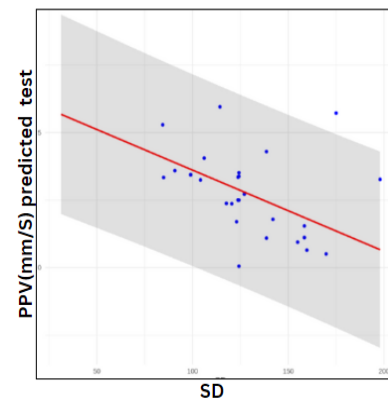
- **Mean Squared Error (MSE)** : L'erreur quadratique moyenne sur l'ensemble de test est de **4.0954**. Comparée à celle de l'ensemble d'entraînement, une MSE plus élevée sur le test set suggère que le modèle est moins précis sur des données non vues.
- **Mean Absolute Error (MAE)** : L'erreur absolue moyenne est de **1.3233** sur l'ensemble de test. Cela montre une augmentation par rapport au MAE de l'entraînement, ce qui est normal, mais cette différence indique que le modèle pourrait bénéficier d'améliorations supplémentaires pour mieux généraliser.

- **R-squared (R^2)** : Le coefficient de détermination sur l'ensemble de test est de **0.4249** (**figure 3.3.1**). Ce score montre que **42,49 %** indiquent une performance modérée en généralisation.
- En complément de ces métriques, une évaluation a été réalisée en utilisant un intervalle de confiance à **95%** **figure 2.2** pour vérifier la robustesse du modèle. Il a été observé que seulement **3.3%** des points prédits se trouvent en dehors de cet intervalle de confiance (**figure 3.3.2**), ce qui est inférieur au seuil de **5%**, rendant le modèle validé pour le train set.

Fig(3.3.1) : Les Points prédits par rapport aux points mesurés (Test)



Fig(3.3.2) : Les Points prédits dans l'intervalle de confiance (Test)



3.1.3 Test du modèle sur de nouvelles données

Le modèle de réseau de neurones a été testé sur de nouvelles données mesurées directement sur le site. Les résultats de ces tests montrent les valeurs réelles de PPV par rapport aux valeurs prédites par le modèle. Voici un **tableau 3.2** résumant ces résultats :

Tab(3.2) : Test du modèle

PPV(mm/s) mesurée	PPV(mm/s) prédite
5.19	3.34
1.32	1.79
5.38	3.35
3.789	1.98
1.16	1.14
1.5	3.11
01.02	6.20
1.14	3.15

3.2 Interprétation des résultats

Le tableau ci-dessus illustre la performance du modèle lorsqu'il est appliqué à des données inédites, montrant la différence entre les valeurs de PPV observées et celles prédites.

Après le test du modèle sur de nouvelles données, des écarts notables entre les valeurs réelles et les valeurs prédites ont été observés. Ces écarts soulignent la nécessité de continuellement affiner et améliorer le modèle. Pour aborder ce défi, l'augmentation des données a été adoptée comme une approche pour améliorer la performance du modèle..

Chapitre 4 : Augmentation du jeu de données, développement et déploiement web des modèles prédictifs

4.1 Techniques de l'augmentation de données utilisées

4.1.1 Interpolation

- **Détails** : L'interpolation crée de nouvelles données en combinant des paires de données existantes. Cela renforce la continuité des données en générant des points intermédiaires entre les valeurs réelles.
- **Principe** : Pour deux points de données x_1 et x_2 , un nouveau point x_{new} est généré tel que

$$x_{new} = \alpha \cdot (x_1) + (1 - \alpha) \cdot x_2 \text{ (équ. 17)}$$

où α est un coefficient compris entre 0 et 1.

- **Avantages** :
 - **Densification des données** : Augmente la densité des points de données, améliorant la capacité du modèle à généraliser.
 - **Préservation de la continuité** : Maintient les relations linéaires ou non linéaires entre les variables.
- **Limites** :
 - **Uniformité** : Une surabondance de données interpolées peut introduire une uniformité non représentative des données réelles.

4.1.2 Jittering

- **Détails** : Le jittering ajoute un bruit aléatoire aux variables indépendantes, rendant le modèle moins sensible aux variations mineures et augmentant sa robustesse.
- **Principe** : Pour chaque point de données x , un nouveau point x_{jitter} est généré tel que :

$$x_{jitter} = x + \epsilon \text{ (équ. 18)}$$

où ϵ est une petite valeur aléatoire tirée d'une distribution gaussienne des données.

- **Avantages** :

- **Robustesse accrue** : Aide à éviter le surajustement aux données spécifiques d'entraînement.
- **Simplicité** : Facile à mettre en œuvre et peu coûteux en termes de calcul.
- **Limites** :
 - **Risque de suréchantillonnage** : Un bruit excessif pourrait introduire des points non représentatifs des données réelles.

4.1.3 Combinaison des deux méthodes

Avec environ 150 lignes de données réelles, la combinaison des méthodes d'*interpolation* et de *jittering* permet une augmentation significative des données :

- **Interpolation** : On peut raisonnablement générer environ 150 à 600 nouvelles lignes de données par série d'interpolations.
- **Jittering** : En appliquant le *jittering* 5 fois à chaque ligne, environ 750 nouvelles lignes peuvent être générées. Typiquement, créer entre 2 et 10 nouvelles copies par ligne d'origine donne entre 300 et 1 500 nouvelles lignes.

En combinant les deux méthodes (*interpolation* et *jittering*), on peut potentiellement augmenter les données de manière significative. Par exemple, si on génère 300 nouvelles lignes par interpolation et appliquant le *jittering* 5 fois sur l'ensemble, cela pourrait nous donner 2 700 nouvelles lignes.

4.2 Développement et optimisation d'un Réseau de Neurones Artificiels (DL)

Dans cette section, nous allons présenter les résultats du modèle de réseau de neurones artificiels (ANN) après avoir appliqué les techniques d'augmentation de données décrites dans la section 4.1. L'architecture du modèle a été détaillée dans la section 3.1. Pour cette étape, nous nous concentrons uniquement sur l'affichage des résultats obtenus avec les données augmentées.

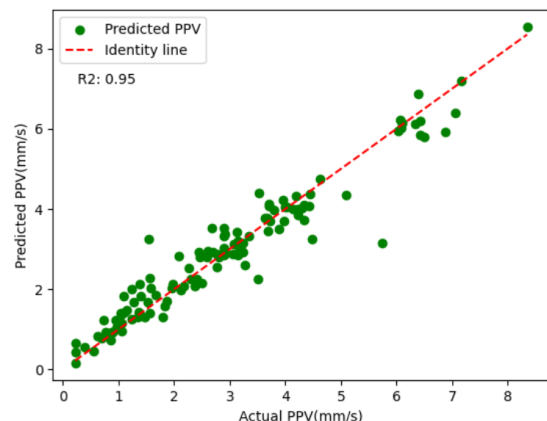
4.2.1 Évaluation et validation du modèle de Réseau de Neurones (Train set)

L'évaluation du modèle de réseau de neurones artificiels sur le jeu de données augmenté est essentielle pour comprendre comment le modèle se comporte après l'application des techniques d'augmentation de données. Les résultats suivants, basés sur les métriques de performance clés, reflètent l'amélioration apportée par l'augmentation des données.

- **Mean Squared Error (MSE)** : L'erreur quadratique moyenne sur l'ensemble d'entraînement est de **0.480**. Cette métrique mesure la moyenne des carrés des écarts entre les valeurs prédites et les valeurs réelles.
- **Mean Absolute Error (MAE)** : L'erreur absolue moyenne sur l'ensemble d'entraînement est de **0.268**. Cette mesure indique la moyenne des erreurs absolues, offrant une perspective claire sur l'exactitude des prédictions du modèle.

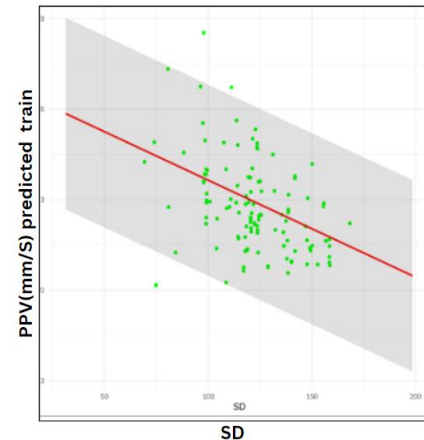
- **R-squared (R^2)** : Le coefficient de détermination est de **0.946** pour l'ensemble d'entraînement (**figure 4.1.1**). Ces valeurs élevées indiquent que le modèle explique **94,5 %** de la variance dans les données d'entraînement montrant ainsi une excellente capacité à généraliser et à capturer la structure des données.

Fig(4.1.1) : Les Points prédits par rapport aux points mesurés (Train)



- En complément de ces métriques, une évaluation a été réalisée en utilisant un intervalle de confiance à **95%** (**figure 2.2**) pour vérifier la robustesse du modèle. Il a été observé que seulement **2.5%** des points prédits se trouvent en dehors de cet intervalle de confiance (**figure 4.1.2**), ce qui est inférieur au seuil de **5%**, rendant le modèle validé pour le train set.

Fig(4.1.2) : Les Points prédits dans l'intervalle de confiance (Train)

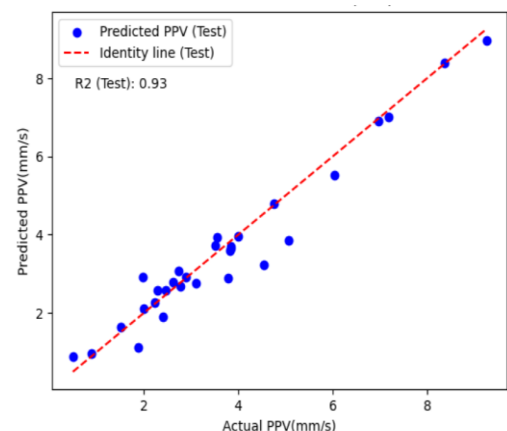


4.2.2 Évaluation et Validation du Modèle de Réseau de Neurones (Test set)

L'évaluation du modèle de réseau de neurones sur l'ensemble de test est cruciale pour évaluer la performance du modèle sur des données non vues. Voici les métriques de performance clés obtenues pour l'ensemble de test :

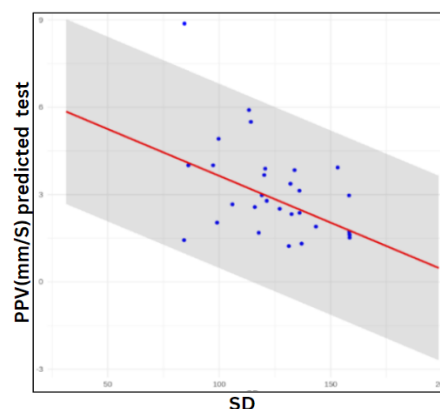
- **Mean Squared Error (MSE)** : L'erreur quadratique moyenne pour l'ensemble de test est de **0.521**. Cette métrique mesure la moyenne des carrés des écarts entre les valeurs prédites et les valeurs réelles, reflétant ainsi la dispersion des erreurs.
- **Mean Absolute Error (MAE)** : L'erreur absolue moyenne sur l'ensemble de test est de **0.314**. Le MAE indique la moyenne des erreurs absolues, fournissant une mesure intuitive de l'exactitude des prédictions du modèle.
- **R-squared (R^2)** : Le coefficient de détermination pour l'ensemble de test est de **0.933** (**figure 4.2.1**). Une capacité élevée du modèle à généraliser et à capturer les variations des données.

Fig(4.2.1) : Les Points prédits par rapport aux points mesurés (Train)



- En complément de ces métriques, une évaluation a été réalisée en utilisant un intervalle de confiance à **95%** (**figure 2.2**) pour vérifier la robustesse du modèle. Il a été observé que seulement **3.3%** des points prédits se trouvent en dehors de cet intervalle de confiance (**figure 4.2.2**), ce qui est inférieur au seuil de **5%**, rendant le modèle validé pour le train set.

Fig(4.2.2) : Les Points prédits dans l'intervalle de confiance (Test)



4.2.3 Test du modèle sur de nouvelles données

Le modèle de réseau de neurones a été testé sur de nouvelles données mesurées directement sur le site. Les résultats de ces tests montrent les valeurs réelles de PPV par rapport aux valeurs prédites par le modèle. Voici un **tableau 4.1** résumant ces résultats :

Tab(4.1) : Test du modèle

PPV(mm/s) mesurée	PPV(mm/s) prédite
5.19	5.21
1.32	1.29
5.38	5.37
1.16	1.22

4.2.4 Comparaison entre les résultats du modèle avant et après augmentation

Pour comparer les performances du modèle avant et après l'augmentation des données, nous avons calculé l'erreur en pourcentage pour chaque prédiction en utilisant la formule suivante :

$$erreur = \left| \frac{PPV \text{ mesurée} - PPV \text{ prédite}}{PPV \text{ mesurée}} \right| \times 100 \quad (\text{éq. 19})$$

Tab(4.2) : Comparaison entre les deux modèles

PPV(mm/s) mesurée	PPV Prédite (Avant Augmentation)	PPV Prédite (Après Augmentation)	Erreur Avant Augmentation (%)	Erreur Après Augmentation (%)
5,19	3.34	5.21	64.87	0.38
1,32	1.79	1.29	35.38	2.08
5,38	3.35	5.37	62.12	0.15
3,78	1.98	1.22	52.89	67.68
1,16	1.14	5.21	1.62	5.73
1,5	3.11	1,99	87.62	32.97
1,02	6.20	0,68	509.91	32.56
1,14	3.15	3,16	176.16	1.23

→ La moyenne de la réduction de l'erreur après l'application des techniques d'augmentation de données est de 33,36 %. Ce résultat montre que l'augmentation des données a globalement amélioré la précision des prédictions du modèle, même si certaines instances ont présenté une augmentation de l'erreur. Néanmoins, ces techniques ont permis de réduire considérablement l'erreur dans la majorité des cas, attestant de leur efficacité pour améliorer la performance du modèle prédictif. Cependant, il est important de noter que l'erreur ne peut être réduite de manière significative que par l'ajout de données réelles. L'augmentation des données, bien que

bénéfique, n'est pas une solution radicale et doit être utilisée en complément de données authentiques pour obtenir des résultats optimaux.

4.3 Déploiement web des modèles prédictifs

4.3.1 Framework: Flask

Flask est un micro-framework web en Python qui est couramment utilisé pour le développement de petites à moyennes applications web. Il est particulièrement apprécié pour sa simplicité, sa flexibilité et son extensibilité. Contrairement à d'autres frameworks plus robustes, comme Django, Flask n'impose pas de structures rigides, ce qui permet aux développeurs de construire des applications web selon leurs propres besoins et préférences.

Pourquoi Flask ?

- **Simplicité et légèreté** : Flask est minimaliste, ce qui le rend facile à apprendre et à utiliser. Il offre seulement les outils nécessaires pour créer une application web, sans ajouter de complexité inutile.
- **Flexibilité** : Grâce à son architecture modulaire, Flask permet aux développeurs d'intégrer uniquement les composants dont ils ont besoin. Cette flexibilité est particulièrement utile pour des projets spécifiques, comme le déploiement de modèles prédictifs.
- **Extensibilité** : Bien que Flask soit léger, il est hautement extensible. Il dispose d'une large gamme d'extensions qui permettent d'ajouter des fonctionnalités supplémentaires, telles que la gestion des bases de données, la validation des formulaires, et plus encore.

4.3.2 Développement de l'application web

Pour développer l'application web, j'ai utilisé le framework *Flask* en Python, qui est léger et bien adapté pour créer des applications web simples et efficaces. Voici les principales étapes du développement de l'application.

- **Structure de l'application :**

L'application *Flask* est construite autour de deux principaux composants :


- **Modèle de prédiction:** Deux modèles de réseaux de neurones ont été pré-entraînés et sauvegardés pour prédire le PPV (vibration des particules). Le premier modèle utilise les données d'entraînement d'origine, tandis que le second modèle a été entraîné sur des données augmentées.
- **Interface utilisateur:** Une page web simple permet aux utilisateurs de saisir les valeurs des différentes variables nécessaires pour prédire le PPV. Les utilisateurs peuvent également choisir le modèle qu'ils souhaitent utiliser pour la prédiction. La **Figure 4.3** ci-dessous représente l'interface de l'utilisateur.

Fig(4.3) : Interface utilisateur web

Please Enter your inputs HERE -->

It's time to predict PPV using AI

Predicted result depends on choice of model



PPV Predicted is:
3.95 mm/s

Predict Peak Particule Velocity

Trous
125

Rangées
9

Profondeur
4.9

Bourage
3.1

Charge
50

Diamètre
225

Distance
600

Model AUGMENTED Data

Predict

Conclusion et perspectives

Ce projet démontre le potentiel impressionnant de l'intelligence artificielle dans le domaine de la prédiction des vibrations induites par les tirs à l'explosif, un aspect crucial en Génie Civil. En développant des réseaux de neurones artificiels et en appliquant des techniques d'augmentation des données, nous avons non seulement amélioré la performance des modèles prédictifs, mais aussi facilité leur utilisation à travers une interface web conviviale.

La démarche suivie a consisté, en premier lieu, à mener une étude de cas dans la littérature pour comprendre les approches existantes. Ensuite, une analyse statistique des données par régression linéaire a été réalisée, permettant de déterminer l'intervalle de confiance, assurant ainsi la robustesse des analyses. Le prétraitement des données, incluant le nettoyage et la normalisation, a été une étape cruciale avant le développement du réseau de neurones. Après la division des données en ensembles d'entraînement et de test, l'architecture du réseau a été définie, entraînée et évaluée, notamment en comparant les résultats avec la droite d'identité ($y=x$) et en tenant compte de l'intervalle de confiance.

Pour améliorer la précision du modèle, des techniques d'augmentation des données, telles que l'interpolation et le jittering, ont été appliquées, suivies d'un réentraînement du modèle sur ces données enrichies. Enfin, une interface web a été développée à l'aide de *Flask* pour permettre aux utilisateurs de saisir de nouvelles données et d'obtenir des prédictions facilement.

Cependant, l'utilisation de l'intelligence artificielle dans des domaines sensibles, comme le génie civil, nécessite une évaluation rigoureuse pour garantir la fiabilité et la précision des modèles. La performance des modèles IA doit être continuellement surveillée et affinée, surtout lorsque des décisions critiques en dépendent. Ce projet illustre non seulement les possibilités offertes par l'IA, mais aussi l'importance d'une approche responsable dans son application.

Les perspectives futures incluent l'amélioration continue de la précision du modèle par l'ajout de données réelles et l'exploration d'autres techniques d'augmentation. En outre, l'intégration de nouvelles sources de données et de modèles d'apprentissage plus sophistiqués pourrait renforcer les résultats

Références

- [1] Amiri, Maryam & Bakhshandeh Amnieh, Hassan & Hasanipanah, Mahdi & Khanli, Leili. (2016). A new combination of artificial neural network and K-nearest neighbors models to predict blast-induced ground vibration and air-overpressure. *Engineering with Computers*. 32. doi:(<https://doi.org/10.1007/s00366-016-0442-5>)
- [2] Álvarez-Vigil, A. E., González-Nicieza, C., Gayarre, F. L., & Álvarez-Fernández, M. I. (2012). Predicting blasting propagation velocity and vibration frequency using artificial neural networks. *International Journal of Rock Mechanics and Mining Sciences*, 55, 108-116. doi:(<https://doi.org/10.1016/j.ijrmms.2012.05.002>)
- [3] Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer. doi:([10.1117/1.2819119](https://doi.org/10.1117/1.2819119))
- [4] Elevli B., Arpaz E., (2010). Evaluation of parameters affected on the blast induced ground vibration (BIGV) by using relation diagram method (rdm), *Acta Montanistica Slovaca* 15 (4) (2010) 261.
- [5] Drucker, Harris & Burges, Christopher & Kaufman, Linda & Smola, Alexander & Vapnik, V. (1997) Support vector regression machines. *Adv Neural Inform Process Syst*. 28. 779-784.
- [6] Duvall. W.I., Petkof B., (1958). Spherical propagation of explosion-generated strain pulses in rock, Tech. rep., Bureau of Mines.
- [7] Ghosh A., Daemen J.J., (1985). Statistics—a key to better blast vibration predictions, in: *Proceedings of the 26th US Symposium on Rock Mechanics*, Rapid City, 1985, pp. 1141–1149.
- [8] Gupta R., Roy P.P. & Singh B., (1988). On a blast induced blast vibration predictor for efficient blasting, in: *Proceedings of the 22nd international conference of safety in Mines*, Beijing, China, 1988, pp. 1015–1021.

- [9] H. Koch, Zur Möglichkeit der Abgrenzung von Lademengen bei Steinbruchsprengungen nach festgestellten Erschütterungsstärken, Nobel Hefte 24 (1958) 92–96.
- [10] LeCun, Yann & Bengio, Y. & Hinton, Geoffrey. (2015). Deep Learning. *Nature*. 521. 436–44. doi:([10.1038/nature14539](https://doi.org/10.1038/nature14539))
- [11] Müller B., Hausmann J. & Niedzwiedz H., (2007). Comparison of different methods of measuring and calculating blast vibrations in rock masses, in: The 4th EFEE World Conference, Vienna, 2007.
- [12] Monjezi M., Ghafurikalajahi M. & Bahrami A., (2011). Prediction of blast-induced ground vibration using artificial neural networks. *Tunnelling and Underground Space Technology*, 2011, 26 (1)46–50. doi:(<https://doi.org/10.1016/j.tust.2010.05.002>)
- [13] Quinlan, J.R. (1986) Induction of decision trees. *Machine Learning* 1, 81–106. doi:(<https://doi.org/10.1007/BF00116251>)
- [14] Roy, P. (1993). Putting ground vibration predictions into practice, *Colliery Guardian*, 241 (2), 1993, 63–7.
- [15] Shang, Y. & Nguyen, Hoang & Bui, Xuan-Nam & Hieu, Tran & Moayed, Hossein. (2019). A Novel Artificial Intelligence Approach to Predict Blast-Induced Ground Vibration in Open-Pit Mines Based on the Firefly Algorithm and Artificial Neural Network. *Natural Resources Research*. 29. 1–15. doi: (<https://doi.org/10.1007/s11053-019-09503-7>)
- [16] Seber, G. A. F., & Lee, A. J. (2012). *Linear Regression Analysis*. Wiley Series in Probability and Statistics Wiley. doi:([10.1002/9780471722199](https://doi.org/10.1002/9780471722199))
- [17] Zeng, J.; Roussis, P.C.; Mohammed, A.S.; Maraveas, C.; Fatemi, S.A.; Armaghani, D.J.; Asteris, P.G. (2021) Prediction of Peak Particle Velocity Caused by Blasting through the Combinations of Boosted-CHAID and SVM Models with Various Kernels. *Appl. Sci.* 2021, 11, 3705. doi:([10.3390/app11083705](https://doi.org/10.3390/app11083705))

[18] Velasquez, Luis F., Prediction of Blast-Induced Ground Vibrations: A Comparison Between Empirical and Artificial-Neural-Network Approaches (2023). Theses and Dissertations--Mining Engineering. 75. doi: (<https://doi.org/10.13023/etd.2023.015>)

Outils et Technologies Utilisés

Bootstrap: Utilisé pour créer des interfaces utilisateur réactives et modernes.

Canva: Utilisé pour la création de visuels et de présentations pour le projet.

Google Colab: Outil utilisé pour l'exécution de notebooks Python dans le cloud, facilitant l'expérimentation et l'entraînement des modèles.

Google Docs et Google Sheets: Utilisés pour la documentation, la collaboration et l'organisation des données.

GitHub: Utilisé pour la gestion de version du code et la collaboration. Le repository du projet est disponible à l'adresse suivante : [<https://github.com/YassineOurara/Explosion-induced-vibrations-AI-Prediction>].

Flask: Framework web léger utilisé pour développer l'application web de prédiction des vibrations.

JavaScript: Langage de programmation utilisé pour ajouter des fonctionnalités interactives à l'application web.

Python: Utilisé pour le développement des modèles prédictifs et pour les calculs de traitement des données.

R: Utilisé pour l'analyse statistique, la vérification des résultats et l'évaluation des intervalles de confiance.

Scikit-Learn (sklearn): Employé pour le prétraitement des données, la modélisation et l'évaluation des performances des modèles.

Tailwind CSS: Framework CSS utilisé pour styliser l'interface web de manière personnalisée et efficace.

Visual Studio Code: Environnement de développement intégré (IDE) utilisé pour coder, déboguer et tester le projet.



Année Universitaire – 2023-2024 –
Master Intelligence Artificielle et Analyse des Données – MIAAD –
Rapport – Stage PFE –

