DATA MINING TOOL FOR JUDICIARY

Dr. Omda Kumar

Abstract- Information extraction process to determine patterns and facts buried in the large databases is being done by using Data Mining, these days. Data mining may help decision makers of Judiciary, in a well-organized manner to help in making quick and more appropriate decisions. The whole paper is divided into six sections. Section 1 describes the tool for judiciary. In section 2 Data Flow Diagram of Judiciary data mining tool has been shown. Section 3 describes algorithm used in the Judiciary data mining tool. Working of the data mining tool for Judiciary has been discussed in the Section 4. Results drawn by the data mining tool has been shown in the Section 5 and section discusses conclusion.

Keywords - Data Mining, Judiciary, Minimum Support, Confidence

1. INTRODUCTION
Data mining ensures optimization of resources and can be utilized in real time. Essential legal texts can be transformed into machine-readable data sets and the outcome / patterns, so discovered by using data mining, help in making opinion for any decision. Relevant provisions of the constitution and past precedents are considered for making judgments in the Judiciary. These are based on the provisions of codes, statutes and earlier judgments of similar nature. Data of Judiciary is available in the text format. A very huge amount of textual data needs to be processed for decision making, which is very difficult. In the judiciary, all this is being done manually, presently.

2. TOOL FOR JUDICIARY
A data mining tool has been designed to eliminate the above problem, which may help the judiciary in decision making. The legal data which is available in the text format, can be converted in electronic form, so that the same could be utilized by data mining tool for making any opinion for Judgments. Association technique has been used in present tool. This tool will help judiciary in discerning inferences from the data of judiciary. It works on databases having transactional datasets. Datasets of occurrences of various instances in judicial database are used in drawing of inference.

3. DATA FLOW DIAGRAM OF DATA MINING TOOL FOR JUDICIARY
The tool has been displayed in the following data flow diagram as shown in the figure-1, below:

Figure 1. Data flow diagram of Data Mining Tool for Judiciary

1 Chaudhary Devi Lal University, Sirsa, Haryana, India
Work flow of the data mining tool for judiciary has been shown in the above data flow diagram. The database of judgments, minimum support and confidence has been shown as input to the data mining tool. Then, the process of scanning the judicial database has been described. Process of calculation of minimum support and confidence has been outlined and in the last it shows the output in the shape of the best rules.

4. STEPS OF ALGORITHM USED IN THE JUDICIARY DATA MINING TOOL

Step 1: Begin
Step 2: Load Judgments database in form of XML file
Step 3: Input minimum support S
Step 4: Input minimum Confidence
Step 5: For item I = 1st item to I = Last Item
    Get the support of item i
    (End of for loop)
Step 6: If s >= Supportmin
    (Begin of If)
    6.1 Add to Frequent1 ITEMSETS, L1
    6.2 Use L(k-1) Join L(k-1) to generate a set of candidate k-itemsets
    6.3 Scan the judicial database to get the support S of each candidate k-itemset
    6.4 If (S >= SUPPORTmin)
        (Begin of If)
        6.4.1 Add to K-frequent itemset
        6.4.2 if (Generated Set = NULL)
            (Begin of If)
            6.4.2.1 For each frequent itemset L, generate all nonempty subsets of L
            6.4.2.2 For each nonempty subset S of L, find confidence C of S
            6.4.2.3 If (C >= CONFIDENCE)
                Add to Best Rules
            (End of if)
        (End of if)
    (End of if)
    Else
    { 
        Go to Step 6.2
    }
(End of if)
Step 7: End

The database of judgments is provided as input to data mining tool in the form of XML file, in the present problem. Then minimum support and confidence is provided as input. The tool scans the judicial data to get the support S of each item. It calculates the minimum confidence of itemsets and generates the best rules from the database.

5. WORKING OF THE JUDICIARY DATA MINING TOOL

To describe the working of the tool as shown in the Table-1, below, three attributes namely CauseOfClaim, SexOfPersonHurt and OffendingVehicle have been taken for the database of judiciary:

<table>
<thead>
<tr>
<th>Name of Attribute</th>
<th>Category</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CauseOfClaim</td>
<td>String</td>
<td>Death, Injuries</td>
</tr>
<tr>
<td>SexOfPersonHurt</td>
<td>String</td>
<td>Male, Female</td>
</tr>
<tr>
<td>OffendingVehicle</td>
<td>String</td>
<td>Motorcycle, Jeep</td>
</tr>
</tbody>
</table>

Number of instances taken for example = 5
Minimum Support = 2
Minimum Confidence = 40%
The dataset derived from judgments of various claim cases under Motor vehicle Act 1988 decided by the local court of Sirsa, has been analyzed through the above tool to illustrate the effectiveness of data mining in Judiciary. Microsoft Excel worksheet has been utilized to normalize the data and it has been converted to “.xml” file. C# has been used to implement the algorithm. For using the data mining tool, five attributes as mentioned in Table-3 have been taken from the dataset of judiciary. Abstract of dataset has been revealed in Table-2, given below:

Table-2: Abstract of dataset of Judgments of claims

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geographical Area of dataset</td>
<td>Sirsa district</td>
</tr>
<tr>
<td>Period</td>
<td>2004-2011</td>
</tr>
<tr>
<td>Total No. of instances</td>
<td>100</td>
</tr>
<tr>
<td>Missing values</td>
<td>No</td>
</tr>
</tbody>
</table>

Table-3: Attributes of Judgments of Claim’s

<table>
<thead>
<tr>
<th>Name of Attribute</th>
<th>Category</th>
<th>Range/Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CauseOfClaim</td>
<td>String</td>
<td>Death, Injuries and disabilities</td>
</tr>
<tr>
<td>SexOfPersonHurt</td>
<td>String</td>
<td>Male, Female</td>
</tr>
<tr>
<td>OffendingVehicle</td>
<td>String</td>
<td>OffendingMotorcycle, OffendingCar, OffendingJeep, OffendingTractor, OffendingTractor, OffendingBus, OffendingTractorTrolley, OffendingThreeWheeler</td>
</tr>
<tr>
<td>TimeofAccident</td>
<td>Numeric</td>
<td>1-24 (31-130 = 1, 131-230 = 2, 231-330 = 3, …….. 2331-0030 = 24,)</td>
</tr>
<tr>
<td>IllfatedVehicle</td>
<td>String</td>
<td>Motorcycle, Scooter, Car, Jeep, Tractor, Truck, Canter, Bus, TractorTrolley, ThreeWheeler, Bicycle, Pedestrians, AutoRikshaw, NeelGaiBlueBull, AceMagic</td>
</tr>
</tbody>
</table>

(Source: District Courts, Sirsa) [10]

In the above data, five attributes have been taken. The values as death, injuries & disabilities have been assigned to cause of the claim. Motorcycle, scooter, car, jeep, tractor, truck, canter, bus, tractor-trolley and three-wheeler are offending vehicle causing an accident. Time of accident has been observed round the clock and normalized as 0:31-1:30 = 1, 1:31-2:30 = 2, 2:31-3:30 = 3, …….. 23:31-00:30 = 24. Ill-fated vehicle/property affected in the accident are motorcycle, scooter, car, jeep, tractor, truck, canter, bus, tractor-trolley, three-wheeler, bicycle, pedestrians, auto-rikshaw, neelgai-bluebull, ace-magic.

6. RESULTS
Figure 3. Snapshot of the results drawn from the datasets of judiciary

Source: Primary.

The results drawn from the data of 100 claim cases decided by the courts in the district Sirsa under Motor Vehicle Act 1988 has been demonstrated in Figure-3 which depicts association among attributes. It shows minimum support as 10 and minimum confidence as 70 per cent.

Table-4: Analysis of results drawn from the datasets of Judgments

<table>
<thead>
<tr>
<th>UniqueID</th>
<th>Analysis</th>
<th>Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>65</td>
<td>Death,OffendingTruck --&gt; Male</td>
<td>95.45%</td>
</tr>
<tr>
<td>52</td>
<td>Death,OffendingCar --&gt; Male</td>
<td>94.12%</td>
</tr>
<tr>
<td>82</td>
<td>Death --&gt; Male</td>
<td>93.55%</td>
</tr>
<tr>
<td>26</td>
<td>Car,Death --&gt; Male</td>
<td>92.31%</td>
</tr>
<tr>
<td>77</td>
<td>OffendingCar --&gt; Male</td>
<td>92.31%</td>
</tr>
<tr>
<td>22</td>
<td>12 --&gt; Male</td>
<td>91.67%</td>
</tr>
<tr>
<td>3</td>
<td>Bicycle --&gt; Male</td>
<td>90.91%</td>
</tr>
<tr>
<td>10</td>
<td>Pedestrians --&gt; Male</td>
<td>90.91%</td>
</tr>
<tr>
<td>79</td>
<td>OffendingTruck --&gt; Male</td>
<td>90.91%</td>
</tr>
<tr>
<td>61</td>
<td>Car --&gt; Male</td>
<td>90%</td>
</tr>
<tr>
<td>40</td>
<td>OffendingJeep --&gt; Male</td>
<td>88.24%</td>
</tr>
<tr>
<td>80</td>
<td>InjuriesAndDisablity --&gt; Male</td>
<td>84.21%</td>
</tr>
<tr>
<td>1</td>
<td>19 --&gt; Male</td>
<td>83.33%</td>
</tr>
<tr>
<td>13</td>
<td>Death,OffendingJeep --&gt; Male</td>
<td>83.33%</td>
</tr>
<tr>
<td>25</td>
<td>OffendingJeep --&gt; Death</td>
<td>70.59%</td>
</tr>
<tr>
<td>64</td>
<td>Male,OffendingTruck --&gt; Death</td>
<td>70%</td>
</tr>
<tr>
<td>7</td>
<td>Jeep --&gt; Male</td>
<td>100%</td>
</tr>
<tr>
<td>43</td>
<td>Death,Motorcycle --&gt; Male</td>
<td>100%</td>
</tr>
<tr>
<td>75</td>
<td>Motorcycle --&gt; Male</td>
<td>100%</td>
</tr>
</tbody>
</table>

Source: Primary.

Rules derived from 100 judgments of claims under Motor Vehicle Act 1988 have been shown in Table 4. It illustrates that Motorcycle is the highest death causing ill-fated vehicle for male showing 100 per cent confidence followed by car (92.31%), while Truck is the highest death-causing / offending vehicle for male showing 95.45per cent confidence followed by Car(94.12%).
7. CONCLUSION AND SUGGESTIONS
Judiciary should be geared up for utilizing its huge textual data by using data mining for appropriate and immediate decisions. It may help the judiciary to save time and resources also. Exploring data by using data mining can play an important role. It may put forward judiciary by providing lawful suggestions.

8. REFERENCES
[10] District Courts, Sirsa
[22] Haryana Police, District Sirsa