Application BASTRI
Fiches Equipes

CELESTE (SR0873YR)

statistique mathématique et apprentissage
SELECT (SR0866IR) CELESTE

Statut: Décision signée

Responsable : Sylvain Arlot

Mots-clés de "A - Thèmes de recherche en Sciences du numérique - 2024" : A3.1.1. Modélisation, représentation , A3.1.8. Données massives (production, stockage, acheminement) , A3.3. Analyse de données et de connaissances , A3.3.3. Analyse de données massives , A3.4. Apprentissage et statistiques , A3.4.1. Apprentissage supervisé , A3.4.2. Apprentissage non supervisé , A3.4.3. Apprentissage par renforcement , A3.4.4. Optimisation pour l'apprentissage , A3.4.5. Méthodes bayésiennes , A3.4.6. Réseaux de neurones , A3.4.7. Méthodes à noyaux , A3.4.8. Apprentissage profond , A3.5.1. Analyse de grands graphes , A6.1. Outils mathématiques pour la modélisation , A9.2. Apprentissage

Mots-clés de "B - Autres sciences et domaines d'application - 2024" : B1.1.4. Génétique et génomique , B1.1.7. Biologie computationnelle , B2.2.4. Maladies infectieuses, Virologie , B2.3. Epidémiologie , B4. Energie , B4.4. Acheminement , B4.5. Consommation , B5.2.1. Véhicules routiers , B5.2.2. Chemin de fer , B5.5. Matériaux , B5.9. Maintenance , B7.1. Gestion des traffics , B7.1.1. Piétons, foules , B9.5.2. Mathématiques , B9.8. Recherche reproductible , B9.9. Ethique

Domaine : Mathématiques appliquées, calcul et simulation
Thème : Optimisation, apprentissage et méthodes statistiques

Période : 01/06/2019 -> 31/12/2027
Dates d'évaluation : 01/12/2022

Etablissement(s) de rattachement : CNRS, UNIV. PARIS-SACLAY
Laboratoire(s) partenaire(s) : LMO (UMR8628)

CRI : Centre Inria de Saclay
Localisation : UP Saclay - Laboratoire de Mathématiques d'Orsay (LMO)
Code structure Inria : 111089-0

Numéro RNSR : 201923222N
N° de structure Inria: SR0873YR

Présentation

La science des données — un vaste domaine comprenant la statistique, l'apprentissage automatique, le traitement du signal, la visualisation des données et les bases de données — fait désormais la une des journaux et a un impact potentiel majeur sur la société, au-delà du rôle important qu'elle a joué en sciences pendant de nombreuses décennies. Au sein de ce domaine, la communauté statistique a une longue expérience de comment déduire des connaissances à partir de données, sur des bases mathématiques solides. Le domaine plus récent de l'apprentissage automatique a également donné lieu à des réalisations majeures, en combinant statistique et optimisation, et en utilisant un point de vue nouveau provenant d'applications où la prédiction est plus importante que la construction de modèles.

Le positionnement de l'équipe-projet CELESTE est à l'interface entre statistique et apprentissage. Nous sommes statisticien(ne)s, membres d'un laboratoire de mathématiques, avec une solide formation en mathématiques, et sommes intéressé(e)s par les interactions entre théorie, algorithmes et applications. En effet, les applications conduisent aux problèmes théoriques les plus intéressants, tandis que la théorie peut jouer un rôle clé pour (i) comprendre comment et pourquoi les algorithmes d'apprentissage statistique efficaces fonctionnent — et donc les améliorer — et (ii) construire de nouveaux algorithmes sur des bases de statistique mathématique.

Nous étudions ainsi plusieurs des défis de l'apprentissage automatique sous l'angle de la statistique mathématique, en gardant toujours à l'esprit que les jeux de données modernes sont souvent de grande dimension et/ou de grand volume, ce qui doit être pris en compte dès la conception des algorithmes d'apprentissage. Il y a en particulier souvent un compromis à trouver entre

Contact

- **Responsable :** Sylvain Arlot
- **Tél :**
- **Secrétariat Tél :**

En savoir plus

- Site de l'équipe
- Site sur inria.fr
- Site du responsable
- Derniers Rapports d'Activité : 2019 , 2020 , 2021 , 2022 , 2023 , 2024

Documents sur la structure

- Intranet
- Privés

Décisions

- 13569 (23/04/2019) : création
- 15814 (29/11/2022) : prolongation
- 16728 (04/01/2024) : prolongation

Localisation

- **Adresse postale :** Laboratoire de Mathématiques d'Orsay Université Paris-Saclay Bâtiment 307 rue Michel Magat 91405 Orsay France
- **Coordonnées GPS :** 48.701137, 2.177226

précision statistique et complexité algorithmique, que nous souhaitons éclairer autant que possible.

Finalement, un ingrédient clé de notre programme de recherche est de confronter nos résultats théoriques et méthodologiques à la réalité de (nombreuses) applications. C'est pourquoi une grande partie de notre travail est consacrée à la modélisation de données industrielles ou médicales, en lien avec différents problèmes issus de nos collaborations de longue date avec plusieurs partenaires, ainsi que de diverses collaborations ponctuelles opportunistes.

Axes de recherche

Notre travail consiste à analyser, du point de vue de la statistique mathématique, les algorithmes d'apprentissage statistique couramment utilisés, et à en développer de nouveaux en nous fondant sur notre expertise statistique et mathématique. Nos principaux axes de recherches méthodologiques et théoriques correspondent à des problèmes relevant aussi bien de la statistique mathématique (par exemple la sélection de modèles ou les tests d'hypothèses, multiples ou non) que de l'apprentissage automatique (par exemple, l'apprentissage en ligne, les réseaux de neurones profonds, la classification non supervisée ou l'inférence de graphes); en particulier, la question de l'équité algorithmique est un thème de Celeste que nous souhaitons mettre en avant.

Un ingrédient clé de notre programme de recherche est de maintenir nos résultats théoriques et méthodologiques en lien étroit avec la réalité concrète de la science des données appliquée. Les membres de Celeste travaillent avec des entreprises industrielles et avec des institutions médicales ou en santé publique, afin de développer des méthodes et appliquer des résultats statistiques — que ce soit pour des problèmes de prédiction, d'agrégation, de détection d'anomalies ou de prévision de séries temporelles, par exemple —, en lien avec des besoins industriels ou sociétaux d'actualité.

Relations industrielles et internationales

Celeste collabore régulièrement avec le département R&D d'EDF, ainsi qu'avec de nombreuses autres entreprises (via des thèses CIFRE, par exemple).

Celeste a des collaborations académiques avec des chercheurs de nombreuses institutions à travers le monde, dont Italian Institute of Technology, University College London, EPFL, University of Potsdam, Cornell University, Brown University et University of Washington.