

Liquid Neural Network

Prise de notes pour l'article LNN

Les réseaux de neurones récurrents avec des états cachés en temps continu déterminés par des équations différentielles ordinaires (EDO) sont des algorithmes efficaces pour modéliser les données temporelles, souvent utilisés dans les contextes médicaux, industriels et commerciaux. L'état d'un neural ODE (ODE neuronale), noté $x(t) \in \mathbb{R}^D$, est défini par la solution de l'équation suivante (Chen et al. 2018) :

$$\frac{dx(t)}{dt} = f(x(t), I(t), t, \theta)$$

où f est un réseau de neurones paramètre par θ . On peut calculer l'état de ce réseau en utilisant un solveur numérique d'EDO, puis entraîner le réseau en appliquant la différentiation automatique en mode inverse (Rumelhart, Hinton, et Williams 1986). Cela peut se faire soit par **descente de gradient à travers le solveur** (Lechner et al. 2019), soit en considérant **le solveur comme une boîte noire** (Chen et al. 2018 ; Dupont, Doucet, et Teh 2019 ; Gholami, Keutzer, et Biro 2019) et en appliquant **la méthode de l'adjoint** (Pontryagin 2018).

Les questions ouvertes sont les suivantes : dans leur formalisation actuelle, quel est le niveau de capacité des neural ODEs à représenter de manière expressive des données ? Peut-on améliorer leur structure pour permettre un apprentissage de représentations plus riche et augmenter leur expressivité ?