

Article

Prediction of Urinary Tract Infection in IoT-Fog Environment for Smart Toilets Using Modified Attention-Based ANN and Machine Learning Algorithms

Abdullah Alqahtani ^{1,*}, Shtwai Alsubai ², Adel Binbusayyis ¹, Mohemmed Sha ¹, Abdu Gumaei ² and Yu-Dong Zhang ^{3,*}

¹ Department of Software Engineering, College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia; a.binbusayyis@psau.edu.sa (A.B.); ms.mohamed@psau.edu.sa (M.S.)

² Department of Computer Science, College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia; sa.alsubai@psau.edu.sa (S.A.); a.gumaei@psau.edu.sa (A.G.)

³ School of Computing and Mathematical Sciences, University of Leicester, Leicester LE1 7RH, UK

* Correspondence: aq.alqahtani@psau.edu.sa (A.A.); yudong.zhang@leicester.ac.uk (Y.-D.Z.)

Abstract: UTI (Urinary Tract Infection) has become common with maximum error rates in diagnosis. With the current progress on DM (Data Mining) based algorithms, several research projects have tried such algorithms due to their ability in making optimal decisions and efficacy in resolving complex issues. However, conventional research has failed to attain accurate predictions due to improper feature selection. To resolve such existing pitfalls, this research intends to employ suitable ML (Machine Learning)-based algorithms for predicting UTI in IoT-Fog environments, which will be applicable to a smart toilet. Additionally, bio-inspired algorithms have gained significant attention in recent eras due to their capability in resolving complex optimization issues. Considering this, the current study proposes MFB-FA (Modified Flashing Behaviour-based Firefly Algorithm) for feature selection. This research initializes the FF (Firefly) population and interchanges the constant absorption coefficient value with the chaotic maps as the chaos possesses an innate ability to evade getting trapped in local optima with the improvement in determining global optimum. Further, GM (Gaussian Map) is taken into account for moving all the FFs to a global optimum in an individual iteration. Due to such nature, this algorithm possesses a better optimization ability than other swarm intelligence approaches. Finally, classification is undertaken by the proposed MANN-AM (Modified Artificial Neural Network with Attention Mechanism). The main intention for proposing this network involves its ability to focus on small and significant data. Moreover, ANNs possess the ability for learning and modelling complex and non-linear relationships, in which the present study considers it. The proposed method is compared internally by using Random Forest, Naive Bayes and K-Nearest Neighbour to show the efficacy of the proposed model. The overall performance of this study is assessed with regard to standard performance metrics for confirming its optimal performance in UTI prediction. The proposed model has attained optimal values such as accuracy as 0.99, recall as 0.99, sensitivity as 1, precision as 1, specificity as 0.99 and f1-score as 0.99.

Keywords: urinary tract infection; smart toilet; firefly algorithm; artificial neural network; attention mechanism; Gaussian Map



Citation: Alqahtani, A.; Alsubai, S.; Binbusayyis, A.; Sha, M.; Gumaei, A.; Zhang, Y.-D. Prediction of Urinary Tract Infection in IoT-Fog Environment for Smart Toilets Using Modified Attention-Based ANN and Machine Learning Algorithms. *Appl. Sci.* **2023**, *13*, 5860. <https://doi.org/10.3390/app13105860>

Academic Editor: Krzysztof Koszela

Received: 1 April 2023

Revised: 29 April 2023

Accepted: 6 May 2023

Published: 9 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

IoT brings more changes to the life of humans and increases the quality of life [1] by integrating Artificial Intelligent (AI) and cloud computing [2,3]. In recent trends, the Internet of Things (IoT) is considered to be the more attractive and revolutionary technology that plays a vital role in various fields such as healthcare, house automation, smart cities [4,5],

wearable devices and so on. IoT greatly influences the healthcare industry which in turn helps various applications of smart healthcare [6]. IoT acts as a base for innovative ideas such as Cyber-physical Systems (CpS) [7], smart cities [5] and Industry 4.0 [8]. In the future, fog computing [9] is used to address the demand for real-time service, which is increasing. Fog computing acted as a platform that provides more storage space, real-time computation power and network services among data centres and end users. By using IoT-Fog computing, many time-sensitive data and services are carried out which include emergency health services and medical diagnosis. Smart healthcare includes various applications such as intelligent patient monitoring, wireless tracking of health and mobile healthcare. Many cities aim to establish the idea of smart city healthcare by means of conventional equipment and devices that integrate the resources of healthcare and smart solutions [10].

Generally, Machine Learning (ML) models consist of many predictor variables and require less a priori hypothesis in comparison to statistical models [11]. In the ML-based approach, SVM (Support Vector Machine) and RF (Random Forest) are used to allow fewer assumptions and many difficult relations between predictors. The algorithms are combined with RFE (recursive feature elimination) to provide the optimal predictors for UTI by using the biomarkers present in urine [12]. To monitor the well-being and health of people who suffered from dementia, IoT has been used, and it also combines ML techniques. The model helps in providing more effective preventive care, which reduces hospitalization [13].

UTI [14] is a particular kind of infection that occurs in various parts of the urinary system, which includes the kidneys, urinary bladder and ureters. After kidney transplantation, many people suffer due to UTI [15]. UTI has been found in both genders irrespective of age [15]. At the age of 24, one out of three women has at least one symptom of UTI, and while considering the lifetime, nearly half a population is affected [16]. A survey by the WHO (World Health Organization) confesses that over 1 million people were hospitalized in 2011 due to UTI [17]. However, quick detection of UTI can prevent people from permanent kidney problems, blood loss and severe pain. Several tests are available to test for UTIs, but home-based UTI testing and detection provide an efficient way for precautionary measures. However, the diagnosis of UTI is still challenging, and current techniques are often prone to more errors. To overcome the issue, several studies have explored the use of ML algorithms to predict UTIs. However, the accuracy of these algorithms is limited due to the improper selection of features. Therefore, the present study aims to develop an ML-based algorithm for UTI prediction in an IoT-Fog environment based on the application of a smart toilet. Additionally, the present study aims to enhance the accuracy rate of UTI prediction, which plays a significant role in patient care.

Further, many IoT devices are available and have an increased capability in acquiring the specific parameters of urine and the determination of UTIs. Especially, IoT includes internet-enabled sensors that have the ability to obtain universal data and transfer the data to a remote location [18]. Additionally, the devices have used technologies that are easily available and hence can be used in a smart toilet system.

IoT technology provides a considerable revolution in the recent healthcare industry, which helps in monitoring remote patients, smart health services and mobile healthcare. Real-time healthcare services have been boosted by using IoT technology and assisted in the development of communication protocols such as CoAP [19], MQTT [20], next-generation IPv6 version [21] and AMQP [22] of inter-network. The IoT has enhanced the home-centric environment by extending healthcare services. IoT sensors help in expanding the acquisition of health data, and various technologies such as data mining results in automatic analysis of health conditions without the presence of a concerned doctor. Furthermore, the large use of fog and cloud computing results in providing enhanced services in healthcare applications, which includes time-sensitive data management and analysis. In the modern world, improved IoT and fog-cloud computing [23] serves many applications in real-time healthcare.

The Firefly Algorithm (FA) [24] has increased importance in swarm intelligence [24] and is used in the area of both optimization and engineering. FA and its variants are used

to solve various problems from different areas. The ideal behaviour of FA is its flashing feature [25], which helps in efficient feature selection. Though the basic Firefly Algorithms are comparatively efficient, the outcomes have been changed when it approaches the optima. The success rate of FA depends on the balance that exists between exploitation and exploration. CGFA (Chaotic Gaussian Firefly Algorithm) possesses improved numerical stability and an enhanced convergence rate [26].

The present study has used MFB-FA (Modified Flashing Behaviour-based Firefly Algorithm) for feature selection by utilizing IoT in a fog environment. Gaussian Chaotic Maps (GCMs) in FA have been used to reveal the optimal fit that helps in exploring chaotic maps. Gauss Map (GM) avoids the trapping of local optima by moving towards the optimal solution and possesses better optimization ability in comparison to swarm intelligent approaches. Hence, the proposed method utilizes GM for attaining improved optimization ability. The present study has used MANN-AM (Modified ANN with Attention Mechanism) for classification. ANN possesses more advantages such as parallel processing, the ability to perform with partial knowledge, storing data within the network and possessing memory distribution and fault tolerance. Similarly, the attention mechanism imitates cognitive attention. MANN-AM performs so that the network has to devote high focus to low but significant data parts. The present study aims to detect UTI by using MFB-FA and MANN-AM based on the following objectives:

- To perform feature selection by using MFB-FA (Modified Flashing Behaviour-based Firefly Algorithm) for improvising classification.
- To perform classification by using MANN-AM (Modified Artificial Neural Network with an Attention Mechanism) to efficiently prognosticate the existence/absence of UTI.
- To evaluate the efficacy of the proposed model with standard-performance metrics for exposing its performance in predicting UTI.

1.1. Novelty

The present study focuses on the relatively new field of research that used bio-inspired and ML algorithms in IoT-Fog environments for predicting UTIs, in which the MFB-FA for feature selection is a revised version of the Firefly Algorithm that utilized chaotic maps for avoiding local optima and assists in improving the global optimization. MANN-AM is a revised version of ANN that utilized an attention mechanism for focusing small and significant data. Therefore, the integration of the novel approaches offers an optimal solution for predicting UTIs by using a smart toilet.

1.2. Contribution

The present research uses an ideal approach that combines both ML and bio-inspired algorithms that highly contributed to the research. MFB-FA utilizes the flashing behaviour depending on the optimal solution of the population. In MFB-FA, each firefly attracts the optimal firefly from the neighbourhood, which depends on a modified attractiveness function. However, MFB-FA possesses limited performance as it becomes trapped in local optima, which may result in a suboptimal feature selection. In order to overcome the limitation, the proposed method utilizes GM for improved optimization ability. By using the GM technique, in each iteration, the position of the fireflies is shifted towards the global optimum. Thereby, GM avoids getting trapped in local optima by moving towards the optimal solution. Moreover, the proposed method swaps the value of the constant absorption coefficient with chaotic maps in MFB-FA. The ability of the chaotic maps evades getting trapped in local optima, introducing randomness into the algorithm, which leads to a highly diverse search of feature space. The combination of GM and chaotic maps in MFB-FA improves the ability of optimization, which results in an exact and efficient selection of features for UTI prediction. MANN-AM utilizes an attention mechanism for scoring and computing the significance of each feature in the prediction of UTI. During classification, the weights of the input features are adjusted based on the scores, which mainly focused

on highly informative features for predicting UTI. Moreover, MANN-AM is modified by including additional layers to perform feature fusion for integrating the selected features. The additional layer combines the selected features by using the computed mean and standard deviation that is given as input to the network. Thus, the additional layers assist in capturing the non-linear and complex relationship among the selected features and in UTI predictions. Therefore, MANN-AM assists in the effective and accurate classification of UTI in an IoT-Fog environment.

1.3. Limitation

Though the proposed system has the ability to attain optimal accuracy, it possesses some limitations and challenges. The proposed system extensively depends on the availability and quality of the data and it remains challenging to use high-quality UTI-related features. The implementation of smart toilets with IoT sensors is more challenging and costlier.

Section 2 of this paper deals with the review of the existing literature for feature selection and classification through various ML models. The problem identified from the existing literature has been also discussed. Section 3 deals with the proposed flow along with the pseudo codes. Section 4 deals with the dataset description, performance metrics, experimental results, performance analysis and comparative results. Section 5 deals with the conclusions and future recommendations of the work.

2. Review of Existing Work

In recent times, the application of smart healthcare attains more popularity. Various ML methods have been used in UTI for feature selection and classification. The section has reviewed the conventional models briefly along with the gaps identified from the state-of-the-art works.

IoT in fog environments has been used for quick prediction and diagnosis of urine infection. Initially, multiple sensors have been used to deliver the data and by using XGBoost [27] ensemble a learning approach, IRF (Infectious Risk Factor) has been calculated at the fog layer. The suggested study [28] used real-time patient data. The results of accuracy, f-score, sensitivity, precision and specificity have been given as 91.45%, 90.12%, 84.79%, 95.49% and 95.96%, respectively.

For monitoring and prediction of UTIs, the existing paper [29] has suggested an IoT-inspired framework in a home-centric environment. By using ID (Infection Degree), probabilistic classification [30] (Bayesian Probabilistic Model) of UTI has been performed in accordance with both infection and non-infection categories. Temporal-Artificial Neural Network (t-ANN) technique, which incorporates IIM (Infection Index Measure), has been used for the detection and prediction of UI. Kohonen Self Organized Mapping Technique [31] (SOM) has been used for the color visualization of information. Additionally, the development of AI has brought several improvements in the healthcare sector. Thereby, ML and DL methods have also been widely used in detecting several life-threatening diseases such as cervical cancer [32] and COVID-19 [33,34] and also in predicting PCOS [35].

Diagnosis and monitoring of clinical voiding dysfunction have been assisted in the recommended study [36]. While urinating, the vibration signals have been detected with the help of an accelerometer and they transformed into MFCC (Mel-Frequency Cepstrum Coefficient) [37]. The CNN model with pooling and normalization layers with ReLU (Rectified Linear Unit) [38] has acted as active functional layers. Classification has been done through the Softmax function [39]. UMAP (Uniform Manifold Approximation and Projection) [40] dimensionality reduction has been used to predict and analyse uroflowmetry and also helps in void dysfunction diagnosis. The result of accuracy has been found to be 98%.

An integrated platform has been developed and ML models have been used for the analysis of UTI risk and agitation in the suggested study. A huge dataset has been collected, by which a new DL (Deep Learning) model has been evaluated that utilized rational and attention mechanisms. Rational layers have passed through the LSTM layer, and LSTM [41] has processed the data available in the attention block. For prediction, classifiers have used

time steps. In detecting the risk of UTIs and agitation, the outcome values of the suggested model [42] for precision and recall have been given as 83% and 91%.

Noise, explainability, high dimensionality, large variability and label scarcity are some challenges present in anomaly detection. The efficiency of MP (Matrix Profile) and SOTA has improved by using CMP (Contextual Matrix Profile), which interprets and visualizes the anomalies present in the data of the noisy real world. Graph-based anomaly detection has been used to detect anomalies by using a real-world dataset in healthcare. The recommended method [43] has possessed an enhanced alert rate, recall and generalizability.

Adaptive DNN (Deep Neural Network) [44] and supervised parabolic model have trained the model of unlabelled data by utilizing unsupervised techniques. DNN has utilized auto-encoder in terms of learning unlabelled data representation. The trained model has been applied to the labelled data of a smaller set, which helps in the development of a supervised classifier with improved reliability. DL and probabilistic models have been used to analyse UTI risks. The classification of data has been carried out by using SVM [45] (Support Vector Machine with polynomial kernels and RBF), LR (Logistic Regression) [46], DT (Decision Tree) [47] and KNN (K-Nearest Neighbour) [48]. The suggested ML model [49] has detected the risk of UTI with a recall value of 85% and precision of 86% with the help of in-home monitoring data. In the existing study, UTI in pregnant women and children less than age 10 has been considered. ML approach has tested three algorithms such as RF [50], NN [51] and extreme gradient boosting; at the same time, factoring of independent variables has been executed. A sensitivity of 95% has been achieved by using the suggested model [52].

Similarly, the existing model [53] has aimed to develop an ML model that predicts the risk of MDR UTI (Multidrug Resistant Urinary Tract Infection) for patients after hospitalization. Catboost has been used to exhibit the results, and the outcomes are found to be 0.909 as Matthews's correlation coefficient (MCC), 0.809 as f1-score and 0.717 as accuracy. The predictive model has used an ML algorithm and provided an efficient performance that facilitates UTI. The suggested model [54] has predicted the RUTI (Recurrent Urinary Tract Infection) that has been associated with vesico-ureteral reflux and with a value of 0.761 as the AUC (Area Under the Curve) in the test set. ML algorithms have been used in the prediction of identifying the incidence of stroke and post-stroke prognosis for immobile stroke patients. The ensemble learning model in the recommended study has provided an efficient performance in the prediction of UTI risk.

A Knowledge Discovery and Data Mining (KDDM) approach using ML methods has been implemented to find the knowledge regarding HA-CAUTI (Hospital-Acquired Catheter-Associated Urinary Tract Infection) from various sources of data and helps in predicting the risk of HA-CAUTI. ML techniques such as DT, LR and SVM have been described in the existing study [55]. The cutting-edge ML approach has been introduced in electronic health records, which improves secondary utilization. The accuracy of DT, SVM and LR has been found to be 75.87%, 71.50% and 75.83%, respectively.

The test characteristics of ML models have been evaluated to find the results of SBI (Serious Bacterial Infections) in infants and the models have been compared with the recognised clinical prediction rules. The four supervising models such as single-hidden layer neural network, LR, SVM and RF have been developed and validated. The methods have been used for the exact identification of patients without UTI and the recommended model [56] has attained a sensitivity of 98%. Statistical models have utilized EHR (Electronic Health Record) for identifying the complications of surgery such as UTI, sepsis, bleeding and surgical infections. The elastic-net model has been used to calculate the regression coefficient and to process variable selection. The recommended model [57] has achieved 83% of sensitivity, 88% of specificity, an area under the curve of 0.93, 52% of positive predictive value and 97% of negative predictive value.

In order to improve predictive accuracy, alternative ML approaches (DL methods and tree-based methods) have been implemented and explored in the recommended study [58]. To predict the risk of an AE (Adverse Effect) including UTI, the selection operator regulation

method, least absolute shrinkage and LR have been applied. For overall AE, the AUC of the prediction model has been found to be 0.7. However, numerous studies have shown AUC values between 0.6 and 0.78 in predicting the complications of UTI after spine surgery. Initially, ML methods have trained both huge records of patients and measurements of antibiotic resistance. The existing study [59] has provided a way to use ML in decision-making, thereby it helps in the recommendation of antibiotics over UTIs, which has reduced the risk of failure of treatment.

By using Fourier transform infrared, bacterial susceptibility has been examined for first-culture *E. coli* colonies for UTI patients. The objective of examining the infrared microscopy in which multivariate analysis has been combined is to cut short the time taken for identifying *E. coli* susceptibility to antibiotics and to find the response of bacteria to the optimal choice of antibiotics. The outcomes of the suggested study [60] have shown that it has been evident to find the optimal choice of antibiotic in which the sensitivity has been found to be 89%, which has followed the first culture. Linear SVM has been used to implement multidimensional classification. When ML classification algorithms have combined with infrared spectroscopy for pattern recognition, it has been considered a powerful tool for determining the susceptibility of *E. coli* bacteria, which is a commonly used antibiotic.

For identifying pathogenic microorganisms, MALDI-TOF MS (MALDI-TOF Mass Spectrometry) has been utilised in the diagnosis of infectious diseases. The DL model has been used in analysing the data, and the DL model performance has been compared with conventional and supervised ML algorithms. The suggested study [61] has illustrated a huge dataset that contains bacterial species in relation to UTIs. By selecting five different species, an accuracy of 85% has been attained.

The ability of infrared spectroscopic microscopy has been evaluated, which includes the multivariate analysis for quick detection of ESBL (extended-spectrum β -lactamase), which produces *Escherichia coli* (*E. coli*) and has been separated from UTI samples. The outcomes of the suggested study [62] have shown that it has been evident to find the ESBL-producing bacteria with 99% of sensitivity, 97% of success rate and 94% of specificity for samples that have been tested with less time by succeeding the first culture.

Least Absolute Shrinkage and Selection Operator (LASSO) and LR have been executed to recognise and select the variables for analysis and its results with the potential model. By using the numerical outcomes and clinical criteria, the final model has been selected. Developing ML algorithms for a huge dataset has been made possible, which delivers beneficial tools for counselling patients and in the assessment of surgical risk. The value of AUC has been found to be 0.70 with a Brier score of 0.01. The results of the predictive model for UTI in accordance with specificity and sensitivity have been given as 95% and 97%, respectively [63].

Problem Identification

Various problems identified from the extensive literature have been discussed as follows:

- The collected data from the system have limitations in facilitating large databases and various sources of information and also in voiding dysfunction diagnosis algorithms. The existing algorithm has fewer limitations in accessing the diagnosis information and sometimes has the probability to get diagnostic errors [36].
- By using ML methods, risk factors have generally been neglected by classical statistical models. The data have been extracted on a small scale, and hence the findings possess some limitations in terms of generalizability [53].
- For practical examination, a more precise scoring tool based on the ML model has been needed to improve the outcomes of the model [11].

3. Proposed Methodology

This study tries to predict the existence or absence of UTI (Urinary Tract Infection) based on ML methods. In spite of the endeavours undertaken by conventional works to accomplish this, they lacked in accordance with classification rates. To attain better prediction performance, this study proposes suitable feature selection and classification algorithms as depicted in Figure 1. In this case, the UTI dataset is initially loaded. Following this, the data is pre-processed. Typically, the pre-processing of data is crucial prior to its actual usage. It is the process of converting raw data into a clean dataset. Pre-processing is undertaken to check the missing values, noise and other inconsistencies prior to the implementation of the algorithm. After checking the missing values, categorical encoding is performed to convert the categorical data into integer form by which data with the converted category values could be afforded to models for attaining enhanced prediction. As data must be in a suitable format for ML, pre-processing is done. Further, suitable features are selected for making the process accurate. This improves the prediction ability of algorithms through the selection of critical variables, thereby eliminating irrelevant and redundant features. To achieve this, the study regards MFB-FA (Modified Flashing Behaviour-based Firefly Algorithm). In this process, GCMs (Gaussian Chaotic Maps) are used with FA for selecting suitable features. This is fed into the train and test split with 80% training and 20% testing. Lastly, the classification process is undertaken by MANN-AM (Modified Artificial Neural Network with an Attention Mechanism).

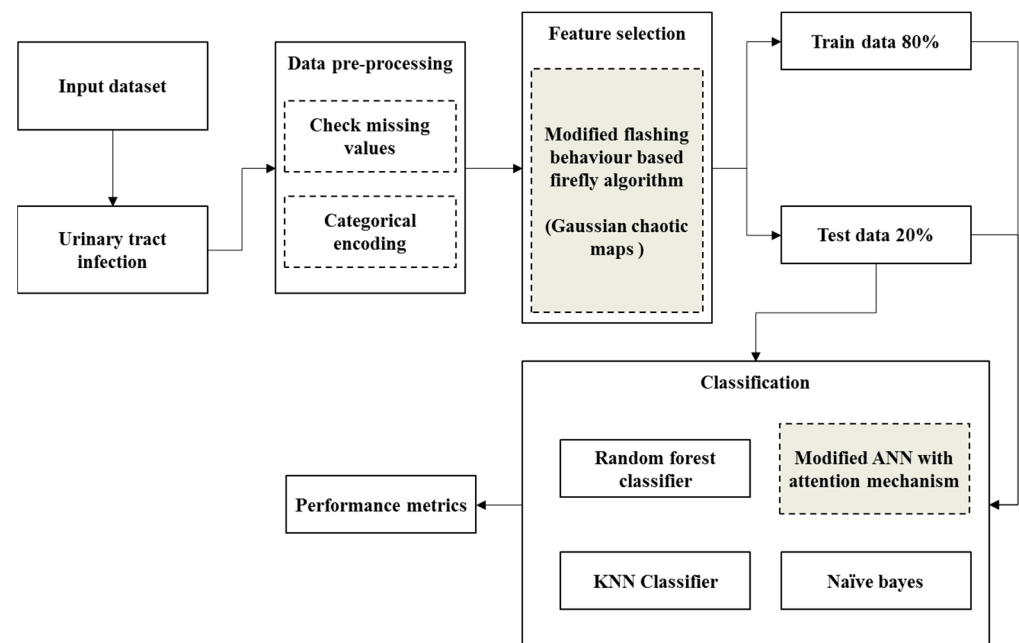


Figure 1. Overall view of the proposed system.

In this case, algorithms, namely, KNN, RF and NB are regarded for internal comparison with the proposed system. The overall performance is assessed with regard to performance metrics.

3.1. Feature Selection MFB-FA (Modified Flashing Behaviour-Based Firefly Algorithm)

FA (Firefly Algorithm) is stimulated by the FF's (Firefly's) flashing behaviour. Three assumptions exist in FA as follows:

- (i) All FFs (Fireflies) are uni-sexual. Each FF gets attracted to each of the other FF.
- (ii) Attraction is proportional to the brightness of FF. For any 2 FFs, the less bright FF will get attracted by the brighter FF, and the brightness will minimize as the distance enhances.

(iii) When no FFs brighter than a specific FF exists, it starts to move in a random manner.

Typically, light intensity (LI(r)) is inversely proportional to the distance from the light source. Hence, when the light passes via a medium with the light absorption coefficient of (β), then LI(r) differs with the distance (r) as specified in Equation (1):

$$LI(r) = I_0 e^{-\beta \text{dist}^2} \quad (1)$$

In Equation (1), I_0 represents the intensity at the source-point. As the computation of $LI(\text{dist}) = \frac{1}{1+\beta \text{dist}^2}$ is simpler than $e^{-\beta \text{dist}^2}$, LI(r) could be computed as follows:

$$I(\text{dist}) = \frac{I_0}{\beta \text{dist}^2} \quad (2)$$

Each FF possesses its unique attractiveness (α) and could be claimed as per Equation (3):

$$\alpha(\text{dist}) = \alpha_0 e^{-\beta \text{dist}^2} \quad (3)$$

Likewise, it could be stated as follows:

$$\alpha(\text{dist}) = \frac{\alpha_0}{1 + \beta \text{dist}^2} \quad (4)$$

In Equation (4), α_0 denotes the attraction at dist as 0.

A FF positioned at the movement (q_i) gets attracted to another bright FF positioned at (q_m) and is given by Equation (5):

$$q_{i+1} = q_i + \alpha_0 e^{-\beta \text{dist}^2} (q_m - q_i) + \zeta \epsilon \quad (5)$$

The second term is accredited to attraction, while the third term gets randomization with ($\zeta(0 \leq \zeta \leq 1)$) and ϵ . For numerous practical issues, a constant value ($\zeta = 0.2$) can be used. In this case, ϵ denotes the random variable vector retrieved from the Gaussian distribution. On the contrary, chaos exists as a stochastic motion mapped by a deterministic equation. It varies from disorder and irregularity. Chaos possesses a fine internal structure. It comprises three characters: random, regularity and ergodic. The ergodic property could search all the states with its formulae within a specific range. Thus, chaos turns into an efficient approach to avoid getting trapped in local optima with the enhancement in determining global optimum.

The Gaussian function is given by Equation (6):

$$q_{i+1} = \exp(-\zeta q_i^2) \quad (6)$$

From Equation (6), it is revealed that the present study initializes the FF population and interchanges the constant absorption coefficient value with the chaotic maps. In this research, GM (also termed mouse-map) is taken into account for moving all the FFs to global optimal in individual iteration. Further, for stabilizing the movement of a FF, a distinct behaviour is proposed for directing the movement of FFs to the global best solution when no ideal solution around it exists.

When a FF is utilized for optimizing the multiple peak function, it could get effortlessly trapped in local minima. This results in less convergence speed. In addition, it is complex for attaining an ideal outcome without using a better search methodology. Chaos is typically a non-linear phenomenon that encompasses features of ergodicity, regularity and randomness as its exquisite internal construction. In this study, MFB-FA was combined with GCM for enhancing the precision and convergence quality of standard FA. Common defects prevailing in FA are discussed below.

The evaluation of the defects of FA in its search process is discussed below.

- (1) Random Initialization: Though random initialization could assure initial FFs are homogeneously distributed in solution space, the solution quality seems to be uncertain as the FFs seem to go far from the global optimum. When initial FFs seem to be not only homogeneously distributed but also with high quality, it would assist in the betterment of the mass of FFs, thereby preventing the approach of getting prematurely trapped in local optima to a certain extent.
- (2) The light-absorption coefficient corresponding to β exists as the constant value. In this case, the β value finds the attraction of all FFs. Generally, $\beta \in [0, 10]$ might be recommended, and it is convenient. Nevertheless, using fixed values for all issues is irrational. The absorption coefficient has to differ with iterations in the searching space. Thus, in this research, β is tuned with the chaotic maps without using a constant value.

In stochastic-searching optimization approaches, the methodologies using chaotic variables rather than random variables are termed chaotic optimization algorithms. In such algorithms, due to ergodicity and non-repeatability of chaos, it could accomplish complete searches at a higher speed than the stochastic searches that rely on probabilities. Thus, in this study, a chaotic map is used for initializing the FF population and tuning the absorption coefficient. The MFB-FA Algorithm 1 is presented below.

Algorithm 1 MFB-FA: Modified Flashing Behaviour-based Firefly Algorithm

```

Input : Size of the population
Output : Best optimal solution (bestsubset)
Begin
set params for Modified_FA();
initpopulation;
Det of light intensity_I();

Objec(fun) = f(q), q = (q1, ..., qd)T
Generate an initial chaotic population of fireflies
qi, i = 1, 2, ... n

compute the 'I' so that it is associated with f(q)

while(T ≤ Max Iteration)
Define absorption coefficient β with chaos Gaussian Map
for m = 1 : i (i ff)
for n = 1 : i (i ff)
if Im > In
move ff towards n;
end if
Vary attractiveness with distance Dist via exp−βDist2
Evaluation of new solutions and updating of light intensity (LI)
end for
end for
Firefly ranking and determination of best solution;
T = T + 1
end while
end

```

3.2. Classification of MANN-AM (Modified Artificial Neural Network with Attention Mechanism)

ANN (Artificial Neural Network) is a term that is used to describe a biologically stimulated sub-area of AI (Artificial Intelligence) and is modelled after a succeeding brain. Usually, ANN is a computational network relying on biological networks that build the structure of the human brain. Identical to the human brain, it possesses neurons interconnected with one another. ANNs also comprise neurons that are associated with one another

in several network layers. Such neurons are termed nodes. The main merits of utilizing ANN involve its parallel processing ability, storing data within the network and the ability to function with partial knowledge. It also possesses memory distribution. It is also fault tolerant. Concurrently, an attention mechanism exists with the idea of permitting a decoder to use relevant input sequences flexibly through weighted integration of all encoded inputs, with related vectors being credited with maximum weights. In ANNs, attention is meant for imitating cognitive attention. The effect improves certain input data parts while alleviating other parts. The motivation for using MANN-AM prevails in the fact that the network has to devote a high focus to less, but significant, data parts. The overall model of MANN-AM is shown in Figure 2, wherein the model performs layer-wise learning.

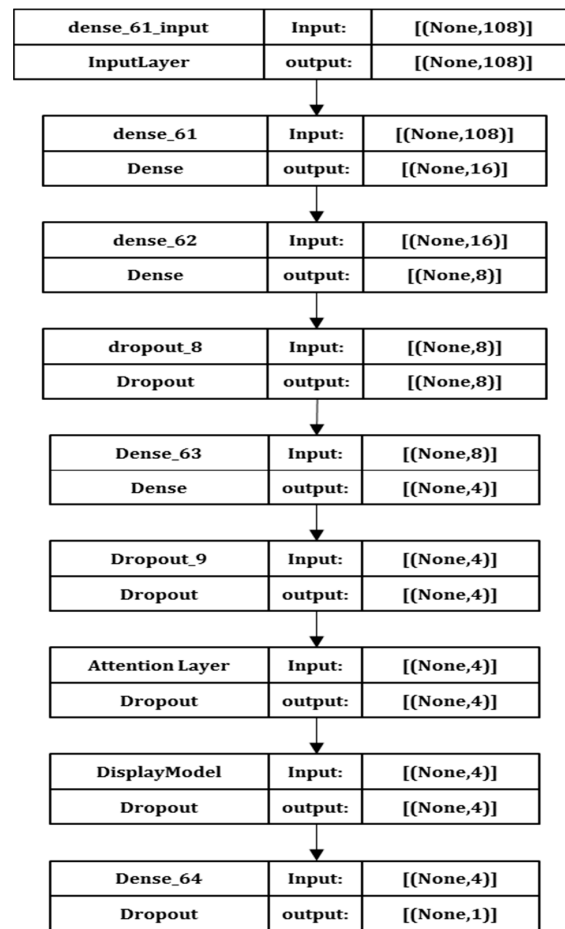


Figure 2. Modified Artificial Neural Network with Attention Mechanism Model.

The proposed model is designed for predicting the absence or existence of UTI. This is undertaken by normalization (Norm), Euclidean distance (d), Nearest Neighbour ($u_i(r)$) and Sigmoid activation-function ($\text{sigm}(x)$).

(i) **Normalization**

Normalization ensures that both negative and positive values considered as input for the subsequent layer exist. This makes learning flexible, and the network considers all the input features to a certain extent while learning. It is given by Equation (7):

$$\text{Norm} = e - \frac{E_{\min}}{E_{\max}} - E_{\min} \quad (7)$$

In Equation (7), e indicates the data, $E_{\min}$ represents the minimum values in each of the columns and $E_{\max}$ represents the maximum values in each of the columns.

(ii) **Euclidean distance**

The Euclidean distance computes the distance amongst the vectors with two real values. It is given by Equation (8):

$$d = \sqrt{\sum_{i=1}^v (P_{1i} - P_{2i})^2} \quad (8)$$

In Equation (8), d represents the distance, v indicates the length and P_{1i} and P_{2i} denote the Euclidean vectors.

(iii) **Nearest-Neighbour**

The Nearest Neighbour is given by Equation (9):

$$u_i(r) = \frac{\sum_{j=1}^K u_{ij} (1/||d(r, r_j)^2||)}{\sum_{j=1}^K (1/||d(r, r_j)^2||)} \quad (9)$$

In Equation (9), u_i indicates the membership-values in the i th class, r_j represents the neighbour residue of (r), r indicates the present residue, which is allocated to the membership values, K denotes the Nearest Neighbour and d represents the distance amongst residues.

Rather than considering all high and low weights, the above equation is utilized for picking only the least Nearest Neighbour that relies on the K value, which affords outcomes effectively in comparison to a conventional classifier.

(iv) **Sigmoid-Activation Function**

The sigmoid-activation function is given by Equation (10):

$$\text{sigm}(x) = \frac{1}{1 + e^{-x}} \quad (10)$$

In Equation (10), e^{-x} indicates the exponential function that is utilized for computing the activation function comprising a small real value ranging from -1 to 1 . The algorithm corresponding to MANN-AM is shown in Algorithm 2.

Algorithm 2 MANN-AM (Modified Artificial Neural Network with Attention Mechanism)

Input : Input : dataset df , x_{train} , x_{test} – independent variable, y_{train} , y_{test} target variable

Output : target y_{preds} with tract infection of UTX_DIAGNOSIS(0, 1)

scale the datasets using // input layer

divide x_{train} , x_{test} and y_{train} , y_{test} //input layer

W is calculated based on dist using //hidden layer

set value for parameter k

Estimating the distance between train and test

Sorting dist in an ascending pattern

select the best k_{neigh}

Repeating steps 2–4 until the algorithm is over

W matrix is saved as result

Simulation involves using //hidden layer

Saving results

Retrieving ANN with attention model

Setting values for number of input, output and hidden layers

Primary weighing of existing neurons in input, output and hidden layers

Calculating the output (y) for each neuron in output layer

Updating ANN parameters

Repeating steps 3–4 until the algorithm is over

Saving results

End of hybrid model

Displaying results

End

The RF algorithm is shown in Algorithm 3.

Algorithm 3 Random Forest

Input : dataset df , x_{train} , x_{test} – independent variable, y_{train} , y_{test} target variable
 output : target y_{preds} with tract infection of UTX_DIAGNOSIS
 A is training set $TS := (a_1, b_1, \dots, (a_n, b_n))$
 training set(TS) = images
 F – feature,
 $N(t)$ = nooftrees
 Forest = E
 begin
 function RandomForest(H, R)
 $I \leftarrow \emptyset$
 for $i \in 1, \dots, E$ do
 $H^{(i)} \leftarrow$ A bootstrap sample from TS
 $a_i \leftarrow$ Randomized Tree Learn ($H^{(i)}, R$)
 $I \leftarrow I \cup (a_i)$
 end for
 return I
 end function
 function RandomizedTreeLearn(H, R)
 At each node :
 $F \leftarrow$ very small subset of R
 split on best feature in F
 Return the learned tree
 end function
 End for
 End

The overall algorithm of K-NN is given in Algorithm 4.

Algorithm 4 K-Nearest Neighbour

Input : dataset df , x_{train} , x_{test} – independent variable, y_{train} , y_{test} target variable
 Input : dataset df , x_{train} , x_{test} – independent variable, y_{train} , y_{test} target variable
 output : target y_{preds} with tract infection of UTX_DIAGNOSIS
 begin
 for each y belong df do

 calculate the Dist $D(x, y)$ bet y and x
 end

 subset(N) from the data frame (df)
 N contains m training samples which are the knn of the test samples x
 cal categorical_{variables} of x :

$$y_x = \arg_{y \in Y} \max \sum_{y \in Y} I(y = \text{class}(\text{categorical}_{\text{variables}}))$$

 end

The NB algorithm is based on the conditional probability principle as per Bayes theorem. The following steps are to be followed (as exposed in Algorithm 5).

Algorithm 5 Naive Bayes classifier

```

Input : dataset df,  $x_{\text{train}}$ ,  $x_{\text{test}}$  – independent variable,  $y_{\text{train}}$ ,  $y_{\text{test}}$  target variable
output : target  $y_{\text{preds}}$  with tract infection of UTX_DIAGNOSIS
Begin
data splitting = 80 : 20
training – set = 80; testing – set = 20 (in %)
training – set
tot(t) = samples in the training set df
prob = probability; and Frequency = fre;
Classi is the class in 'ts'
compute the prob of individual class
 $\text{Prob}(C_j) = \frac{\text{fre}(C_j)}{t}$ 
compute 'm' and 'std'
mean(m) =  $\mu$  and standard deviation(sd) =  $\sigma$ ;
compute m, sd for each features in each class in df and save it.

Testing set
A sample in the testing set df
Probability density function(pdf)
pdf(G) at  $C_j$  for values of features of G. it exist in the df
 $p(G_i|C_j)$ 
Calculating conditional probability((cp)(G)) at  $C_j$ 
result get by below equation
 $P(G|C_j) = \prod_{i=1}^n P(h_i|C_j)$ 
Calculating posterior probability((pp)(G)),  $p(C_j|G)$ 
((pp)(G)),  $p(C_j|G) \rightarrow$  probability of samples at  $C_j$ 
Assign class labels to the class of G by  $\text{maxi}(p) = (C_j|G)$ 
Return class (UTXdiagnosis(1) or UTXdiagnosis(0))
end

```

4. Results and Discussion

The results that have been obtained by the implementation of this work are included in this section, with its dataset description, standard-performance metrics, experimental results, performance and comparative analysis.

4.1. Dataset Description

Predictor variables that include laboratory results, urine dipstick results, urinalysis, past medical history, structure historical findings, physical exam findings, chief complaints and demographic information were available in the dataset that includes a reduced set of 10 variables and a full set of 211 variables. The dataset has been taken from the IoT-based fog environment. The dataset has been collected from https://figshare.com/articles/dataset/Predicting_urinary_tract_infections_in_the_emergency_department_with_machine_learning/5959417?file=10679215 (accessed on 8 March 2018).

The dataset consists of 80,387 attributes and 219 rows, and some of them are given below:

Age: Shows the patient's age.

Sex: Represents the patient's gender.

Diabetes: The presence of diabetes is shown by using a binary output.

Hypertension: This shows the existence of hypertension by using a binary variable.

UTI history: This shows the presence of UTI or not.

Fever: A binary variable shows whether the patient has a fever or not.

Dysuria: A binary output indicates whether the patient has any discomfort or pain during urination.

Urgency: A binary output indicates whether the patient feels the need to urinate urgently or not.

Frequency: A binary output that denotes whether the frequency of the urination is increased over normal times.

Hematuria: A binary output that denotes whether the patient has blood in their urine or not.

WBC_count: This represents the count of the white blood cells, which is measured in cells per microliter.

Bacteria_count: This represents the bacteria count present in the urine of the patient, which is measured in colony-forming units per milliliter.

Nitrite: A binary output that denotes the existence of nitrite content in a patient's urine.

Leukocyte: A binary output that denotes the existence of leukocytes in a patient's urine.

Diagnosis: A binary output that indicates whether the patient has UTI or not.

The above-mentioned attributes are some of the attributes that help in training and testing the ML algorithms in predicting the UTIs by using the medical history and symptoms.

Pre-processing has been carried out to identify and handle the missing values. Categorical data have been encoded with the values 0 and 1 and check for the missing values.

4.2. Performance Metrics

The performance metrics have become a part of each ML. Metrics, namely, precision, accuracy, f1-score, specificity, recall and sensitivity are utilized to analyse the classification models.

4.3. Experimental Results

The experimental results attained when using the proposed method MANN-AM have been shown in Figure 3, in which the red color denotes 0 and the blue color denotes 1. The green line denotes the target value that classifies the target data.

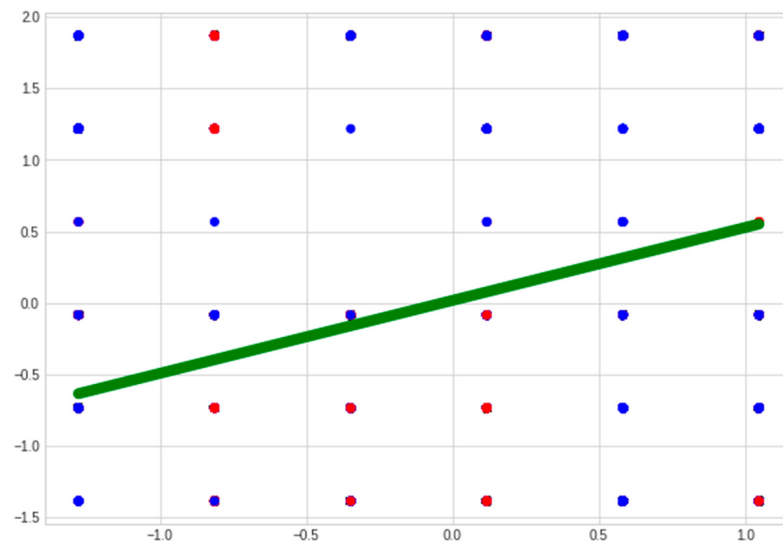


Figure 3. ANN display model.

Figure 4 shows the feature importance of the proposed model based on the number of iterations. From Figure 4, it is clearly found that during the initial iteration, the proposed model has taken a long time for feature selection. After several iterations, the proposed model has learned to select features more efficiently. Selecting the appropriate features is an important step for predicting UTI. Based on the selected features, the proposed model performs a classification by using MANN-AM, which efficiently prognosticates the existence/absence of UTI.

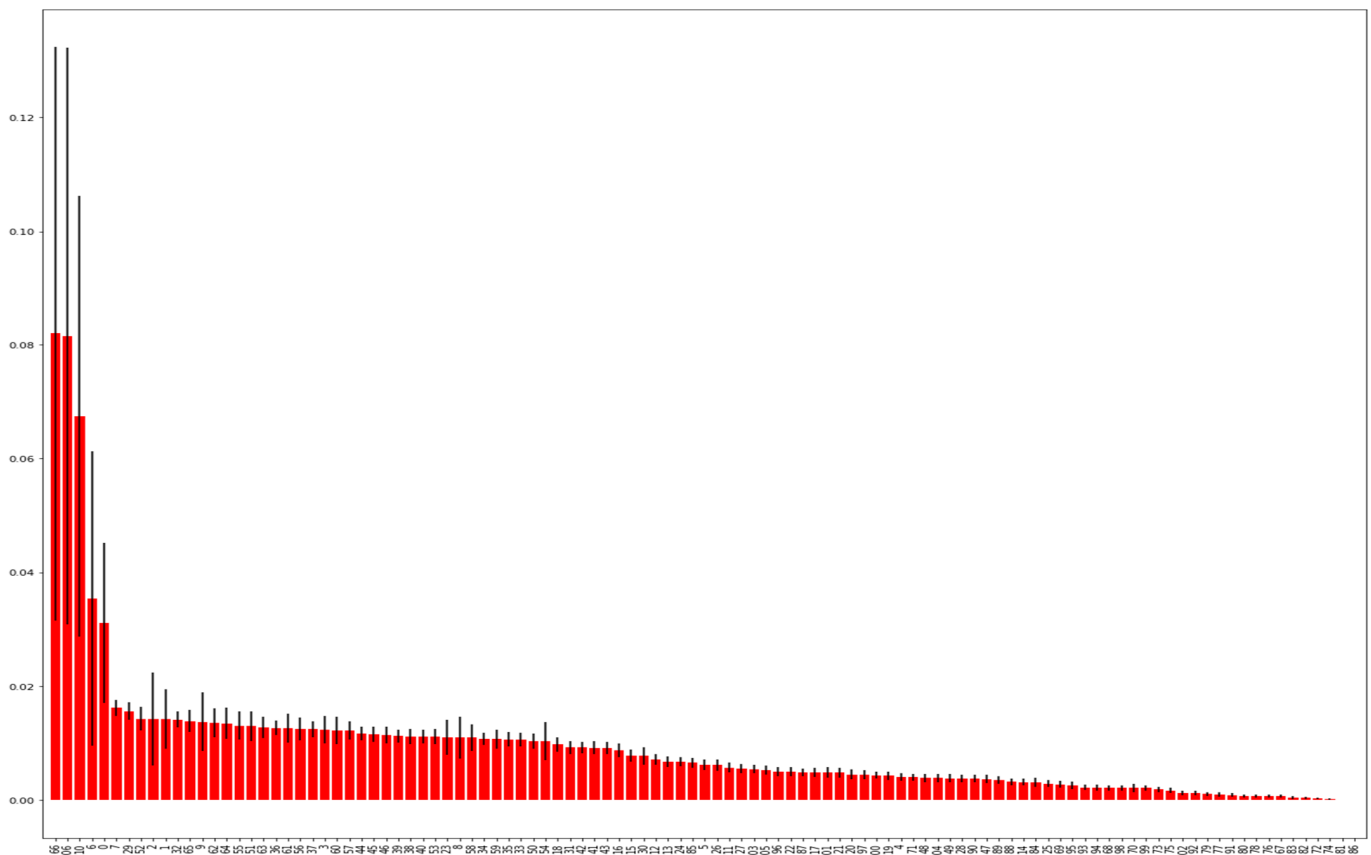


Figure 4. Feature importance of the proposed model.

4.4. Performance Analysis

Performance analysis of the proposed study that has used the proposed MANN-AM classifier, KNN, NB and RF has been discussed; it gives detailed information and helps in a better understanding of the proposed model. Figure 5 shows the AUC curve for the proposed MANN-AM classifier, Naive Bayes, K-Nearest Neighbour and Random Forest, where TPR refers to True Positive Rate and FPR refers to False Positive Rate.

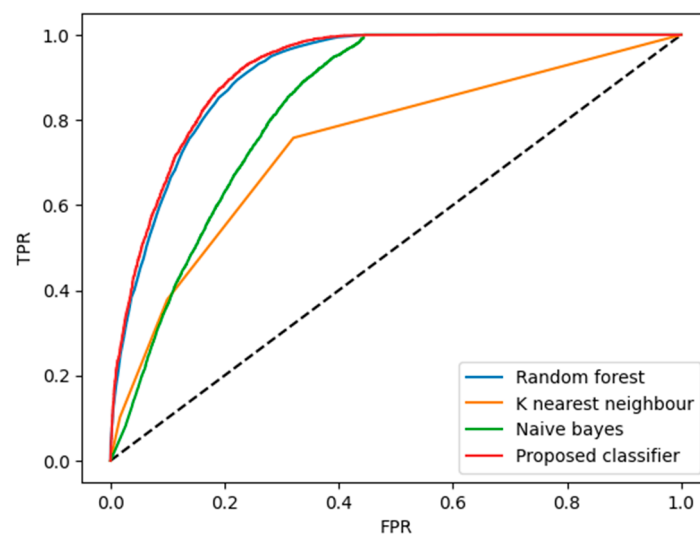


Figure 5. AUC curve (proposed system, Naive Bayes, K-Nearest Neighbour and Random Forest).

AUC shows a collective measure of performance across all possible classifications. From Figure 5, it is observed that the value of AUC for the proposed MANN-AM classifier

has been given as 0.99961, which is closer to 1. For the KNN classifier, the value of AUC has been given as 0.64831. For the Naive Bayes classifier, the value of AUC has been given as 0.77063. For the Random Forest classifier, the value of AUC has been given as 0.72489. Hence, it is revealed that the proposed model has better efficiency in accordance with classification.

The confusion matrix of the proposed system, KNN, NB and RF, is shown in Figures 6–9, respectively. From Figure 6, it is clear that the proposed method has a majority of correct estimates of UTI, which is represented in the diagonal portion (0, 0 and 1, 1). Though it estimates a few wrong predictions, which are represented in black color, when comparing, the proposed method has more accurate estimates than the wrong ones. Hence, the proposed method is more efficient. From Figure 7, it is clear that the KNN classifier has 11,406 true positives, 1257 true negatives, 2063 false negatives and 1352 false positives. The performance analysis of the Naive Bayes classifier has been shown in Figure 10 by using the AUC curve.

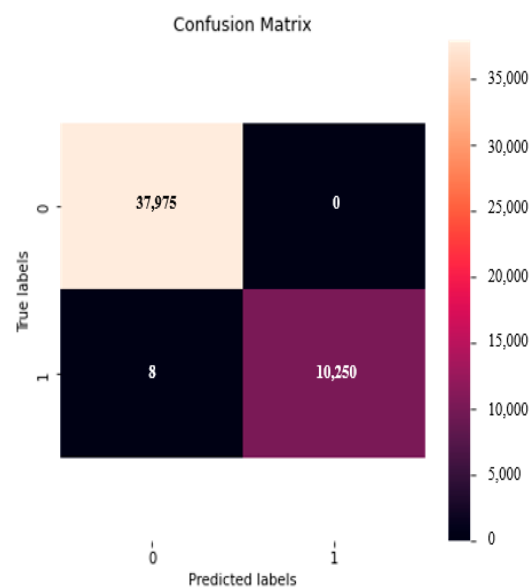


Figure 6. Confusion matrix (proposed system).

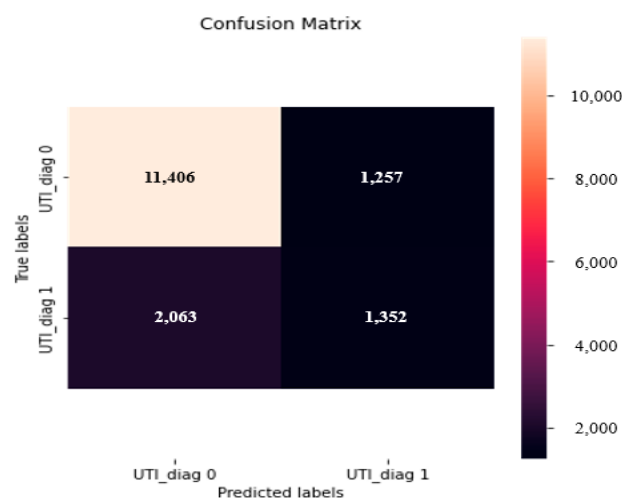


Figure 7. Confusion matrix for the KNN classifier.

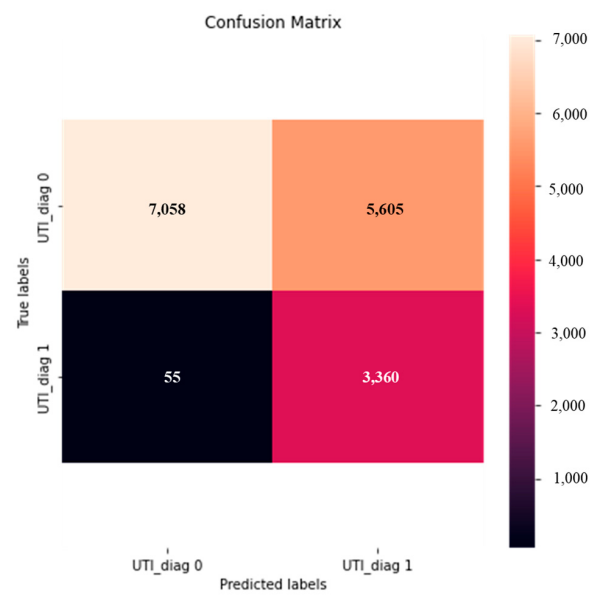


Figure 8. Confusion matrix for the Naive Bayes classifier.

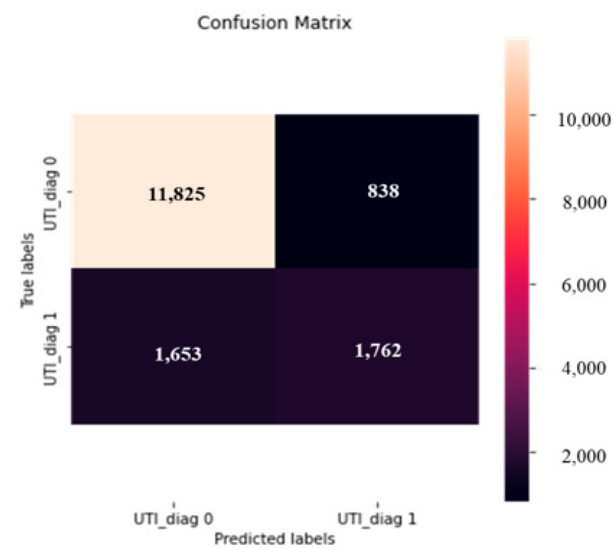


Figure 9. Confusion matrix for the Random Forest classifier.

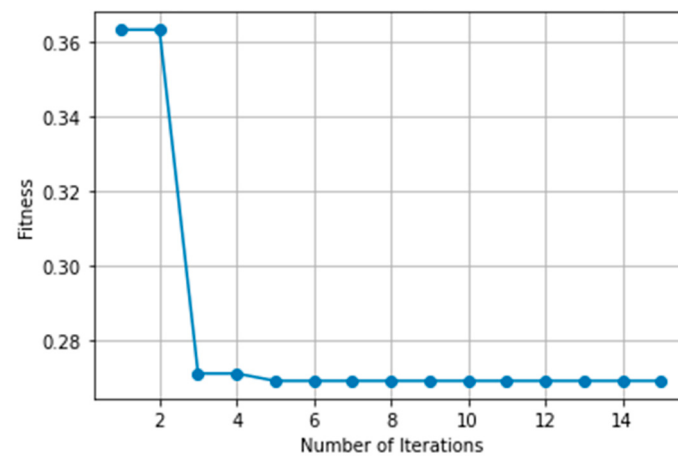


Figure 10. Firefly Algorithm feature selection.

From Figure 8, it is clear that the Naive Bayes classifier has a majority of correct estimates of UTI, which are 7058 true positives, 5605 true negatives, 55 false negatives and 3360 false positives. From Figure 9, it is clear that the Random Forest classifier has 11,825 true positives, 838 true negatives, 1653 false negatives and 1762 false positives.

The Modified Flashing Behaviour-based Firefly Algorithm using a Gaussian Map has been shown in Figure 10.

During the initial iterations, the proposed model attained a higher fitness value. When the number of iterations increased, the proposed model learned to select features more efficiently and steadily, which is shown in Figure 10.

4.5. Performance Analysis of Classifiers

The proposed model Modified ANN-AM has been compared internally with KNN, Naive Bayes and Random Forest classifiers to show that the proposed model has improved efficiency among other classifiers. The performance analysis of Modified ANN-AM, NB, KNN and RF is shown in Table 1 and Figure 11.

Table 1. Performance analysis of Modified ANN-AM, NB, KNN and RF.

	Accuracy	Precision	Recall	F1-Score	Sensitivity	Specificity
Modified ANN with AM	0.99	1	0.99	0.99	1	0.99
Naive Bayes Classifier	0.65	0.68	0.77	0.63	0.55	0.98
KNN Classifier	0.79	0.68	0.65	0.66	0.9	0.39
Random Forest Classifier	0.85	0.78	0.72	0.75	0.93	0.51

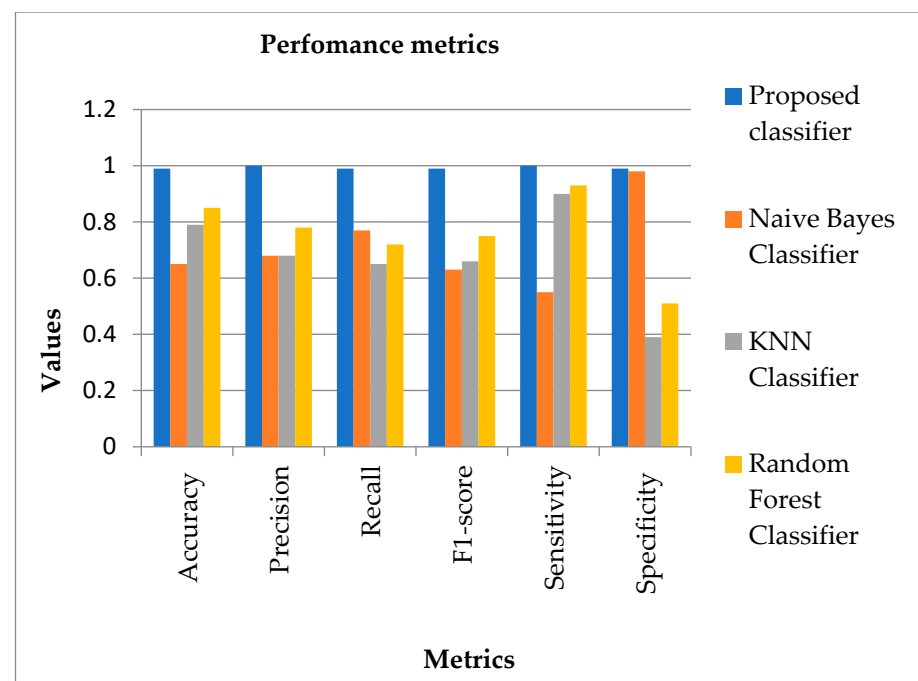


Figure 11. Performance analysis of Modified ANN-AM, NB, KNN and RF.

The proposed model has been analysed by using six performance metrics, and the values are given in Table 1. The proposed model has an accuracy of 0.99, recall of 0.99, precision of 1, f1-score of 0.99, specificity of 0.99 and sensitivity of 1. The Naive Bayes has an accuracy of 0.65, recall of 0.77, precision of 0.68, F1-score of 0.63, specificity of 0.98 and

sensitivity of 0.55. The KNN classifier has an accuracy of 0.79, recall of 0.65, precision of 0.68, f1-score of 0.66, specificity of 0.39 and sensitivity of 0.9. The Random Forest classifier has an accuracy of 0.85, recall of 0.72, precision of 0.78, f1-score of 0.75, specificity of 0.51 and sensitivity of 0.93.

From Table 1 and Figure 11, it is clear that the proposed model of Modified ANN with AM is more efficient and has better values in all performance metrics. Hence, it can be concluded that the proposed model shows an enhanced performance in UTI detection.

Table 2 shows the feature ranking by Firefly Algorithm, in which chloride refers to the chloride content that exists in urine, ua_ketones refers to the ketones level in urine, ua_clarity refers to the clarity level of the urine, ua_color refers to the color of the urine, ua_glucose refers to the glucose level in urine, Temp_Mean refers to the body temperature of the patient and the fever attribute shows whether the patient has a fever or not.

Table 2. Feature ranking by Firefly Algorithm.

Feature Importance	Feature Ranking by Firefly Algorithm
Chloride	0.089883
HR_Min	0.074717
Chloride	0.072249
ua_ketones	0.033614
ua_clarity	0.031833
ID	0.015908
ua_color	0.015579
Chloride	0.014618
fever	0.014009
lang	0.013988
abd_distended2	0.01389
ua_glucose	0.013608
Temp_Mean	0.013354
HR_First	0.013126
UCX_abnormal	0.013082

By using the proposed method, the feature ranking for chloride, HR_Min and chloride again are the three most important features for predicting UTI, and the feature importance scores are given to be 0.089883, 0.074717 and 0.072249, respectively. The other features such as ua_ketones, ua_clarity, ID, ua_color, fever, lang, abd_distended2, ua_glucose, Temp_Mean, HR_First and UCX_abnormal have lower feature importance scores, which indicates that they have a relatively lower contribution in the prediction of UTI. In Table 2, it is observed that chloride has a higher ranking of 0.089883.

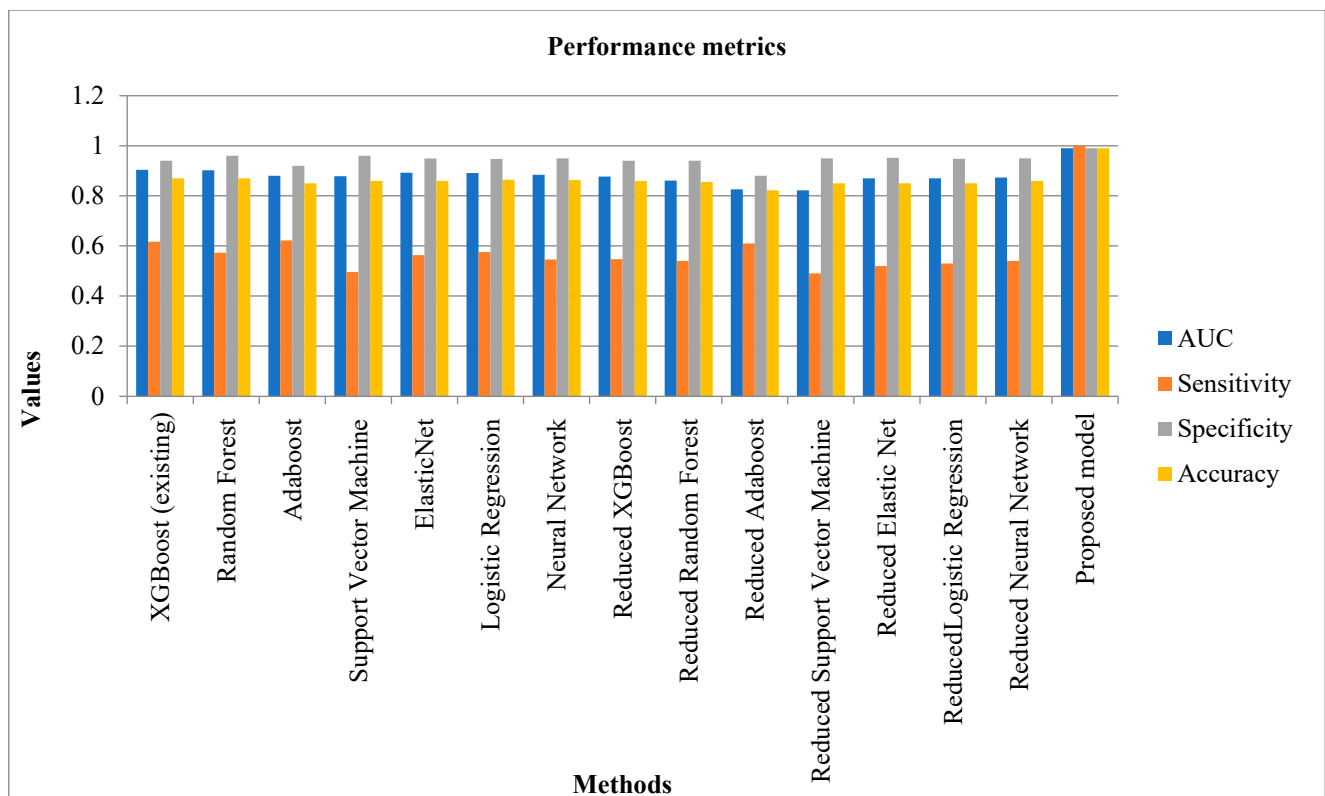
4.6. Comparative Analysis

The proposed model has been compared with two traditional studies and the respective results have been discussed. The dataset of the proposed study has used an IoT-based fog environment, hence the existing studies which have similar types of datasets have been considered in the comparative analysis. Table 3 shows the comparative analysis of various models.

Table 3. Comparison of performance metrics of classifiers [64].

Models	AUC	Sensitivity	Specificity	Accuracy
XGBoost (existing) [64]	0.9	0.617	0.94	0.87
Random Forest	0.9	0.573	0.96	0.87
Adaboost	0.88	0.622	0.92	0.85
Support Vector Machine	0.88	0.496	0.96	0.86
ElasticNet	0.89	0.563	0.949	0.86
Logistic Regression	0.89	0.575	0.947	0.864
Neural Network	0.88	0.546	0.95	0.863
Reduced XGBoost	0.88	0.547	0.94	0.859
Reduced Random Forest	0.86	0.54	0.94	0.855
Reduced Adaboost	0.83	0.61	0.88	0.822
Reduced Support Vector Machine	0.82	0.49	0.95	0.85
Reduced Elastic Net	0.87	0.52	0.952	0.85
Reduced Logistic Regression	0.87	0.53	0.948	0.85
Reduced Neural Network	0.87	0.54	0.95	0.859
Proposed model	0.99	1	0.99	0.99

For extensive analysis, the proposed model has been compared with the existing study [64], which has a dataset similar to the proposed study and used various classification models. In the existing methods, XGBoost has a higher accuracy of 0.9 compared to other models. In Table 3 and Figure 12 it is revealed that the proposed model possesses ideal values in every performance metric and has attained the highest accuracy of 0.99.

**Figure 12.** Comparison of performance metrics of classifiers [64].

The comparative analysis of performance metrics of the proposed model with the existing study [64] is shown in Figure 12. Table 4 shows the comparative analysis of various models in the existing study [32].

Table 4. Comparison of performance metrics of classifiers [32].

Model	Accuracy	Precision	F1-Score	Specificity	Sensitivity
XGBoost (existing) [32]	0.9145	0.9549	0.9012	0.9596	0.8479
Random Forest	0.8149	0.886	0.8126	0.8945	0.7446
MLP	0.8395	0.8546	0.8478	0.8446	0.8349
K-NN	0.7718	0.9125	0.7508	0.9456	0.6256
SVM	0.8515	0.8945	0.8564	0.8945	0.8143
Proposed system	0.99	1	0.99	0.99	1

For broad analysis, the comparative analysis has been done with the existing study [32], which has a dataset similar to the proposed study and used various classification models.

The comparative analysis of the performance metrics of the proposed model has been shown in Figure 13. In the existing methods, XGBoost has a higher accuracy of 0.9145 compared to other models. In Table 4 and Figure 13, it is revealed that the proposed system has better values in every performance metric and has attained the highest accuracy of 0.99. The proposed MANN-AM classifier has greater potential in focusing on small and significant data. The MFB-FA model for feature selection has helped in attaining a global optimum. Hence, from the experimental results, performance analysis, internal analysis and comparative analysis, it is revealed that the proposed study has better efficacy in every performance metric in detecting UTI.

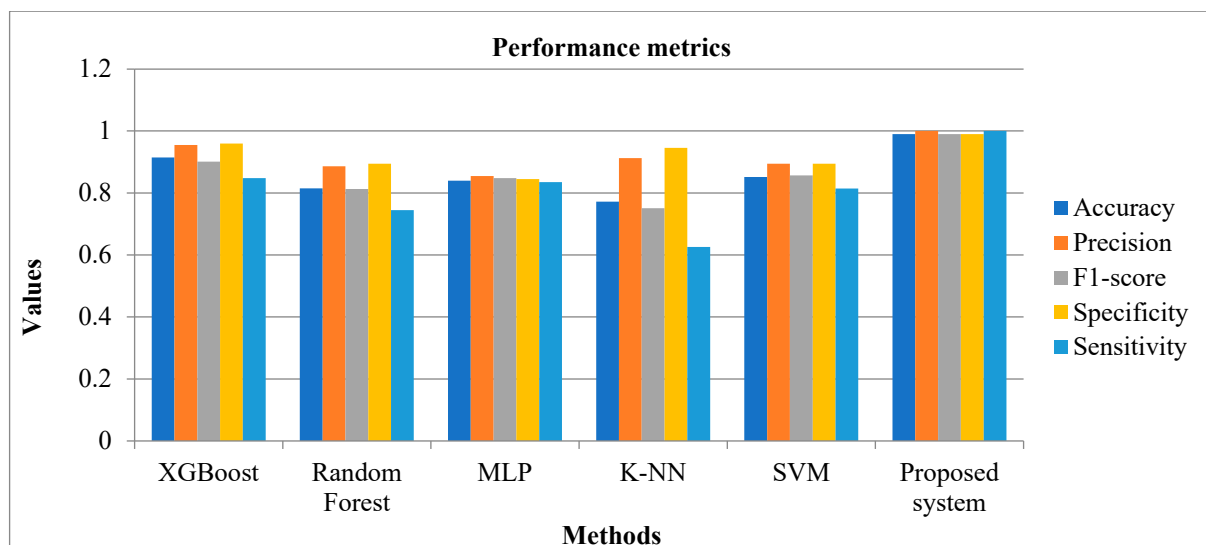


Figure 13. Comparison of performance metrics of classifiers [32].

5. Conclusions

This research aimed to identify the existence or absence of UTIs using effective ML-based algorithms. This study used MFB-FA for feature selection and MANN-AM for classification. Significant features chosen by the MFB-FA were also presented to afford an idea about the features selected based on ranking. Experimental results were presented to reveal the outcomes of the proposed system in dividing the target value, thereby classifying the data. To evaluate the performance of this work, an internal comparison was performed with ML algorithms, namely, NB, RF and KNN. The accuracy rate of the proposed system was found to be 0.99, NB exposed 0.65 as the accuracy rate, RF showed 0.85 as the accuracy rate and KNN explored 0.79 as the accuracy rate. From the internal comparison, the

proposed methodology exposed a higher accuracy rate than the other three algorithms. Further, a confusion matrix was attained to showcase the correct classification rates of the proposed system. In addition, for evaluating the efficacy of the proposed system, a comparison was performed with two conventional studies. The comparative analysis outcomes confirmed the proposed work was more effective than existing studies, with 0.99 as the accuracy rate. Due to the high prediction rate, the study possesses the ability to be applied in real-time. Moreover, in the future, various other hybrid ML algorithms could be regarded to enhance the prediction rate. Additionally, the proposed system can be deployed in real-world scenarios such as public restrooms, hospitals and offices, which will increase the effectiveness and feasibility of diagnosing UTIs.

Author Contributions: Conceptualization, A.A., S.A., A.B., M.S., A.G. and Y.-D.Z.; methodology, A.A., S.A., A.B., M.S., A.G. and Y.-D.Z.; software, A.A., S.A., A.B., M.S., A.G. and Y.-D.Z.; validation, A.A., S.A., A.B., M.S., A.G. and Y.-D.Z.; formal analysis, A.A., S.A., A.B., M.S., A.G. and Y.-D.Z.; investigation, A.A., S.A., A.B., M.S., A.G. and Y.-D.Z.; resources, A.A., S.A., A.B., M.S., A.G. and Y.-D.Z.; data curation, A.A., S.A., A.B., M.S., A.G. and Y.-D.Z.; writing—original draft preparation, A.A., S.A., A.B., M.S., A.G. and Y.-D.Z.; writing—review and editing, A.A., S.A., A.B., M.S., A.G. and Y.-D.Z.; visualization, A.A., S.A., A.B., M.S., A.G. and Y.-D.Z.; supervision, A.A., S.A., A.B., M.S., A.G. and Y.-D.Z.; project administration, A.A., S.A. and A.B.; funding acquisition, A.A., S.A., A.B., M.S. and A.G. All authors have read and agreed to the published version of the manuscript.

Funding: The authors extend their appreciation to the Deputyship for Research and Innovation, Ministry of Education in Saudi Arabia, for funding this research work through the project number (IF2/PSAU/2022/01/22929).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bansal, M.; Sirpal, V.; Choudhary, M.K. Advancing e-Government using Internet of Things. In *Mobile Computing and Sustainable Informatics*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 123–137.
2. Lin, Y.W.; Lin, Y.B.; Liu, C.Y. Altalk: A tutorial to implement AI as IoT devices. *IET Netw.* **2019**, *8*, 195–202. [\[CrossRef\]](#)
3. Haji, L.M.; Ahmad, O.M.; Zeebaree, S.; Dino, H.I.; Zebari, R.R.; Shukur, H.M. Impact of cloud computing and internet of things on the future internet. *Technol. Rep. Kansai Univ.* **2020**, *62*, 2179–2190.
4. Aceto, G.; Persico, V.; Pescapé, A. Industry 4.0 and health: Internet of things, big data, and cloud computing for healthcare 4.0. *J. Ind. Inf. Integr.* **2020**, *18*, 100129. [\[CrossRef\]](#)
5. Qian, Y.; Wu, D.; Bao, W.; Lorenz, P. The Internet of Things for Smart Cities: Technologies and Applications. *IEEE Netw.* **2019**, *33*, 4–5. [\[CrossRef\]](#)
6. Dhanvijay, M.M.; Patil, S.C. Internet of Things: A survey of enabling technologies in healthcare and its applications. *Comput. Networks* **2019**, *153*, 113–131. [\[CrossRef\]](#)
7. Özdemir, V. The Big Picture on the “AI Turn” for Digital Health: The Internet of Things and Cyber-Physical Systems. *OMICS J. Integr. Biol.* **2019**, *23*, 308–311. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Khan, I.H.; Javaid, M. Role of Internet of Things (IoT) in Adoption of Industry 4.0. *J. Ind. Integr. Manag.* **2022**, *7*, 515–533. [\[CrossRef\]](#)
9. Mutlag, A.A.; Ghani, M.K.A.; Arunkumar, N.; Mohammed, M.A.; Mohd, O. Enabling technologies for fog computing in healthcare IoT systems. *Futur. Gener. Comput. Syst.* **2019**, *90*, 62–78. [\[CrossRef\]](#)
10. Alshamrani, M. IoT and artificial intelligence implementations for remote healthcare monitoring systems: A survey. *J. King Saud Univ.-Comput. Inf. Sci.* **2022**, *34*, 4687–4701. [\[CrossRef\]](#)
11. Zhu, C.; Xu, Z.; Gu, Y.; Zheng, S.; Sun, X.; Cao, J.; Song, B.; Jin, J.; Liu, Y.; Wen, X.; et al. Prediction of post-stroke urinary tract infection risk in immobile patients using machine learning: An observational cohort study. *J. Hosp. Infect.* **2022**, *122*, 96–107. [\[CrossRef\]](#)
12. Gadalla, A.A.H.; Friberg, I.M.; Kift-Morgan, A.; Zhang, J.; Eberl, M.; Topley, N.; Weeks, I.; Cuff, S.; Wootton, M.; Gal, M.; et al. Identification of clinical and urine biomarkers for uncomplicated urinary tract infection using machine learning algorithms. *Sci. Rep.* **2019**, *9*, 19694. [\[CrossRef\]](#) [\[PubMed\]](#)

13. Enshaeifar, S.; Zoha, A.; Skillman, S.; Markides, A.; Acton, S.T.; Elsaleh, T.; Kenny, M.; Rostill, H.; Nilforooshan, R.; Barnaghi, P. Machine learning methods for detecting urinary tract infection and analysing daily living activities in people with dementia. *PLoS ONE* **2019**, *14*, e0209909. [\[CrossRef\]](#) [\[PubMed\]](#)
14. de Llano, D.G.; Moreno-Arribas, M.V.; Bartolomé, B. Cranberry Polyphenols and Prevention against Urinary Tract Infections: Relevant Considerations. *Molecules* **2020**, *25*, 3523. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Arienzo, A.; Cellitti, V.; Ferrante, V.; Losito, F.; Stalio, O.; Murgia, L.; Marino, R.; Cristofano, F.; Orrù, M.; Visca, P.; et al. A new point-of-care test for the rapid detection of urinary tract infections. *Eur. J. Clin. Microbiol. Infect. Dis.* **2020**, *39*, 325–332. [\[CrossRef\]](#)
16. Yelin, I.; Snitser, O.; Novich, G.; Katz, R.; Tal, O.; Parizade, M.; Chodick, G.; Koren, G.; Shalev, V.; Kishony, R. Personal clinical history predicts antibiotic resistance of urinary tract infections. *Nat. Med.* **2019**, *25*, 1143–1152. [\[CrossRef\]](#)
17. Bhatia, M.; Kaur, S.; Sood, S.K. IoT-inspired smart toilet system for home-based urine infection prediction. *ACM Trans. Comput. Healthc.* **2020**, *1*, 1–25. [\[CrossRef\]](#)
18. MacDermott, A.; Kendrick, P.; Idowu, I.; Ashall, M.; Shi, Q. Securing Things in the Healthcare Internet of Things. In Proceedings of the 2019 Global IoT Summit (GIoTS), Aarhus, Denmark, 17–21 June 2019; pp. 1–6. [\[CrossRef\]](#)
19. Balakrishna, S.; Thirumaran, M.; Solanki, V.K. IoT Sensor Data Integration in Healthcare using Semantics and Machine Learning Approaches. In *A Handbook of Internet of Things in Biomedical and Cyber Physical System*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 275–300.
20. Tareq, R.W.; Khaleel, T.A. Implementation of MQTT Protocol in Health Care Based on IoT Systems: A Study. *Int. J. Electr. Comput. Eng. Syst.* **2021**, *12*, 215–223. [\[CrossRef\]](#)
21. Gohar, M.; Anwar, S.; Ali, M.; Choi, J.-G.; Alquhayz, H.; Koh, S.-J. Partial bicasting with buffering for proxy mobile IPV6 mobility management in CoAP-based IoT networks. *Electronics* **2020**, *9*, 598. [\[CrossRef\]](#)
22. Islam, M.; Nooruddin, S.; Karray, F.; Muhammad, G. Internet of Things: Device Capabilities, Architectures, Protocols, and Smart Applications in Healthcare Domain. *IEEE Internet Things J.* **2022**, *10*, 3611–3641. [\[CrossRef\]](#)
23. Ijaz, M.; Li, G.; Lin, L.; Cheikhrouhou, O.; Hamam, H.; Noor, A. Integration and Applications of Fog Computing and Cloud Computing Based on the Internet of Things for Provision of Healthcare Services at Home. *Electronics* **2021**, *10*, 1077. [\[CrossRef\]](#)
24. Nayak, J.; Naik, B.; Dinesh, P.; Vakula, K.; Dash, P.B. Firefly Algorithm in Biomedical and Health Care: Advances, Issues and Challenges. *SN Comput. Sci.* **2020**, *1*, 311. [\[CrossRef\]](#)
25. Nemnich, M.A.; Debbat, F.; Slimane, M. Hybridizing Bees Algorithm with Firefly Algorithm for Solving Complex Continuous Functions. *Int. J. Appl. Metaheuristic Comput. (IJAMC)* **2020**, *11*, 27–55. [\[CrossRef\]](#)
26. Kaveh, A.; Javadi, S. Chaos-based firefly algorithms for optimization of cyclically large-size braced steel domes with multiple frequency constraints. *Comput. Struct.* **2019**, *214*, 28–39. [\[CrossRef\]](#)
27. Li, X.; Zhou, K.; Wang, J.; Guo, J.; Cao, Y.; Ren, J.; Guan, T.; Sheng, W.; Zhang, M.; Yao, Z.; et al. Diagnostic Value of the Fimbriae Distribution Pattern in Localization of Urinary Tract Infection. *Front. Med.* **2021**, *8*, 589409. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Gupta, A.; Singh, A. Early Urine Infection Prediction Framework using XGBoost Ensemble Model in IoT-Fog Environment. *Res. Sq.* **2022**. [\[CrossRef\]](#)
29. Bhatia, M.; Kaur, S.; Sood, S.K. IoT-inspired smart home based urine infection prediction. *J. Ambient. Intell. Humaniz. Comput.* **2020**, *1*–15. [\[CrossRef\]](#)
30. Bafna, P.; Deepanjali, S.; Mandal, J.; Balamurugan, N.; Swaminathan, R.P.; Kadiravan, T. Reevaluating the true diagnostic accuracy of dipstick tests to diagnose urinary tract infection using Bayesian latent class analysis. *PLoS ONE* **2020**, *15*, e0244870. [\[CrossRef\]](#)
31. Yang, Q.; Tian, G.-L.; Qin, J.-W.; Wu, B.-Q.; Tan, L.; Xu, L.; Wu, S.-Z.; Yang, J.-T.; Jiang, J.-H.; Yu, R.-Q. Coupling bootstrap with synergy self-organizing map-based orthogonal partial least squares discriminant analysis: Stable metabolic biomarker selection for inherited metabolic diseases. *Talanta* **2020**, *219*, 121370. [\[CrossRef\]](#)
32. Chadaga, K.; Prabhu, S.; Bhat, V.; Sampathila, N.; Umakanth, S.; Chadaga, R. A Decision Support System for Diagnosis of COVID-19 from Non-COVID-19 Influenza-like Illness Using Explainable Artificial Intelligence. *Bioengineering* **2023**, *10*, 439. [\[CrossRef\]](#)
33. Khanna, V.V.; Chadaga, K.; Sampathila, N.; Prabhu, S.; Bhandage, V.; Hegde, G.K. A Distinctive Explainable Machine Learning Framework for Detection of Polycystic Ovary Syndrome. *Appl. Syst. Innov.* **2023**, *6*, 32. [\[CrossRef\]](#)
34. Pradhan, A.; Prabhu, S.; Chadaga, K.; Sengupta, S.; Nath, G. Supervised Learning Models for the Preliminary Detection of COVID-19 in Patients Using Demographic and Epidemiological Parameters. *Information* **2022**, *13*, 330. [\[CrossRef\]](#)
35. Chadaga, K.; Prabhu, S.; Sampathila, N.; Chadaga, R.; Swathi, S.K.; Sengupta, S. Predicting cervical cancer biopsy results using demographic and epidemiological parameters: A custom stacked ensemble machine learning approach. *Cogent Eng.* **2022**, *9*, 2143040. [\[CrossRef\]](#)
36. Pong, Y.-H.; Tsai, V.F.; Hsu, Y.-H.; Lee, C.-H.; Wang, K.-C.; Tsai, Y.-T. Application of a Deep Learning Neural Network for Voiding Dysfunction Diagnosis Using a Vibration Sensor. *Appl. Sci.* **2022**, *12*, 7216. [\[CrossRef\]](#)
37. Liu, X.; Sahidullah; Kinnunen, T. A Comparative Re-Assessment of Feature Extractors for Deep Speaker Embeddings. *Interspeech* **2020**, *2020*, 3221–3225. [\[CrossRef\]](#)
38. Vadde, V.; Muralidharan, B.; Sharma, A. Orthogonal Spin Current Injected Magnetic Tunnel Junction for Convolutional Neural Networks. *arXiv* **2022**, arXiv:2207.14603.

39. Hong, Y.-W.; Leu, J.-S.; Faisal, M.; Prakosa, S.W. Analysis of Model Compression Using Knowledge Distillation. *IEEE Access* **2022**, *10*, 85095–85105. [\[CrossRef\]](#)
40. Corkins, M.E.; Achieng, M.; DeLay, B.D.; Krneta-Stankic, V.; Cain, M.P.; Walker, B.L.; Chen, J.; Lindström, N.O.; Miller, R.K. A comparative study of cellular di-versity between the *Xenopus* pronephric and mouse metanephric nephron. *Kidney Int.* **2022**, *103*, 77–86. [\[CrossRef\]](#)
41. Iriya, R. Computer Vision Methods for Urinary Tract Infection Diagnostics. Ph.D. Thesis, Arizona State University, Tempe, AZ, USA, 2020.
42. Li, H.; Rezvani, R.; Kolanko, M.A.; Sharp, D.J.; Wairagkar, M.; Vaidyanathan, R.; Nilforooshan, R.; Barnaghi, P. An attention model to analyse the risk of agitation and urinary tract infections in people with dementia. *arXiv* **2021**, arXiv:2101.07007.
43. Bijlani, N.; Maldonado, O.M.; Kouchaki, S. G-CMP: Graph-enhanced Contextual Matrix Profile for unsupervised anomaly detection in sensor-based remote health monitoring. *arXiv* **2022**, arXiv:2211.16122.
44. Chen, M.; Yang, J.; Lu, J.; Zhou, Z.; Huang, K.; Zhang, S.; Yuan, G.; Zhang, Q.; Li, Z. Ureteral calculi lithotripsy for single ureteral calculi: Can DNN-assisted model help preoperatively predict risk factors for sepsis? *Eur. Radiol.* **2022**, *32*, 8540–8549. [\[CrossRef\]](#)
45. Chawla, K.; Bankapur, A.; Acharya, M.; D’Souza, J.S.; Chidangil, S. A micro-Raman and chemometric study of urinary tract infection-causing bacterial pathogens in mixed cultures. *Anal. Bioanal. Chem.* **2019**, *411*, 3165–3177.
46. Lee, A.C.; Mullany, L.C.; Koffi, A.K.; Rafiquallah, I.; Khanam, R.; Folger, L.V.; Rahman, M.; Mitra, D.K.; Labrique, A.; Christian, P.; et al. Urinary tract infections in pregnancy in a rural population of Bangladesh: Population-based prevalence, risk factors, etiology, and antibiotic resistance. *BMC Pregnancy Childbirth* **2020**, *20*, 1. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Tandan, M.; Timilsina, M.; Cormican, M.; Vellinga, A. Role of patient descriptors in predicting antimicrobial resistance in urinary tract infections using a decision tree approach: A retrospective cohort study. *Int. J. Med. Inform.* **2019**, *127*, 127–133. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Fu, Q.; Zhang, Y.; Wang, P.; Pi, J.; Qiu, X.; Guo, Z.; Huang, Y.; Zhao, Y.; Li, S.; Xu, J. Rapid identification of the resistance of urinary tract pathogenic bacteria using deep learning-based spectroscopic analysis. *Anal. Bioanal. Chem.* **2021**, *413*, 7401–7410. [\[CrossRef\]](#) [\[PubMed\]](#)
49. Li, H.; Kolanko, M.A.; Enshaeifar, S.; Skillman, S.; Markides, A.; Kenny, M.; Soreq, E.; Kouchaki, S.; Jensen, K.; Cameron, L.; et al. Deep Representation for Connected Health: Semi-supervised Learning for Analysing the Risk of Urinary Tract Infections in People with Dementia. *arXiv* **2020**, arXiv:2011.13916.
50. Colborn, K.L.; Bronsert, M.; Hammermeister, K.; Henderson, W.G.; Singh, A.B.; Meguid, R.A. Identification of urinary tract infections using electronic health record data. *Am. J. Infect. Control.* **2019**, *47*, 371–375. [\[CrossRef\]](#)
51. Checcucci, E.; De Cillis, S.; Granato, S.; Chang, P.; Afyouni, A.S.; Okhunov, Z.; Uro-technology and SoMe Working Group of the Young Academic Urologists Working Party of the European Association of Urology. Applications of neural networks in urology: A systematic review. *Curr. Opin. Urol.* **2020**, *30*, 788–807. [\[CrossRef\]](#)
52. Burton, R.J.; Albur, M.; Eberl, M.; Cuff, S.M. Using artificial intelligence to reduce diagnostic workload without compro-mising detection of urinary tract infections. *BMC Med. Inform. Decis. Mak.* **2019**, *19*, 171. [\[CrossRef\]](#)
53. Mancini, A.; Vito, L.; Marcelli, E.; Piangerelli, M.; De Leone, R.; Pucciarelli, S.; Merelli, E. Machine learning models predicting multidrug resistant urinary tract infections using “DsaaS”. *BMC Bioinform.* **2020**, *21*, 347. [\[CrossRef\]](#)
54. Advanced Analytics Group of Pediatric Urology; ORC Personalized Medicine Group. Targeted workup after initial febrile urinary tract infection: Using a novel machine learning model to identify children most likely to benefit from voiding cystourethrogram. *J. Urol.* **2019**, *202*, 144–152. [\[CrossRef\]](#)
55. Park, J.I.; Bliss, D.Z.; Chi, C.-L.; Delaney, C.W.; Westra, B.L. Knowledge discovery with machine learning for hospi-tal-acquired catheter-associated urinary tract infections. *CIN: Comput. Inform. Nurs.* **2020**, *38*, 28–35. [\[PubMed\]](#)
56. Ramgopal, S.; Horvat, C.M.; Yanamala, N.; Alpern, E.R. Machine Learning To Predict Serious Bacterial Infections in Young Febrile Infants. *Pediatrics* **2020**, *146*, e20194096. [\[CrossRef\]](#) [\[PubMed\]](#)
57. Bronsert, M.; Singh, A.B.; Henderson, W.G.; Hammermeister, K.; Meguid, R.A.; Colborn, K.L. Identification of postoperative complications using electronic health record data and machine learning. *Am. J. Surg.* **2020**, *220*, 114–119. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Han, S.S.; Azad, T.D.; Suarez, P.A.; Ratliff, J.K. A machine learning approach for predictive models of adverse events following spine surgery. *Spine J.* **2019**, *19*, 1772–1781. [\[CrossRef\]](#) [\[PubMed\]](#)
59. Didelot, X.; Pouwels, K.B. Machine-learning-assisted selection of antibiotic prescription. *Nat. Med.* **2019**, *25*, 1033–1034. [\[CrossRef\]](#) [\[PubMed\]](#)
60. Sharaha, U.; Rodriguez-Diaz, E.; Sagi, O.; Riesenber, K.; Salman, A.; Bigio, I.J.; Huleihel, M. Fast and reliable determination of *Escherichia coli* susceptibility to antibiotics: Infrared microscopy in tandem with machine learning algorithms. *J. Biophotonics* **2019**, *12*, e201800478. [\[CrossRef\]](#) [\[PubMed\]](#)
61. Papagiannopoulou, C.; Parchen, R.; Rubbens, P.; Waegeman, W. Fast Pathogen Identification Using Single-Cell Matrix-Assisted Laser Desorption/Ionization-Aerosol Time-of-Flight Mass Spectrometry Data and Deep Learning Methods. *Anal. Chem.* **2020**, *92*, 7523–7531. [\[CrossRef\]](#) [\[PubMed\]](#)
62. Sharaha, U.; Rodriguez-Diaz, E.; Sagi, O.; Riesenber, K.; Lapidot, I.; Segal, Y.; Bigio, I.J.; Huleihel, M.; Salman, A. Detection of Extended-Spectrum β -Lactamase-Producing *Escherichia coli* Using Infrared Microscopy and Machine-Learning Algorithms. *Anal. Chem.* **2019**, *91*, 2525–2530. [\[CrossRef\]](#)

63. Fatima, N.; Zheng, H.; Massaad, E.; Hadzipasic, M.; Shankar, G.M.; Shin, J.H. Development and Validation of Machine Learning Algorithms for Predicting Adverse Events After Surgery for Lumbar Degenerative Spondylolisthesis. *World Neurosurg.* **2020**, *140*, 627–641. [[CrossRef](#)]
64. Taylor, R.A.; Moore, C.L.; Cheung, K.-H.; Brandt, C. Predicting urinary tract infections in the emergency department with machine learning. *PLoS ONE* **2018**, *13*, e0194085. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.