

# Prédiction du taux de pollens dans l'air à partir de mesures photoélectriques

lundi 18 mars 2024 09:20 (20 minutes)

Il s'agit de prédire le taux de pollens dans l'air à partir de mesures photoélectriques obtenues par le capteur développé par l'entreprise LifyAir. Ce capteur prend 40 mesures photoélectriques toutes les 10 secondes. Cependant, à cause des différents facteurs de l'environnement extérieur (vents, pluies, sables, pollution, etc.) ces données contiennent beaucoup de bruits. Pour ce problème, nous bénéficions d'une vérité terrain qui est obtenue à partir d'un autre capteur (beaucoup plus coûteux et ne permettant pas l'obtention de données en temps réel, d'où l'enjeu de notre recherche) et qui donne la concentration de chaque type de pollens dans l'air toutes les deux heures. La première étape est donc de nettoyer les données du capteur de LifyAir. À ces fins, j'ai eu l'idée, d'après mes lectures (surtout [ZP17]), de développer un AutoEncoder régularisé. L'idée est de minimiser le terme suivant :

min

$\|L - DpEpLq\|_{\lambda_1} + \lambda_2 \|S\|_{\lambda_2} + \|FpLq - H\|$

où  $L$  et  $DpEpLq$  sont nos données photoélectriques,  $D$  et  $E$  sont les deux parties d'un

AutoEncoder,  $F$  est la fonction qui nous permet d'obtenir la concentration totale de particules dans l'air toutes les 2 heures à partir des données photoélectriques (c'est une méthode qui provient de la physique) et  $H$  est la concentration totale de pollens dans l'air toutes les 2 heures obtenue à partir de la vérité terrain. Dans l'idée  $L$  ne contiendrait que les données qui comptent véritablement des pollens.

## Thématiques

**Auteur principal:** M. HOUÉDRY, Pierre (IRMAR)

**Orateur:** M. HOUÉDRY, Pierre (IRMAR)

**Classification de Session:** Statistiques appliquées