

Université Nouveaux Horizons

**Apprentissage supervisé, la reconnaissance des
plaques d'immatriculation utilisant l'algorithme
de descente de gradient stochastique.**

(Supervised learning, License Recognition using the
stochastic gradient descent algorithm)

par TSHELEKA KAJILA HASSAN

Faculté des Sciences Informatiques 2022
Calcul Scientifique

Mémoire présenté à la Faculté des Sciences Informatique en vue de
l'obtention du grade de Licencié en Sciences informatique. Filière : Calcul
Scientifique

@ TSHELEKA Hassan, 8 janvier 2022

Résumé

Au cours de la dernière décennie, la taille des données a augmenté plus rapidement que la vitesse des processeurs. Dans ce contexte, faire un traitement de reconnaissance des formes sur des vidéos en temps réel, les ensembles de données d'entraînement pour les problèmes de détection d'objets sont généralement très volumineux et les capacités des méthodes d'apprentissage automatique statistique sont limitées par le temps de calcul plutôt que par la taille de l'échantillon.

Le cas à des problèmes d'apprentissage grande échelle implique la complexité de calcul de l'algorithme d'optimisation sous-jacent de manière non triviale. Des algorithmes d'optimisation improbables tels que la **descente de gradient stochastique** (en anglais : **Stochastic Gradient Descent** ou SGD) montre des performances étonnantes pour les problèmes à grande échelle, lorsque l'ensemble d'apprentissage est volumineux. En particulier, le gradient stochastique du second ordre et la SGD moyenné, n'utilisent qu'un seul nouvel échantillon d'apprentissage à chaque itération, sont asymptotiquement efficaces après un seul passage sur l'ensemble d'apprentissage.

Ce travail vise à proposer une méthode intelligente, basé sur l'intelligence artificielle, qui permet aux ordinateurs et aux systèmes de dériver des informations significatives à partir d'images numériques, de vidéos et d'autres entrées visuelles, dans notre contexte la reconnaissance des plaques d'immatriculation des véhicules à l'aide du classificateur de descente de gradient stochastique Ridge-Adaline (en anglais : Ridge Adaline Stochastic Gradient Descent ou RASGD). Pour minimiser la **fonction coût** du classificateur, le RASGD adopte un modèle d'optimisation sans contrainte. De plus, pour augmenter la vitesse de convergence du classificateur, le classificateur de descente de gradient stochastique Adaline, (Adaline Stochastic Gradient Descent) est intégré à la Ridge Regression.

Mots clés : Supervised Learning, Computer Vision, Stochastic Gradient Descent, Pattern recognition.

Abstract

Recent cationic technological advancements in the field of information and communication have introduced intelligent ways to manage various aspects of life.

In this context, doing pattern recognition processing on videos in real time, the capabilities of statistical machine learning methods are limited by computing time rather than sample size. A more precise analysis reveals qualitatively different trade-offs for the case of small and large-scale learning problems. The large-scale case involves the computational complexity of the underlying optimization algorithm in a non-trivial way. Unlikely optimization algorithms such as stochastic gradient descent show amazing performance for large scale problems. In particular, the second order stochastic gradient and the averaged stochastic gradient are asymptotically efficient after a single pass over the training set. This work aims to provide an artificial intelligence (AI) -based intelligent system for vehicle license plate recognition using the Ridge Stochastic Gradient Descent Adaline (aka RASGD) Ridge Stochastic Gradient Descent Classifier. To minimize the cost function of the classifier, the RASGD adopts an unconstrained optimization model. In addition, to increase the speed of convergence of the classifier, the Adaline Stochastic Gradient Descent classifier is integrated with Ridge Regression. Finally, to validate the efficiency of intelligent system, the results of the proposed scheme were compared with advanced machine learning algorithms such as Support Vector Machine (SVM) and logistic regression methods. The intelligent RASGD system achieves 92% accuracy, which is better than other selected classifiers.

Keywords : Suppevised Learning, Computer Vision, Stocastic Gradient Descent, Objects recognition.

Remerciements

Je remercie...

Rappelez-vous de remercier L^AT_EX

Table des matières

Résumé	i
Abstract	i
Remerciements	ii
0 Introduction	2
0.1 Généralités	2
0.2 Contexte de notre recherche	3
0.3 Problématique	4
0.4 Objectifs	4
1 Les bases mathématiques pour le Machine Learning	6
1.1 Eléments de calcul différentiel	6
1.1.1 Convexité	6
1.1.2 Développement limité	7
1.1.3 Fonctions dérivables	7
1.2 Statistique & probabilité	11
1.3 Série statistique	11
1.3.1 Echantillonnage	11
1.3.2 Analyse bayésienne	11
1.4 Méthode d'optimisation et de minimisation d'erreur	12
1.4.1 Erreur et fonction coût	12
1.4.2 Moindres carrés linéaires	13
1.4.3 Descent de gradient	13
1.4.4 Descente de gradient stochastique	14
1.5 Apprentissage de modèles linéaires	14
1.5.1 Régression Linéaire	14
1.5.2 Régression Logistique	15
1.6 Réseaux de neurones	15
1.6.1 Perceptron	15

1.6.2	Neurones	16
1.6.3	CNN	16
1.6.4	RNN	16
2	Méthodologie	17
2.1	Méthode de détection et de reconnaissance d'objets	17
2.2	Détection d'objets en mouvement	17
2.2.1	Représentation d'une vidéo	17
2.2.2	Extraction des régions d'intérêt	17
3	Expérimentation (Implémentation)	18
3.1	Materiel utilisé pour l'implémentation	18
3.2	Élaboration de la base de données	18
3.3	Construiction d'un model d'entrainement	18
3.4	Résultats d'expérimentation	19
4	Conclusion	20
4.1	First class test	20
	Annexe A : Adaline SGD Training Code	20
	Annexe B : Result of Test	21
	Annexe C : Dataset & Model	22
	References	23

Chapitre 0

Introduction

0.1 Généralités

La reconnaissance d'objet dans les vidéos est un problème important dans la vision par ordinateur (Computer Vision¹) et en traitement d'images. Cette tâche est très utile vue l'accroissement du nombre de vidéos générées par des smartphones, des systèmes de sécurité, des caméras de circulation et autres dispositifs dotés d'instruments visuels. La reconnaissance automatique des objets en vidéos peut ainsi renforcer la sécurité, faciliter la gestion des vidéos ainsi que permettre de nouvelles applications en interaction personne/machine.

Par ailleurs, les images numériques et la vidéo sont devenues indispensables pour divers domaines d'application, tels que la détection d'intrusions pour la sécurité, la surveillance du trafic routier, la médecine pour l'imagerie médicale, ou encore lors des événements sportifs (ex., renforcement de l'arbitrage, création automatique de résumés). Des contraintes d'exploitation découlent des observations citées ci-dessus, parmi lesquelles nous citerons celles qui sont liées à la reconnaissance des objets en mouvement dans les vidéos. Par exemple, de nos jours, un très grand nombre de caméras est déployé exclusivement pour la surveillance vidéo.[1] Souvent, le contenu de ces vidéos est interprété par des opérateurs humains qui engendrent des coûts exorbitants pour le suivi et l'analyse du contenu, sans mentionner les erreurs qui peuvent être induites par la fatigue et l'inattention humaine. Un des problèmes importants abordés dans la surveillance vidéo est la reconnaissance des types d'objets en mouvement et leurs actions, afin de détecter, par

1. La vision par ordinateur est un domaine de l'intelligence artificielle (IA) qui permet aux ordinateurs et aux systèmes de dériver des informations significatives à partir d'images numériques, de vidéos et d'autres entrées visuelles, et de prendre des mesures ou de faire des recommandations sur la base de ces informations.

exemple, des menaces potentielles (ex., vols, attentats, accidents), ou tout simplement pour des fins de statistiques (ex., compter le nombre d'individus, de voitures dans une entrée de parc).

Les applications du monde réel démontrent l'importance de la vision par ordinateur pour les entreprises, les secteurs du divertissement, des transports, des soins de santé et dans la vie quotidienne. L'un des principaux moteurs de la croissance de ces applications est le flot d'informations visuelles provenant des médias numériques (ex., internet, la télévision, les vidéos personnelles, la surveillance vidéo). Ces données pourraient jouer un rôle majeur dans les opérations de toutes les industries, mais elles sont aujourd'hui inutilisées. Ces informations constituent un banc d'essai pour la formation des applications de vision par ordinateur et une rampe de lancement pour leur intégration dans toute une série d'activités humaines :(IBM, Computer Vision)

- IBM a utilisé la vision par ordinateur pour créer My Moments pour le tournoi de golf Masters 2018. IBM Watson a regardé des centaines d'heures d'enregistrements filmés de Masters et a pu identifier les images (et les sons) des plans importants. Ces moments clés ont été organisés et livrés aux fans sous forme de séquences personnalisées.
- Google Translate² permet aux utilisateurs de pointer la caméra d'un smartphone vers un panneau dans une autre langue et d'obtenir presque immédiatement une traduction du panneau dans la langue de leur choix.
- Le développement des véhicules autonomes repose sur la vision par ordinateur, qui donne un sens aux données visuelles fournies par les caméras et autres capteurs de la voiture. Il est essentiel d'identifier les autres voitures, les panneaux de signalisation, les marqueurs de voie, les piétons, les vélos et toutes les autres informations visuelles rencontrées sur la route.
- IBM applique la technologie de vision par ordinateur avec des partenaires comme Verizon afin d'amener l'IA intelligente à la périphérie et d'aider les constructeurs automobiles à identifier les défauts de qualité avant qu'un véhicule ne quitte l'usine.

??? compléter la liste des existants ?

0.2 Contexte de notre recherche

Au cours de la dernière décennie, la taille des données a augmenté plus rapidement que la vitesse des processeurs. Dans ce contexte, faire un traite-

2. Google Translate est un service de traduction automatique fourni par Google, qui permet de traduire un texte ou une page Web dans une autre langue.

ment de reconnaissance des formes sur des vidéos en temps réel, les ensembles de données d'entraînement pour les problèmes de détection d'objets sont généralement très volumineux et les capacités des méthodes d'apprentissage automatique statistique sont limitées par le temps de calcul plutôt que par la taille de l'échantillon.[7]
??? sur quoi se base mes recherche ?

0.3 Problématique

Pour le système de vision humain, la reconnaissance des objets est une tâche simple et triviale. La vision par ordinateur fonctionne de la même manière que la vision humaine, sauf que les humains ont une longueur d'avance. La vue humaine a l'avantage de pouvoir s'entraîner à distinguer les objets, à en déterminer la distance, à savoir s'ils sont en mouvement et si quelque chose ne va pas dans une image. L'humain est capable de faire la distinction, d'une part, entre des objets et l'arrière-plan d'une image et d'autre part, entre plusieurs objets présents dans une scène de vidéo.

La vision par ordinateur a besoin de beaucoup de données. Elle exécute des analyses de données encore et encore jusqu'à ce qu'elle perçoive des distinctions et reconnaisse finalement les images. Par exemple, pour entraîner un ordinateur à reconnaître des plaques d'immatriculation de voiture, elle doit recevoir de grandes quantités d'images de plaque d'immatriculation et d'éléments liés aux plaques pour apprendre les différences et reconnaître une plaque, en particulier un pneu sans défaut.

??? qu'est-ce que ça résout comme problème ?

??? c'est quoi ma retouche, pour le problème, par rapport à l'existant ?

0.4 Objectifs

Ce travail vise à proposer une méthode intelligent, basé sur l'intelligence artificielle (IA), qui permet aux ordinateurs et aux systèmes de dériver des informations significatives à partir d'images numériques, de vidéos et d'autres entrées visuelles, dans notre contexte la reconnaissance des plaques d'immatriculation des véhicules à l'aide du classificateur de descente de gradient stochastique Ridge-Adaline (en anglais : Ridge Adaline Stochastic Gradient Descent ou RASGD). Pour minimiser la fonction de coût du classificateur, le RASGD adopte un modèle d'optimisation sans contrainte. De plus, pour augmenter la vitesse de convergence du classificateur, le clas-

sificateur de descente de gradient stochastique Adaline, (Adaline Stochastic Gradient Descent) est intégré à Ridge Régression.[10]

??? pourquoi faire un système de reconnaissance des formes ?

Chapitre 1

Les bases mathématiques pour le Machine Learning

1.1 Éléments de calcul différentiel

1.1.1 Convexité

Définition : (Ensemble convexe) Une partie $\mathcal{C} \subset \mathbb{R}^n$ est dite convexe si et seulement si pour tout $(x, y) \in \mathcal{C}^2$, et pour tout $\alpha \in [0, 1]$,

$$\alpha x + (1 - \alpha)y \in \mathcal{C}$$

combinaison convexe.

Définition : (Fonction convexe) Une fonction f d'un intervalle réel $I \in \mathcal{C}$ est dite fonction convexe lorsque, $\forall (x, y)$ de I tel que $(x, y) \in \mathcal{C}^2$ et tout $\alpha \in [0, 1]$ on a :

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y) \quad (1.1)$$

et si

$$f(\alpha x + (1 - \alpha)y) < \alpha f(x) + (1 - \alpha)f(y) \quad (1.2)$$

on dit que la fonction est strictement convexe dans \mathcal{C}

Ex :

- La fonction $f(x) = x^2$ est convexe.
- La fonction $f(x) = x^T x$ est convexe.
- La fonction $f(x) = x^T A x$ est convexe, ssi A est symétrique semi-définie positive.

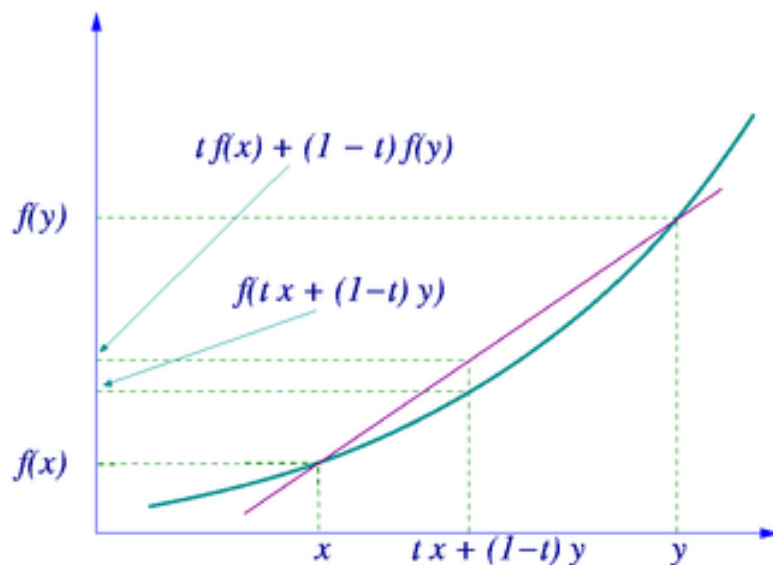


FIGURE 1.1 – Fonction convexe (image de Wikipédia)

1.1.2 Développement limité

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.

A Différentiabilité au sens de Frechet

??? parler de son implication dans le gradient Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

1.1.3 Fonctions dérivables

A Gradient

Définition : Le gradient d'une fonction de plusieurs variables en un certain point est un vecteur qui caractérise la variabilité de cette fonction au voisinage de ce point. Défini en tout point où la fonction est différentiable, il définit un champ de vecteurs, également dénommé gradient. Le gradient

est la généralisation à plusieurs variables de la dérivée d'une fonction d'une seule variable.

Définition mathématique : Dans un système de coordonnées cartésiennes, le gradient d'une fonction $f(x_1, x_2, \dots, x_n)$ est le vecteur de composantes $\partial f / \partial x_i$ ($i = 1, 2, \dots, n$), c'est-à-dire les dérivées partielles de f par rapport aux coordonnées.

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix} \in \mathbb{R}^n$$

Gradient sous forme de développement limité : Si une application admet un gradient en un point, alors on peut écrire ce développement limité du premier ordre (voir le point 1.1.2)

$$f(x+h) = f(x) + \langle \nabla f(x) | h \rangle + o(h)$$

ou

$$f(x-h) = f(x) - \langle \nabla f(x) | h \rangle + o(h)$$

Numériquement, il est très intéressant de faire ensuite la demi-différence des deux développements pour obtenir la valeur du gradient et on note que celui-ci ne dépend pas en fait de la valeur de la fonction au point $x : f(x)$. Cette formule a l'avantage de tenir compte des gradients du 2e ordre et est donc beaucoup plus précise et numériquement robuste. L'hypothèse est, en pratique, de connaître les valeurs "passé" et "futur" de la fonction autour d'un petit voisinage du point x .

Définition numérique : Une fonction multivariée (à variable vectorielle) $f(x) : \mathbb{R}^n \rightarrow \mathbb{R} : x \rightarrow f(x)$ définie sur un ouvert $O \in \mathbb{R}^n$ est dite dérivable (au sens de Fréchet ??) en x ssi il existe un vecteur noté $\nabla f(x) \in \mathbb{R}^n$ tel que

$$f(x+h) = f(x) + \nabla f(x)^T h + o(\|h\|) \quad (1.3)$$

$\nabla f(x) \in \mathbb{R}^n$ et où l'on a posé que le reste $o(\|h\|) = \|h\| \epsilon(h) \in \mathbb{R}^n$, avec $h \in \mathbb{R}^n$

$$\epsilon(h) : \mathbb{R}^n \rightarrow \mathbb{R}, \quad \lim_{\|h\| \rightarrow 0} \epsilon(h) = 0.$$

Le vecteur $\nabla f(x)$ est unique et s'appelle **gradient** de $f(x)$ en x . Le gradient s'adresse aux fonctions scalaires à variables vectorielles.

A propos de la notation $o(\|h\|)$: La notation de Landau $o(\|h\|)$ traduit le comportement d'une fonction de h qui [est ??] tend vers 0 d'un ordre de grandeur plus vite que $\|h\|$.

Elle est infiniment plus petit que h dans le voisinage de 0.

B Hessienne

Définition mathématique : Etant donnée une fonction f à valeurs réelles

$$f : \mathbb{R}^n \rightarrow \mathbb{R}; (x_1, \dots, x_n) \mapsto f(x_1, \dots, x_n)$$

dont toutes les dérivées partielles secondes existent, le coefficient d'indice i, j de la **matrice hessienne**¹ $H(f)$ vaut $H_{ij}(f) = \frac{\partial^2 f}{\partial x_i \partial x_j}$.

Autrement dit,

$$H(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}.$$

Définition numérique : Supposons que $f : \mathbb{R}^n \rightarrow \mathbb{R}$ définie sur un ouvert $\mathcal{O} \in \mathbb{R}^n$. La fonction $f(x)$ est dite 2 fois continûment dérivable (au sens de Fréchet ??) si en tout $x \in \mathcal{O}$ on a

$$f(x+h) = f(x) + \nabla f(x)^T h + \frac{1}{2} h^T \nabla^2 f(x) h + o(\|h\|^2) \quad (1.4)$$

avec $\nabla f(x) \in \mathbb{R}^{n \times n}$ et où on a posé que le reste $o(\|h\|^2) = \|h\| \epsilon(h) \in \mathbb{R}$ avec $\lim_{\|h\| \rightarrow 0} \epsilon(h) = 0$. La matrice carrée symétrique $\nabla^2 f(x)$ appelée **Hessien** de $f(x)$ en x . Remarque :

$$\lim_{\|h\| \rightarrow 0} \frac{o(\|h\|^2)}{\|h\|} = 0 \in \mathbb{R}$$

La Hessienne s'adresse aux fonctions scalaires à variables vectorielles.

1. En mathématiques, la matrice hessienne (ou simplement la hessienne) d'une fonction numérique f est la matrice carrée, notée $H(f)$, de ses dérivées partielles secondes.

C Jacobienne

Définition mathématique : Soit F une fonction d'un ouvert de \mathbb{R}^n à valeurs dans \mathbb{R}^m ($F : \mathbb{R}^n \rightarrow \mathbb{R}^m$). Une telle fonction est définie par ses m fonctions composantes à valeurs réelles :

$$F : \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix}.$$

Les dérivées partielles de ces fonctions en un point M , si elles existent, peuvent être rangées dans une matrice à m lignes et n colonnes, appelée **matrice jacobienne**² de F :

$$J_F(M) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}.$$

La case sur la ligne i et la colonne j contient $\frac{\partial f_i}{\partial x_j}$ qui est la dérivée partielle de f_i selon la variable x_j . Cette matrice est notée :

$$J_F(M), \quad \frac{\partial (f_1, \dots, f_m)}{\partial (x_1, \dots, x_n)} \quad \text{ou} \quad \frac{D(f_1, \dots, f_m)}{D(x_1, \dots, x_n)}$$

Pour $i = 1, \dots, m$, la i -ème ligne de cette matrice est la transposée du vecteur **gradient** (voir le point A) au point M de la fonction f_i , lorsque celui-ci existe. La matrice jacobienne est également la matrice de la différentielle de la fonction, lorsque celle-ci existe.

Définition numérique : Soit $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ définie sur un ouvert $\mathcal{O} \subset \mathbb{R}$. On dit que $f(x)$ est dérivable (au sens de Fréchet) en x , si chacune des composantes $f_i(x)$ est dérivable en x . On a alors

$$f(x+h) = f(x) + D_f(x)h + o(\|h\|) \quad (1.5)$$

avec $D_f(x) \in \mathbb{R}^{n \times m}$ et/où $o(\|h\|) = \|h\|\epsilon(h) \in \mathbb{R}^m$ avec $\lim_{\|h\| \rightarrow 0} \epsilon(h) = 0$.

Remarque :

$$\lim_{\|h\| \rightarrow 0} \frac{o(\|h\|^2)}{\|h\|} = 0 \in \mathbb{R}$$

2. En analyse vectorielle, la matrice jacobienne est la matrice des dérivées partielles du premier ordre d'une fonction vectorielle en un point donné.

$$\text{Soient } x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n \text{ et } f(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{bmatrix} \in \mathbb{R}^m$$

$$D_f(x) = \begin{bmatrix} \frac{\partial f_1(x)}{\partial x_1} & \dots & \frac{\partial f_1(x)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m(x)}{\partial x_1} & \dots & \frac{\partial f_m(x)}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \nabla f_1(x)^T \\ \nabla f_2(x)^T \\ \vdots \\ \nabla f_m(x)^T \end{bmatrix} \in \mathbb{R}^{n \times m},$$

La matrice $D_f(x) \in \mathbb{R}^{n \times m}$ est appelée **Jacobienne** de $f(x)$ en x . La Jacobienne s'adresse aux fonctions vectorielles à variables vectorielles.

Note : Lorsque $m = 1$ la Jacobienne est la même que le gradient car il s'agit d'une généralisation du gradient.

1.2 Statistique & probabilité

1.3 Série statistique

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

1.3.1 Echantillonnage

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

1.3.2 Analyse bayésienne

La statistique bayésienne est une théorie dans le domaine des statistiques basée sur l'interprétation bayésienne de la probabilité où la probabilité exprime un degré de croyance en un événement. Le degré de croyance peut être

basé sur des connaissances antérieures sur l'événement, telles que les résultats d'expériences précédentes, ou sur des croyances personnelles sur l'événement. Cela diffère d'un certain nombre d'autres interprétations de la probabilité, telles que l'interprétation fréquentiste qui considère la probabilité comme la limite de la fréquence relative d'un événement après de nombreux essais.

Les statistiques bayésiennes portent le nom de Thomas Bayes³, qui a formulé un cas spécifique du théorème de Bayes dans un article publié en 1763.

Theorem 1 (Théorème de Bayes). *Le théorème de Bayes est utilisé dans les méthodes bayésiennes pour mettre à jour les probabilités, qui sont des degrés de croyance, après avoir obtenu de nouvelles données. Compte tenu de deux événements A et B , la probabilité conditionnelle de A étant donné que B est vrai s'exprime comme suit :*

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)} \quad (1.6)$$

où $\mathbb{P}(B) \neq 0$ Bien que le théorème de Bayes soit un résultat fondamental de la théorie des probabilités, il a une interprétation spécifique dans les statistiques bayésiennes.

1.4 Méthode d'optimisation et de minimisation d'erreur

1.4.1 Erreur et fonction coût

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

A Erreur d'apprentissage

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo

3. Thomas Bayes était un Anglais statisticien, philosophe et ministre presbytérien qui est connu pour la formulation d'un cas spécifique du théorème qui porte son nom : théorème de Bayes.

consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

Quis autem vel eum iure reprehenderit qui in ea voluptate velit esse quam nihil molestiae consequatur, vel illum qui dolorem eum fugiat quo voluptas nulla péariatur.

B Fonction cout ℓ cas de la régression linéaire

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

C Fonction cout ℓ cas de la classification

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

1.4.2 Moindres carrés linéaires

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

1.4.3 Descent de gradient

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim

veniam,

$$A = \begin{pmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1n} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & x_{m3} & \cdots & x_{mn} \end{pmatrix}$$

quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

1.4.4 Descente de gradient stochastique

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

1.5 Apprentissage de modèles linéaires

1.5.1 Régression Linéaire

A Le probleme de la régression linéaire

Lorem $O(h^3)$ ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

B Le cas de le régression linéaire

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo

consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

C La fonction d'erreur

1.5.2 Régression Logistique

A Le problème de classification

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.

B Le cas séparable

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.

C Le cas non séparable

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.

D Le modèle de la régression logistique

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

1.6 Réseaux de neurones

1.6.1 Perceptron

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam,

1.6.2 Neurones

1.6.3 CNN

1.6.4 RNN

Chapitre 2

Méthodologie

2.1 Méthode de détection et de reconnaissance d'objets

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

2.2 Détection d'objets en mouvement

2.2.1 Représentation d'une vidéo

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

2.2.2 Extraction des régions d'intérêt

xxx ...

Chapitre 3

Expérimentation (Implémentation)

3.1 Matériel utilisé pour l'implémentation

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

3.2 Élaboration de la base de données

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

3.3 Construction d'un model d'entraînement

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

lum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

3.4 Résultats d'expérimentation

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

At vero eos et accusamus et iusto odio dignissimos ducimus qui blanditiis praesentium voluptatum deleniti atque corrupti quos dolores et quas molestias excepturi sint occaecati cupiditate non provident, similique sunt in culpa qui officia deserunt mollitia animi, id est laborum et dolorum fuga. Et harum quidem rerum facilis est et expedita distinctio. Nam libero tempore, cum soluta nobis est eligendi optio cumque nihil impedit quo minus id quod maxime placeat facere possimus, omnis voluptas assumenda est, omnis dolor repellendus. Temporibus autem quibusdam et aut officiis debitis aut rerum necessitatibus saepe eveniet ut et voluptates repudiandae sint et molestiae non recusandae. Itaque earum rerum hic tenetur a sapiente delectus, ut aut reiciendis voluptatibus maiores alias consequatur aut perferendis doloribus asperiores repellat

Chapitre 4

Conclusion

4.1 First class test

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Sed ut perspiciatis unde omnis iste natus error sit voluptatem accusantium doloremque laudantium, totam rem aperiam, eaque ipsa quae ab illo inventore veritatis et quasi architecto beatae vitae dicta sunt explicabo. Nemo enim ipsam voluptatem quia voluptas sit aspernatur aut odit aut fugit, sed quia consequuntur magni dolores eos qui ratione voluptatem sequi nesciunt. Neque porro quisquam est, qui dolorem ipsum quia dolor sit amet, consectetur, adipisci velit, sed quia non numquam eius modi tempora incidunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim ad minima veniam, quis nostrum exercitationem ullam corporis suscipit laboriosam, nisi ut aliquid ex ea commodi consequatur? Quis autem vel eum iure reprehenderit qui in ea voluptate velit esse quam nihil molestiae consequatur, vel illum qui dolorem eum fugiat quo voluptas nulla pariatur

Annexe A : Adaline SGD Training Code

Annexe B : Result of Test

Annexe C : Dataset & Model

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Bibliographie

- [1] Yaovi Ahadjitse. *Reconnaissance d'objets en mouvement dans la vidéo par description géométrique et apprentissage supervisé*. PhD thesis, Université du Québec en Outaouais, 2013.
- [2] T Syed Akheel, V Usha Shree, and S Aruna Mastani. Stochastic gradient descent linear collaborative discriminant regression classification based face recognition. *Evolutionary Intelligence*, pages 1–15, 2021.
- [3] Shun-ichi Amari. Backpropagation and stochastic gradient descent method. *Neurocomputing*, 5(4-5) :185–196, 1993.
- [4] Vincent Barra Antoine Cornuéjols, Laurent Michet. *Apprentissage automatique : Deep learning, concepts et algorithmes*. Eyrolles, 3rd edition, 2018.
- [5] Unkown Author. *Python Machine Learning*. Unkown Publisher, 2008.
- [6] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer-Verlag New York, first edition, 2006.
- [7] Léon Bottou. Large-scale machine learning with stochastic gradient descent. In *Proceedings of COMPSTAT'2010*, pages 177–186. Springer, 2010.
- [8] Léon Bottou. Stochastic gradient descent tricks. In *Neural networks : Tricks of the trade*, pages 421–436. Springer, 2012.
- [9] IBM Corporation. What is computer vision ?, 2017.
- [10] Natarajan Deepa, B Prabadevi, Praveen Kumar Maddikunta, Thippa Reddy Gadekallu, Thar Baker, M Ajmal Khan, and Usman Tariq. An ai-based intelligent system for healthcare analysis using ridge-adaline stochastic gradient descent classifier. *The Journal of Supercomputing*, 77 :1998–2017, 2021.
- [11] Kary Främling. Scaled gradient descent learning rate. *Reinforcement Learning With Light-Seeking Robot, Proceedings of ICINCO*, pages 1–8, 2004.

- [12] Thilo-Thomas Frieß and Robert F Harrison. A kernel based adaline. In *ESANN*, volume 72, pages 21–23. Citeseer, 1999.
- [13] Rob GJ Wijnhoven and PHN de With. Fast training of object detection using stochastic gradient descent. In *2010 20th International Conference on Pattern Recognition*, pages 424–427. IEEE, 2010.