



📍 ENS-2S14

MEMBRE DU CENTRE BORELLI

Nicolas VAYATIS

ENSEIGNANT-CHERCHEUR

PU - ENS PARIS SACLAY

Statut : Enseignant-Chercheur

☎ 33 1 81 87 53 89

🌐 Site internet (<http://nvayatis.perso.math.cnrs.fr/>)

@ Courriel

Thématique de recherche

Spécialiste de science des données et de machine learning, Nicolas Vayatis anime un groupe de recherche qui développe des travaux en modélisation et algorithmie prédictive, science des réseaux, traitement du signal, en interface constante avec l'industrie et le domaine biomédical.

Applications

La question du passage à l'échelle des applications issues de la science des données et du machine learning - de la preuve de concept à l'industrialisation - est dans la pratique un problème ouvert pour beaucoup de domaines d'applications. Nicolas Vayatis en a fait son principal axe de réflexion, que ce soit dans le domaine industriel ou dans le champ biomédical.

Pourquoi le Centre Borelli ?

«Venant des mathématiques, les projets interdisciplinaires que je mène au Centre Borelli, avec des collègues issus d'autres disciplines, en interaction permanente avec des terrains d'application, m'ouvrent de nouvelles directions scientifiques par leurs questionnements et leurs contraintes.»

Recherche

› Thème 1 - Fondements mathématiques de l'apprentissage statistique.

Depuis la fin du XXe siècle, Nicolas Vayatis considère la question des fondements théoriques comme un élément essentiel de la recherche en apprentissage statistique dans le cadre conceptuel introduit par Vladimir Vapnik, enrichi par des outils probabilistes sophistiqués que sont les inégalités de déviation et de concentration pour des processus empiriques.

Principaux résultats obtenus : inégalités de Vapnik-Chervonenkis à taux exponentiel exact (grandes déviations), consistance du boosting régularisé et vitesses rapides sous conditions de marge à la Tsybakov, analyse théorique des algorithmes de descente miroir pour la minimisation de risques convexifiés pour les méthodes d'ensemble, consistance des algorithmes maximisant l'AUC comme U-statistique en scoring ou ranking biparti.

› Thème 2 - Développement de méthodes et d'algorithmes de machine learning

Dans la confrontation à des applications, il apparaît très souvent que les algorithmes d'apprentissage enseignés dans les livres requièrent des modifications pour prendre en compte les contraintes ou les critères d'évaluation des utilisateurs finaux des règles de décision produites par ces algorithmes. Ces modifications peuvent parfois amener à reconsidérer en profondeur le cadre théorique ou bien faire émerger des innovations de rupture sur le plan méthodologique.

Principales contributions méthodologiques et algorithmiques : algorithme itératif de type descente miroir pour la calibration des poids optimaux dans une méthode d'ensemble en classification, algorithmes à base d'arbres de décision maximisant la courbe ROC pour le scoring (TreeRank, Ranking Forests, version multi-catégories), tests d'homogénéité en grande dimension pour la comparaison d'échantillons, méthodes d'optimisation globale pour la calibration d'hyper-paramètres en machine learning ou pour les plans d'expérience en calcul scientifique (GP-UCB-PE, LIPO, ...).

> Thème 3 – Machine learning « in real life »

Ces dernières années, Nicolas Vayatis a été témoin de l'explosion incontrôlée de son domaine de recherche dans le monde scientifique et technologique. Les nombreuses interactions qu'il a pu avoir avec des «consommateurs» d'algorithmes, que ce soit dans les sciences (physique, santé publique, biologie...) ou dans les entreprises, lui ont révélé le décalage important qui existe entre l'état de l'art scientifique et la réalité technologique tirée par des attentes sociétales ou économiques. En essayant de comprendre l'origine de ce décalage, il lui est apparu qu'il y avait deux directions à développer pour le corriger : a) la prise en compte de la structure sous-jacente (temporelle, spatiale, relationnelle, modèle « physique ») dans les données d'apprentissage ; b) la prise en compte des usages reposant sur les technologies intégrant les règles de décision produites par les algorithmes.

Principales directions de recherche : machine learning sur des signaux temporels et signaux sur graphes (détection de ruptures, apprentissage de représentations, reconnaissance de formes), étude de processus de propagation et de contrôle sur les graphes et applications (applications en épidémiologie et recherche opérationnelle).

Enseignements

- > Cours M2 MVA : Introduction to Statistical Learning.
- > Cours M2 CHPS : Science des données et apprentissage.
- > Cours M1 Parcours IA : Fondements de l'intelligence artificielle et du machine learning.

Diffusion du savoir

- Conférences grand public en 2019 auprès des sociétaires MAIF (Perpignan, Vélizy).
- Création d'un média en ligne, le Data Analytics Post, dont Nicolas Vayatis assure la direction éditoriale en s'appuyant sur une équipe de journalistes scientifiques. L'objectif est de sensibiliser plus largement les professionnels et les étudiants sur les enjeux de la science des données, des algorithmes et de l'IA. Le lectorat régulier est d'environ 10 000 utilisateurs.
- Lancement d'une école d'été annuelle franco-allemande sur des thèmes autour de la science des données et du machine learning. Cette initiative réunit chercheurs académiques et industriels en partenariat avec la région Île-de-France et le Land de Bavière et avec le soutien de l'Université Franco-Allemande. Les dernières éditions se sont tenues à l'ENS Cachan (2018), l'université de Passau en Allemagne (2019) et visio (2020).
- Accompagnement de start-up (spin-offs, parrainage, advisory board) : DessIA, CrossingMinds, Scibids, Hyperlex, Geo4Cast, Valgorythm.
- Mission France IA 2017 auprès de la secrétaire d'État en charge du Numérique, responsable du groupe de travail sur le thème de la formation en IA.
- Point focal national auprès de la Direction générale des entreprises du ministère de l'Économie et des Finances pour la conférence sur l'IA dans le cadre des travaux du G7 en 2018 à Montréal - animateur et rapporteur d'un atelier sur le thème «IA et travail».
- Interventions auprès du Conseil national de l'industrie sur le thème «emploi et compétences», de l'Institut national de la consommation sur «IA et consommation», du Comité économique et social européen sur le thème «politiques publiques et IA».

Parcours

- Depuis 2020 : cofondateur et directeur du Centre Borelli (UMR 9010)
- 2012-2019 : directeur du CMLA (Centre de mathématiques et leurs applications, UMR CNRS 8536)
- Depuis 2007 : professeur des universités, ENS Paris-Saclay (ex-ENS Cachan)
- 2007 : professeur invité School of Mathematics, Georgia Institute of Technology, Atlanta
- 2006 : habilitation à diriger des recherches (HDR), université Paris 6
- 2002-2007 : maître de conférences LPMA et UFR de mathématiques, université Pierre et Marie Curie
- 2000-2002 : postdoctorat bourse Marie-Curie, Universitat Pompeu Fabra, Barcelone
- 2000 : doctorat de mathématiques appliquées à l'École Polytechnique ; 1996 : assistant scientifique Marine nationale
- 1995 : DEA de sciences cognitives à l'École Polytechnique
- 1995 : diplôme d'ingénieur de l'École centrale Paris.

2024

Journal articles

[Arm-CODA: A Data Set of Upper-limb Human Movement During Routine Examination](#)

Sylvain Combettes, Paul Boniol, Antoine Mazarguil, Danping Wang, Diego Vaquero-Ramos, Marion Chauveau, Laurent Oudre, Nicolas Vayatis, Pierre-Paul Vidal, Alexandra Roren, Marie-Martine Lefèvre-Colau

[hal-04473267](#)



[Heterogeneities of the perceptual-motor style during locomotion at height](#)

Danping Wang, Ioannis Bargiotas, Jiuwen Cao, Nicolas Vayatis, Laurent Oudre, Pierre-Paul Vidal

[hal-04467446](#)



Preprints, Working Papers, ...

[Stein Boltzmann Sampling: A Variational Approach for Global Optimization](#)

Gaëtan Serré, Argyris Kalogeratos, Nicolas Vayatis

[hal-04442217](#)



[Deep Anti-Regularized Ensembles provide reliable out-of-distribution uncertainty quantification](#)

Antoine de Mathelin, Francois Deheeger, Mathilde Mougeot, Nicolas Vayatis

[hal-04455937](#)

