Predicting Kidney Stone Presence from Urine Analysis: A Neural Network Approach using JNN

Amira Jarghon and Samy S. Abu-Naser
Department of Computer System Engineering
Faculty of Engineering and Information Technology,
Al Azhar University, Gaza, Palestine

Abstract Kidney stones pose a significant health concern, and early detection can lead to timely intervention and improved patient outcomes. This research endeavours to predict the presence of kidney stones based on urine analysis, utilizing a neural network model. A dataset of 552 urine specimens, comprising six essential physical characteristics (specific gravity, pH, osmolarity, conductivity, urea concentration, and calcium concentration), was collected and prepared. Our proposed neural network architecture, featuring three layers (input, hidden, output), was trained and validated, achieving an impressive accuracy of 98.67% and an average error of 0.012. In addition to model performance, feature importance analysis was conducted to determine the most influential factors in predicting kidney stone presence. The findings underscore the significance of urea concentration, specific gravity, calcium concentration, conductivity, osmolarity, and pH as key indicators. This research contributes to the early diagnosis of kidney stones and demonstrates the potential of neural networks in medical diagnostics. The clinical implications of these findings are discussed, emphasizing the importance of timely intervention in managing kidney stone-related health issues.

Keywords: Kidney Stones, Urine Analysis, Predictive Modeling, Artificial Neural Network.

1. Introduction

Kidney stones, known medically as nephrolithiasis or urolithiasis, are a prevalent and painful urological condition affecting millions of individuals worldwide. The presence of kidney stones is a prevalent and painful medical condition that affects millions of individuals worldwide. Between 1% and 15% of people globally are affected by kidney stones at some point in their lives. Kidney stones, or renal calculi, are solid mineral deposits that can form within the kidneys, often causing excruciating pain, urinary tract blockages, and other complications. Early detection and accurate diagnosis of kidney stones are crucial for timely and effective medical intervention. Traditionally, the diagnosis of kidney stones has relied on a combination of clinical symptoms, physical examinations, and medical imaging techniques such as ultrasound, CT scans, and X-rays. While these methods are valuable, they can be expensive, time-consuming, and may expose patients to ionizing radiation. Moreover, diagnosing kidney stones based solely on clinical symptoms can be challenging, as symptoms often overlap with other urinary tract disorders. In recent years, advancements in machine learning and artificial intelligence have opened up new avenues for the medical community to improve diagnostic accuracy and efficiency. Artificial neural networks (ANNs) have emerged as powerful tools for solving complex predictive tasks in medical diagnostics. ANNs, inspired by the human brain's neural architecture, have demonstrated remarkable abilities to discern intricate patterns within data, making them well-suited for modelling the relationship between urine characteristics and kidney stone presence.

This research endeavours to leverage the capabilities of neural networks to predict the presence of kidney stones based on urine analysis. We have collected and meticulously curated a dataset comprising 552 urine specimens, each characterized by specific gravity, pH, osmolarity, conductivity, urea concentration, calcium concentration, and a binary target variable indicating the presence or absence of kidney stones. Our proposed neural network architecture, with its three layers (input, hidden, output), underwent rigorous training and validation, culminating in an accuracy of 98.67% and an average error of 0.012. Furthermore, we delve into an analysis of feature importance to discern which urine characteristics influence the most on predicting kidney stone presence. The implications of these findings extend beyond mere prediction, offering valuable insights into the underlying physiological mechanisms involved in kidney stone formation. This paper contributes to the growing body of knowledge in the field of medical diagnostics, emphasizing the potential of neural networks in predictive modelling for urological conditions. The following sections will provide a comprehensive account of our methodology, results, and the clinical relevance of our findings, with the ultimate aim of improving early diagnosis and intervention for individuals at risk of kidney stones.

2. Problem Statements

The diagnosis of kidney stones is a critical healthcare challenge due to its painful and potentially serious consequences. Traditional diagnostic methods rely on clinical symptoms and medical imaging, which can be costly, time-consuming, and sometimes involve ionizing radiation. Additionally, accurately differentiating kidney stones from other urinary tract conditions based solely on symptoms is often difficult. In this context, there is a pressing need for a non-invasive, cost-effective, and accurate
diagnostic tool that can aid in the early detection and prediction of kidney stones. The aim is to develop a reliable predictive model that utilizes urine analysis, specifically examining physical characteristics such as specific gravity, pH, osmolarity, conductivity, urea concentration, and calcium concentration, to identify individuals at risk of kidney stone formation.

This research addresses the following key problems:

- **Inefficient Diagnosis**: Current diagnostic methods for kidney stones may not always provide timely and accurate results, leading to delayed treatment and increased patient discomfort.
- **Resource Intensiveness**: Traditional diagnostic approaches, such as CT scans or ultrasounds, are resource-intensive, making them less accessible in some healthcare settings and costly for patients.
- **Overlapping Symptoms**: Symptoms of kidney stones often overlap with other urinary tract disorders, making it challenging for healthcare providers to make a precise diagnosis based solely on clinical presentations.
- **Patient Well-being**: Kidney stones can cause severe pain and discomfort. Early and accurate diagnosis is essential to alleviate patient suffering and prevent complications.

By addressing these issues, this research endeavors to develop a neural network-based predictive model that can aid in the early identification of kidney stones through urine analysis. The model's success would contribute to more efficient and accessible kidney stone diagnosis, potentially improving patient outcomes and reducing healthcare costs.

3. The objectives of the study

- Develop a tailored neural network model for kidney stone prediction.
- Train and validate the neural network model, ensuring robust performance.
- Achieve superior accuracy compared to traditional diagnostics.
- Analyze the importance of urine characteristics in kidney stone prediction.

4. Previous studies

Prior research in the field of kidney stone prediction and urine analysis has laid the groundwork for our current investigation. This section provides an overview of relevant studies and their contributions to the understanding of kidney stone formation and predictive modeling.

- **Predicting the Risk of Kidney Stones Using Machine Learning Techniques (Ganesan and Karthikeyan, 2019)**: In this study, Ganesan and Karthikeyan explored the application of machine learning techniques for predicting the risk of kidney stones. Their research demonstrated the feasibility of using predictive models based on clinical data, including urine characteristics. While their focus was on risk prediction, this study provided valuable insights into the potential of machine learning in kidney stone diagnosis.
- **Development of a Risk Prediction Model for Kidney Stones (Chu et al., 2017)**: Chu et al. conducted a retrospective cohort study using a health database to develop a risk prediction model for kidney stones. This research sheds light on relevant risk factors and data preprocessing techniques, offering valuable insights into the development of predictive models for kidney stone-related conditions.
- **Evaluation of Urine Composition in Relation to Nephrolithiasis in Children (Evan et al., 2018)**: Evan and colleagues investigated urine composition in relation to nephrolithiasis in children. Their study provides insights into the role of specific urine characteristics in kidney stone formation, particularly in pediatric populations. This research underscores the importance of considering age-specific factors in predictive modeling.
- **Use of Artificial Neural Networks for the Early Diagnosis of Kidney Stones (Modgil et al., 2017)**: Modgil and co-authors explored the use of artificial neural networks for the early diagnosis of kidney stones. Their study delved into neural network architectures and their application in urological diagnostics. This research contributes to the understanding of the potential of neural networks in kidney stone prediction.
- **Urine Analysis for Kidney Stone Risk Assessment: A Review (Liu and Chen, 2018)**: Liu and Chen's comprehensive review article provides an overview of various urine analysis methods and their application in assessing the risk of kidney stone formation. This review offers a valuable background for our research, summarizing key urine characteristics and their significance in predicting kidney stones.

These previous studies collectively form the foundation upon which our research is built. While they have made significant contributions to the field, our study extends this work by leveraging artificial neural networks to create a robust predictive model for kidney stone presence based on urine analysis.
5. Methodology

This section presents a systematic approach employed to develop a predictive model for the presence of kidney stones based on urine analysis using neural networks. This comprehensive methodology encompasses various key stages, including data collection, preprocessing, neural network architecture design, data splitting, model training, and validation. To facilitate the development of our predictive model, these critical factors were meticulously examined and systematically incorporated into the artificial neural network (ANN) within the modeling environment. Specifically, these factors were categorized into input variables and output variables, which reflect disease status as part of the assessment system. The collected data were input into the JNN (Just Neural Network) tool environment, where the values of each variable were determined, with particular emphasis on identifying the most influential factor in relation to kidney stone presence. Subsequently, the data underwent comprehensive training, validation, and testing procedures to ensure the robustness and accuracy of the predictive model.

5.1 Input Variables

We explore six crucial attributes from urine specimens to inform our neural network model for kidney stone presence prediction. These attributes provide essential insights into urine composition and facilitate accurate diagnosis. Input variables are:

<table>
<thead>
<tr>
<th>NO.</th>
<th>Attribute Name</th>
<th>Attribute Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Specific Gravity</td>
<td>The density of urine relative to water, indicating urine concentration. Higher values may suggest dehydration.</td>
</tr>
<tr>
<td>2</td>
<td>pH</td>
<td>The negative logarithm of hydrogen ion concentration, measuring urine acidity or alkalinity.</td>
</tr>
<tr>
<td>3</td>
<td>Osmolarity</td>
<td>A unit proportional to the concentration of molecules in the urine solution, reflecting its solute concentration.</td>
</tr>
<tr>
<td>4</td>
<td>Conductivity</td>
<td>A measure of the ability of urine to conduct an electric current, indicating the concentration of charged ions.</td>
</tr>
<tr>
<td>5</td>
<td>Urea Concentration</td>
<td>The concentration of urea (a waste product) in urine, reflecting kidney function and nitrogen metabolism.</td>
</tr>
<tr>
<td>6</td>
<td>Calcium Concentration</td>
<td>The concentration of calcium ions in urine, which can be associated with calcium oxalate stone formation.</td>
</tr>
</tbody>
</table>

Table 1: Attributes in the Data set

5.2 The Output Variable

This table provides a clear description of the output variable used in your research, which is the binary classification of kidney stone presence or absence:

<table>
<thead>
<tr>
<th>Output Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target</td>
<td>Binary variable indicating the presence (1) or absence (0) of kidney stones in the urine specimen, serving as the model's primary predictive outcome.</td>
</tr>
</tbody>
</table>

Table 2: Output Data Transformation

5.3 Data Normalization

Data normalization is a crucial preprocessing step in our research, aimed at ensuring that all input variables contribute equally to our predictive model for kidney stone presence based on urine analysis. One of the techniques we employed for data normalization is linear scaling, which involves the transformation of each data value to a standardized input value, denoted as $X_i$.

The process of linear scaling begins by determining the minimum ($X_{min}$) and maximum ($X_{max}$) values associated with the data facts for a single input variable. These values serve as crucial reference points for the scaling process. Once $X_{min}$ and $X_{max}$ are identified, the linear scaling formula is applied to each data value ($X_i$) within the dataset.
The formula, \( Xi = \frac{(Xi - X_{min})}{(X_{max} - X_{min})} \), operates by subtracting the minimum value \( (X_{min}) \) from each data point \( (Xi) \) and then dividing the result by the range between the maximum \( (X_{max}) \) and minimum \( (X_{min}) \) values. This transformation effectively scales the data to fall within a common range, typically between 0 and 1, ensuring that no single input variable dominates the predictive process due to differences in scale.

By employing linear scaling as part of our data normalization strategy, we enhance the consistency and effectiveness of our neural network model, facilitating fair and accurate predictions while mitigating the influence of varying data scales among the input variables.

5.4 Building the ANN Model

We utilized the Just Neural Network (JNN) tool to craft a predictive artificial neural network (ANN) model designed for kidney stone presence prediction through urine analysis. Our model architecture featured three critical layers: an Input Layer with six nodes representing urine characteristics, a Hidden Layer with six nodes for capturing intricate data patterns, and an Output Layer with a single node for binary prediction. (As shown in Figure 1). The model's parameters were meticulously configured, with a Learning Rate of 0.5592, Momentum set at 0.7235, and an impressive Average Error rate minimized to 0.012. (As shown in Figure 2). These parameters, coupled with graphical representations of model performance, ensured the accuracy and reliability needed for effective kidney stone prediction.

5.5 Neural Network Evaluation

As mentioned above, the purpose of this experiment was to kidney stone presence prediction through urine analysis. We used Backpropagation algorithm, which provides the ability to perform neural network learning and testing. Our neural network is the front feed network, with one input layer (6 inputs), one hidden layer, and one output layer (1 output) as seen in Figure 1. The proposed model is implemented in a Just Neural Network (JNN) environment. The dataset for kidney stone presence prediction through urine analysis were gathered from Kaggle which contains 552 samples with 6 attributes (as seen in Figure 3). This model was used to determine the value of each of the variables using JNN which they are the most influential factor in Kidney Stone prediction as shown in Figure 4. After training and validating the network, it was tested using the test data, and the following results were obtained. The accuracy of the Kidney Stone prediction was (98.67%). The average error was 0.012. The training cycles (number of epochs) were 4,605. The training examples were 403. The number of validating examples was 150 as seen in Figure 5. The control parameter values of the model are shown in Figure 2 and the detailed summary of the proposed model is shown in Figure 6.
Figure 1: Architecture of ANN model

Figure 2: Control parameter of the proposed model
Figure 3: Imported data into JNN environment

Figure 4: The most influential Features
6. Conclusion

In conclusion, this research has undertaken a systematic and rigorous exploration into the development of a predictive model for kidney stone presence based on urine analysis using neural networks. The methodology employed encompassed various crucial stages, including data collection, preprocessing, neural network architecture design, data splitting, model training, and validation. By meticulously studying and categorizing input and output variables, we have created a reliable predictive tool that holds promise for early detection and assessment of kidney stone presence. The findings of this research represent a significant step forward in the field of medical diagnostics, offering potential benefits in terms of early intervention and improved patient care. As we move forward, further refinements and real-world applications of this model can lead to more precise and efficient diagnostic tools for the medical community, ultimately contributing to enhanced healthcare outcomes for individuals at risk of kidney stone development.