

Titre de la thèse : Apprentissage de représentations du contexte. Réseaux de neurones non supervisés pour les flux de données hétérogènes

Encadrant : Ludovic Denoyer – LIP6

ED de rattachement : EDITE

Liens avec les axes du LABEX SMART :

Le projet de thèse qui vise à construire des modèles d'apprentissage de représentations pour les flux de données hétérogènes s'insère dans le Labex SMART principalement dans l'axe **Humans at the convergence of digital and real environments, 1) Large scale complex data processing**. En effet, l'objectif du travail proposé est le développement de méthodes permettant le traitement de flux multi-modaux de données. Ces flux peuvent provenir par exemple de capteurs (sur un robot, dans un véhicule, ...) ou de réseaux de capteurs, mais aussi de traces d'interactions produites par un utilisateur (email, calendrier, activité sur le Web, etc...) Cette thèse implique donc la possibilité de traiter **de grandes quantités de données hétérogènes**, en temps réels si possible. Du point de vue des applications, le sujet a été conçu avec l'idée de développer des algorithmes pour le *context awareness*, c'est-à-dire capable de conserver, depuis un flux temporel, toutes les informations pertinentes permettant de caractériser le contexte à l'instant t et de prendre des décisions basées sur ce contexte. Les applications visées sont donc d'une part des applications pour lesquelles la caractérisation du contexte est essentielle pour la prise de décision (robotique, véhicule intelligent, ...), mais aussi des applications d'interactions homme/machine pour lesquelles les modèles permettront de caractériser l'interaction en cours. Le sujet rentre donc en écho avec l'axe « **Interfaces and Interaction with humans** » et notamment le sous-axe « **Multimodal interactions** »

Projet

Note activité quotidienne laisse plus en plus de traces numériques. Ces traces sont de différentes natures, et de différente dynamique : données de profil, de calendrier, d'interactions (emails), de déplacements, capteurs, etc.. Ensemble, elles constituent une information très riche permettant de caractériser, à l'instant t , le contexte dans lequel un utilisateur se trouve. La compréhension et la caractérisation de ce contexte est un point clef pour le développement de nouveaux services aux utilisateurs mais aussi un problème complexe tant du point de vue formel que algorithmique. Il implique en effet le traitement de grandes masses de données hétérogènes et hautement dynamiques pour lesquelles les méthodes existantes ne sont pas adaptées. L'objectif de cette thèse est donc de développer de nouveaux modèles de caractérisation et d'exploitation de ces données "contextuelles" centrées sur les utilisateurs. L'apprentissage de représentations [Goodfellow et al. 2016] est un domaine émergent de l'apprentissage statistique. Le domaine est au cœur de plusieurs voies de recherche en apprentissage et est porté par le récent succès des réseaux de neurones profonds (Deep learning). La création récente d'une conférence entièrement dédiée à cette thématique et de laboratoires privés de recherche (Facebook AI, Google Deepmind, etc...) témoigne de son importance. Cette communauté est aujourd'hui très active et vise particulièrement le développement de modèles automatiques d'apprentissage permettant le traitement de grandes quantités de données complexes (images, texte, etc...), et plus particulièrement d'en découvrir quels en sont les facteurs latents (ou représentations). Cependant, les travaux existants sont principalement focalisés sur des problèmes simples, soit en termes de type de données d'entrée, soit en termes de tâches à résoudre (classification, régression). Les modèles du domaine sont limités et incapables aujourd'hui de traiter les données hétérogènes et dynamiques qui nous intéressent dans ce travail de thèse.

Enjeux

Organisation du travail

Dans cette thèse, nous proposons de nous intéresser simultanément à deux aspects critiques des systèmes d'apprentissage de représentations mentionnés précédemment, à savoir, d'une part le traitement de données temporelles, c'est-à-dire l'apprentissage d'une représentation basée sur des flux continus d'informations, et d'autre part l'intégration dans ces flux temporels d'une grande diversité de types d'information. Le projet de recherche sera donc organisé selon ces deux tâches.

Apprentissage de représentations pour les flux temporels

Si les travaux récents dans le domaine du deep learning se focalisent de plus en plus sur le traitement de données séquentielles [Sutskever et al. 2014] – que ce soit pour des problématiques de type reconnaissance de la parole, traduction automatique, ou bien traitement de vidéos – les solutions existantes – à base de réseaux de neurones récurrents – sont loin d'être satisfaisantes. Particulièrement, elles n'ont pas été développées pour le traitement de flux continus d'informations. Nous visons ici à créer de nouveaux modèles permettant l'intégration « en continu » de nouvelles données issues de capteurs. L'idée principale de ce travail sera basée sur le développement de nouveaux modèles non-supervisés. Nous nous appuyerons particulièrement sur la notion de « forward models » qui visent à

modéliser l'évolution dans le futur des flux temporels étudiés [Stadi et al. 2015, Wahlstrom et al. 2014, Contardo et al. 2014] . Ces approches qui ont été particulièrement étudiées il y a plusieurs années dans le cadre de l'apprentissage par renforcement (e.g « predictive states » [Littman et al. 2002]) suscitent aujourd'hui un regain d'intérêt de la part de la communauté machine learning. Une piste particulièrement prometteuse est le développement de modèles génératifs continus (de type variational auto-encoders [Watter et al 2015]) sur les données dynamiques, ainsi que l'utilisation de représentations gaussiennes – qui permettront de modéliser une distribution sur le futur— au lieu des représentations vectorielles classiques.

Flux temporels hétérogènes

La seconde étape du travail de thèse visera à étendre les solutions proposées pour pouvoir traiter des flux d'informations hétérogènes, c'est-à-dire constitués de données de différentes natures (signal, données sémantiques) et disponibles avec des temporalités différentes. Ce travail consistera à étendre les algorithmes d'apprentissage de représentations non supervisés temporels pour leur permettre le traitement de données multi-modales en y intégrant des modèles d'estimation de densités sur des données structurées. Pour cela, le candidat étudiera particulièrement des approches de prédiction structurées séquentielles qu'il intégrera dans les architectures temporelles précédentes, dans l'inspiration de ce qui a été fait dans l'équipe sur des données statiques [Maes et al. 2009].

Références

[Contardo2014] Gabriella Contardo, Ludovic Denoyer, Thierry Artières, Patrick Gallinari : Learning States Representations in POMDP. ICLR 2014

[Goodfellow et al. 2016] Ian Goodfellow, Yoshua Bengio, and Aaron Courville : Deep Learning, Book in preparation for MIT Press

[Littman et al. 2002] Littman, Michael L. ; Richard S. Sutton ; Satinder Singh (2002). "Predictive Representations of State"). Advances in Neural Information Processing Systems 14 (NIPS). pp. 1555–1561

[Maes et al. 2009] Francis Maes, Ludovic Denoyer, Patrick Gallinari : Structured prediction with reinforcement learning. Machine Learning 77(2-3) : 271-301 (2009)

[Stadie et al. 2015] Bradly C. Stadie, Sergey Levine, Pieter Abbeel : Incentivizing Exploration In Reinforcement Learning With Deep Predictive Models. CoRR abs/1507.00814 (2015)

[Sutskever et al. 2014] Ilya Sutskever, Oriol Vinyals, Quoc V. Le : Sequence to Sequence Learning with Neural Networks. NIPS 2014 : 3104-3112

[Wahlstrom et al. 2014] Niklas Wahlström, Thomas B. Schön, Marc Peter Deisenroth : Learning deep dynamical models from image pixels. CoRR abs/1410.7550 (2014)

[Watter et al 2015] Manuel Watter, Jost Tobias Springenberg, Joschka Boedecker, Martin A. Riedmiller : Embed to Control : A Locally Linear Latent Dynamics Model for Control from Raw Images. NIPS 2015 : 2746-2754

Ouverture à l'international

Le projet de thèse amènera des collaborations internationales, tant avec des académiques que avec des partenaires industriels (Ecole Centrale Marseille, INRIA, Facebook, ...)