

Classification et identification de moisissures pathogènes

Thèmes : Science des données, programmation, sciences du vivant

Période : mars/avril 2018 à août/septembre 2018

Organismes :

- Laboratoire de Probabilités, Statistique et Modélisation (UPMC), Institut des Sciences du Calcul et des Données (UPMC)
- Hôpital la Pitié Salpêtrière, Assistance Publique-Hôpitaux de Paris et Faculté de médecine, Sorbonne Université

Contexte : Les infections provoquées par des moisissures (aspergilloses, fusarioses, mucormycoses, etc.) sont un problème de santé émergent, responsable de milliers de décès chaque année en France. Ces infections qui surviennent surtout chez des sujets fragilisés (patients greffés, ou sous chimiothérapie anticancéreuse par exemple) sont difficiles à diagnostiquer. En particulier l'identification des moisissures pathogènes nécessite une expertise spécifique dont seuls disposent quelques laboratoires hospitalo-universitaires très spécialisés. Une nouvelle technique diagnostique vient d'émerger, basée sur l'analyse des spectres de masse, une sorte de code-barres obtenu après avoir transféré de l'énergie à un échantillon au moyen d'un laser.

Pour faciliter le diagnostic des infections dues aux moisissures, l'équipe du laboratoire de parasitologie et de mycologie médicale de l'hôpital de la Pitié-Salpêtrière a développé un prototype d'une application en ligne permettant de comparer ce code-barres spectral à une base de données de plus de 22.000 spectres et ainsi, d'identifier l'échantillon testé. L'application donne des résultats très prometteurs, mais nécessite d'être revue pour en améliorer les performances et en faciliter l'utilisation. Il sera important, en particulier, de tester plusieurs algorithmes d'identification afin de sélectionner celui qui donne les meilleurs résultats en termes de rapidité et de précision. Dans un avenir proche, cette application est destinée à être appliquée à divers autres volets de la biologie, comme la parasitologie, l'entomologie, l'hygiène alimentaire, ouvrant ainsi la voie à une biologie connectée et participative.

Objectifs : La mission principale du stagiaire sera de mener à bien les tâches suivantes :

- reprogrammer toute la méthode d'analyse et d'identification de spectres en Python ;
- tester différents algorithmes d'identification et en évaluer les performances ;
- rédiger une notice détaillée décrivant chaque étape de la méthode d'analyse et d'identification de spectres ;
- proposer de nouvelles évolutions pour élargir le champ d'application de ce logiciel en ligne.

Etudiants concernés :

Master 2 en Ingénierie Mathématique, en Statistique ou en Informatique ; Ecoles d'ingénieurs

Compétences requises :

Programmation Python
Statistique :

- statistique descriptive
- régression

Compétences souhaitées :

Statistique :

- estimation non-paramétrique, lissage de courbes
- apprentissage (clustering)

Intérêt pour les sciences du vivant

Contacts :

- nathalie.akakpo@upmc.fr
- charlotte.dion@upmc.fr
- maud.thomas@upmc.fr
- Pr Renaud Piarroux, Service de parasitologie et de mycologie, Pavillon Laveran, Hôpital de la Pitié-Salpêtrière (01 42 16 01 00)

Localisation du stagiaire :

UPMC

Hôpital de la Pitié-Salpêtrière