Jérémie Béliveau-Lefebvre – 04 494 470

Sébastien Leblanc – 18 206 273

TP4 - solution

1. Le gradient de perte par rapport à  est donné en dérivant la formule suivante   
   par rapport à W :

Sachant que et

Nous pouvons trouver par la règle de dérivée en chaîne. Ainsi nous pouvons décortiquer le problème tel qu’illustré ici :

Et si nous simplifions, nous pouvons effectuer le travail suivant :

Où :

Et B =

Ainsi

<=>

par définition

<=>

par dérivée

=

<=>

par définition

<=>

selon

<=>

par associativité

<=>

arythmétique

<=>

arythmétique

<=>

arythmétique

<=>

par définition

=

<=>

par dérivée

Il ne reste qu’à remettre les solutions trouvés dans la définition de dérivée en chaine plus haut :

<=>

par définition

<=>

arythmétique

<=>

arythmétique

<=>

par définition

<=>

par distributivité

<=>

par définition d’un « one-hot »

CQFD

* 1. Une distribution de vraisemblance est la distribution de probabilité conditionnelle d’observer un paramètre dans une classe donnée, ex : soit la probabilité d’observer x étant donné t. Pour la calculer, nous pourrions supposer que chaque classe de véhicule suit une distribution gaussienne et ainsi calculer une gaussienne par classe. Ainsi, dans chaque classe t, nous aurons une formule gaussienne qui nous permet d’en connaitre la répartition x.
  2. La distribution à priori est la distribution de probabilité d’observer une classe quelconque, ex : Pour chaque catégorie, nous pourrions calculer le nombre de véhicule divisé par le nombre total de véhicule.
  3. Oui, puisque les voitures sport sont de même catégorie et ont plus de chance de consommer tous beaucoup d’essence et peu de chance qu’une de ces voitures consomme très peu d’essence.

1. Si nous partons des principes de dynamique qui s’appliquent sur un corps en mouvement, nous avons les paramètres suivants : la masse d’un objet, la gravité qui agit sur celui-ci, la pression du sol sur ce corps, la pente du sol ainsi que la friction entre la masse et le sol. . Les formules suivantes décrivent le mouvement de la masse (en 1D) en fonction du temps :

où

et où a : accélération

v : vitesse

d : position

: variation de pente de vitesse

Les formules de descente de gradient de type momentum se définissent comme suit :

Où w : un certain poid

v : vitesse

 : constante de contrôle de vitesse initiale

constante de contrôle de variation de vitesse

t : temps donné

: gradient

Pour mettre en valeur les liens entre la descente de gradient de type momentum et la position, vitesse et accélération, il suffit de prendre les équations de vitesse et d’accélération, d’isoler et et d’étudier les similarités. Nous obtiendrons les équations suivantes :

Par rapport aux formules du momentum  suivantes :

* Les tout comme peuvent d’abord être interprétés comme la position de la masse sur une pente en un temps .
* peut être vue comme étant la variation de l’itération à laquelle on regarde la vitesse.
* peut être vu comme la friction qui ralenti la masse
* et l’accélération (a) peuvent être vu comme la force qui fait varier la position de la masse dans la pente (la gravité).
* La différences des signes + et – s’explique par la direction d’application des variations de vitesses et de gradient sur la masse. Dans le gradient, + signifie que le poid remonte alors que dans la vitesse, un + signifie que la masse accélère vers le bas et prend plus de distance.

C’était les similitudes entre la descente de gradient de type momentum et la position, vitesse et accélération.