

Sujet de thèse : Utilisation d'approches d'Intelligence Artificielle pour exploiter les enquêtes en population pour la surveillance du COVID-19

Directeur de thèse : Vincent Breton, Directeur de Recherches au CNRS

Laboratoires ou UR : Laboratoire de Physique de Clermont

Email et téléphone : vincent.breton@clermont.in2p3.fr

Co-encadrant(s) :

- **Professeur Laurent Gerbaud, Institut Pascal et Service de Santé Publique du CHU de Clermont-Ferrand**
- **Professeur Anne-Françoise Yao, Laboratoire de Mathématiques Blaise Pascal**

Résumé (10 lignes max.) : la pandémie de COVID-19 a mis en évidence en France le besoin de disposer de dispositifs de surveillance et d'alarmes ne reposant pas exclusivement sur un dépistage biologique du virus. Sur la base des données d'une étude conduite sur l'histoire du virus et de sa transmission dans le foyer épidémique du département du Haut-Rhin au printemps 2020, l'objectif de la thèse est d'explorer l'utilisation de méthodes d'intelligence artificielle pour un système de surveillance des maladies infectieuses émergentes couplant des enquêtes en population à des informations hospitalières. Celles-ci seront utiliser notamment pour identifier de nouveaux indicateurs de déclenchement d'alerte précoce, analyser l'émergence de l'épidémie et sa diffusion et proposer des simulations de systèmes de surveillance.

Profil du candidat :

De formation scientifique, le candidat doit avoir un solide bagage en statistiques, analyse de données et connaissance des méthodes d'intelligence artificielle. Curieux, il doit pouvoir s'adapter à une thématique de recherche très multidisciplinaire et s'intéresser aux aspects médicaux et mathématiques du sujet. Travailleur et motivé, il devra être capable de conduire ses travaux de façon autonome et les expliquer avec clarté.

Mots clés : COVID-19, enquête en population, Intelligence Artificielle, science citoyenne, agrégats

Description détaillée (1 page max) :

La crise du COVID-19 a mis en évidence l'impréparation de nos sociétés occidentales face à l'émergence d'une maladie infectieuse. Elle a montré aussi l'importance de la détection et de la prise en charge adaptée du signal précoce d'une Maladie Infectieuse Emergente. En situation

de maladies émergentes, les enquêtes en population peuvent jouer un rôle important pour déclencher une alerte sanitaire en complément des informations biologiques et hospitalières.

Une enquête a ainsi permis de mettre en évidence un décalage entre le démarrage de l'épidémie de COVID-19 dans le département du Haut-Rhin et le déclenchement de l'alerte sanitaire [1].

Lancée le 22 avril 2020, cette enquête a permis de collecter des informations auprès d'environ 1400 foyers représentant plus de 4000 personnes dont plus de 1500 se sont déclarées avoir présenté des symptômes compatibles avec une infection COVID-19. Une première analyse des données met en évidence un début d'épidémie pouvant être donnée plus d'un mois avant l'alerte sanitaire qui a été déclenchée.

Ces observations sont à mettre en regard des autres sources d'informations notamment hospitalières. La mise en relation de toutes ces observations, grâce notamment aux nouvelles technologies numériques et à l'Intelligence Artificielle, doit permettre d'accélérer le déclenchement rapide d'alerte [3].

L'objectif de la thèse est de s'appuyer sur les données collectées dans l'étude du foyer épidémique du Haut-Rhin pour explorer la faisabilité d'un système de surveillance des maladies infectieuses émergentes en couplant des enquêtes en population à des informations hospitalières.

Nous proposons d'explorer différentes pistes dans le cadre de cette thèse pluridisciplinaire :

- Analyser l'émergence de l'épidémie et sa diffusion
- Identifier de nouveaux indicateurs de déclenchement d'alerte précoce, notamment basés sur la détection d'agrégats selon des scanners spatio-temporels tels que proposés par Kulldorf [2]
- Améliorer l'extraction de connaissances des enquêtes par l'automatisation du processus de genèse de thésaurus (maladies, traitement, activités sociales,...) à partir de réponses libres
- Proposer des scénarios (ou simulation) de systèmes de surveillance à partir des trois points précédents

Références (1/2 page max) :

[1] Breton, Vincent and Guiguet-Auclair, Candy and Odoul, Joséphine and Peterschmitt, Jonathan and Ouchchane, Lemlih and Gerbaud, Laurent, Population Based Survey of the COVID-19 Outbreak in the Haut-Rhin Department from January to April 2020 (May 11, 2020). Available at SSRN: <https://ssrn.com/abstract=3601684>

[2] Kulldorff M., A spatial scan statistic. Communications in Statistics - Theory and Methods. 26 (6) (1997): 1481–1496. doi:10.1080/03610929708831995

[3] P. Le Turniera, C. Leport, P. Martine, C. Jadand, B. Hoenf, J.-F. Guegan, Multi-sectorial research is paramount for preventing and controlling emerging infectious diseases, , Revue d'Epidémiologie et de Santé Publique 68 (2020) 133–136

Comment candidater ?

Contactez le directeur de thèse