

Subrogation Probability Prediction

Ravi Teja | siddharth | karishma

# Introduction

In this report, we describe the process of building various models for the sake of recommending weather or not subrogation should be the course of action for a particular insurance claim. We look at historical data to try and find out patterns in individual cases. In these cases, the subrogated cases have certain characteristic features. These will be taken into consideration to recommend the course of action for any future case.

Since all the cases are medical ones, case specifics like the details of the incident and entire medical snapshot are noted down. In order to make sense of these notes, text mining is utilized. This increases the accuracy with which prediction can be made for future cases. Designing this model removes a lot of human effort that goes into understanding each and every individual case and hence resources of the insurance firm can be diverted towards actually suing the third-party vendors. Thus, operational efficiency is increased and at the same time, money spent in hiring people to do the job of examining individual cases is saved.

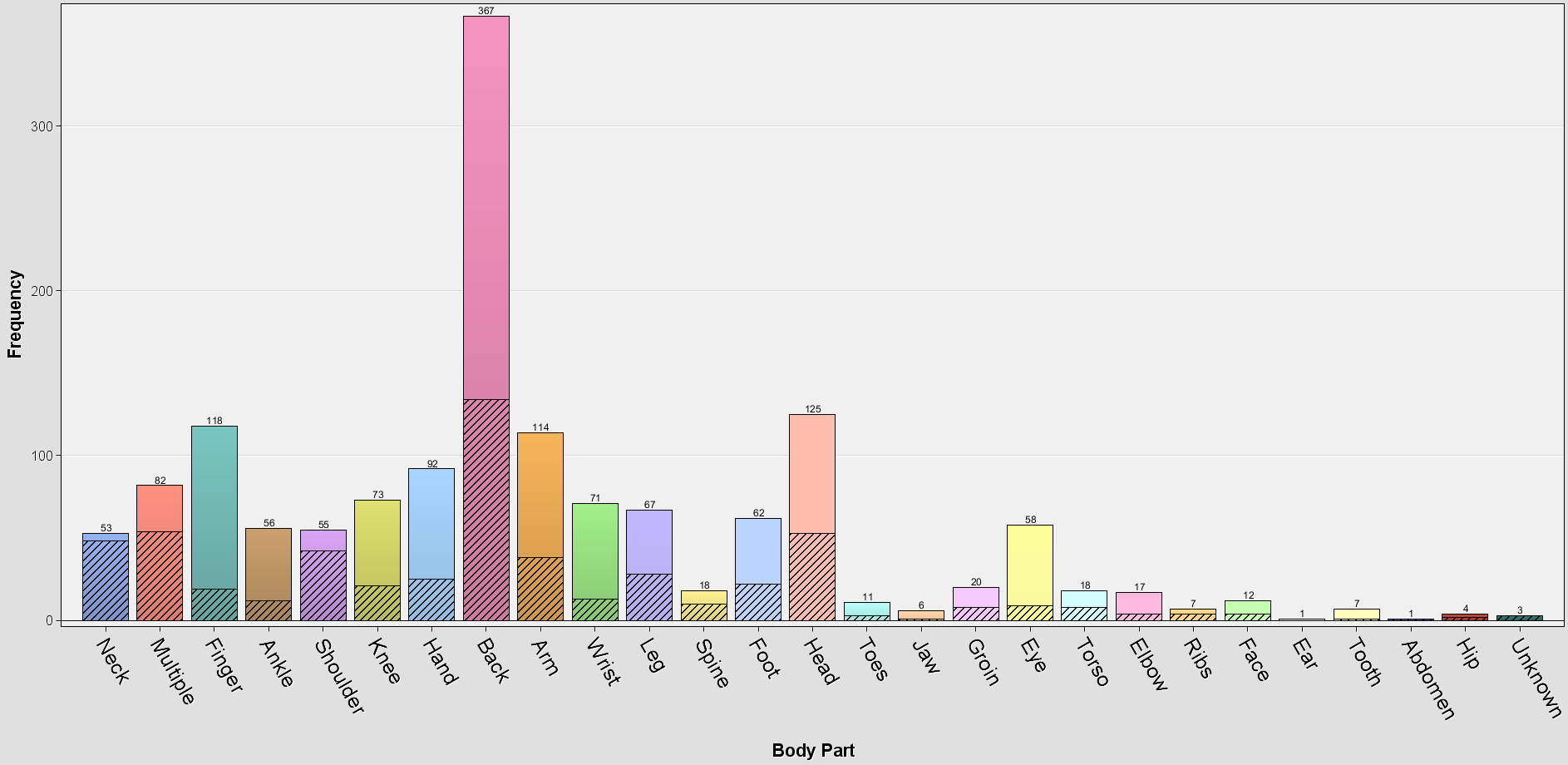
The models developed will be discussed in the report, of which the best one is chosen which has an accuracy of 74%. This model comprises of text mining in which keywords will be clustered into different "topics" or categories. For example, they key word "Motor vehicle" is almost always preset in situations where there is a greater probability of wrongdoing of the third party and hence greater chance of subrogation. This way, keywords are clustered into groups based on the type of incident.

Hence, by utilizing this tool, we can understand the different kinds of scenarios and the chance of earning money in those scenarios. The model, which gains this intelligence based on historical data is now capable enough to recommend the course of action for any future case.

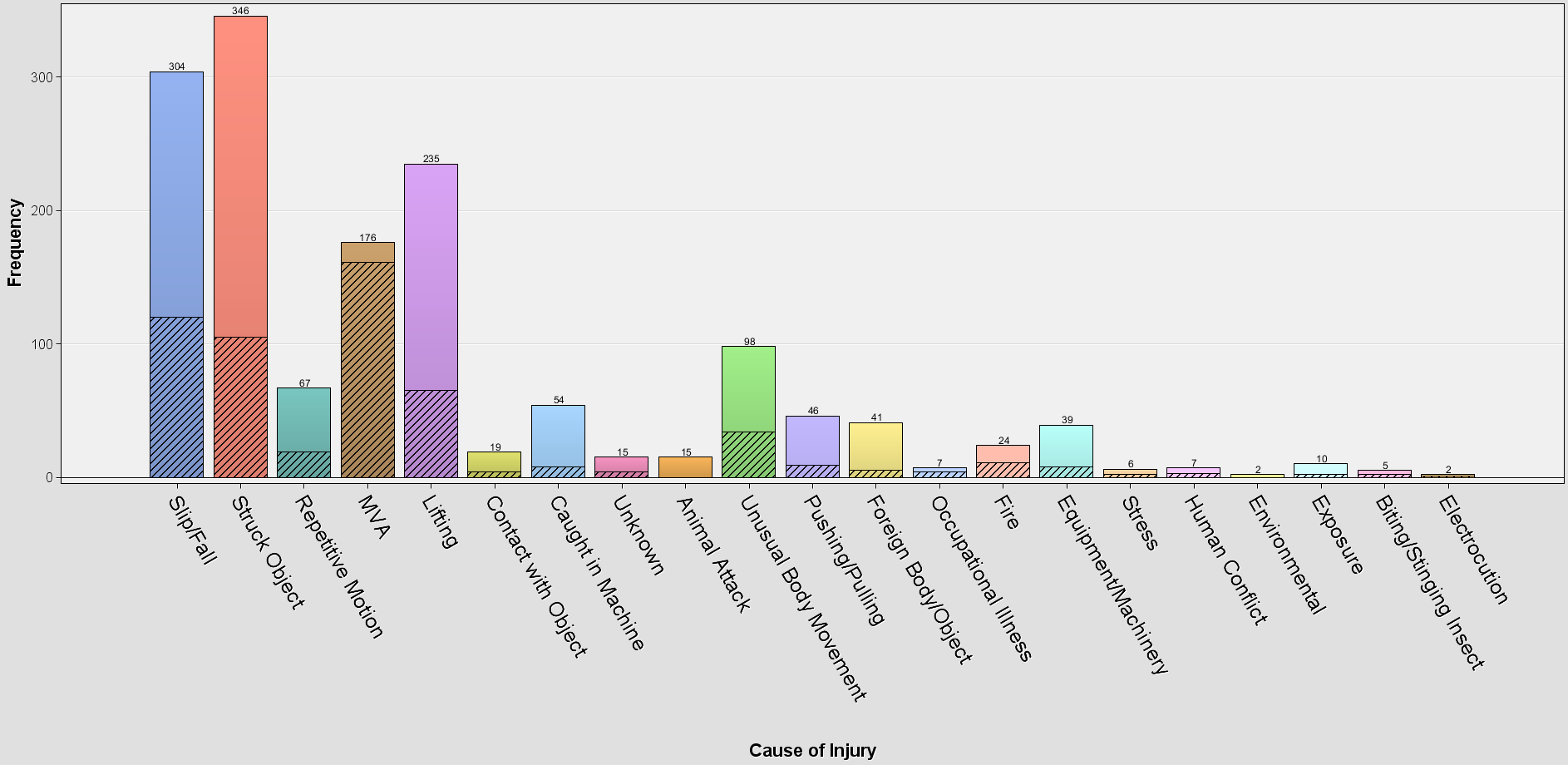
## Meet the data

* Claim Number – Claim number of the Insurance Policy
* Adjuster Notes – Gives a brief description of claim
* Body Part - Body part involved if the insurer in claim
* Nature of injury - Nature of damage to the person
* Cause of Injury - cause of the claim for which user claims
* Vehicle flag - whether vehicle involved in the accident or not
* Subrogation – Target Variable

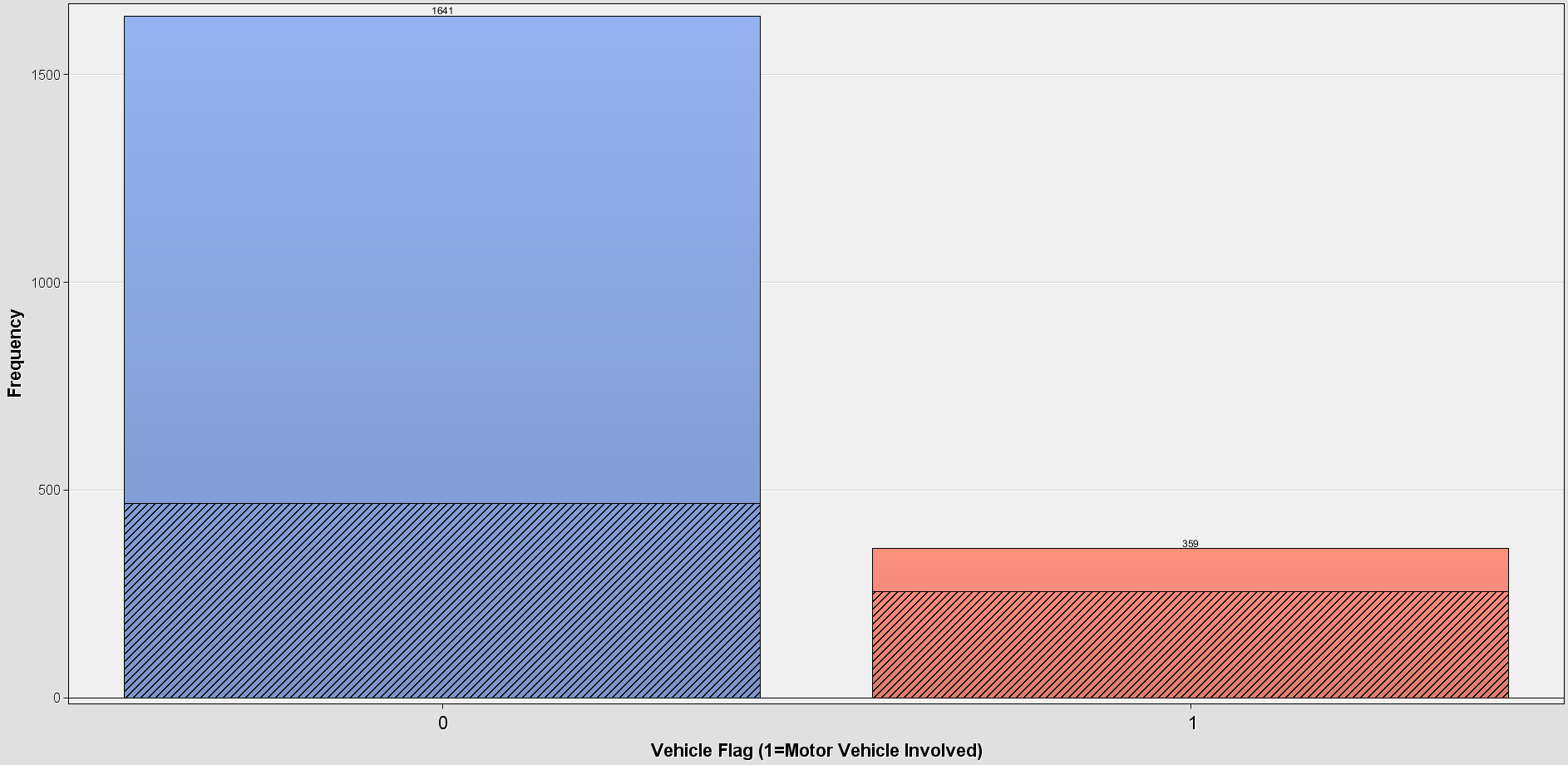
## Data analysis



This plot provides the information about the body part involved in the claim and the chance of subrogation when it is involved in the claim. From the above plot, we can see that there is 95 % chance when a Neck is involved in the claim. As well as when back part is involved in the claim there is 40% chance of subrogation the claim.



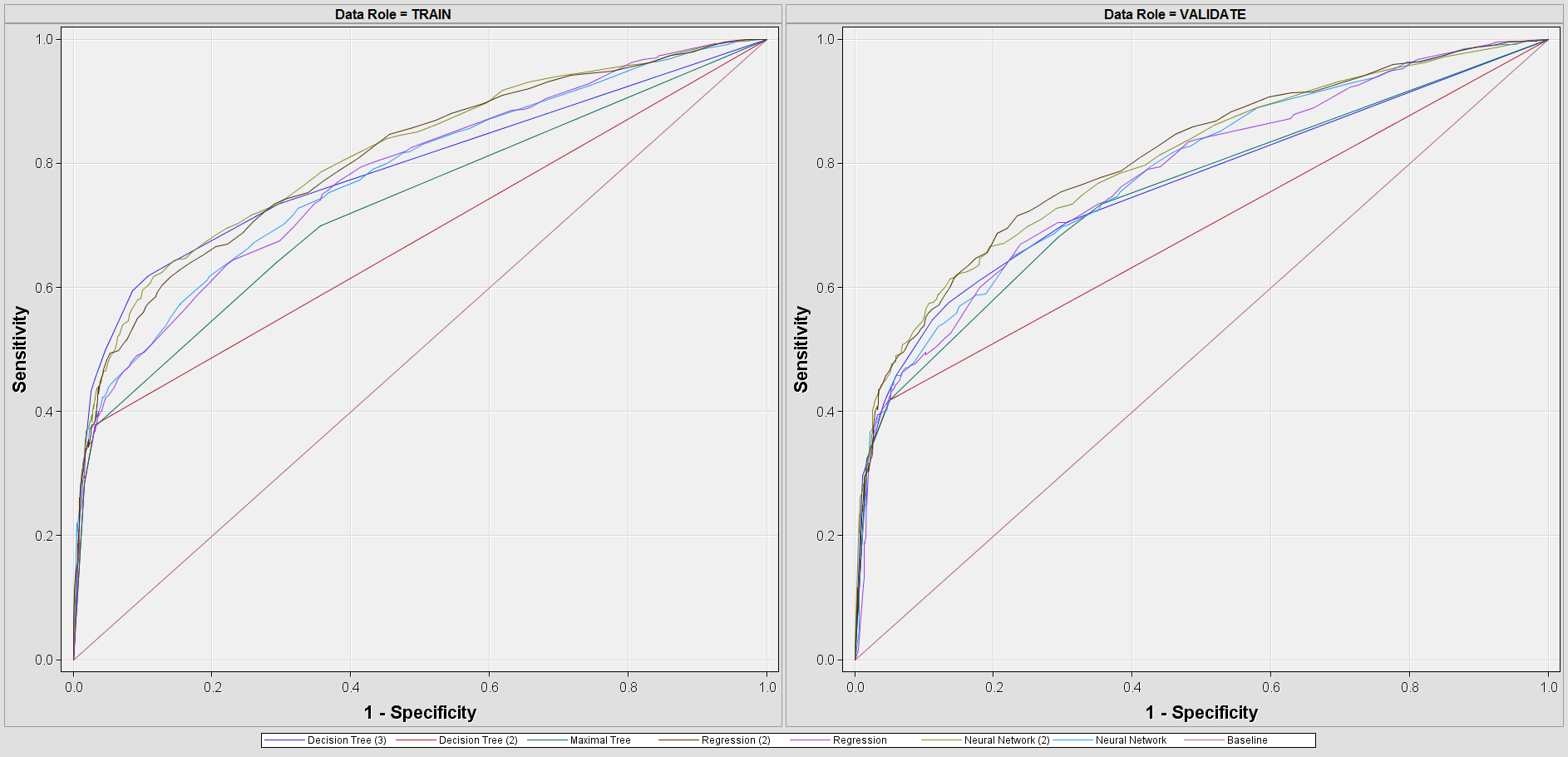
This plot provides the information about the cause of injury involved in the claim and the chance of subrogation when it is involved in the claim. From the above plot, we can see that there is 95 % chance when a Moving Vehicle Accident is involved in the claim. As well as when slip/Fall is the reason for the incident there is 40% chance of subrogation the claim



Above plot provides the information about the vehicle involvement in the incident in the claim and the chance of subrogation when it is involved in the claim. From the above plot, we can see that there is 90 % chance when a Vehicle is involved in the claim. On the other side there is 40% chance of subrogation the claim when no vehicle is involved in the incident.

## model BUILDING

Over all of all the models there are 7 models where maximal, optimal trees, regression and neural network model were built before Text Mining and decision tree, regression, and neural network models were built after Text Mining and in all those neural network being the best model.



From the Roc chart we can see that the Neural network is selected as the best model among all models and there is a 74 % accuracy from the model.

|  |  |  |
| --- | --- | --- |
|  | Predicted Yes | Predicted No |
| Actual Yes | TRUE Positive | False Positives |
|  | 213 | 30 |
| Actual No | False Negative | TRUE Negative |
|  | 354 | 921 |

## recommendations

* Chance of subrogation is high when Motor Vehicle Accident involves in claim.
* Odds of subrogation are very high if there is any Neck or Multiple Injuries.
* Chance of Subrogation is very less when there is animal involved in the claim.