

Génération de séries temporelles à partir de facteurs contextuels démêlés

Perrine Cribier-Delande^{1,2}, Raphael Puget², Vincent Guigue¹, Ludovic Denoyer¹ *¹

¹ MLIA, Sorbonne Université, CNRS, LIP6, F-75005 Paris, France

²Renault, DEA-IR, Technocentre, 1 avenue du Golf 78084 Guyancourt , France

La question de la modélisation des séries temporelles doit être abordée en tenant compte du coût de l’acquisition de ces données. C’est pourquoi les datasets couvrent rarement toutes les situations observées dans le monde réel. Nous ne disposons en l’occurrence de mesures précises que sur une courte période ou à un lieu précis. Le domaine de l’apprentissage de représentations démêlées, où chaque élément appris correspond à un concept sémantique réel, a fait l’objet d’une attention croissante ces dernières années. Ces architectures offrent de nouvelles possibilités pour améliorer la modélisation des séries grâce à l’introduction d’espaces latents démêlés.

La modélisation des séries temporelles peut-être abordée à travers différentes approches. La plus classique est la question de la prédiction. Longtemps abordée via des modèles statistiques, les méthodes d’apprentissage profond ont représenté une avancée majeure dans le domaine. L’utilisation des réseaux récurrents a offert de nouvelles possibilités.

Nous avons choisi d’aborder la modélisation des séries temporelles en transformant le problème de prédiction en problème de génération : compte tenu de deux facteurs contextuels étiquetés dans le dataset, nous sommes en mesure de générer les séries correspondant à n’importe quelle combinaison de facteurs.

Notre proposition repose sur une architecture d’encodeur-décodeur. L’encodeur agit comme extracteur de facteurs : il permet de représenter les facteurs contextuels associés aux séries temporelles. Le décodeur est le module de génération : à partir des représentations des deux facteurs, il génère une série temporelle.

Au stade de l’inférence, certaines questions se posent sur la manière d’agrèger les représentations des facteurs latents. Nous avons étudié et comparé plusieurs

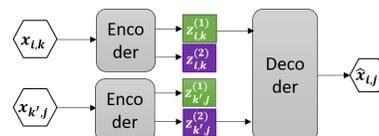


FIGURE 1 – Notre architecture encodeur-décodeur stratégies, dont une reposant sur un mécanisme d’attention.

Nous exploitons notre architecture dans deux tâches différentes : (i) la génération de séries manquantes à partir des valeurs des facteurs contextuels déjà présentes dans l’ensemble d’entraînement et (ii) la génération de séries temporelles correspondant à des valeurs de facteurs jamais observées dans l’ensemble d’entraînement et une fois seulement au moment de l’inférence.

Nous testons notre architecture sur quatre datasets différents (dont trois disponibles sur Kaggle), chacun étant composé de séries de mesures prises en différents lieux et à chaque heures (séries de taille 24). Les deux facteurs contextuels à démêler sont donc le temps et le lieu. Ces datasets mesurent des grandeurs de nature extrêmement différente, le premier mesure l’affluence dans des stations de métro, le deuxième des consommations d’énergie et les deux derniers des concentrations de polluants dans l’air.

Nos résultats montrent que notre architecture est plus performante que les baselines sur nos deux tâches. L’introduction de bruit dans l’ensemble d’entraînement (*via* la modification d’une fraction des étiquettes des facteurs) nous permet de démontrer que notre mécanisme d’attention minimise l’impact des valeurs aberrantes et préserve une performance quasi optimale. Quelques résultats qualitatifs permettent de mettre en valeur la structure acquise par les espaces latents et l’importance du mécanisme d’attention dans cette structuration.

Mots-clé : Apprentissage de représentations ; disentanglement

*This work is partially supported by the European Union’s Horizon 2020 Research and Innovation Programme under grant agreement No 780754, ”Track & Know”.