

Analyse des Données sur l'occupation d'une Pièce en fonction des paramètres environnementaux

Introduction

Contexte

Cette étude vise à analyser les facteurs environnementaux influençant l'occupation d'une pièce. Dans cette étude nous avons sept variables mais, les variables qui feront l'objet dans notre étude sont : la température, l'humidité, la lumière, le CO2, et le ratio d'humidité.

Description des données

Les données sont constituées de plusieurs variables environnementales mesurées dans un environnement fermé. Temperature : Température en degrés Celsius. Humidity : Humidité relative en pourcentage. Light : Niveau de lumière en lux. CO2 : Concentration de dioxyde de carbone en parties par million (ppm). HumidityRatio : Ratio d'humidité dans l'air. La variable cible (Occupancy) est binaire, où :0 signifie que la pièce n'est pas occupée, 1 signifie que la pièce est occupée.

Objectifs

L'objectif est d'identifier les variables clés permettant de prédire si une pièce est occupée ou non et proposer un modèle simple pour prédire l'occupation.

Méthodes Utilisées

Dans le cadre notre étude nous avons utilisé deux méthodes pour répondre à nos objectifs: Régression LASSO et l'Arbre de décision. La méthode de Régression LASSO nous permettra d'identifier les variables les plus importantes et celle de l'Arbre de Décision nous aidera à créer un modèle interprétable.

Présentation des Données

La présentation des données se penche sur l'approche des statistiques descriptives. La base de données comprennent 20 560 observations et 7 variables, les résultats se trouvent dans le tableau ci-dessous.

Variable	Moyenne	Min	1er Quartile	Médiane	3e Quartile	Max
Température (°C)	20.91	19.00	20.20	20.70	21.52	24.41
Humidité (%)	27.66	16.75	24.50	27.29	31.29	39.50
Lumière (Lux)	130.8	0.0	0.0	0.0	301.0	1697.2
CO2 (ppm)	690.6	412.8	460.0	565.4	804.7	2076.5
Ratio d'Humidité	0.0042	0.0027	0.0037	0.0042	0.0048	0.0064
Occupation (0 = non occupé, 1 = occupé)	0.231	0	0	0	0	1

D'après les résultats ci-dessus seulement 23.1 % des observations indiquent une pièce occupée.

Méthodes d'Analyse

Modèle LASSO

Le modèle de régression logistique pénalisée (LASSO) a été utilisé pour identifier les variables influentes. Parmi les variables importantes, nous avons : la lumière comme principal prédicteur ; le CO2 que nous pouvons considérer comme facteur secondaire et la température comme impact marginal. Le modèle nous donne une précision globale de 90.19%. On considère qu'une pièce est occupée lorsque : La lumière dépasse un certain seuil et quand le niveau de CO2 est élevé.

Arbre de Décision

L'arbre de décision a été construit pour fournir des règles simples de prédiction. Comme règles principales : Si la lumière est inférieure à 388.17 nous déduisons que la pièce est non occupée. Sinon, si la température est inférieure à 21.21 ou le CO2 est inférieur à 771.17 nous pouvons considérer que la pièce est non occupée. Dans le cas contraire nous considérerons que la pièce est occupée.

Importance des variables :

Lumière (55%)

La lumière est la variable la plus influente car une pièce non occupée a généralement une faible intensité lumineuse. Lorsqu'une pièce est occupée, des activités humaines augmentent généralement le niveau d'éclairage (éclairage naturel ou artificiel).

Par conséquent, le seuil de 388 Lux identifié dans l'arbre de décision est un point critique pour déterminer si une pièce est occupée ou non.

Explication intuitive : Si la pièce est sombre (lumière faible), elle est très probablement vide. À l'inverse, un niveau de lumière élevé suggère une présence humaine.

Température (22 %)

Pourquoi c'est important ?

La température joue un rôle secondaire car les pièces occupées tendent à maintenir une température stable ou légèrement plus élevée due à la présence de personnes ou d'appareils en fonctionnement. Les variations de température peuvent donc être des indices, mais elles sont moins fiables que la lumière.

Exemple concret : Une pièce non occupée peut être plus froide, surtout si elle est dans un environnement non chauffé. Une légère hausse de la température peut indiquer une activité humaine.

CO2 (19 %)

Pourquoi c'est important ?

Le niveau de CO2 est directement lié à la respiration humaine. Lorsqu'une pièce est occupée, le CO2 tend à augmenter progressivement en raison de l'expiration des occupants. Si le niveau de CO2 est faible, cela suggère que la pièce est inoccupée ou bien ventilée.

Lien avec l'occupation : Faible CO2 : Indique une pièce vide ou bien ventilée. CO2 élevé : Confirme une présence prolongée dans la pièce.

Résultats et Discussion

Résultats Clés

La lumière est le facteur déterminant dans la prédiction de l'occupation. Le CO2 et la température améliorent légèrement la précision du modèle. Le modèle LASSO et l'arbre de décision montrent une précision supérieure à 90 %. Une faible lumière indique généralement une pièce non occupée. Un niveau élevé de CO2 est un signe d'occupation prolongée.

Conclusion et Recommandations

Conclusion

L'étude a permis d'identifier les variables clés influençant l'occupation d'une pièce. La lumière est le principal facteur, suivie du CO2 et de la température.

Recommandations

Surveillance automatisée : Installer des capteurs de lumière et de CO2 pour prédire l'occupation. Optimisation énergétique : Ajuster les systèmes d'éclairage et de ventilation en fonction des prédictions.