

Résumé

Le paludisme est l'une des maladies les plus mortelles en Afrique. En effet il manque très souvent de personnel qualifié pour réaliser un diagnostic fiable et précis d'où la nécessité de trouver des outils automatisés pour aider les professionnels de la santé dans la prise des décisions. Dans ce papier nous proposons un algorithme de diagnostic du paludisme basé sur les signes et symptômes du patient en plus du TDR (Test de Diagnostic Rapide). Pour la prédiction du paludisme la régression logistique est utilisée. Cependant pour le traitement des données et l'extraction des features les logiciels OpenRefine, R et python sont utilisés alors que pour la production de l'algorithme le Logiciel Python a été choisi.

Mots clés : Diagnostic, Paludisme, Prédiction, Signes et symptômes, TDR

Introduction :

Le paludisme est l'une des maladies les plus mortelles notamment dans des pays subaharien comme le Sénégal. Elle est due à la piqure de l' une maladie potentiellement mortelle due à des parasites transmis à l'homme par des piqûres de moustiques femelles infectés.

Selon le dernier Rapport sur le paludisme dans le monde, publié en novembre 2017, il y a eu 216 millions de cas de paludisme en 2016, contre 211 millions en 2015. On estime à 445 000 le nombre de décès dus au paludisme en 2016, un chiffre similaire à celui de l'année précédente (446 000). La Région OMS de l'Afrique supporte une part disproportionnée de la charge mondiale du paludisme. En 2016, 90% des cas de paludisme et 91% des décès dus à cette maladie sont survenus dans cette région. 80% de la charge de morbidité due au paludisme pesaient sur une quinzaine de pays – tous situés en Afrique subsaharienne, sauf l'Inde.

Malgré les efforts consentis par les Etats et organismes non gouvernementaux, le paludisme demeure toujours un facteur majeur de mortalité à travers le monde.

Cependant, la plus part de ces décès sont dus à un diagnostic incomplet qui ne révèle pas le type exact de paludisme.

La Stratégie technique mondiale de lutte contre le paludisme 2016-2030, adoptée par l'Assemblée mondiale de la Santé en mai 2015, constitue un cadre technique pour tous les pays d'endémie. Elle repose sur 3 piliers principaux:

- garantir l'accès universel à la prévention, au diagnostic et au traitement du paludisme,
- accélérer les efforts vers l'élimination et vers l'obtention du statut exempt de paludisme,
- faire de la surveillance du paludisme une intervention de base.

En outre le diagnostic précoce d'un paludisme grave permettra de réduire le nombre de décès au paludisme.

Dans cette étude nous proposons un outil de diagnostic du paludisme basé sur les signes et les symptômes du paludisme en plus du TDR (Test de Diagnostic Rapide). L'algorithme utilisé est une régression logistique et l'objectif est de fournir une aide au spécialiste au moment de faire un diagnostic.

Related Work

Le paludisme est l'une des maladies les plus mortelles notamment dans des pays subsaharien comme le Sénégal. Elle est due à la piqure de l'anophèle femelle dont la plus dangereuse est le Plasmodium falciparum. Ainsi beaucoup d'études ont porté sur la répartition et l'évolution

de ce moustique mais dans le diagnostic et la détection de la gravité de la maladie. Le machine learning est l'une des techniques les plus utilisées ces dernières années. La plus part de ces techniques sont basés sur l'analyse des données sanguine obtenues à partir de microscope de haut définition comme celle proposé par [1]. Cette méthode est un algorithme d'apprentissage automatique pour apprendre, détecter et déterminer les types de cellules infectées du sang. Elle consiste à la quantification du nombre de parasite plasmodium dans un frottis sanguin. 'The Jordan-Elman neural networks classifier' [2] quant à lui permet de détecter rapidement le paludisme mais aussi se déterminer sa gravité. Elle est basée sur une analyse des caractéristiques des données sanguine des patients. Une méthode semi-automatique de quantification et de classification des érythrocytes infectés par des parasites du paludisme sur des images microscopiques est proposée par (DÍAZ, Gloria, GONZÁLEZ, Fabio A., et ROMERO, Eduardo)[5]. Cette étude présente une méthode originale de quantification et de classification des érythrocytes dans de minces pellicules sanguines infectées par *Plasmodium falciparum*.

D'autres méthodes basées sur l'étude des signes et des symptômes du paludisme avec l'utilisation du machine learning ont été développées. Ainsi une technique basée sur les arbres de décisions est mise au point au Nigéria [3]. Cette approche est une méthode de classification basé sur les données de diagnostic des patient pour fournir une prédiction du paludisme. Notre contribution consiste en un modèle de prédiction en appliquant une régression logistique sur les signes et les symptômes des patients souffrant de paludisme et considérant le diagnostic comme variable de prédiction

Références :

[1] Kunwar, Suman, Manchana Shrestha, and Rojesh Man Shikhrakar. "Malaria Detection Using Image Processing and Machine Learning." *arXiv preprint arXiv:1801.10031* (2018).

[2] CHIROMA, Haruna, ABDUL-KAREEM, Sameem, IBRAHIM, Umar, *et al.* Malaria severity classification through Jordan-Elman neural network based on features extracted from thick blood smear. *Neural Network World*, 2015, vol. 25, no 5, p. 565.

[3] Ugwu, C., N. L. Onyejebu, and I. C. Obagbuwa. "The application of machine learning technique for malaria diagnosis." *International Journal of Green Computing (IJGC)* 1.1 (2010): 68-77.

[4] SAVKARE, S. S. et NAROTE, S. P. Automatic system for classification of erythrocytes infected with malaria and identification of parasite's life stage. *Procedia Technology*, 2012, vol. 6, p. 405-410.

[5] DÍAZ, Gloria, GONZÁLEZ, Fabio A., et ROMERO, Eduardo. A semi-automatic method for quantification and classification of erythrocytes infected with malaria parasites in microscopic images. *Journal of Biomedical Informatics*, 2009, vol. 42, no 2, p. 296-307.

[6] STEKHOVEN, Daniel J. et BÜHLMANN, Peter. MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics*, 2011, vol. 28, no 1, p. 112-118.

[7] CHEN, Hong, HAILEY, David, WANG, Ning *et al.* Examen des méthodes d'évaluation de la qualité des données pour les systèmes d'information de la santé publique. *Revue internationale de recherche sur l'environnement et la santé publique*, 2014, vol. 11, no 5, p. 5170-5207.

[8] JAMAL, Salma, PERIWAL, Vinita, et SCARIA, Vinod. Predictive modeling of anti-malarial molecules inhibiting apicoplast formation. *BMC bioinformatics*, 2013, vol. 14, no 1, p. 55.