Classification of congestion problems in a telecommunications network using its performance indices

Borel SONNA  Laure OBAYA  Lauren MOUKODI  André NYEMB
BEL’S AI INITIATIVE Deep Learning Indaba 2023

We are BEL’S AI INITIATIVE

BEL’S AI INITIATIVE for BE EDUCATED and LEARN SKILLS in ARTIFICIAL INTELLIGENCE INITIATIVE is an association based in Cameroon dedicated to demystifying artificial intelligence in Africa. Our primary objective is to provide specialized content and training in artificial intelligence while initiating various projects aimed at promoting this discipline on the African continent.

UNDERSTANDING THE MACHINE LEARNING PROBLEM IN THIS PROJECT

• Congestion in a telecommunications network occurs when the amount of traffic or data in circulation exceeds the maximum capacity of the network to handle it efficiently. This leads to degraded performance, transmission delays and disruptions to data delivery, affecting the quality of service offered to users. In short, congestion occurs when the network is overloaded and cannot handle the volume of data flowing through it.
• That said, congestion problems depend on the performance indices of a network at a given time. So the data we will use to train our machine learning model on how to classify congestion problems in the network will be a big dataset of network’s performance indices when there is a congestion problem.
• Our dataset will be a set of data labelled for each record with the associated type of congestion. This is a supervised learning machine learning algorithm.
• Since the type of congestion is a qualitative data, we’re going to use a supervised learning classification algorithm.

In short, this is a supervised machine learning algorithm for multiclass classification of congestion problems in a telecommunications network using network performance indices.

DESCRIPTION OF THE DATASET

This dataset has been downloaded from the github platform via the following link https://github.com/CallMeAmartya/IITKgp-Interhall-Data-Analytics. Its features are as follows:
• In our dataset, we have 39 attributes or columns and 78560 rows or records.
• In our dataset we have 39 attributes, 37 of which are integers (these values are numeric) and the other two are objects that need to be numbered.
• No missing values
We have a fairly clean dataset with no missing values and with almost all the attributes corresponding to the performance indices already in numerical form. It should be noted that only the prediction variable, i.e., the type of congestion, is not numeric; nor is the variable corresponding to the equipment manufacturer who designed the telecommunication equipment on which we have collected the data.

The performance of the model is given by this table.

<table>
<thead>
<tr>
<th>Description</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>91.17%</td>
</tr>
<tr>
<td>Precision</td>
<td>90.78%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>90.78%</td>
</tr>
<tr>
<td>Specificity</td>
<td>90.78%</td>
</tr>
<tr>
<td>F1-score</td>
<td>90.78%</td>
</tr>
<tr>
<td>Recall</td>
<td>90.78%</td>
</tr>
<tr>
<td>Balanced accuracy</td>
<td>90.78%</td>
</tr>
</tbody>
</table>

In conclusion, BEL’S AI INITIATIVE is an association committed to promoting artificial intelligence in Africa. At this Deep Learning Indaba 2023 conference, we have decided to present one of our projects we worked on with our students. The project involves the design and implementation of a machine learning model that classifies the congestion experienced by telecommunications equipment in the network, using performance indices. We obtain a score of about 90%. It thus offers the possibility of ensuring predictive and preventive maintenance of telecommunications equipment for network operators.

Figure 1: Website link

Figure 2: Data repartition

Figure 3: learning curve

Figure 4: performance of the model