# 

|  |  |
| --- | --- |
| **Name:** | OLUBODE OLAWALE OLUDARE |
| **Programme:** | M.Sc DATA SCIENCE |
| **Student ID:** | 22073847 |
| **Project Title:** | *Substituting PSA and Biopsy with Machine learning Algorithm for clinical diagnosis of Prostate Cancer* |
| **Problem statement:** | *Prostate cancer is one of the most prevalent forms of cancer among men, and early and accurate diagnosis is crucial for effective treatment and improved patient outcomes. However, the current diagnostic methods, such as PSA testing and biopsies, have limitations in terms of accuracy, invasiveness, and potential economic burden. These limitations often result in unnecessary biopsies, delays in diagnosis, and subsequent adverse effects on patients' quality of life.* |
| **Aim and scope:** | *The aim of this study is to investigate the potential of machine learning algorithms in substituting PSA testing and biopsies for the clinical diagnosis of prostate cancer. This study aims to develop accurate prediction models that leverage various patient data sources, such as demographics, medical imaging, and genetic information, to provide reliable and non-invasive diagnostic alternatives.*  *The study will involve collecting relevant datasets that include patient demographics, medical imaging etc. The data will be collected from clinical databases, research repositories, or collaborating healthcare institutions. The collected data will undergo pre-processing steps to handle missing values, outliers, and data inconsistencies.*  *Machine Learning Algorithms: I would focus on exploring various machine learning algorithms suitable for prostate cancer diagnosis, which may include but is not limited to classification models, ensemble methods and deep learning approaches. Evaluation metrics such as accuracy, sensitivity, specificity, precision, and area under the receiver operating characteristic curve (AUC-ROC) will be used to assess the performance of the models.*  *I intend to discuss the clinical relevance and practical implications of the developed machine learning models for prostate cancer diagnosis.* |
| **Project objectives:** | *This will include the followings:*  *i) Investigate the challenges and limitations associated with current prostate cancer diagnostic methods, including PSA testing and biopsies.*  *ii) Review existing literature on machine learning applications in healthcare and cancer diagnosis, with a specific focus on prostate cancer.*  *iii) Identify gap in the current literature regarding the use of machine learning algorithms to substitute PSA testing and biopsies for clinical diagnosis of prostate cancer.*  *iv) Develop a machine learning framework that integrates multiple data sources to accurately predict prostate cancer diagnosis.*  *v) Evaluate the performance of the proposed machine learning models and compare them with traditional diagnostic methods.*  *vi) Assess the clinical relevance, interpretability, and ethical considerations of the machine learning models in a clinical setting.*  *vii) Provide recommendations for the implementation and future research directions in the field of using machine learning for prostate cancer diagnosis.* |
| **Expected project outcomes:** | *Designing of machine learning-based diagnostic models specifically for clinical diagnosis of prostate cancer. This can take the form of a software application or a predictive model, which will leverage on various patient data sources, to provide accurate and reliable prediction, by addressing the challenge of false positives and false negatives in diagnosis, leading to more precise and reliable identification of prostate cancer cases.*  *Also, the study will tend to improve diagnostic accuracy, of non-invasive and cost-effective alternatives to invasive procedures like biopsies, using machine learning algorithms and integrating multiple data sources, to reduce patient discomfort and potential complications associated with invasive procedures. Moreover, it has the potential to alleviate the economic burden placed on patients and healthcare systems by minimizing the need for costly diagnostic tests and procedures. By providing accurate predictions, this will facilitate early detection of prostate cancer, enabling timely initiation of treatment and potentially improving patient outcomes and survival rates.*  *The developed models will serve as a support tool, assisting healthcare professionals in making informed decisions regarding prostate cancer diagnosis, treatment planning, and monitoring. With reliable and interpretable predictions, this will empower clinicians to make evidence-based decisions and optimize patient care. To add with, ethical considerations would be prioritized in the design, ensuring patient data privacy, informed consent, and fairness in predictions.* |
| **Brief review of relevant literature:** | *Prostate cancer is the third most common cancer globally and the fifth leading cause of cancer-specific death in men (Bray et al., 2018). Despite relatively high survival rates, the significant number of prostate cancer-related deaths recorded in 2012 highlights the need for improved diagnostic methods. Age, family history, and race, particularly among black men, are established risk factors for prostate cancer incidence and mortality (Ferlay et al., 2018).*  *The prostate-specific antigen (PSA) diagnostic technique is widely used for screening and detecting prostate cancer (Pinsky et al., 2017). However, the reliability of PSA can be influenced by false positives caused by inflammations or prostatic hyperplasia, leading to inaccurate prognostic reports (Pinsky et al., 2017). Transrectal ultrasound-guided biopsy, while considered the most precise diagnostic method, has limitations such as potential infections, invasiveness, and associated costs (U.S Prev. Ser, 2019; Wang et al., 2018; PSA Fact Sheet, 2019).*  *Recent developments in artificial intelligence, particularly in machine learning techniques, have shown promise in addressing the challenges of prostate cancer diagnosis. Although several studies have attempted to predict and diagnose prostate cancer using machine learning individually (Eleni et al., 2010; Takuechi and Hospital, 2018) low prediction accuracy and sensitivity remain significant issues. These limitations may be attributed to factors such as inadequate feature selection, small sample sizes, or suboptimal model architectures (Leydon et al., 2015).*  *This literature review highlights the limitations of existing diagnostic methods and the need for innovative solutions to enhance the accuracy, sensitivity, and specificity of prostate cancer detection, ultimately leading to improved patient outcomes and reduced mortality rate.* |
| **References:** | *Bray, Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J. Clin. 2018, 68, 394–424. [CrossRef] [PubMed]*  *Ferlay J, Soerjomataram I, Ervik M, et al. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11. International Agency for Research on Cancer, Lyon. 2013. http:// globocan.iarc.fr. Accessed 24 May 2018.*  *Schroder FH, van der Cruijsen-Koeter I, de Koning HJ, Vis AN, Hoedemaeker RF, Kranse R. Prostate cancer detection at low prostate specific antigen. J Urol. 2000;163(3):806-812.*  *Pinsky PF, Prorok PC, Yu K, et al. Extended mortality results for prostate cancer screening in the PLCO trial with median follow up of 15 years. Cancer. 2017;123(4):592-599*  *U.S Preventive services Task Force. Final Update Summary: Prostate Cancer: Screening; U.S. Preventive Services Task Force: Rockville, MD, USA, 2018. Data 2019, 4, 129 14 of 15*  *Wang, G.; Teoh, J.Y.; Choi, K. Diagnosis of prostate cancer in a Chinese population by using machine learning methods. In Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 17–21 July 2018.*  *Prostate-Specific Antigen (PSA) Test. [4/10/2019]. Available online: https://www.cancer.gov/types/prostate/ psa-fact-sheet (accessed on 8 June 2019).*  *Eleni Alexandratou \*, Vassilis Atlamazoglou and Trias Thireou George Ag. Evaluation of machine learning techniques for prostate cancer diagnosis and Gleason grading Int. J. Comput. Intell. Syst. Biol., vol. 1, no. 3, pp. 298–315, 2010.*  *T. Takeuchi and K. R. Hospital, “Prediction of prostate cancer by deep learning with multilayer artificial neural,” no. August, 2018.*  *P. Leydon, F. Sullivan, and F. Jamaluddin, “Machine Learning in Prediction of Prostate Brachytherapy Rectal Dose Classes at Day 30,” in Proceedings of the 17th Irish Machine Vision and Image Processing Conference, 2015, pp. 105–109.* |
| **Testing and Evaluation:** | *Cross-validation techniques will be applied to assess the capability and compare its performance against traditional diagnostic methods. The interpretability and explain ability will be examined, considering factors influencing its predictions. Robustness and generalization will be tested using external datasets or prospective validation studies. Ethical considerations will be evaluated to ensure fairness and unbiased application.* |
| **Signature** |  |