**Application des méthodes de Machine Learning pour l'évaluation de la potabilité de l'eau**

**Motivation :**

L'Afrique souffre d'un manque d'infrastructures de distribution et d'assainissement qui permettraient à la population d'avoir accès à l'eau potable. L'envie de contribuer à la lutte contre ce problème m'a conduit à mettre en place un projet permettant d'analyser la potabilité de l'eau.

(44 mots)

**L’ancrage :**

Dans le cadre du projet Tipe sous, thème "Santé et prévention", je proposerai des modèles mathématiques et informatiques permettant d'analyser la potabilité de l'eau avec des méthodes modernes basées sur l'apprentissage automatique.

(32 mots)

**Positionnement thématique :**

*MATHEMATIQUES* (Algèbre Linéaire), *MATHEMATIQUES* (Statistique), *MATHEMATIQUES* (Analyse), *INFORMATIQUE* (Intelligence Artificielle).

**Mots-clés :**

**Mots-Clés** (en anglais) **Mots-Clés** (en français)

Potable water L'eau potable

Machine Learning Apprentissage automatique

Classification Classification

Artificial intelligence L'intelligence artificielle

Water analysis l'analyse de l'eau

**Bibliographie commentée :**

Bien que les trois quarts de la surface de la terre soient recouverts d'eau, seul environ 1% est constitué d'eau douce et la qualité de l'eau a un impact direct sur la santé publique et l'environnement. L'eau est utilisée pour diverses pratiques, comme la boisson, l'irrigation, l'agriculture et l'industrie [1]. Mais l'augmentation de la population mondiale et l'industrialisation entraînent également la prolifération des polluants dans les masses d'eau. On estime que la mauvaise qualité de l'eau est la principale cause des maladies, l’eau contaminée et le manque d’assainissement entraînent la transmission de maladies comme le choléra, la diarrhée, la dysenterie, l’hépatite A, la typhoïde et la poliomyélite…[2]

Il est donc nécessaire de surveiller en permanence la qualité de l'eau provenant de sources naturelles en général et des eaux de surface en particulier. La surveillance et l'évaluation de la qualité de l'eau sont réalisables par la détermination de paramètres biologique ou des paramètres physico-chimiques. Grâce à la chimie moderne, nous pouvons détecter des milliers de produits chimiques dans l'eau, même à des concentrations extrêmement faibles.

Après avoir identifié les différents formats de tests, la question suivante est : que faut-il tester ? L'UNICEF recommande de donner la priorité à la surveillance chimique du fluor, de l'arsenic et du nitrate. Dans les régions où la terre est naturellement riche en minéraux contenant du fluor et de l'arsenic, les niveaux dans l'eau de puits peuvent être suffisamment élevés pour qu'une exposition chronique soit dangereuse pour la santé humaine [3]. La liste de plus en plus des tests disponibles peut sembler écrasante, alors que la grande majorité des méthodes nécessitent des installations de laboratoire de pointe, et beaucoup d'équipements et de multiples tests. Car un test standard de potabilité de l'eau prend en compte plusieurs facteurs.

Il existe des entreprises et des laboratoires qui proposent de tester l'eau du robinet ou l'eau de puits en ligne [4]. Il faut compter environ 100 euros pour un test de potabilité de l'eau du robinet ou un test de qualité générale de l'eau d'un forage par exemple. L'entreprise analyse l'eau en laboratoire et les résultats sont disponibles en ligne après une dizaine de jours. Ce qui est très difficile et nécessitant beaucoup de temps et de l’argent. Alors, comment exploiter des analyses chimiques simples (pH, Conductivité, …) sans faire appel à l’aide d’un spécialiste professionnel ? Autrement dit comment classifier des prélèvements d'eau selon la potabilité d’une manière automatique en se basant sur des simples analyses ?

Pour répondre à cette question, Le machine Learning intervient avec ses différentes méthodes pour trouver une solution et développer des applications et apprendre à la machine comment juger la potabilité ou la non potabilité d’une masse l’eau, en créant, entrainant, utilisant et évaluant le modèle sur différents échantillons de masses d’eau caractérisés par différentes variables [5].

Le machine Learning peut être classé en trois grandes catégories (Machine Learning  
avec supervision, sans supervision et renforcement). Dans notre cas, il s’agit de problème de machine Learning avec supervision (Apprendre par l’exemple) avec un modèle **de classification binaire (juger la potabilité ou la non** **potabilité de l’eau).** Cet apprentissage supervisé consiste en des variables d’entrée (X) et une variable de sortie (Y). Nous utilisons un algorithme pour apprendre la fonction de mapping de l’entrée à la sortie. Le but est d’appréhender si bien la fonction de mapping que, lorsque nous avons de nouvelles données d’entrée (X), nous pouvons prédire les variables de sortie (Y) pour ces données avec précision.

(569 mots)

**Problématique retenue :**

Les méthodes expérimentales dans les abrogatoires pour l’analyse de la potabilité de l’eau prend beaucoup de temps et d’investissement. De plus les payer africains qui souffres de ce problème généralement n’ont pas les moyennes techniques avancer d’analyse. Ce qui rend ces méthodes inutiles.

(38 mots)

**Objectifs du TIPE :**

L’objectif principale de ce sujet et d’essayer de trouver une méthode efficace non coûteux qui permet la détermination de la potabilité de l’eau avec une très grande précision en se basant sur des modèles de machine Learning.

(37 mots)

**Références bibliographiques :**

**Sites web :**

[1] : <https://www.cieau.com/connaitre-leau/leau-dans-la-nature/eau-douce-tout-savoir/#:~:text=L'eau%20douce%20utilisable%20par,les%20eaux%20souterraines%20(aquif%C3%A8res).&text=1%20%25%20de%20l'eau%20douce,fleuves%2C%20les%20lacs%2C%20etc>

**Sites web :**

[2] : <https://www.who.int/fr/news-room/fact-sheets/detail/drinking-water>

**Livre** (FEUILLE DE ROUTE)**:**

[3] : FEUILLE DE ROUTE POUR L’EAU DE l’UNICEF

Accès universel à des services durables d’approvisionnement en eau à l’horizon 2030

UNICEF

**Sites web :**

[4] : <http://environnement.wallonie.be/de/esu/laboeau.pdf>

**Thèse de doctorat :**

[5] : See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/346039290> Artificial intelligence for surface water quality monitoring and assessment: a systematic literature analysis Article in  Modeling Earth Systems and Environment · June 2021.