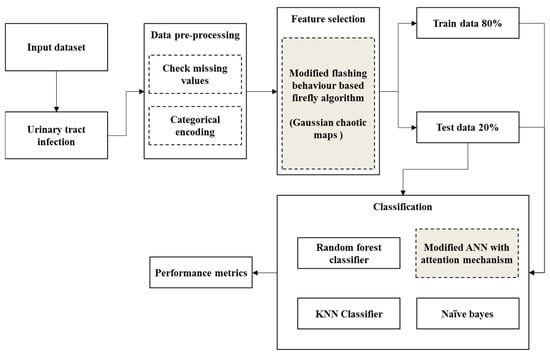
SMART PUBLIC RESTROOM

ALGORTHM:

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**](https://www.mdpi.com/2076-3417/13/10/5860#fig_body_display_applsci-13-05860-f001). 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 improvises 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 (LIr) is inversely proportional to the distance from the light source. Hence, when the light passes via a medium with the light absorption coefficient of (�), then LIr differs with the distance (r) as specified in Equation (1):

LIr=I0e−�dist2

(1)

In Equation (1), I0 represents the intensity at the source-point. As the computation of LI(dist)=11+�dist2 is simpler than e−�dist2, LIr could be computed as follows:

Idist=I0�dist2

(2)

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

∝dist=∝0e−�dist2

(3)

Likewise, it could be stated as follows:

�dist=�01+�dist2

(4)

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

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

qi+1=qi+�0e−�dist2qm−qi+Ϛ�

(5)

The second term is accredited to attraction, while the third term gets randomization with (Ϛ(0≤Ϛ≤1)) and �. For numerous practical issues, a constant value (Ϛ=0.2) can be used. In this case, � 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):

qi+1=exp(−Ϛqi2)

(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 � exists as the constant value. In this case, the � value finds the attraction of all FFs. Generally, �∈[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, � 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(); |
|  |
| Objecfun=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 � 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 withdistance Dist via exp−�Dist2 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 |