

EXAM - Maths - PSB Msc DM

JEREMIE SAYAG

January 2021

Criteres : Qualite de la presentation, pedagogie, decouverte, utilisation, simplicité

1 Analysis of a PhD thesis

Soukaina EL GHALDY

<https://github.com/soukainaElGhaldy/PSB-X/blob/main/Mathematics/maths.pdf>

1.1 What's inside

L'auteur nous presente une analyse de la these de doctorat de Baptiste Barreau et en particulier son travail sur un algorithme de Machine Learning appele "Expert Network". Il s'agit de permettre une prediction des interets futurs de clients sur les marches financiers.

La problematique tres interessante a laquelle s'attaque Baptiste Barreau est: "À une date t , quel investisseur est intéressé par l'achat/vente de quel actif financier ?"

1.2 Focus

Concentrons nous sur la formule suivante :

$$V_{(t=n)} = P_e * \left(1 + \frac{I_r}{p}\right)^{np}$$

Cette equation permet de calculer la valeur d'une obligation a un certain nombre d'annees ecoules n , à périodicité p , à un prix d'émission P_e et un taux d'intérêt annuel I_r .

L'auteur s'est rendu compte qu'au fur et a mesure des annees la courbe des taux d'interet de l'obligation aura une forme exponentielle. Cette periodicite exponentielle est appele "Continuous Compounding".

1.3 V Criterias

- Qualité de la presentation: Bonne presentation, utilisation de schema et graphique.
- Pedagogie: L'auteur tente de nous faire comprendre un sujet complique pour les novices en finances. Peut etre, l'utilisation de plus d'exemples concret nous aurait aide.
- Decouverte: Aimant la finance, la decouverte totale de cette these m'a particulierement plu.
- Utilisation: Cela est interessant mais particulierement utile surtout pour les personnes travaillant dans le monde de la fincance.
- Simplicité: Dure a comprendre pour les novices en finances, il faut aller chercher des explications plus poussees pour mieux comprendre.

1.4 The End

C'est un sujet qui m'interesse particulierement,et avec l'essor des fintech la data et la finance font un bon mix.

2 Algorithme Genetique

COMLAN Florine HOUNTONDI Ramya

https://github.com/fcom-stack/PSBX/blob/main/Algorithme_genetique/Algo_genetique.pdf

2.1 What's inside

Une introduction et explication d'un type d'algorithme particulier que sont les algorithmes genetiques.

Ces algorithmes ont la particularite d'avoir un mode de fonctionnement similaire au corps humain. Les auteurs nous detaillent dans un sujet tres complet le fonctionnement de ces algorithmes genetiques.

2.2 Focus

Expliquons une fonction qu'on appelle "Fonction de Fitness" :

0	1	2	3	4	5	6	Item Number
0	1	0	1	1	0	1	Chromosome
2	9	8	5	4	0	2	Profit Values
7	5	3	1	5	9	8	Weight Values

Knapsack capacity = 15
Total associated profit = 18
Last item not picked as it exceeds knapsack capacity

La fonction de fitness est une fonction qui permet de mesurer la performance d'une solution candidate par rapport au probleme propose.

C'est une fonction de performance qui prend en valeur d'entree une solution candidate et dont le calcul de la valeur de performance est effectue de maniere repete dans un algorithme genetique et qui doit etre suffisamment rapide.

2.3 V Criterias

- Qualité de la presentation: Presentation tres complete qui permet de cerner les enjeux des algorithmes genetiques;
- Pedagogie: Utilisation de beaucoup de schemas d'explications qui nous permettent de comprendre;
- Decouverte: Je ne connaissais pas du tout l'existence de ce genre d'algorithme, tres interessant ! ;
- Utilisation: Il est necessaire d'avoir plusieurs outils a porte de mains pour utiliser le plus approprie en fonction de notre problematique;
- Simplicité: Essayez de se rapprocher du fonctionnement de cette formidable machine qu'est le corps humain est difficile mais l'auteur tente de nous faire comprendre ce type d'algorithme de maniere tres didactique;

2.4 The End

Presentation tres complete des algorithmes genetiques et leur fonctionnement.

3 EPARS Early Prediction of At-risk Students with Online and Offline Learning Behaviors

AUFRERE Thuy, MAZZUCATO Claire, REN Claude

<https://github.com/clairemazzucato/PSBX>

3.1 What's inside

Les auteurs nous proposent une reflexion interessante sur un papier de recherche qui porte sur la prédiction du comportement d'apprentissage des étudiants à risque (STAR) afin d'intervenir à temps en cas d'abandon scolaire.

2 observations ont ete faites :

1- Les étudiants à risque ne disposent pas d'une routine d'étude régulière et claire.

2- L'environnement des etudiants a risque sont en general a risque eux memes

3.2 Focus

Considerons l'equation suivante :

$$\max_f \sum_{u \in V} \log \left(\prod_{v_i \in N_s(u)} \frac{\exp(f(u) \cdot f(v_i))}{\sum_{v \in V} \exp(f(u) \cdot f(v))} \right)$$

Le fait d'avoir parcouru le voisinage des noeuds permettent donner un ensemble qui est appelé $N_s(u)$. Les fonctions $f(.)$ sont des fonctions de représentation des noeuds en question.

3.3 V Criterias

- Qualité de la presentation: Presentation succinte
- Pedagogie: Pas assez d'exemple d'utilisation
- Decouverte : Totalement nouveau pour moi
- Utilisation: utilisation par des personnes specifiques (enseignants, chercheurs, etc.)
- Simplicité: Complique a utiliser, notions mathematiques complexes.

3.4 The End

Malgre la difficulte de comprehension que j'ai pu avoir j'ai adore decouvrir ce papier de recherche tres interessant.

4 4

ABBES Ahmed, BEN YOUSSEF Salah, BENSALEM Akram

<https://github.com/Ahmed-Abbes5/psbx>

4.1 What's inside

Une introduction aux algorithmes d'apprentissage automatique qui est le fondement du machine learning et de l'intelligence artificielle.

Ces algorithmes se fondent sur des approches mathématiques et statistiques pour donner aux ordinateurs la capacité d'apprendre à partir de données, c'est-à-dire d'améliorer leurs performances à résoudre des tâches sans être explicitement programmés pour chacune.

4.2 Focus

Focalisons nous sur les algorithmes de régression :

La régression concerne la modélisation de la relation entre les variables qui est affinée de manière itérative à l'aide d'une mesure d'erreur dans les prédictions faites par le modèle.

Les méthodes de régression sont une bête de somme des statistiques et ont été cooptées dans l'apprentissage automatique statistique. Cela peut prêter à confusion car nous pouvons utiliser la régression pour désigner la classe du problème et la classe de l'algorithme.

4.3 V Criterias

- Qualité de la presentation: Bonne presentation mais manque d'exemple
- Pedagogie: Le manque d'exemple nous fait défaut pour parfaitement comprendre les concepts expliqués ici.
- Decouverte: Nous connaissions les principes de ces algorithmes mais le travail des auteurs nous a permis de mieux comprendre ces concepts.
- Utilisation: Ces algorithmes sont à la base du machine learning et de l'intelligence artificielle et donc nécessaires.
- Simplicité: Plutôt simple de prime abord mais plus compliqué lorsque l'on doit les utiliser.

4.4 The End

Nous remercions les auteurs de nous avoir permis de nous replonger dans les algorithmes de base du machine learning et de l'I.A qui nous a permis de tout réviser.

5 Naïve Bayes Classifier

Jordy HOUNSINOU

https://github.com/Jordyhsn/PSB_Hounsinou

5.1 What's inside

On profite tous même sans s'en rendre compte d'un algorithme d'apprentissage automatique qui se base sur la classification de "Naïve Bayes Classifier" et sur le théorème de Bayes.

En effet, il s'agit du fameux filtre anti-spam présent aujourd'hui dans tous nos comptes emails.

5.2 Focus

Regardons la fameuse loi de Bayes:

$$P(A | B) = \frac{P(B | A) * P(A)}{P(B)}$$

Il s'agit d'une loi statistique, qui permet de connaître la probabilité que l'événement A se réalise en fonction de l'événement B.

Le théorème de Bayes vise à calculer les probabilités a posteriori d'un événement en fonction des probabilités a priori de cet événement. A priori et a posteriori s'entendent par rapport à la connaissance d'une information. L'exemple typique est celui du diagnostic : a priori on juge que le patient a une telle probabilité d'avoir la maladie M.

Cette loi permet de répondre à cette question:

Que devient a posteriori cette probabilité lorsque l'on apprend le résultat de tel examen clinique ou biologique ?

5.3 V Criterias

- Qualité de la présentation: Présentation très complète
- Pédagogie: Utilisation d'exemple concret par l'auteur, très didactique
- Découverte: Nous connaissions ce théorème de Bayes depuis le lycée, il est intéressant de le revoir.
- Utilisation: Très utilisé en statistique
- Simplicité: Assez simple à utiliser

5.4 The End

L'auteur nous permet de nous replonger dans nos souvenirs de lycée en nous présentant la fameuse loi de Bayes.

6 Les reseaux de neurones antagonistes generatifs

Jeremie SAYAG

<https://github.com/Jeremiesayag/PSBX-R>

6.1 What's inside

Une introduction aux GAN "Generative Adversarial Nets" ou reseaux de neurones antagonistes generatifs en français, qui permet entre autre la realisation de ce site :thispersondoesnotexist.com.

Ce reseau de neurones permet de de construire des modeles d'intelligence artificielle capable de generer un contenu artificiel realiste.

6.2 Focus

Le principe sous jacent de ces neurones est le suivant :

Considerons 2 entites : Un generateur et un discriminateur.

Le generateur a comme role de creer un contenu realiste

Le discriminateur a comme role de verifier si un contenu est reel ou artificiel.

Les GAN ont comme objectifs de creer un generateur tellement efficace qu'il cree une image fausse en faisant penser au discriminateur que l'image est vraie.

6.3 V Criterias

- Qualité de la presentation: Presentation plutot complete mais a completer avec des exemples concret.
- Pedagogie: Manque de visuel qui pourrait permettre une meilleure pedagogie
- Decouverte: Nous avons adore le concept du site internet, il a donc ete interessant de comprendre comment cela etait possible.
- Utilisation: Utilisation de plus en plus courante dans le futur avec l'essor des deep fake.
- Simplicité: Concept mathematiques plutot complique au debut mais apres quelques heures de recherche cela devient un peu plus simple.

6.4 The End

Nous avons essaye de trouver un sujet interessant, et de l'expliquer de la maniere la plus simple possible. Les GAN sont un sujet qui nous interesse particuliere-ment a la lumiere de toutes les fake news qui sont autour de nous dans le context actuel.

7 The Real End

Après de longues heures de recherches mes camarades et moi même avons essayé de présenter un sujet qui nous intéresse et intéresse les autres. Certains de mes camarades ont utilisé beaucoup de schéma et d'image ainsi que d'exemple concret qui je le pense est un réel avantage pour une bonne compréhension.

En effet, un sujet théorique aussi facile qu'il soit, pour bien le comprendre rien de mieux qu'une mise en pratique. Cela est encore plus vrai avec ces sujets difficiles.

Concernant, le choix des sujets mon coup de cœur est pour le sujet sur l'Expert Network car cela concerne la finance, qui m'intéresse.

Nous aurions pu avoir une meilleure pédagogie d'ensemble en créant des diaporamas ou vidéos explicatives de type tuto nous permettant de capter mieux notre auditoire et permettre une meilleure compréhension.