**Projet MA S8 – Brouillon**

*Quentin Douzery*

**02 février**

Modélisation numérique, directe et inverse, des écoulements de fleuves & rivières dont la surface est mesurée par satellite. L'objectif est d’estimer le débit des rivières (m3/s) à l'échelle planétaire à partir des seules observations altimétriques (hauteur et largeur de la surface d'eau).

Bases de données :

* HYDRoSWOT
* PEPSI

Variables d’intérêt pour l’explication de la variabilité du débit des rivières :

* L’aire drainée
* Les aires ajoutées
* La largeur
* Le taux de limon
* Sinuosity & Meandwave

Réseaux de neurones :

* Vecteur de poids associé à chaque neurone
* Couche : ensemble de neurones n’ayant pas de connexions entre eux
* Perceptron : un neurone d’une couche cachée est connecté en entrée à chacun des neurones de la couche précédente et en sortie à chaque neurone de la couche suivante.
* Estimation des poids par apprentissage (utilisation d’algorithmes d’optimisation)
* Choix des paramètres : nombre de couches cachées, nombre de neurones par couche cachée, nombre maximum d’itérations, erreur maximale tolérée, terme de régularisation, taux d’apprentissage, taille des ensembles d’observations
* Difficile de se faire une idée de la qualité du modèle
* Fixer le nombre de neurones et d’itérations assez grand, et jouer sur le réglage de la pénalisation
* Librairie e1071

Deep learning :

* ConvNet > analyse d’image
* LSTM > dimension temporelle à prendre en compte
* Diabolo > détection d’anomalies
* Empilement de couches de neurones

Questions

* Qu’est-ce que la *Meandwave* ?
* Qu’est-ce que sont les *sections mouillées A (x, t)* ? Profondeur, oui mais laquelle ?

Problème initial

Bathymétrie, débit, frictions

* Débit, section mouillée (A0 inconnue)

Problème inverse

Modèles physiques (équations de mouvement et de la masse) , largeur d’eau, aire drainée, hauteur d’eau (pas A0)

* Débit, section mouillée

Réseaux de neurones

* Architecture donne le modèle
* Sortie comparée avec la base de données GRDC (minimisation en modifiant les poids des neurones)
* LSTM permet la corrélation temporelle

Deux bases de données : 1 est grande mais avec des données « fausses », 1 est petite mais avec des « vraies » données.

*A faire*

*~~Cours LSTM (+ lecture autour ANN/RNN/LSTM avec Aggarwal)~~*

*~~Analyser base de données point de vue stats (notebooks Kevin)~~*

*Jeudi 11 février, 16h*

**11 février**

Database la même que celle de Manon ?

Donnée ‘U’ ?

* vitesse moyenne de l’écoulement (PEPSI :1km longitudinal, sur toute la section de la rivière ; HydroSWOT : mesures réelles avec ADCP, très localisé). Max vitesse à 20% de la profondeur

Aire drainée ? Pourquoi ne varie-t-elle pas ?

* Totalité de l’aire du bassin versant

A quoi correspondent les différentes mesures pour une même station ? Différentes heures d’une journée ? Différents jours ?

* Pepsi : tous les jours
* Pas de mesure particulière, des jours quelconques

Q5, Q10, etc. ?

* Quantiles du débit (estimations, sortie d’un modèle de Machine Learning)

Saison ? Lune ? Température extérieure ?

CNN

Image

* Couches de convolution (avec des Kernel, matrice 3x3 par exemple avec des 0 et des 1, qui effectuent la convolution ; Kernel auto apprennent))
* Pooling (max, average, etc.)
* Couches de convolution (nombre augmente)
* Pooling (réduit au fur et à mesure l’image)
* Etc.
* Flat layers (récupère taille originale de l’image)
* Réseau de neurones (pour la classification)
* Possibilité de rajouter des couches de dropout (mise à 0 de certains pixels, pourcentage fixé dans les paramètres)

Mercredi 17 février 9h00 – Skype

**17 février**

Mse : pourcentage d’erreur

Mae : on le relie à la mesure physique du débit

Utiliser flowacc

Enlever U

Rajouter Pente (dans Pepsi)

S : pente de surface libre (Pepsi)

Pente des berges dans HydroSwot

Rajouter histogramme des sorties (pas de normalisation)

En entrée du réseau de neurones : sélection des variables d’entrée (pas toutes)

Par exemple, rajouter Pa (pluie)

*Revenir sur le notebook ANN*

*Stats descriptives pour trouver les variables d’entrée intéressantes*

**23 février – 9h30**