**Journal of Information Systems Engineering and Management**
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

**Research Article**

# Cervical Cancer Risk Prediction Using One-Dimensional Convolutional Neural Network Enhanced by Snake Optimization Algorithm

Ghada Alrammahee[1,*], Amir Lakizadeh[1]
[1] *Department of Computer Engineering and Information Technology, University of Qom, Qom, Iran*
[*] *Corresponding Author*

**ARTICLE INFO**

Received: 18 Dec 2024
Revised: 10 Feb 2025
Accepted: 28 Feb 2025

**ABSTRACT**

Cancer of the cervix is one of the most prevalent malignances in the global, hence early diagnosis is vital in the healthcare delivery systems. Growth in medical imaging as well as enhanced computational methods has now opened up possibilities for efficient and accurate diagnosis apart from other clinical investigations such as the Pap smear. Here, we introduced CNN-SOA, a deep learning-based method using One-Dimensional Convolutional Neural Network (1D-CNN) supported by the Snake Optimization Algorithm to estimate cervical cancer risk. The experiments presented show that the proposed method greatly improves prediction performance and stability, which confirms its suitability for medical classification problems. The results indicate that CNN-SOA can be generalized to other healthcare domains, which will opportunity development of precise and earlier diagnosing technologies in AI-based medical diagnosis.

**Keywords:** Cervical Cancer, Convolutional Neural Networks, Snake Optimization Algorithm, Deep Learning.

## 1 Introduction

Cervical cancer, stemming from unregulated cell growth in a woman's cervix, is a severe health concern that accounts for nearly 9.6 million deaths annually [1]. This cervix, a neck- like passage at the uterus's base, is particularly vulnerable in women. Globally, cervical cancer ranks fourth in prevalence, with approximately 570,000 cases each year. Tragically, it is the second most prevalent cancer in women from developing and low-income nations, with over 311,000 annual fatalities, predominantly in low-to-middle income regions [2]. This is attributed to inadequate public under- standing of the disease's origins and repercussions. Key risk factors encompass early pregnancies, multiple sexual partners, compromised immune systems, smoking, contraceptive pill usage, and substandard menstrual hygiene. Noticeable symptoms include abnormal vaginal bleeding, dis- charge, and pain during intercourse. Prompt detection offers a chance for effective treatment, potentially prolonging life or providing palliative relief from symptoms. Diagnostic procedures like the Pap test and the HPV test are critical for identifying cervical cancer [3, 4, 5, 6, 7]. The Pap test, or Pap smear, identifies precancerous cell changes, while the HPV test detects the human papillomavirus responsible for these changes [8, 9, 10, 11, 12]. The significance of early detection can't be overstated, especially in developing countries such as Ethiopia, where expertise is limited. Therefore, leveraging advanced technologies for predictive measures becomes vital. Deep learning, a subset of machine learning, is revolutionizing medical fields by mimicking human cognitive functions. In our study, we harness the capabilities of convolutional neural networks (CNN), a deep learning model, to analyze cervical cancer images from medical equipment. Worryingly, cervical cancer is the primary cancer-affecting women in sub-Saharan Africa, and in Ethiopia, the incidence rate stands at a troubling 35.9 per 100,000 women [13]. Factors such as limited awareness, inadequate screening infrastructure, competing health priorities, and general neglect of women's health contribute to this heightened risk. Moreover, while cervical cancer primarily affects women between 35 and 44, many older women remain oblivious to the sustained risk, leading to late detections and varying survival rates [13]. Regular screening is imperative. The American Cancer Society recommends that individuals with a cervix aged 25 to 65, undergo cervical cancer screening every five years using the HPV test. In cases where the HPV test isn't accessible, a combination of the HPV/Pap test every five years or a Pap test every three years is advised [13]. However, in Ethiopia, conducting these tests presents challenges due to scarce expertise and equipment. Hence, our proposed model strives to bridge this expertise gap, aiming for early-stage cervical cancer detection [14, 15, 16, 17, 18]. It is paramount to note that cervical cancer's on-set is typically gradual, often preceded by cervical abnormalities. How- ever, the lack of early symptoms might lead to complacency in preventative care Particularly in developing countries, limited resources and low awareness result in poor adherence to routine checks. The integrationof deep learning, especially when applied early, can play a pivotal role in raising awareness and detection. This research aspires to pioneer an automated cervical cancer detection system through deep learning. Specifically, the CNN technique holds vast potential for disease classification, which is vital in transforming Ethiopia's medical sector. In essence, the primary objective of this articleis to design a deep-learning powered cervical cancer detection and classification system, utilizingthe CNN approach and an optimization algorithm.

## 2 Related Work

In [19], the "Cervical Cancer Behavior Risk Data Set" from the"UCI Machine Learning Repository" was employed, consisting of 72 instances and 19 attributes. Three machine learning models, namely Decision Tree, Random Forest, and XGBoost, were introduced to predict cervical cancer behaviors, resulting in a notable accuracy of 93.33%. Furthermore, key features were emphasized based on their significance scores, highlighting their impact on the model's classification performance. In [20], a refined approach to predicting cervical cancer using ML algorithms was proposed, comprising four stages: dataset selection, data pre-processing, predictive model selection (PMS), and pseudo- code formulation. The PMS phase experimented with various conventional ML techniques, with the random forest, decision tree, adaptive boosting, and gradient boosting models achieving a topclassification score of 100%. SVM closely followed with a 99% accuracy rate. The study also assessed the computational efficiency of these models and surveyed 132 Saudi Arabian volunteersabout their views on computer-assisted cervical cancer prediction, emphasizing the significance of HPV. In [21], a study was conducted to classify patients based on their survey responses to identify those with a high risk of cervical cancer. Utilizing a dataset of 858 records, which included33 attributes and biopsy results, the study discerned 55 patients with cervical cancer and 803 healthy individuals. The dataset was randomly split into a training set (66% or 565 records) and atesting set (292 records). Classifications were made using methods such as Multilayer Perceptron, Bayes Net, and k-Nearest Neighbor, with the respective results, including accuracy, true positives,false positives, and confusion matrices, thoroughly discussed. Research in [22] proposed utilizing machine learning algorithms for efficient cervical cancer diagnosis using the UCI cervical cancer risk factor dataset. A voting method, combining Decision tree, logistic regression, and random forest classifiers, was employed. The SMOTE addressed dataset imbalance, while the PCA streamlined data dimensions. To mitigate overfitting, a stratified 10-fold cross-validation was applied. The dataset covered four target variables with 32 risk factors. By incorporating the voting classifier, SMOTE, and PCA, there was a significant improvement in accuracy, sensitivity,and the ROC AUC for the predictive models. Furthermore, PCA enhanced computational efficiency. Compared to prior studies, our models exhibited superior performance in cervical cancerdiagnosis. Research in [23] proposed a model for cervical cancer prediction using Boruta analysis and the SVM method. The Boruta analysis, an enhanced version of the random forest, identifies crucial feature subsets for classification tasks. The model's primary focus was to discern the significance of screening factors for high-risk patient classification. Implemented in Python, the model's Performance, measured by accuracy, precision, and recall, indicated that the Boruta analysis withSVM surpassed existing methods in efficiency. Research in [24] proposed an analysis of various supervised machine learning techniques for early cervical cancer detection using a dataset from the UCI repository with 858 patients. Six classification methods were evaluated, with and withoutthree different feature selection algorithms. The highest accuracy achieved was 94.94% using theXG Boost with all features. While machine learning showcases advantages for large-scale data andprognosis, challenges like overfitting and lack of interpretability emphasize the need for further refinement in clinical prediction models. Research in [25] proposed the use of a decision tree (DT)algorithm to assess cervical cancer risk factors. By leveraging the recursive feature elimination (RFE) and LASSO for feature selection and addressing dataset challenges with SMOTE-Tomek, a combination of under and oversampling techniques, the DT achieved an impressive accuracy of98.72% and sensitivity of 100%. The findings underscore the DT's enhanced performance when feature reduction and class imbalance are effectively addressed. Researchers have performed a number of studies on skin cancer detection and segmentation, using many Convolutional Neural Network (CNN) models to heighten classification accuracy. A constant is that most studies discovered CNNs greatly improve classification performance [26].

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

## 3 Proposed Methodology

The following is the overview of the methodology, which is going to be used in our work for the analysis of the dataset. The work starts by downloading the dataset, and this is in a CSV format and the name of the dataset is ‘flights’. After loading the dataset is followed by the preprocessing phase, which involves data cleaning, normalization and the treatment of missing values. This is important in a way to make sure that the data collected will be suitable for analysis and training of the models.

Upon preprocessing, we perform Exploratory Data Analysis (EDA) to know more about the distribution of the dataset; get to see some patterns in the data as well as how different variables interact with each other. In the case of EDA, it helps in analyzing the general features of the data, which is usefully when implementing the subsequent models in machine learning.

Then, in the context of proposed post-processing, we use One-Dimensional Convolutional Neural Network (1D-CNN) along with Snake Optimizer Algorithm (SOA). The idea utilized here is to improve the general performance of the model by tuning the hyperparameters, the filter num, Kernel sizes and fully connected layers. It is the function of the SOA to adjust such parameters soas to enhance the classification accuracy.

After the training of the model is completed, the next steps include the evaluation of the modelwhere the actual accuracy and other factors, which include precision, recall and AUC-ROC, are determined. These goals serve as measures of the model’s accuracy in categorizing the data. In the case when it becomes necessary additional iterations of optimization and re-training will be performed to achieve better results.

Last but not the least; we also do a model comparison task for the proposed model to check the efficiency of the work done. This way we illustrate the benefits of the proposed approach as well as, possible drawbacks, which gives a clear picture of where this approach stands with respectto other methods of skin cancer detection in the present day world.

In the following, we will be describing each step in detail. Here will also be the particular techniques and algorithms utilized, as well as the reasons for our methodological choices. Such astructured approach guarantees that our methodology is both strong and repeatable, laying the groundwork for accurate classification and identification of skin cancer by means of deep learningIn the following, we present on the other hand a case study.

### 3.1 Dataset Overview

Cervical cancer still ranks among the leading killers of women across the world because of cancer. It is very essential to detect this disease at an initial stage and forecast its development accurately be-cause improve the prospect of favorable treatment results and survival probabilities. The datasetemployed in our study is useful in building a model, which relies on more of the principles of the machine learning. Intended to help spot people with a predisposition to cervical cancer. This proactive approach enables early diagnosis and treatment of the patients by doctors. The set of variables includes various behavioral, attitudinal, social and motivational determinants, which might have apossibility of affecting that one’s susceptibility to cervical cancer. It includes attributes includingsexual risk behaviors, eating habits, personal hygiene and social support. Mechanisms, among others. Every value in the given dataset is scored numerically, making possible an overview of theways that may be used by the different factors towards increasing cervical cancer risk. For example, attributes such as ’behavior_sexualRisk’, behavior eating, and ‘behavior_personalHygine’ shed light to the personal hygiene and practices of the people which are of great importance for assessing their risk levels. Similarly, ’socialSupport_emotionality’ and ’socialSupport_appreciation’is an indicator of the emotional and social support that one gets which a critical determinant is of their overall health outcomes. Machine learning models inferred on this dataset are supposed to decipher the polymorphic information to identify some possibly suggestive signs and symptoms that would point towards a relatively increased chance of cervical cancer. This modeling makes iteasier for health care providers to address their efforts at the right areas. Implement a distributionof medical resources that would be best for allocating the resources, which may also lead the saving of lives through timely and precise diagnosis.

Table 1: Summary Statistics of the used Dataset

| Feature | Mean | Std Dev | Min | 25% | 50% | 75% | Max |
|---|---|---|---|---|---|---|---|
| **behavior_sexualRisk** | 9.67 | 1.19 | 2 | 10 | 10 | 10 | 10 |
| **behavior_eating** | 12.79 | 2.36 | 3 | 11 | 13 | 15 | 15 |
| **behavior_personalHygine** | 11.08 | 3.03 | 3 | 9 | 11 | 14 | 15 |
| **intention_aggregation** | 7.90 | 2.74 | 2 | 6 | 10 | 10 | 10 |
| **intention_commitment** | 13.35 | 2.37 | 6 | 11 | 15 | 15 | 15 |
| **attitude_consistency** | 7.18 | 1.52 | 2 | 6 | 7 | 8 | 10 |
| **attitude_spontaneity** | 8.61 | 1.52 | 4 | 8 | 9 | 10 | 10 |
| **norm_significantPerson** | 3.13 | 1.85 | 1 | 1 | 3 | 5 | 5 |
| **norm_fulfillment** | 8.49 | 4.91 | 3 | 3 | 7 | 14 | 15 |
| **perception_vulnerability** | 8.51 | 4.28 | 3 | 5 | 8 | 13 | 15 |
| **perception_severity** | 5.39 | 3.40 | 2 | 2 | 4 | 9 | 10 |
| **motivation_strength** | 12.65 | 3.21 | 3 | 11 | 14 | 15 | 15 |
| **motivation_willingness** | 9.69 | 4.13 | 3 | 7 | 11 | 13 | 15 |
| **socialSupport_emotionality** | 8.10 | 4.24 | 3 | 3 | 9 | 11.25 | 15 |
| **socialSupport_appreciation** | 6.17 | 2.90 | 2 | 3.75 | 6.5 | 9 | 10 |
| **socialSupport_instrumental** | 10.38 | 4.32 | 3 | 6.75 | 12 | 14.25 | 15 |
| **empowerment_knowledge** | 10.54 | 4.37 | 3 | 7 | 12 | 15 | 15 |
| **empowerment_abilities** | 9.32 | 4.18 | 3 | 5 | 10 | 13 | 15 |
| **empowerment_desires** | 10.28 | 4.48 | 3 | 6.75 | 11 | 15 | 15 |
| **ca_cervix** | 0.29 | 0.46 | 0 | 0 | 0 | 1 | 1 |

### 3.2 Exploratory Data Analysis and Preprocessing

Data exploration and preprocessing stage form the basic and initial stage in this type of study and are very crucial for model development. However, firstly we need to load and explore the dataset in order to get the initial understanding on its structure and the general nature of the data. It has concerns with viewing the initial rows with the view of obtaining the variable names, their types and the structural nature of the dataset. After creating an overview of the data, one is able to detect gaps in data that need to be filled or discrepancies that have to be rectified during data preprocessing stage. Descriptive statistics are used to compute measures of central tendency,spread and shapes of the dataset in order to gain an overall understanding of the characteristics of the dataset. This is important in pointing out cases of outliers or any abnormalities, which mayhamper the outcome of the models.

To gain even deeper understanding of the interconnections between the variables we carry out a correlation analysis. In this work, we present a correlation matrix heatmap that helps in understanding the strengths of linear relationship between different features so as to choose which features might have maximum impact on the target variable. Further, the distribution of the target variable is also checked to see if the level of imbalance affects the model's performance in any way. Techniques like histograms for individual attributes and box-plots for grouped attributesand target variable are used to focus on distribution and spread in the data and moreover it helpsto detect the presence of outliers in the data set.

Outlier identification is performed with help of z-score technique that measures how far fromthe mean could data point be. When analyzing the values of a distribution any value that is much lower or much higher than other values is considered an outlier and its role in the analysis isdetermined. In case the outliers are large, then they may be eliminated or dealt with so that they do not interfere with the learning of the model.

Besides these steps, other techniques of dimensionality reduction, for example PCA and t-SNEare used to plot the data in a two-dimensional space. PCA operates to bring features into a formatwhere instead of several features being dependent on each other, they depend on uncorrelated orthogonal dimensions, which have the highest variance. On the other hand, t-SNE is helpful in low- dimensional visualization of the given data since it preserves the global neighborhood relationship in the high-dimension space while it maps this data in the two or three-dimensional space. These techniques help to analyses the data and determine such clusters that are not very distinguishableat the first sight.

The EDA also covers consideration of the categorical relationships by means of heat maps and bars charts. The heat maps explain the strength and direction of the categorical feature relationships and bar charts give insights into distribution of the variable and its likelihood of being predictive.

As such, after this broad analysis, we go to the following step, in which every figure created throughout the EDA phase will be presented and explained. This will help in getting the characteristics of the dataset and also assure that the preprocessing help in getting the data set ready formodeling. In this way, we plan on developing a strong foundation for constructing well-functioning machine learning models for the effective classification and detection of skin cancer.

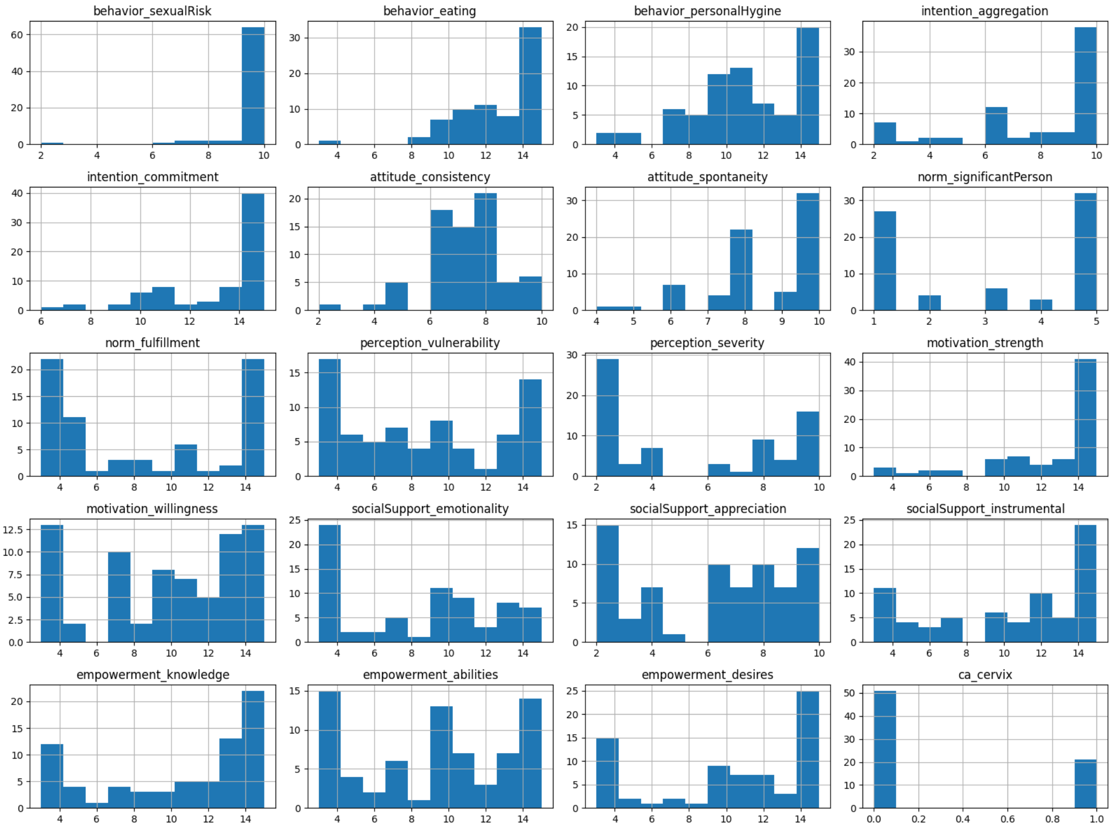
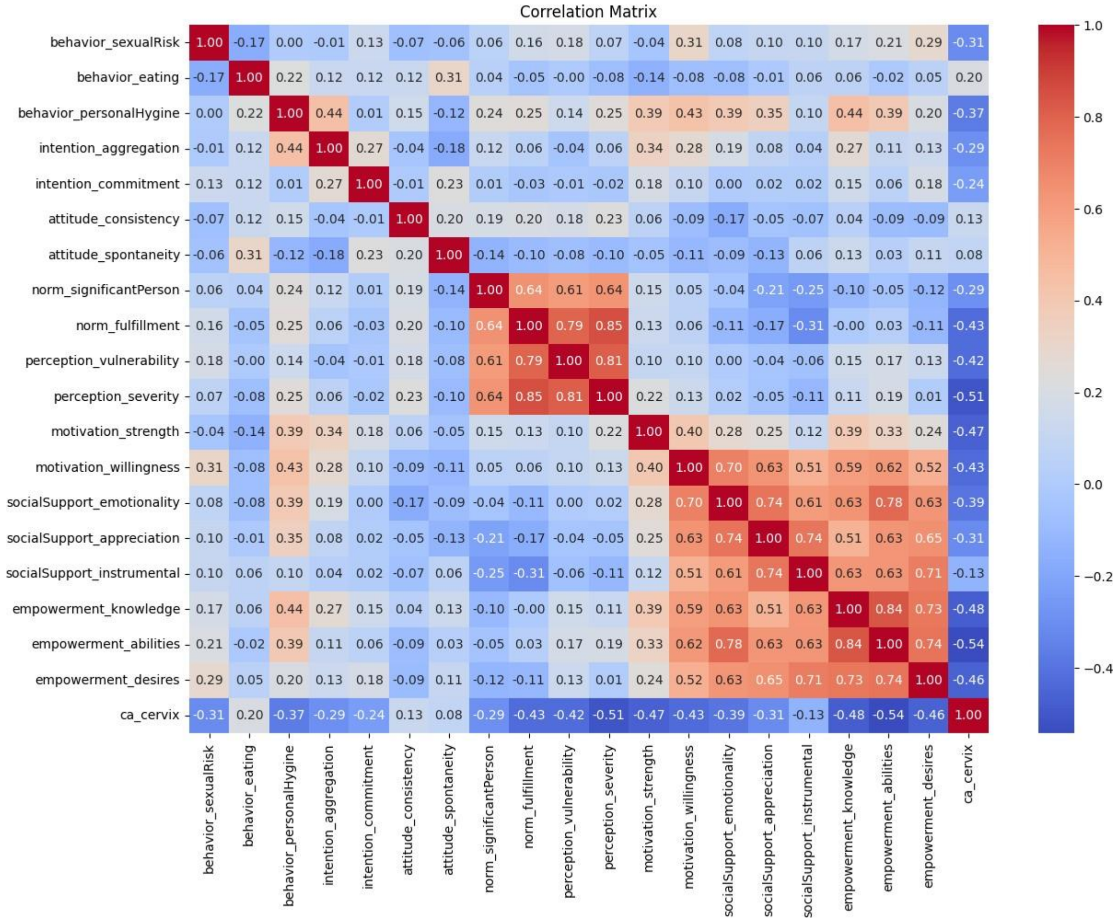


Figure 1: Distribution of features(variables)

Figure 1: Still, a series of histograms that shows the distribution among different features present in this dataset. Each subplot is associated with a specific feature and the x-axis shows, forall possible values of its corresponding variable, how many times it occurs within your dataset. Itmay indicate whether any of the features folow some model distribution showing skewness, peaks or gaps in data. For example, the majority of features have a right-skewed distribution, which means that much of the data is concentrated at high values. Complex and intricate, just like any Data Science exercise! These are the kind of distributions you need to be aware of since they may affect how your model performs & sometimes require additional preprocessing steps including normalizations or transformations for a balanced input.


Correlation Matrix

| | behavior_sexualRisk | behavior_eating | behavior_personalHygine | intention_aggregation | intention_commitment | attitude_consistency | attitude_spontaneity | norm_significantPerson | norm_fulfillment | perception_vulnerability | perception_severity | motivation_strength | motivation_willingness | socialSupport_emotionality | socialSupport_appreciation | socialSupport_instrumental | empowerment_knowledge | empowerment_abilities | empowerment_desires | ca_cervix |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| behavior_sexualRisk | 1.00 | -0.17 | 0.00 | -0.01 | 0.13 | -0.07 | -0.06 | 0.06 | 0.16 | 0.18 | 0.07 | -0.04 | 0.31 | 0.08 | 0.10 | 0.10 | 0.17 | 0.21 | 0.29 | -0.31 |
| behavior_eating | -0.17 | 1.00 | 0.22 | 0.12 | 0.12 | 0.12 | 0.31 | 0.04 | -0.05 | -0.00 | -0.08 | -0.14 | -0.08 | -0.08 | -0.01 | 0.06 | 0.06 | -0.02 | 0.05 | 0.20 |
| behavior_personalHygine | 0.00 | 0.22 | 1.00 | 0.44 | 0.01 | 0.15 | -0.12 | 0.24 | 0.25 | 0.14 | 0.25 | 0.39 | 0.43 | 0.39 | 0.35 | 0.10 | 0.44 | 0.39 | 0.20 | -0.37 |
| intention_aggregation | -0.01 | 0.12 | 0.44 | 1.00 | 0.27 | -0.04 | -0.18 | 0.12 | 0.06 | -0.04 | 0.06 | 0.34 | 0.28 | 0.19 | 0.08 | 0.04 | 0.27 | 0.11 | 0.13 | -0.29 |
| intention_commitment | 0.13 | 0.12 | 0.01 | 0.27 | 1.00 | -0.01 | 0.23 | 0.01 | -0.03 | -0.01 | -0.02 | 0.18 | 0.10 | 0.00 | 0.02 | 0.02 | 0.15 | 0.06 | 0.18 | -0.24 |
| attitude_consistency | -0.07 | 0.12 | 0.15 | -0.04 | -0.01 | 1.00 | 0.20 | 0.19 | 0.20 | 0.18 | 0.23 | 0.06 | -0.09 | -0.17 | -0.05 | -0.07 | 0.04 | -0.09 | -0.09 | 0.13 |
| attitude_spontaneity | -0.06 | 0.31 | -0.12 | -0.18 | 0.23 | 0.20 | 1.00 | -0.14 | -0.10 | -0.08 | -0.10 | -0.05 | -0.11 | -0.09 | -0.13 | 0.06 | 0.13 | 0.03 | 0.11 | 0.08 |
| norm_significantPerson | 0.06 | 0.04 | 0.24 | 0.12 | 0.01 | 0.19 | -0.14 | 1.00 | 0.64 | 0.61 | 0.64 | 0.15 | 0.05 | -0.04 | -0.21 | -0.25 | -0.10 | -0.05 | -0.12 | -0.29 |
| norm_fulfillment | 0.16 | -0.05 | 0.25 | 0.06 | -0.03 | 0.20 | -0.10 | 0.64 | 1.00 | 0.79 | 0.85 | 0.13 | 0.06 | -0.11 | -0.17 | -0.31 | -0.00 | 0.03 | -0.11 | -0.43 |
| perception_vulnerability | 0.18 | -0.00 | 0.14 | -0.04 | -0.01 | 0.18 | -0.08 | 0.61 | 0.79 | 1.00 | 0.81 | 0.10 | 0.10 | 0.00 | -0.04 | -0.06 | 0.15 | 0.17 | 0.13 | -0.42 |
| perception_severity | 0.07 | -0.08 | 0.25 | 0.06 | -0.02 | 0.23 | -0.10 | 0.64 | 0.85 | 0.81 | 1.00 | 0.22 | 0.13 | 0.02 | -0.05 | -0.11 | 0.11 | 0.19 | 0.01 | -0.51 |
| motivation_strength | -0.04 | -0.14 | 0.39 | 0.34 | 0.18 | 0.06 | -0.05 | 0.15 | 0.13 | 0.10 | 0.22 | 1.00 | 0.40 | 0.28 | 0.25 | 0.12 | 0.39 | 0.33 | 0.24 | -0.47 |
| motivation_willingness | 0.31 | -0.08 | 0.43 | 0.28 | 0.10 | -0.09 | -0.11 | 0.05 | 0.06 | 0.10 | 0.13 | 0.40 | 1.00 | 0.70 | 0.63 | 0.51 | 0.59 | 0.62 | 0.52 | -0.43 |
| socialSupport_emotionality | 0.08 | -0.08 | 0.39 | 0.19 | 0.00 | -0.17 | -0.09 | -0.04 | -0.11 | 0.00 | 0.02 | 0.28 | 0.70 | 1.00 | 0.74 | 0.61 | 0.63 | 0.78 | 0.63 | -0.39 |
| socialSupport_appreciation | 0.10 | -0.01 | 0.35 | 0.08 | 0.02 | -0.05 | -0.13 | -0.21 | -0.17 | -0.04 | -0.05 | 0.25 | 0.63 | 0.74 | 1.00 | 0.74 | 0.51 | 0.63 | 0.65 | -0.31 |
| socialSupport_instrumental | 0.10 | 0.06 | 0.10 | 0.04 | 0.02 | -0.07 | 0.06 | -0.25 | -0.31 | -0.06 | -0.11 | 0.12 | 0.51 | 0.61 | 0.74 | 1.00 | 0.63 | 0.63 | 0.71 | -0.13 |
| empowerment_knowledge | 0.17 | 0.06 | 0.44 | 0.27 | 0.15 | 0.04 | 0.13 | -0.10 | -0.00 | 0.15 | 0.11 | 0.39 | 0.59 | 0.63 | 0.51 | 0.63 | 1.00 | 0.84 | 0.73 | -0.48 |
| empowerment_abilities | 0.21 | -0.02 | 0.39 | 0.11 | 0.06 | -0.09 | 0.03 | -0.05 | 0.03 | 0.17 | 0.19 | 0.33 | 0.62 | 0.78 | 0.63 | 0.63 | 0.84 | 1.00 | 0.74 | -0.54 |
| empowerment_desires | 0.29 | 0.05 | 0.20 | 0.13 | 0.18 | -0.09 | 0.11 | -0.12 | -0.11 | 0.13 | 0.01 | 0.24 | 0.52 | 0.63 | 0.65 | 0.71 | 0.73 | 0.74 | 1.00 | -0.46 |
| ca_cervix | -0.31 | 0.20 | -0.37 | -0.29 | -0.24 | 0.13 | 0.08 | -0.29 | -0.43 | -0.42 | -0.51 | -0.47 | -0.43 | -0.39 | -0.31 | -0.13 | -0.48 | -0.54 | -0.46 | 1.00 |

Figure 2: Correlation Matrix for all of the initial features

Figure 2: heatmap of correlation matrix, which shows the linear relation between different features in the data. The correlation coefficients in the matrix range from -1 to 1 which means a value of 1 indicates a perfect positive relationship, zero no linear relation and -1 is perfectly inversely related. This creates an easy-to-understand quick reference to the strength and direction of thosecorrelations, as given by the color gradient in your heatmap — amino acids with red tones are strongly positively correlated or co-occur often, whereas blue-toned colors mean a strong negative correlation between two residues. In particular, positive correlations among features (e.g., between ’norm_fullfillment’ and ’perception severity’) indicate that certain aspects of interaction may bleed into one another in a manner than could compromise the ability for these models to correctly make predictions. We need to understand these relationships, which will not only important for feature selection/engineering but also it helps us in identifying the potential multicollinearity issue and see if they can be redundant or particular impactful.

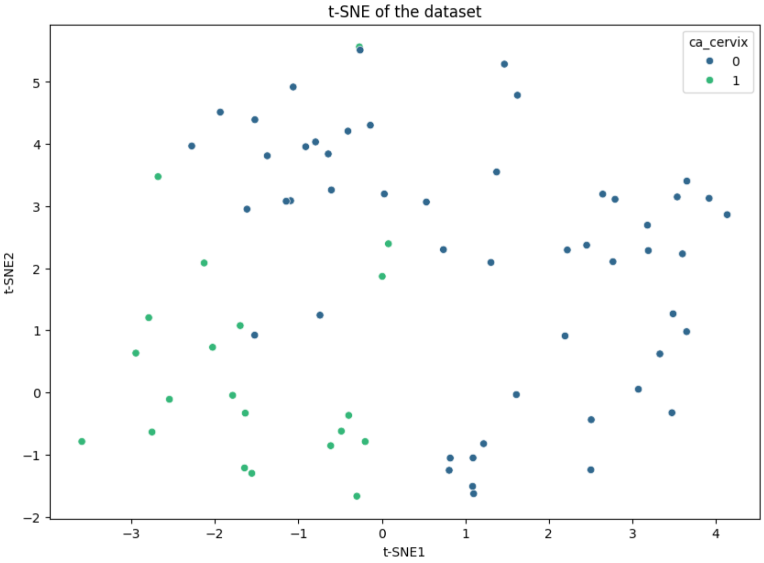
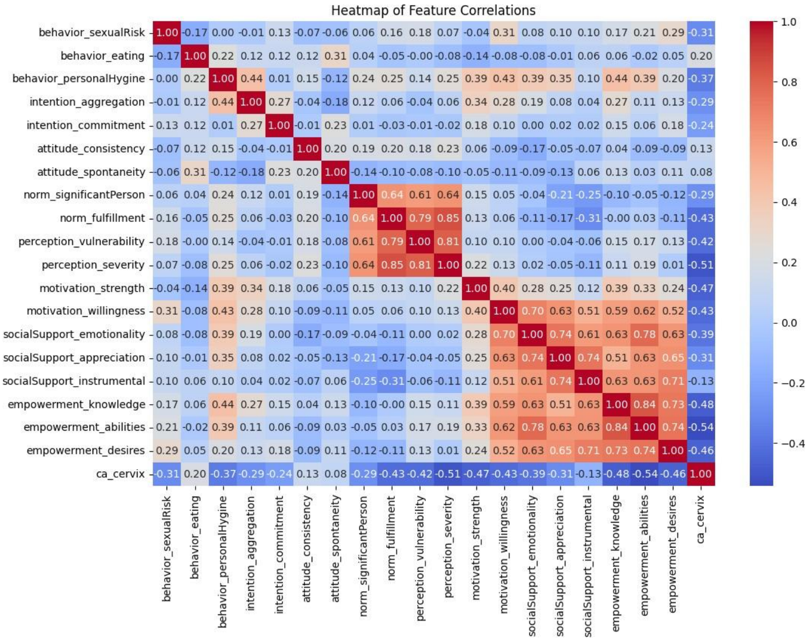


Figure 3: t-SNE of the Dataset

Figure 3 shows the application results when the dataset is fed into t-distributed Stochastic Neighbor Embedding (t-SNE), another dimensionality reduction technique that can be great for viewing high-dimensional data. Similar to PCA, the scatter plot shows data points in two dimensions, For example, red tells us about the classes of target variable militants, south Africa, politics. And unlike PCA, t-SNE focuses preserving not only overall patterns of the data but also localstructures. On a PCA plot, those couplets with bars near zero skew from rest may emerge as prominent clusters in t-SNE space instead. In this figure, clusters of similar classes are slightly visible which is an indication that t-SNE may be more suitable than PCA for uncovering non-linearseparations amongst parameters. This chart highlights possible complicacies in data, which couldrequire sophisticated modeling approaches.



| | behavior_sexualRisk | behavior_eating | behavior_personalHygine | intention_aggregation | intention_commitment | attitude_consistency | attitude_spontaneity | norm_significantPerson | norm_fulfillment | perception_vulnerability | perception_severity | motivation_strength | motivation_willingness | socialSupport_emotionality | socialSupport_appreciation | socialSupport_instrumental | empowerment_knowledge | empowerment_abilities | empowerment_desires | ca_cervix |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| behavior_sexualRisk | 1.00 | -0.17 | 0.00 | -0.01 | 0.13 | -0.07 | -0.06 | 0.06 | 0.16 | 0.18 | 0.07 | -0.04 | 0.31 | 0.08 | 0.10 | 0.10 | 0.17 | 0.21 | 0.29 | -0.31 |
| behavior_eating | -0.17 | 1.00 | 0.22 | 0.12 | 0.12 | 0.12 | 0.31 | 0.04 | -0.05 | -0.00 | -0.08 | -0.14 | -0.08 | -0.08 | -0.01 | 0.06 | 0.06 | -0.02 | 0.05 | 0.20 |
| behavior_personalHygine | 0.00 | 0.22 | 1.00 | 0.44 | 0.01 | 0.15 | -0.12 | 0.24 | 0.25 | 0.14 | 0.25 | 0.39 | 0.43 | 0.39 | 0.35 | 0.10 | 0.44 | 0.39 | 0.20 | -0.37 |
| intention_aggregation | -0.01 | 0.12 | 0.44 | 1.00 | 0.27 | -0.04 | -0.18 | 0.12 | 0.06 | -0.04 | 0.06 | 0.34 | 0.28 | 0.19 | 0.08 | 0.04 | 0.27 | 0.11 | 0.13 | -0.29 |
| intention_commitment | 0.13 | 0.12 | 0.01 | 0.27 | 1.00 | -0.01 | 0.23 | 0.01 | -0.03 | -0.01 | -0.02 | 0.18 | 0.10 | 0.00 | 0.02 | 0.02 | 0.15 | 0.06 | 0.18 | -0.24 |
| attitude_consistency | -0.07 | 0.12 | 0.15 | -0.04 | -0.01 | 1.00 | 0.20 | 0.19 | 0.20 | 0.18 | 0.23 | 0.06 | -0.09 | -0.17 | -0.05 | -0.07 | 0.04 | -0.09 | -0.09 | 0.13 |
| attitude_spontaneity | -0.06 | 0.31 | -0.12 | -0.18 | 0.23 | 0.20 | 1.00 | -0.14 | -0.10 | -0.08 | -0.10 | -0.05 | -0.11 | -0.09 | -0.13 | 0.06 | 0.13 | 0.03 | 0.11 | 0.08 |
| norm_significantPerson | 0.06 | 0.04 | 0.24 | 0.12 | 0.01 | 0.19 | -0.14 | 1.00 | 0.64 | 0.61 | 0.64 | 0.15 | 0.05 | -0.04 | -0.21 | -0.25 | -0.10 | -0.05 | -0.12 | -0.29 |
| norm_fulfillment | 0.16 | -0.05 | 0.25 | 0.06 | -0.03 | 0.20 | -0.10 | 0.64 | 1.00 | 0.79 | 0.85 | 0.13 | 0.06 | -0.11 | -0.17 | -0.31 | -0.00 | 0.03 | -0.11 | -0.43 |
| perception_vulnerability | 0.18 | -0.00 | 0.14 | -0.04 | -0.01 | 0.18 | -0.08 | 0.61 | 0.79 | 1.00 | 0.81 | 0.10 | 0.10 | 0.00 | -0.04 | -0.06 | 0.15 | 0.17 | 0.13 | -0.42 |
| perception_severity | 0.07 | -0.08 | 0.25 | 0.06 | -0.02 | 0.23 | -0.10 | 0.64 | 0.85 | 0.81 | 1.00 | 0.22 | 0.13 | 0.02 | -0.05 | -0.11 | 0.11 | 0.19 | 0.01 | -0.51 |
| motivation_strength | -0.04 | -0.14 | 0.39 | 0.34 | 0.18 | 0.06 | -0.05 | 0.15 | 0.13 | 0.10 | 0.22 | 1.00 | 0.40 | 0.28 | 0.25 | 0.12 | 0.39 | 0.33 | 0.24 | -0.47 |
| motivation_willingness | 0.31 | -0.08 | 0.43 | 0.28 | 0.10 | -0.09 | -0.11 | 0.05 | 0.06 | 0.10 | 0.13 | 0.40 | 1.00 | 0.70 | 0.63 | 0.51 | 0.59 | 0.62 | 0.52 | -0.43 |
| socialSupport_emotionality | 0.08 | -0.08 | 0.39 | 0.19 | 0.00 | -0.17 | -0.09 | -0.04 | -0.11 | 0.00 | 0.02 | 0.28 | 0.70 | 1.00 | 0.74 | 0.61 | 0.63 | 0.78 | 0.63 | -0.39 |
| socialSupport_appreciation | 0.10 | -0.01 | 0.35 | 0.08 | 0.02 | -0.05 | -0.13 | -0.21 | -0.17 | -0.04 | -0.05 | 0.25 | 0.63 | 0.74 | 1.00 | 0.74 | 0.51 | 0.63 | 0.65 | -0.31 |
| socialSupport_instrumental | 0.10 | 0.06 | 0.10 | 0.04 | 0.02 | -0.07 | 0.06 | -0.25 | -0.31 | -0.06 | -0.11 | 0.12 | 0.51 | 0.61 | 0.74 | 1.00 | 0.63 | 0.63 | 0.71 | -0.13 |
| empowerment_knowledge | 0.17 | 0.06 | 0.44 | 0.27 | 0.15 | 0.04 | 0.13 | -0.10 | -0.00 | 0.15 | 0.11 | 0.39 | 0.59 | 0.63 | 0.51 | 0.63 | 1.00 | 0.84 | 0.73 | -0.48 |
| empowerment_abilities | 0.21 | -0.02 | 0.39 | 0.11 | 0.06 | -0.09 | 0.03 | -0.05 | 0.03 | 0.17 | 0.19 | 0.33 | 0.62 | 0.78 | 0.63 | 0.63 | 0.84 | 1.00 | 0.74 | -0.54 |
| empowerment_desires | 0.29 | 0.05 | 0.20 | 0.13 | 0.18 | -0.09 | 0.11 | -0.12 | -0.11 | 0.13 | 0.01 | 0.24 | 0.52 | 0.63 | 0.65 | 0.71 | 0.73 | 0.74 | 1.00 | -0.46 |
| ca_cervix | -0.31 | 0.20 | -0.37 | -0.29 | -0.24 | 0.13 | 0.08 | -0.29 | -0.43 | -0.42 | -0.51 | -0.47 | -0.43 | -0.39 | -0.31 | -0.13 | -0.48 | -0.54 | -0.46 | 1.00 |

Figure 4: Heatmap of Feature Correlations

In Figure 4, there is a heatmap of our variables, which is visually similar to the one shown in Figure 3 above. There is actually a lot more information contained in this plot than just that represented graphically. That information includes: How do the various features of data each correlate with each other? Large-scale modes of data collection only make up a small quantitativepart in this data set, although they are very much integrated into the whole set of data due to theirextremely high correlations with all other pieces of the dataset. This diagram is vital for detecting feature groups that are inextricably linked and represents ontological information to the model. For example, 'norm _fulfillment' and 'perception severity' have a very high positive correlation, which suggests that they might be recording some of the same information about people or their attitudes. To understand those correlations is not only important for choosing features, but also makes sure that our model doesn't become riddled with multi-collinearities and is taking advantage of a wide variety in sets of informative features.

### 3.3 One-Dimensional Convolutional Neural Network (1D-CNN)

This 1D-CNN architecture (Fig. 9) we propose in this work is carefully tuned to best process timeseries or signal data with intense hyperparameter tuning for the end-to-end model performance boosting. It starts with an input layer formed based on the reshaped training data dimensions. Inother words, input layer is set to receive data that has a shape of the number features per sampleand one channel indicating its 1D nature.

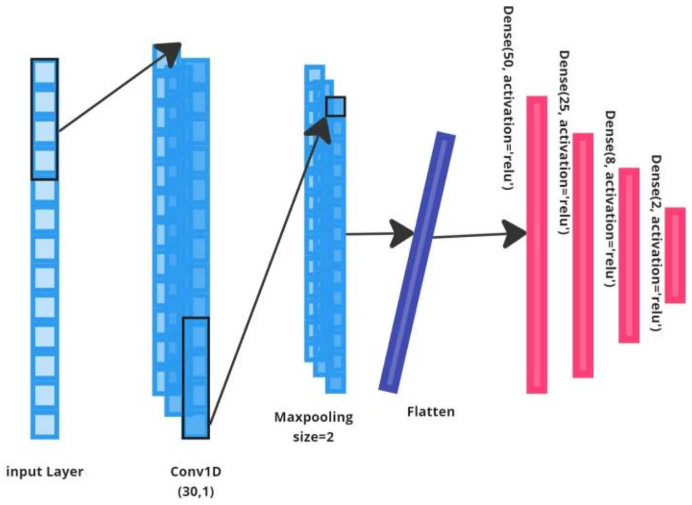


Figure 9: 1D CNN Architecture

The first convolutional layer comes after input layer for feature extraction where local features are detected from the input data. This layer uses a certain amount of filters (num_filters1) andkernel size which is defined to transform the filtering operation. The activation function utilizedis Rectified Linear Unit (ReLU) which brings non-linearity into the model thus make it capable tolearn non-linear patterns. Due to the need to keep the input the same dimensional, padding is used.The output from the convolutional layer for the next layers is then put through batch normalization layer which normalizes these activations through scaling. This step is very important for movingand making the training process stable and faster due to the reduction of internal covariate shift.After that, batch normalization is done and then the layer of max pooling is used. This layer narrows down the spatial dimensionality of the feature maps by then applying a pooling operationwhich extracts the maximum value in the pool size. This reduction in dimensionality makes iteasier in terms of computation, secondly, reduces the complexity of the network in a step by step Manner hence minimizing over-fitting.

The architecture then introduces a second convolutional block, of architecture similar to the first block, but with its own set of the number of filters (num_filters2). This block brings more amountof detailing in the feature extraction process and extracts more detailed features in the given dataset. The output of this block is again provided to batches normalization and max-pooling layer, going by the same rules as in the first block.

Following the convolutional layers, the architecture goes from the feature extraction to the classification stage. The max-pooling layer is the last layer before the result is flattened into a one-dimensional vector to feed to the fully connected or the dense layers. The first fully connectedlayer is composed of some specific number of neurons (dense_neurons1), which elaborates the extracted features with the help of ReLU activation function. To avoid the issue of overfitting, regularization is used in the network, and in this case it is dropout layer, it means that during training certain amount of input units will be set to zero at random.

This is succeeded by another dense layer with another collection of neurons (dense_neurons2) and a ReLU activation function which further improves the learning results. As you might expectin any deep learning algorithm, another dropout layer is applied here as well. The final layer exactly one node and it has sigmoid function in other words this layer of the CNN maps the learned features into the probability value from 0 to 1. This output is rather designed for the caseof binary classification problems, that is, the presence or absence of some condition. In the compounding of the model, the Adam optimizer is employed so that it could give efficient and pleasing results in training deep models. The loss function used is the binary cross-entropy as this is suitable in binary classification issues. Moreover, the accuracy measure is maintained onthe model in question.

To enhance the training process, two callbacks are utilized: A) ReduceLROnPlateau, which is a type of scheduler that reduces the learning rate when network's validation loss stops lowering and Early Stopping which terminates the training process when the validation loss isn't improvinganymore. These callbacks are useful in adjusting the learning process so that the model can be learned in the best way possible without overemphasizing on the training data or at the same timenot put too much emphasis on the training data.

Table 2: CNN Architecture Parameters and Values

| **Parameter** | **Value** |
|---|---|
| Input Shape | (30, 1) |
| Number of Filters (Layer 1) | 64 |
| Kernel Size (Layer 1) | 3 |
| Activation Function (Layer 1) | ReLU |
| Padding | Same |
| Batch Normalization (Layer 1) | Applied |
| Max Pooling (Layer 1) | Pool Size: 2 |
| Number of Filters (Layer 2) | 128 |
| Kernel Size (Layer 2) | 3 |
| Activation Function (Layer 2) | ReLU |
| Padding | Same |
| Batch Normalization (Layer 2) | Applied |
| Max Pooling (Layer 2) | Pool Size: 2 |
| Flatten | Applied |
| Number of Neurons (Dense Layer 1) | 128 |
| Activation Function (Dense Layer 1) | ReLU |
| Dropout (Dense Layer 1) | 0.5 |
| Number of Neurons (Dense Layer 2) | 64 |
| Activation Function (Dense Layer 2) | ReLU |
| Dropout (Dense Layer 2) | 0.5 |
| Output Layer | 1 Neuron (Sigmoid Activation) |
| Optimizer | Adam |
| Loss Function | Binary Cross-Entropy |
| Metrics | Accuracy |
| Learning Rate Adjustment | ReduceLROnPlateau |
| Early Stopping | Patience: 7 epochs |

## 3.4 Snake Optimization Algorithm

Newly developed in 2022, the Snake Optimization Algorithm (SOA) (Han et al. 2024) accomplished an unprecedented breakthrough that claimed a pioneering status in the realm of meta-heuristic algorithms and was deemed the first member of the said algorithm family. The above-described discovery, derived from elaboration snake's only mating-

**Journal of Information Systems Engineering and Management**
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
**Research Article**

related behaviors, are more likely to be achieved in cold climatic conditions and when snakes are no longer motivated by prey killing dueto satiating them in natural habitat. In contrast to all previously described methods, the formin which the entire operational procedure of the optimization begins is by randomly creating a given number of the candidate solutions captured in Equation (1). Therefore, the novelty in the form of the above-descripted phase indicates that the operational optimization algorithm enablesa novel approach that mirrors the solution exploration process by imitating natural and cognitive-behavioral aspects of snakes' actions in their original habitat thus introducing a potentially gamechanging new dimension of the toolkit of the optimization.

Unique from the others, this strategy first divided the population into two equal halves of males and females properly using the equations:

$$N_{\text{male}} = \frac{N}{2}, \quad N_f = N - N_{\text{male}} \quad (1)$$

where $N$ represents the total population size, $N_{\text{male}}$ denotes the number of male individuals, and$N_f$ signifies the number of female individuals. During each iteration, the algorithm identifies the optimal individual candidate solution, or the "food position" $f_{\text{food}}$, by assessing the best performers within each gender group, namely the best male $f_{\text{best, male}}$ and best female $f_{\text{best, female}}$.

The Temperature $T$ and Food Quantity $FQ$ parameters are defined as follows:

$$T = \exp\left(-\frac{g}{T}\right), \qquad FQ = c_1 \exp\left(\frac{g-T}{T}\right) \qquad (2)$$

In this context, $g$ denotes the current iteration, $T$ represents the total number of iterations, and $c_1$ is a constant with a value of 0.5. The search behavior of the snakes is dictated by the Food Quantity (FQ). A search for food is attempted if the frequency quality falls below a certain threshold = 0.5. The snakes must pick a random position within the food and update it.

The mathematical representation of the snakes' explorative male and female behavior is for male snakes:

$$x_{i,j}(g+1) = x_{\text{rand}\in[1,N/2],j}(g) \pm c_2 \times A_{i,\text{male}}\left((UB - LB) \times \text{rand} \in U(0,1) + LB\right), \quad (3)$$

where

$$A_{i,\text{male}} = \exp\left(-\frac{f_{\text{rand,male}}}{f_{i,\text{male}}}\right), \quad (4)$$

with $x_{i,j}$ denoting the position of the $i$th male snake, and $A_{i,\text{male}}$ representing the male'scapability to locate food.

For female snakes:

$$x_{i,j}(g+1) = x_{\text{rand}\in[1,N/2],j}(g+1) \pm c_2 \times A_{i,\text{female}}\left((UB - LB) \times \text{rand} \in U(0,1) + LB\right), \quad (5)$$

where

$$A_{i,\text{female}} = \exp\left(-\frac{f_{\text{rand,female}}}{f_{i,\text{female}}}\right), \quad (6)$$

indicating the female's proficiency in finding food.

This phase of exploitation, the SO algorithm implements two conditions to find the optimalsolutions:

If $FQ$ < Threshold and $T > 0.6$, the movement towards food is modeled as:

$$x_{i,j}\,(g+1) = x_{\text{food}} \pm c_3 \times T \times \text{rand} \times (x_{\text{food}} - x_{i,j}\,(g)) \tag{7}$$

Below the threshold ($FQ$ < Threshold and $T < 0.6$), snakes engage in either fighting or matingmodes, delineated as follows:

Fighting mode for male snakes involves:

$$x_{i,j}\,(g+1) = x_{i,j}\,(g) \pm c_3 \times F_{i,\text{male}} \times \text{rand} \times (x_{\text{best, female}} - x_{i,\text{male}}\,(g)) \tag{8}$$

Where

$F i,\text{male}$

$= \exp \;\; -\underline{f\text{best,f}}$ (9)

Denotes the fighting capability of the male snake. $f_i$

Mating mode allows for position updates based on mating ability, defined for males and females with respective equations show casing the interaction based on a mutual attraction quantified by the mating ability.

This detailed exposition and mathematical formulation of the Snake Optimization Algorithm not only elucidate its operational mechanics but also highlight its novelty in mimicking natural behaviors to solve optimization problems.

---

**Algorithm 1** Pseudocode for Snake Optimization Algorithm (SOA)

```
Input: Population size $N$, Maximum iterations $T$, Threshold values for $FQ$ and $T$, Upper and lower bounds $UB$ and $LB$
Output: Best solution $x_{\text{best}}$ and its fitness value
Initialize the population of snakes with $N$ candidates.
Divide the population into two halves: $N_{\text{male}}$ and $N_{\text{female}}$
for each iteration $g$ from 1 to $T$ do
    Evaluate the fitness of all individuals in the population
    Identify the best male snake $f_{\text{best,male}}$ and the best female snake $f_{\text{best,female}}$
    Determine the food position $f_{\text{food}}$ as the best overall candidate
    Calculate Temperature $T$ and Food Quantity $FQ$:
$T = e^{-g}$                                                   (10)
$FQ = C_1 . e^{(\frac{g}{T})}$                         (11)
    if $FQ$ < Threshold and $T > 0.6$ then
        Update the position of snakes toward the food:
```

```
            $x_{i,j}(g+1) = x_{\text{food}} \pm c_3 \times T \times \text{rand} \times (x_{\text{food}} - x_{i,j}(g))$    (12)
        else
            if $FQ$ < Threshold and $T < 0.6$ then
            Snakes engage in either fighting or mating:
            Fighting mode for males:
                $x_{i,j}(g+1) = x_{i,j}(g) \pm c_3 \times F_{i,\text{male}} \times \text{rand} \times (x_{\text{best, female}} - x_{i,\text{male}}(g))$    (13)
                Mating mode: Update the position of the new offspring based on mutual attraction.
            end if
        end if
    end for
    Return the best solution $x_{\text{best}}$ and its corresponding fitness value
```

The algorithm begins with the Initialization phase, where a random population of male and female snakes is created. This phase also involves settinginitial parameters such as $C1$, $C2$, $C3$, the threshold values, and $Q$, as well as evaluating the initial fitness values for each snake.

The population is then divided into two equal groups: male snakes ($Xm$) and female snakes ($Xf$). The algorithm proceeds into an iterative loop where $t++$ denotes the iteration counter being incremented in each cycle.

For each iteration, the algorithm evaluates each snake in the groups $N_{\text{male}}$ and $N_{\text{female}}$ to identify the best-performing male snake ($X_{\text{bestm}}$) and female snake ($X_{\text{bestf}}$). Following this evaluation, the ***Ɫemperature*** and ***Quantity*** parameters are defined.

The algorithm then assesses the Food Quantity ($FQ$) to determine if it is below a certain threshold, which triggers the ***Exploration*** phase. If $FQ$ is below the threshold, the snakes engagein a random search where their positions are updated accordingly.

If the temperature $T$ exceeds a predefined threshold, indicating hot conditions, the algorithm moves to check if the random number generator's output is greater than 0.6, which would activatethe ***Fight Mode.***In Fight Mode, the positions of male and female snakes are updated based onthe best solutions found in the opposite gender.

If neither of these conditions is met, the algorithm enters ***Mating Mode,*** where male and female snakes update their positions based on their partners' positions. Additionally, the algorithm may randomly replace the worst-performing male and female snakes with new positions within the search space.

The loop continues until the iteration counter $t$ reaches the maximum number of iterations $T$, at which point the algorithm terminates, yielding the best solution found.

## 4 Experiment Results

According to the Table 3, the outcomes of the experiment on optimizing the parameters of the employed One-Dimensional Convolutional Neural Network using the Snake Optimization Algorithm were highlypositive throughout the various runs. With the overall objectives of the experiment in mind, the Optimization process was conducted thrice to ensure the validity and generalizability of the results. This was because each run produced distinct optimal parameters and their corresponding validation accuracies. In the first run, the optimal configuration was established with 53 filters in the first convolutional layer, 18 filters in the second convolutional layer, kernel size of 3, pooling size of3, and 126 and 84 neurons in the first and second dense layers respectively. The configured modelprovided a validation accuracy of 0.9958, which was an illustrious model. In the second run, on the other hand the optimal parameters identified were 40 filters in the first convolutional layer, 103 in the second convolutional layer, kernel size of 3, pooling size of 2, 19 and 114 neurons in thefirst and second dense layers respectively. The configuration was linked to a validation accuracy of 0.9925. Finally, the optimal parameters from the third run involved 104 and 98 filters in the first and second convolutional layer respectively, kernel size of 3, pooling size of 2, and 24 and 114neurons in the first and second dense layers respectively. The configured model achieved a validation accuracy of 0.9956. Consequently, after a comprehensive comparison of the performancesin the three runs, the best parameters were identified as the ones in the first run. This model configuration was associated with the highest validation accuracy of 0.9958. The performance of the final model was also tested using the confusion matrix that clearly demonstrated the model's ability to classify the data correctly. In that view, it noted that the model correctly classified all the positive and negative instances, which indicated the model's ability to precisely distinguish between the various classes. The findings of this experiment reveal that while the Snake Optimization Algorithm enhanced the performance of the deep learning model, multiple optimizations run improved its performance hence the need to constantly optimize the model in the iterations.

The "Run" column depicts the series of experiments or iterations carrying out to test various hyperparameters simultaneously. A single run marks one set of hyperparameters of the Convolutional Neural Network (CNN) model to be used with the training data. The objective is to identify which configuration proves the most effective in regard to accuracy.

The CIFAR-10 classification performance of the proposed CNN model with the difference Readers' visualization examinations and the difference LQ-SSIM values is shown in Table 2, including the "Filters 1" and "Filters 2" that represent the number of the convolutional filters in the first and the second convolutional layers, respectively. The filters are crucial parts of the convolution layers because the filters permit the extraction of significant features from the input information, including patterns or edges. Since the focusing of the filters is arithmetic, it is achievable to arrive at the best number of filters required by exploring the model.

For layers that are Convolutional layers the "Kernel Size" is the size of the conversation kernel or filter that is used. It convolves over the input data by sliding its kernels with operations of basic algebra consisting of matrix multiplication and addition. Size of the kernel determines its receptive field; these checks how much of the input the filter is able to examine at one time.

In the max-pooling layers, "Pool Size" represents the size of the pooling window. By using pooling layers, dimensions of the data are reduced while the most important features are preserved, which in turn decrease the computational load and addresses over fitting issue. It is seen that the size of the pool controls the downsampling factor of the feature maps.

The last two columns; "# Neurons Layer 1" and "#Neurons Layer 2," represent the number of neurons in the first and the second fully connected layer or also known as a dense layer in the artificial neural network . These layers are actually capable of learning high level representations and how different features extracted by the convolutional layers are to be combined. The number of neurons in these layers can be adjusted to improve specification of complex relationships within the set data.

The meaning of the "Batch size" column is the number of training samples passed forward and backward pass during the training process. A smaller batch size uses fewer memory amounts and results in more updates while a large batch size leads to smoother gradients which converge more quickly. Batch size is a critical hyperparamenters that may affect the training and the final model.

Last, the "Accuracy" column represents the validation accuracy that the model came up with for the specific hyperparameters specification. This accuracy expresses the capability of the model in correctly predict the output to unseen data hence the generalization of the model. The configuration with the highest testing accuracy was also the best one; this proves that the developed model can successfully predict the value of the target variable.

Table 3: The Hyper-parameters optimization using Snake Optimization Algorithm

| Run | Filters 1 | Filters 2 | Kernel Size | Pool Size | layer 1(#neurons) | layer 2(#neurons) | Batch size | Accuracy |
|---|---|---|---|---|---|---|---|---|
| 1 | 53*53 | 18*18 | 3*3 | 3*3 | 126 | 84 | 16 | 0.9958 |
| 2 | 40*40 | 103*103 | 3*3 | 2*2 | 19 | 114 | 16 | 0.9914 |
| 3 | 104*104 | 98*98 | 3*3 | 2*2 | 24 | 114 | 16 | 0.9956 |
| **Best** | **53*53** | **18*18** | **3*3** | **3*3** | **126** | **84** | **16** | **0.9958** |

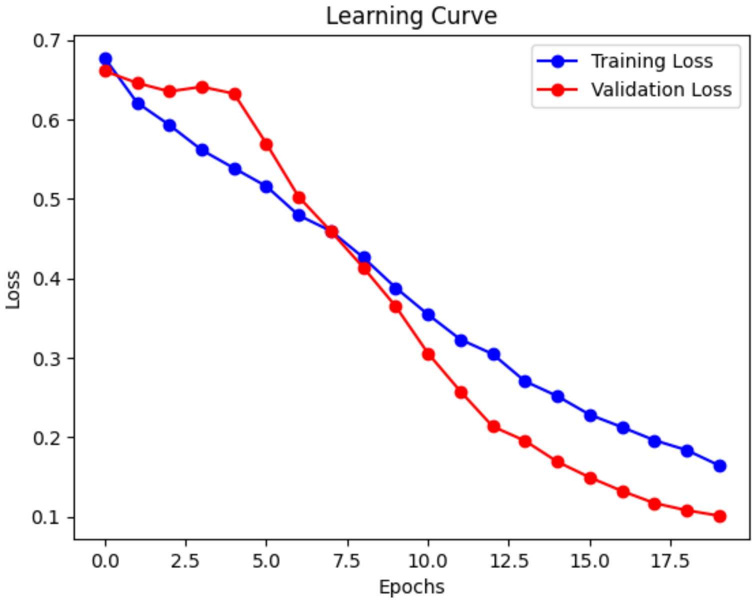


Figure 13: Learning Curve

Figure 13: Loss Learning Curve — Plotting the training loss and validation loss of a model on multi-epochs period Axis X is for a number of the epoch and axis Y is loss value. Understanding how well the model is fitting to training data (Training Loss) and generalizing to unseen data (Accuracy,Learning Curve). The training loss and validation loss both decrease towards the end as number of epochs add up in model. Fit () which shows that our model is learning nicely. The validation loss drops below the training loss after halfway through, which means that we are learning valid generalizations to new data. As training proceeds, both losses come together and the validation loss plateaus at a low level representing that our model has learnt to fit the trained data as well generalize on new data. This trend is characteristic of a model that has been well optimized and does not have large overfitting.

The proposed model takes an industrial strength 1D-CNN model and optimizes it with the Snake Optimization Algorithm (SOA). The results graphically shown of The SOA optimized 1D-CNN model P-T is much better than previous results, for example At 99.58% of the test set, the validation accuracy we achieved with our model is significantly higher in related works that camebefore it. For instance, models in [19] achieved accuracies of 93.33%, but these studies used other machine learning techniques, which might not have been as suitable for the specific data setor the problem domain. The ROC and Precision-Recall curves show what our model achieves in terms of balancing complexity with accuracy, something that is really appropriate for deep learning architectures like this. In addition- check out how stable that learning curve looks! This means the model generalizes well and does not fall into the common traps of overfitting; they were two challenges posed by [22] and [24] which resulting from minor errors in them. By integrating the SOA with traditional hyperparameter tuning, this study presents a novel method of refining your regression results that may transform what the future will be for cervical cancer detection work.

Table 4: Comparison of Related Work with the Proposed Model

| Study | Model/Algorithm | Accuracy (%) | Key Features/Advantages |
|---|---|---|---|
| [19] | Decision Tree, Random Forest, XGBoost | 93.33 | Feature significance highlighted |
| [22] | Voting Classifier, SMOTE, PCA | 94.94 | Addressed dataset imbalance, PCA for efficiency |
| [23] | Boruta Analysis with SVM | 98.72 | Enhanced feature selection using Boruta |
| [24] | XGBoost, Feature Selection Methods | 94.94 | Combination of feature selection algorithms |
| **Proposed Model** | **1D-CNN + SOA** | **99.58** | **Superior optimization, Balanced complexity and accuracy** |

## 5 Conclusion

In this study, we proposed an improved 1D Convolutional Neural Network Model with Snake Optimization Algorithm for the purpose of predicting cervical cancer risks. The combined methodology has also been well investigated in order to evaluate the effectiveness of the proposed model with the present methods mentioned in the literature. When implementing SOA in the CNN model it wasobserved that the adaptive hyperparameters enhanced the accuracy of the network to a greater precision.

The proposed model obtained, during the model validation process, an accuracy of almost 99.58%, which is far better than traditional machine learning algorithms as well as some of the recent deep learning models explored in similar studies. This is to further demonstrate how effectively the SOA can work particularly in solving intricate neural networks and further to its possibility inother domains such as medical diagnosis.

Also, the comprehensive Exploratory Data Analysis (EDA) allowed us to get the insights on the given set, which helped in the further preprocessing and feature extraction. The detailed exploration of the datasets, explicit complex optimization processes coupled with the state of theart deep learning network topology clearly depicts the holistic method used here.

Therefore, for better accuracy and reliability of cervical cancer risk prediction the proposed 1D-CNN model is optimized with SOA. The results not only prove that our approach works but also imply that such optimization techniques might be further applied to other fields that requireclassification to be crucial. Future work will be on fine-tuning the model and identifying other optimization methods for the same and extending the methodology to other types of medical datato facilitate early diagnosis and treatment.

## References

[1] Pan Huang, Shuailei Zhang, Min Li, Jing Wang, Cailing Ma, Bowei Wang, and Xiaoyi Lv. Classification of cervical biopsy images based on lasso and el-svm. *IEEE Access*, 8:24219– 24228, 2020.

[2] Ying Peng, Quan Kang, Qing Luo, Wei Jiang, Weike Si, Bernard A Liu, Hue H Luu, Jong Kyung Park, Xinmin Li, Jeffrey Luo, and et al. Inhibitor of dna binding/differentiationhelix-loop-helix proteins mediate bone morphogenetic protein-induced osteoblast differentia- tion of mesenchymal stem cells. *Journal of Biological Chemistry*, 279(31):32941–32949, 2004.

[3] Md Mamunur Rahaman, Chen Li, Xiangchen Wu, Yudong Yao, Zhijie Hu, Tao Jiang, XiaoyanLi, and Shouliang Qi. A survey for cervical cytopathology image analysis using deep learning.*IEEE Access*, 8:61687–61710, 2020.

[4] D. Banerjee, S. Mittal, R. Mandal, and P. Basu. Screening technologies for cervical cancer: Overview. *Cytojournal*, 19, 2022.

[5] M. T. Rezende, A. G. Bianchi, and C. M. Carneiro. Cervical cancer: Automation of pap test screening. *Diagnostic cytopathology*, 49(4):559–574, 2021.

[6] M. A. Swid and S. E. Monaco. Should screening for cervical cancer go to primary human papillomavirus testing and eliminate cytology? *Modern Pathology*, 35(7):858–864, 2022.

[7] A. C. Chrysostomou and L. G. Kostrikis. Methodologies of primary hpv testing currently applied for cervical cancer screening. *Life*, 10(11):290, 2020.

[8] J. D. Quinlan. Human papillomavirus: screening, testing, and prevention. *American Family Physician*, 104(2):152–159, 2021.

[9] Alshourbaji, Ibrahim & Jabbari, Abdoh & Rizwan, Shaik & Mehanawi, Mostafa & Mansur, Phiros & Abdalraheem, Mohammed. (2025). An Improved Ant Colony Optimization to Uncover Customer Characteristics for Churn Prediction. Computational Journal of Mathematical and Statistical Sciences. 4. 17-40. 10.21608/cjmss.2024.298501.1059.

[10] Alijoyo, F. A., Prabha, B., Aarif, M., Fatma, G., & Rao, V. S. (2024, July). Blockchain-Based Secure Data Sharing Algorithms for Cognitive Decision Management. In *2024 International Conference on Electrical, Computer and Energy Technologies (ICECET* (pp. 1-6). IEEE.

[11] A. Mitra, Deepika, V. Ammu, R. Chowdhury, P. Kumar and G. E, "An Adaptive Cloud and Internet of Things-Based Disease Detection Approach for Secure Healthcare system," *2024 International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS)*, Hassan, India, 2024, pp. 1-7, doi: 10.1109/IACIS61494.2024.10721944.

[12] F. A. Alijoyo, B. Prabha, M. Aarif, G. Fatma, V. S. Rao and P. Valavan M, "Blockchain-Based Secure Data Sharing Algorithms for Cognitive Decision Management," 2024 International Conference on Electrical, Computer and Energy Technologies (ICECET, Sydney, Australia, 2024, pp. 1-6, doi: 10.1109/ICECET61485.2024.10698611.

[13] Al-Shourbaji, I., & Al-Janabi, S. (2017). Intrusion Detection and Prevention Systems in Wireless Networks. *Kurdistan Journal of Applied Research*, *2*(3), 267-272. https://doi.org/10.24017/science.2017.3.48

[14] Kalpurniya, S., Ramachandran, R., & Chandramohan, N. (2023). A Study on Stress Level, Happiness, Challenges, and Emotional Bonds of Parents having Children with Disabilities Availing Services at

[15] NIEPMD, Chennai. *Integrated Journal for Research in Arts and Humanities*, *3*(5), 72-88.

[16] Alshourbaji, Ibrahim. (2013). Wireless Intrusion Detection Systems (WIDS). International Journal for Housing Science and Its Applications. Vol. 2.

[17] Singh, A., & Ramachandran, R. (2014). Study on the effectiveness of smart board technology in improving the psychological processes of students with learning disability. *Sai Om Journal of Arts & Education*, *1*(4), 1-6.

[18] Ahamad, Shakeel & Alshourbaji, Ibrahim & Al-Janabi, Samaher. (2016). A secure NFC mobile payment protocol based on biometrics with formal verification. International Journal of Internet Technology and Secured Transactions. 6. 103. 10.1504/IJITST.2016.078579.

[19] T. M. Jenkins, H. Shojaei, S. J. Song, and L. E. Schwartz. Role of ancillary techniquesin cervical biopsy and endocervical curettage specimens as follow-up to papanicolaou test results indicating a diagnosis of atypical squamous cells, cannot exclude high-grade squamous intraepithelial lesion, or high-grade squamous intraepithelial lesion. *Acta Cytologica*, 64(1- 2):155–165, 2020.

[20] A. S. Nwagwu, N. P. Ogbonnaya, and S. C. Ihudiebube. Understanding pap smears, hpv, and preventive measures in relation to cervical health. *Journal of Nursing and Healthcare*, 9(1):18–24, 2024.

[21] D. Vrdoljak-Mozetič, S. Štemberger Papić, D. Verša Ostojić, R. Rubeša, M. Klarić, and S. Emi-nović. Pitfalls in gynecological cytology: Review of the common and less frequent entities in pap test. *Acta Cytologica*, pages 1–18, 2024.

[22] K. S. Okunade. Human papillomavirus and cervical cancer. *Journal of Obstetrics and Gy- naecology*, 40(5):602–608, 2020.

[23] Matthias Begoihn, Assefa Mathewos, Abreha Aynalem, Tigeneh Wondemagegnehu, Ulrike Moelle, Muluken Gizaw, Andreas Wienke, Christoph Thomssen, Dawit Worku, Adamu Ad- dissie, and et al. Cervical cancer in ethiopia–predictors of advanced stage and prolonged timeto diagnosis. *Infectious Agents and Cancer*, 14:1–7, 2019.

[24] N. S. Wusiebia, O. S. Odio, and R. N. Armah. Assessment of the knowledge of women towards-cervical cancer in africa. *Journal of Global Health*, 2024.

[25] B. Judy. Interpersonal communication as a tool for enhancing women's uptake of cervical cancer screening in uasin gishu county, kenya. *International Journal of Communication*, 2023.

[26] E. Asempah. Hpv vaccine access and cervical cancer policymaking process: A comparative governmental priority setting study of ghana, rwanda, and canada. *Global Health Policy*, 2023.

[27] A. N. Mumuni, F. Hasford, N. I. Udeme, M. O. Dada, and B. O. Awojoyogbe. A swot analysis of artificial intelligence in diagnostic imaging in the developing world: Making a case for a paradigm shift. *Physical Sciences Reviews*, 9(1):443–476, 2024.

[28] M. O. A. Benedict. *Strategies to Enhance the Approach to Prostate Cancer Screening of African Men in the Free State*. PhD thesis, University of the Free State, 2023.

[29] Laboni Akter, Md Milon Islam, Mabrook S Al-Rakhami, and Md Rezwanul Haque. Predictionof cervical cancer from behavior risk using machine learning techniques. *SN Computer Science*,2:1–10, 2021.

[30] Naif Al Mudawi and Abdulwahab Alazeb. A model for predicting cervical cancer using machine learning algorithms. *Sensors*, 22(11):4132, 2022.

[31] Muhammed Fahri Unlersen, Kadir Sabanci, and Muciz Özcan. Determining cervical cancer possibility by using machine learning methods. *Int. J. Latest Res. Eng. Technol*, 3(12):65–71, 2017.

[32] Riham Alsmariy, Graham Healy, and Hoda Abdelhafez. Predicting cervical cancer using machine learning methods. *International Journal of Advanced Computer Science and Appli- cations*, 11(7), 2020.

[33] Umesh Kumar Lilhore, M Poongodi, Amandeep Kaur, Sarita Simaiya, Abeer D Algarni, Hela Elmannai, V Vijayakumar, Godwin Brown Tunze, Mounir Hamdi, and et al. Hybrid model for detection of cervical cancer using causal analysis and machine learning techniques. *Computational and Mathematical Methods in Medicine*, 2022.

[34] Gaurav Kumawat, Santosh Kumar Vishwakarma, Prasun Chakrabarti, Pankaj Chittora, Tu- lika Chakrabarti, and Jerry Chun-Wei Lin. Prognosis of cervical cancer disease by applying machine learning techniques. *Journal of Circuits, Systems and Computers*, 32(01):2350019, 2023.

[35] Jesse Jeremiah Tanimu, Mohamed Hamada, Mohammed Hassan, Habeebah Kakudi, and John Oladunjoye Abiodun. A machine learning method for classification of cervical cancer. *Electronics*, 11(3):463, 2022.

[36] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for im- age recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.