**Jeu de données Agri-food CO₂ emission dataset – Forecasting ML**

**Utilisation du Machine Learning pour la prévision des émissions agricoles de CO₂** – IPCC / FAO

**À propos du jeu de données**

Ce jeu de données sur les émissions agricoles de CO₂ a été construit en fusionnant et retraitant environ une douzaine de jeux de données distincts provenant de la **Food and Agriculture Organization (FAO)** et de données issues de l’**IPCC**.  
Ces données ont été nettoyées, prétraitées et fusionnées pour créer un ensemble cohérent et complet destiné à l’analyse et à la prévision.

Le jeu de données, tel que présenté dans le notebook, décrit les émissions de CO₂ liées à l’agroalimentaire, représentant environ **62 % des émissions mondiales annuelles**.  
Les émissions du secteur agroalimentaire sont un élément majeur dans l’étude du changement climatique, et leur compréhension est essentielle pour réduire l’impact environnemental et développer des pratiques durables.

Dans le notebook associé, une analyse des relations entre les émissions, le changement climatique et la zone géographique est présentée, ainsi qu’un exemple de régression pour prédire les variations en pourcentage des températures.

**Caractéristiques du jeu de données**

* **Savanna fires** : Émissions liées aux incendies dans les écosystèmes de savane.
* **Forest fires** : Émissions issues des incendies en zones forestières.
* **Crop Residues** : Émissions provenant de la combustion ou décomposition des résidus de culture après récolte.
* **Rice Cultivation** : Émissions de méthane générées lors de la culture du riz.
* **Drained organic soils (CO₂)** : Émissions de CO₂ libérées lors du drainage de sols organiques.
* **Pesticides Manufacturing** : Émissions issues de la production de pesticides.
* **Food Transport** : Émissions liées au transport des produits alimentaires.
* **Forestland** : Surface forestière (puits de carbone avec émissions négatives).
* **Net Forest conversion** : Variation nette de la surface forestière (déforestation et reboisement).
* **Food Household Consumption** : Émissions dues à la consommation alimentaire au niveau des ménages.
* **Food Retail** : Émissions des commerces de détail alimentaires.
* **On-farm Electricity Use** : Consommation d’électricité dans les exploitations agricoles.
* **Food Packaging** : Émissions liées à la production et à l’élimination des emballages alimentaires.
* **Agrifood Systems Waste Disposal** : Émissions liées à l’élimination des déchets dans le système agroalimentaire.
* **Food Processing** : Émissions issues de la transformation des produits alimentaires.
* **Fertilizers Manufacturing** : Émissions liées à la production d’engrais.
* **IPPU** : Émissions issues des procédés industriels et de l’utilisation de produits.
* **Manure applied to Soils** : Émissions dues à l’épandage de fumier sur les sols agricoles.
* **Manure left on Pasture** : Émissions dues au fumier laissé sur les pâturages.
* **Manure Management** : Émissions liées à la gestion et au traitement du fumier.
* **Fires in organic soils** : Émissions des incendies de sols organiques.
* **Fires in humid tropical forests** : Émissions des incendies en forêts tropicales humides.
* **On-farm energy use** : Consommation d’énergie dans les exploitations agricoles.
* **Rural population** : Population vivant en zone rurale.
* **Urban population** : Population vivant en zone urbaine.
* **Total Population - Male** : Nombre total d’hommes dans la population.
* **Total Population - Female** : Nombre total de femmes dans la population.
* **total\_emission** : Émissions totales de gaz à effet de serre provenant de toutes les sources listées.
* **Average Temperature °C** : Augmentation moyenne annuelle de la température (en °C).

**Importance et contexte**

Le secteur agricole contribue à environ **62 % des émissions mondiales totales de CO₂**. Ce jeu de données est essentiel pour comprendre et suivre l’impact des activités agricoles sur les émissions.  
En utilisant des techniques de machine learning, il permet de prévoir les émissions futures et d’aider décideurs et chercheurs à concevoir des stratégies ciblées pour des pratiques agricoles durables.

**Notes de l’auteur**

* Les valeurs de CO₂ sont exprimées en **kilotonnes (kt)** : 1 kt = 1 000 kg de CO₂.
* **Average Temperature °C** peut servir de variable cible dans un modèle ML : par exemple, 0.12 signifie une augmentation de 0,12 °C sur un an.
* **Forestland** est la seule variable avec des émissions négatives, car les forêts agissent comme puits de carbone grâce à la photosynthèse.
* Pour convertir en Fahrenheit : °F = (°C × 9/5) + 32.