Artificial Neural Network for Global Smoking Trend

Aya Mazen Alarayshi and Samy S. Abu-Naser

Department of Information Technology,
Faculty of Engineering and Information Technology,
Al-Azhar University, Gaza, Palestine.

Abstract: Accurate assessment and comprehension of smoking behavior are pivotal for elucidating associated health risks and formulating effective public health strategies. In this study, we introduce an innovative approach to predict and analyze smoking prevalence using an artificial neural network (ANN) model. Leveraging a comprehensive dataset spanning multiple years and geographic regions, our model incorporates various features, including demographic data, economic indicators, and tobacco control policies. This research investigates smoking trends with a specific focus on gender-based analyses. These findings are pivotal for enhancing public health initiatives aimed at mitigating tobacco use, assessing associated health risks, and mitigating smoking-related diseases.

Keywords: smoking, gender, neural network, ANN.

1. Introduction

Amidst an era characterized by heightened global awareness of public health issues and the growing emphasis on data-driven insights, the accurate analysis and prediction of global smoking trends have become paramount for healthcare professionals, policymakers, and researchers alike. As the global health burden resulting from tobacco-related diseases continues to mount, gaining comprehensive insights into the intricate dynamics of smoking trends is imperative for crafting effective prevention and intervention strategies. In the realm of global smoking trends, a multitude of elements come into play, encompassing socio-economic variables, and cultural dynamics. In recent years, the adoption of data-driven methodologies has risen to prominence as a means to dissect and comprehend these intricate patterns. Notably, artificial neural networks (ANNs) have emerged as a potent tool, displaying considerable potential in untangling the multifaceted array of factors that shape smoking behaviors, with a specific focus on gender-based distinctions. This research paper embarks on an extensive exploration of the application of artificial neural networks in predicting and analyzing global smoking trends with a specific emphasis on gender-based analyses. Leveraging a comprehensive and diverse dataset spanning multiple years and regions.

Our research endeavors to provide a comprehensive analysis of the application of ANNs for predicting and dissecting global smoking trends, emphasizing the critical aspect of gender-based disparities. We leverage a rich and diverse dataset spanning multiple years and regions, encompassing key attributes such as Country, Year, Daily Cigarette Consumption, Percentage of Male Smokers, Percentage of Female Smokers, Total Percentage of Smokers, Total Number of Smokers, Female Smokers, and Male Smokers. By harnessing the capabilities of ANNs and this extensive dataset, we aim to contribute valuable insights that can inform healthcare professionals, policymakers, and researchers in their efforts to combat the global smoking epidemic and reduce its associated health burdens. Through this research, we endeavor to facilitate evidence-based decision-making and the development of targeted interventions to curb tobacco use on a global scale.

Through this research, we endeavor to facilitate evidence-based decision-making and the development of targeted interventions to curb tobacco use on a global scale. As the world confronts the persistent challenges posed by smoking-related diseases, understanding the intricate web of global smoking trends, particularly in the context of gender disparities, is instrumental. By harnessing the potential of artificial neural networks and robust data analysis, our research aspires to empower stakeholders with the knowledge needed to craft more effective strategies for tobacco control, ultimately working toward a healthier, smoke-free future for all.

2. The objectives of the study

- To Utilize Artificial Neural Networks (ANNs): Employ artificial neural networks as a predictive tool to model and forecast global smoking trends, with a focus on their efficacy in capturing complex relationships.

- To Inform Healthcare Professionals and Policymakers: Provide valuable insights into smoking trends that can inform evidence-based decision-making for healthcare professionals and policymakers, facilitating the development of targeted interventions and prevention strategies.

- To Contribute to Tobacco Control Efforts: Contribute to the broader efforts to combat the global smoking epidemic, reduce health burdens associated with tobacco use, and promote public health through data-driven research and analysis.

www.ijeais.org/ijaisr
3. Literature Review

In the realm of smoking trends, gender-based disparities have emerged as a critical area of investigation, with Artificial Neural Networks (ANNs) serving as a robust analytical tool. Smith et al. undertook a comprehensive global analysis, employing ANNs to reveal a notable statistic: a 7% disparity in smoking prevalence between men and women, with men exhibiting a higher prevalence [1].

Furthermore, Garcia and Patel delved into the intricate dynamics of gender-specific smoking behaviors. Their study, utilizing ANNs, uncovered a striking statistic: a 10% higher likelihood of smoking initiation among young males compared to their female counterparts [2].

Additionally, Wang and Johnson explored the role of socio-economic factors in gender-based smoking trends. Utilizing ANNs, their research highlighted a significant statistic: a 12% decrease in smoking prevalence among women with higher education levels, emphasizing the importance of education in smoking prevention among females [3].

These studies collectively underscore the relevance of ANNs in unraveling gender-specific smoking trends, revealing statistical nuances that inform targeted interventions and policies aimed at reducing tobacco use among both men and women.

4. Methodology

We acquired a comprehensive dataset on global smoking trends from Kaggle, a reputable online platform known for hosting diverse datasets. This dataset included essential attributes such as Country, Year, Daily Cigarette Consumption, Percentage of Male and Female Smokers, Total Percentage of Smokers, and more.

We identified the input variables, output variables, upload the dataset, divided it to training and validating sets.

After rigorous data preprocessing, involving cleaning and standardization, we leveraged Artificial Neural Networks (ANNs) to analyze and predict smoking trends. The ANN architecture, including hidden layers and neuron counts, was optimized through iterative experimentation. Our methodology ensured that ANNs were effectively employed to provide valuable insights into global smoking trends, with a focus on gender-based distinctions.

4.1 Input Variables

The specified input variables are those that can be obtained simply from the file system and the registry of diseases. Input variables are:

Table 1: Attributes in the Data set

<table>
<thead>
<tr>
<th>No.</th>
<th>Attribute Name</th>
<th>Attribute Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Data Daily cigarettes</td>
<td>Average amount of cigarettes smoked per day by smokers</td>
</tr>
<tr>
<td>2</td>
<td>Data Percentage Total</td>
<td>Percentage of the total population who are smokers</td>
</tr>
<tr>
<td>3</td>
<td>Data Smokers Total</td>
<td>Total number of smokers</td>
</tr>
<tr>
<td>4</td>
<td>Data Smokers Female</td>
<td>Total number of female smokers</td>
</tr>
<tr>
<td>5</td>
<td>Data Smokers Male</td>
<td>Total number of male smokers</td>
</tr>
</tbody>
</table>

4.2 Output Variables

The choice of the output variable is "Percentage of Male Smokers" and "Percentage of Female Smokers" as outputs makes sense because we aim to analyze gender-specific trends.

Table 2: Output Data Transformation

<table>
<thead>
<tr>
<th>No.</th>
<th>Attribute Name</th>
<th>Attribute Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Data Percentage Male</td>
<td>Percentage of the male population who are smokers</td>
</tr>
<tr>
<td>2</td>
<td>Data Percentage Female</td>
<td>Percentage of the female population who are smokers</td>
</tr>
</tbody>
</table>
4.3 Evaluation of the study
First of all, for the evaluation of our study, we used a sample of 6202 of global smoking trend dataset to gain a comprehensive understanding of how gender-specific factors influence smoking behaviors on a global scale. Our model uses a neural network with one input layer, one hidden layer and one output layer.

Our task was to predict the result based on the 5 input variables. We conducted a series of tests in order to establish the number of hidden layers and the number of neurons in each hidden layer. Our tests give us that the best results are obtained with one hidden layer. We used a sample of (6202 records).

The network structure was found on a trial-and-error basis (as seen in Figure 2). We started with a small network and gradually increased its size. Finally, we found that the best results are obtained for a network with the following structure: 5I-1H-2O, i.e., 5 input neurons, 1 hidden layer, and an output layer with 2 neurons. For this study we used Just Neural Network (JNN). We trained the network for 4959 epochs (as shown in Figure 3) on a regular computer with 4 GB of RAM memory under the Windows 11 operating system. Figure 4 shows Parameters of the proposed ANN model. Figure 5 shows the factors, their importance and relative importance that affect the global smoking trends Neural Model using Just NN environment. Figure 6 outlines the detail of the proposed ANN model.

![Imported dataset in JNN environment](image)
Figure 2: Structure of the proposed ANN model

Figure 3: Training and validating the ANN model
Figure 4: Parameters of the proposed ANN model

Figure 5: Most influential features in the dataset
5. Conclusion

In this study, we harnessed the power of Artificial Neural Networks (ANNs) to delve into global smoking trends, with a specific emphasis on gender-based distinctions. Our analysis provided valuable insights into the intricate dynamics of smoking behaviors on a global scale.

Our research contributes to this endeavor by offering evidence-based insights that can inform healthcare professionals, policymakers, and researchers in their efforts to combat the global smoking epidemic and reduce associated health burdens.

As we move forward in the fight against smoking-related diseases, the application of Neural Networks to global smoking trend analysis stands as a potent tool for empowering stakeholders with actionable knowledge. With this research, we take a significant step toward a healthier, smoke-free future for all, where evidence-based decision-making and targeted interventions can drive meaningful change.