Erratum to: A hybrid feature selection approach for urinary tract infection detection and prediction in IoT-fog environment

Anisha Nagpal\textsuperscript{a} | Munish Sabharwal\textsuperscript{b} | Rohit Tripathi\textsuperscript{b}

\textsuperscript{a}School of Computing Science and Engineering, Galgotias University, Greater Noida, India.
\textsuperscript{b}Department of Electronics Engineering, JC Bose University of Science and Technology, India.

*Corresponding author: anishanagpal@outlook.com

- **Title:**
  
  The original title “A hybrid feature selection approach for urinary tract infection detection and prediction in IoT-fog environment” was changed to “A hybrid feature selection approach for urinary tract infection detection and prediction”.

- **1. Introduction:**
  
  Topic content was updated by:
  
  Urinary tract infections (UTIs) are widespread bacterial ailments, impacting approximately 150 million people annually. UTI serves as an umbrella term encompassing infections characterized by the infiltration of pathogens throughout the urinary system, including cystitis, pyelonephritis, renal abscess, urethritis, and prostatitis. UTIs result from a range of microorganisms, including both Gram-negative and Gram-positive bacteria, as well as specific fungal strains. Notably, uropathogenic Escherichia coli (UPEC) emerges as the predominant infectious agent for both uncomplicated and complicated UTIs, although the distribution of causative pathogens may exhibit variation (Goździkiewicz, 2022).

  Numerous people across the world suffer from urinary tract infections (UTIs), which are a prevalent medical problem (Bankar, 2021). UTIs are brought on by bacteria that enter and attack the urinary tract, which consists of the bladder, urethra, ureters, and kidneys (Verma, 2023). These bacteria are typically from the digestive system i.e Escherichia coli. The signs and symptoms of this infection can range from unpleasant to possibly dangerous. It can affect people of any age or gender identity, but women are more probable than men to experience them due to the female urethra’s being smaller, which makes it simpler for bacteria to get into the urinary system (Wong, 2023). The use of urinary catheters, anomalies of the urinary tract, pregnancy, menopause, a compromised immune system, and certain medical disorders including diabetes are additional risk factors. A medical expert frequently obtains a urine sample to diagnose UTIs in order to check for bacteria or white blood cells (Silva, 2022). An antibiotic course is typically prescribed as part of the treatment to eradicate the disease. To ensure full elimination of the germs and avoid repeated sickness, it is imperative to finish the entire prescribed course of medicines. UTIs can be prevented by taking proactive steps, such as practicing excellent hygiene, consuming lots of water, urinating regularly, and clearing the bladder both before and after sexual activity (Baijwan, 2023). Further examination by a medical professional may be required for people who have recurring UTIs in order to determine the root cause and create a personalized preventative strategy. The infection can be obstructing but with the right medical care, they are typically manageable (Aggarwal, 2022). People can reduce their risk and seek prompt medical attention when necessary by being aware of the causes, signs, and preventative measures linked to infection (Sulis, 2022).

  The early and efficient management of those with urinary tract infections depends critically on the detection and prediction of these conditions. Machine learning and data-driven methodologies have recently demonstrated considerable potential in supporting medical practitioners in identifying and forecasting UTIs (Bijlani, 2022).

  Machine learning models for predictions are proving to be invaluable tools in clinical practice, as they offer improved guidelines for decision-making in personalized patient care. These models have the capability to self-diagnose a wide range of diseases, aligning their assessments with established clinical guidelines. This integration of machine learning in healthcare empowers medical professionals to make more informed and precise decisions, ultimately leading to better outcomes for individual patients. In addition, IoT has been integrated with ML techniques to monitor the well-being and health of individuals.
affected by dementia. This combined model aids in delivering more effective preventive care, ultimately reducing the need for hospitalization (Enshaeifar 2019, Gadalla 2019).

In the forthcoming years, the significance of fog computing is projected to be of paramount importance in addressing the ever-growing need for instantaneous services. As a platform, fog computing offers increased storage capacity, real-time computational power, and network services, effectively bridging the gap between data centers and end-users. The integration of IoT-Fog computing facilitates the execution of numerous time-sensitive data and services, including emergency health services and medical diagnosis. This convergence of technologies is poised to revolutionize healthcare services, enabling more efficient and responsive healthcare delivery for the benefit of individuals and communities (Bansal 2022, Alshamrani 2022).

The primary objective of this current research is to formulate an innovative algorithm grounded with Machine Learning (ML) technique to enhance the accuracy rate of UTI prediction, as this factor holds significant importance in ensuring better patient care. A hybrid features selection method is one such strategy that utilizes the effectiveness of various feature selection techniques in order to increase UTI detection and predicting reliability and precision (Gehringer 2021). In this study, a novel mixed feature choice methodology for UTI identification and forecasting is presented. The suggested approach makes use of the advantages of several algorithms for choosing features to extract the most useful and pertinent features from a wide range of input parameters. The hybrid technique when compared to conventional feature selection methods, the combination has a number of benefits (Hateet 2022). It can get over the drawbacks associated with separate procedures and provide a more thorough evaluation of the feature space by integrating various approaches (Behzadi 2019).

The generalization of the generated models is improved, and the risk of over fitting is decreased. Additionally, the hybrid feature selection technique is effective in handling duplicate or insignificant characteristics and high-dimensional information, enhancing computational effectiveness and model interpretation. The proposed work presents a hybrid feature selection approach with a Guided Regularized Random Forest (GRRF) classification model to assist with the diagnosis of UTI. Further the UTI prevention and course of action must be optimized via enhanced research and progressive efforts to overcome resistance from antibiotics.

2. Literature review:

The paragraphs below have been added:

The study (Jeng, 2022) develops a machine-learning model to detect recurrent UTIs. The work uses several classifiers for comparison, and the model uses a decision tree and achieves the accuracy of 92% in stage 1 and 94% in stage 2.

The aim (Enshaeifar 2019) is to propose a UTI detection algorithm and use a Non-negative Matrix Factorisation (NMF) technique to extract latent factors from raw observation and use them for clustering and identifying the possible UTI cases. The model achieves an accuracy of 85%.

4.1. Dataset Description:

The paragraphs below have been replaced:

The dataset (Taylor, 2018) used in this study comprises predictor variables encompassing laboratory results, urine dipstick results, urinalysis, past medical history, structural historical findings, physical exam findings, chief complaints, and demographic information. The dataset exists in two versions: one with a reduced set of 10 variables and another with a full set of 211 variables. It can be accessed through the link:


The another dataset (Taylor, 2018) also used in this study comprises predictor variables encompassing laboratory results, urine dipstick results, urinalysis, past medical history, structural historical findings, physical exam findings, chief complaints, and demographic information. The dataset exists in two versions: one with a reduced set of 10 variables and another with a full set of 211 variables. It can be accessed through the link:


Figure 3 was updated:
- **4.4. Accuracy and Loss for Training and Validation**

Where it reads “The accuracy of a urinary tract infection is illustrated in Figure 7, and it refers to how well it can efficiently integrate and preserve pertinent information from the input pictures while eliminating noise, antiques, and irregularities”, it is replaced by:

a) The relationship among accuracy and loss has an important role in training process of machine learning models. As, the high accuracy of the model is favorable but it is also necessary to keep monitoring the errors. This helps in achieving a balance between accuracy and loss. It facilitates the understanding of model learning and its ability of generalizing new data. The figure 7 shows the model’s accuracy for training and validation data.

Where it reads “The difference between the fused image’s quality or fidelity and the original input photos is shown in Figure 8. It shows the degree in which the process of fusion generates artifacts or discrepancies, fails to safeguard important information, or otherwise affects the quality”, it is replaced by:

b) The figure 8 shows the error between the values predicted by the model and the actual target values given in the dataset. The training loss shows the model's effectiveness in learning from the training data, whereas the validation loss examines the model’s capacity to generalize to unfamiliar data. The balance between minimizing training loss and avoiding over-fitting is crucial for the development of the models that consists strong performance across various datasets.

- **5.1. Limitation**

Topic content was updated by:

While the proposed system demonstrates the capability to achieve optimal accuracy, it is not without limitations and challenges. One significant limitation is its heavy reliance on the availability and quality of data, which can pose challenges in obtaining high-quality UTI-related features. Additionally, if ML algorithms need to be integrated with Internet of Things for the detection of UTI then the implementation of smart toilets equipped with IoT sensors presents significant challenges and incurs higher costs, further adding to the complexities of the system.

- **References**

References have been updated:


Behjadi, P., Behjadi, E., & Pawlak-Adamska, E. A. (2019). Urinary tract infections (UTIs) or genital tract infections (GTIs)? It's the diagnostics that count. GMS hygiene and infection control, 14.


The online version of the original article can be found at https://doi.org/10.31893/multiscience.2024080