

Eric Moulines

Ecole Polytechnique, France



Modèles Génératifs Profonds : l'échantillonnage en haute dimension revisité

Résumé

Les modèles génératifs (GM) permettent d'inférer des modèles de loi pour des observations structurées de grande dimension, qui sont typiques de l'IA moderne. Les modèles génératifs peuvent également être utilisés pour échantillonner de nouveaux exemples, en reliant le problème d'inférence à l'échantillonnage.

L'apprentissage de modèles génératifs profonds (MGD) capables de capturer les structures de dépendance complexe de lois à partir de grands ensemble de données dans un cadre non- ou semi-supervisé apparaît aujourd'hui comme l'un des principaux défis de l'IA. Les modèles génératifs profonds ont de nombreuses applications passionnantes pour résoudre la pénurie de données en générant de " nouveaux " exemples, pour préserver la confidentialité en diffusant le modèle génératif à la place des données mais aussi pour détecter les observations aberrantes.

Dans cette présentation, je vais couvrir trois directions de recherche sur lesquelles je travaille actuellement.

Une première approche est basée sur la minimisation de l'entropie croisée (divergence de Kullback-Leibler) entre la distribution des observations et un modèle paramétré soit par des réseaux de neurones profonds, soit par des fonctions d'énergies plus adaptées, reliant les modèles génératifs et les « energy based models » qui ont été introduits pour l'apprentissage non-supervisé (mais dans un cadre non-probabiliste). Cette approche est séduisante mais elle pose des problèmes de calcul difficiles, liés à la nécessité d'estimer la constante de normalisation et son gradient.

Une deuxième approche repose sur les méthodes d'entropie maximale. Cette approche trouve son origine dans les quantités de physique statistique pour apprendre une distribution maximisant l'entropie sous contrainte de moment, qui sont construites à partir d'une représentation issue d'un réseau de neurones profonds

Une troisième approche consiste à utiliser des auto-encodeurs variationnels (Variational Autoencoder, VAE), un cas particulier d'inférence variationnelle. Les VAE apprennent conjointement un algorithme pour générer des échantillons à partir de la distribution ainsi qu'un espace latent qui résume la distribution des observations.

J'illustrerai ces approches par des exemples et je discuterai des challenges théoriques et numériques que ces approches posent.

Mini biographie

Éric Moulines entre à l'École Polytechnique en 1981, puis part étudier à Télécom ParisTech.

Il débute sa carrière au Centre national d'études des télécommunications où il travaille sur la synthèse de parole à partir du texte. Il participe au développement de nouvelles méthodes de synthèse par formes d'ondes, appelées PSOLA (« pitch synchronous overlap and add »).

Après avoir soutenu sa thèse en 1990, il rejoint l'École Nationale Supérieure des Télécommunications comme maître de conférences. Il s'intéresse alors à différents problèmes de traitement statistique du signal. Il contribue notamment au développement des méthodes sous-espaces pour l'identification de systèmes linéaires multivariés et la séparation de sources. Il développe aussi de nouveaux algorithmes d'estimation adaptatives de systèmes.

Il reçoit l'habilitation à diriger les recherches en 2006 et devient professeur à Télécom Paris. Il se consacre alors principalement à l'application des méthodes bayésiennes avec des applications en traitement du signal et en statistique.

Il reçoit par ailleurs la médaille d'argent du CNRS en 2010, le prix France Télécom remis en collaboration avec l'académie des sciences en 2011. Il est nommé « fellow » de l'European Association for Signal Processing en 2012 et de l'Institut of Mathematical Statistics en 2012. Il est ingénieur général du corps des mines.

Éric Moulines a dirigé 21 thèses, a été président de jury pour 9 thèses, a été rapporteur pour 10 thèses, a été membre de jury pour 6 thèses.