CF969-7-SP-CO
Report on Learning about risk: Machine learning for risk assessment

Rahul Kithalamane Basavaraj - 2212064

April 20, 2023

1 Introduction

The paper talks about taking risk assessment approaches based on machine learning due to increasing understanding of risk involved and the emergence of knowledge gained from technological advancements in these safety-critical industries such as petroleum and chemical. With increasing knowledge, the understanding of risk is continuously changing. It also intensifies learning from past lessons to process relevant data to deal with unexpected events which can endanger large number of people at a time. This paper highlights setting up of high reliable risk assessment and management approaches, by certain clear operation procedures and by providing advanced training's for computational management of machines.

2 Algorithms

In this paper the author has developed Deep Neural Network (DNN) model for testing in a real-world scenario involving an oil and gas drilling rig, to show how accurate the predictions are to overcome any risk assessment challenges. Considering the limitations in the model, a proper selection and customisation is necessary for decision making process. The author analyses the notion of risk given by Kaplan and Garrick (1981) which is considered as one of the most renowned definitions. The author in this paper analyses and gives sense into the understanding of risk assessment by various other authors, from the events that went horribly wrong to the reliability in organisational performance, human control on machines, set of rules, training and communication etc for finding potential threats caused by human errors. The analysis also showed how computational management of machines in high-risk environment gives rise to reliable technical operations. The downside of this analysis showed that less data and outdated data leads to overall risk with a possibility of unwanted events or worst-case scenarios taking place.

Addition of risk knowledge to the stated risk formula is an intrinsic feature added to calculate the value of risk. The risk is depicted in a 2- dimensional and 3- dimensional matrix encoded with traffic light colours. The matrix reaches its peak red if the probability and consequence have the highest values and this means a thorough knowledge is required for the cases falling in this area. The conditions which are set at the initial stage reflects a response to risk based on continuous calibration and correction of these analysis repeated over time in turn improving system knowledge.

The author talks about Dynamic Risk Management Framework which is an endless process of collecting data, early warnings and monitoring unknown events from taking place. The main challenge faced here is using the knowledge accumulated and avoiding any kinds of unknowns from taking place. Increased use of technological solutions such as internet of things has resulted in collection of complex data with rising uncertainty about technological capabilities. Dynamic risk analysis helps in preventing real world risk by means of handy software tools which are otherwise incapable of predicting without the use of these models through traditional risk analysis methods.

There is a need for continuous updating of risk assessment through real time data collection and theory-based approaches. These included use of advanced technological software's, live monitoring, and use of preliminary methodologies. The main challenges towards an ideological risk assessment is looking into possible accident scenarios and improving upon them, learning from past events, data processing and analysis, dealing with lack of risk experience and supporting the decision making. The methods used by the author is the concept of safety barriers for modelling which are of 3 levels, technical, operational, and organisational elements for limiting any errors/accidents from taking place. There has been an increase of attention of this model which cover only some determinants of the overall risk but some industries require immense data and processing for risk assessment which is a major challenge.

So the author came up with a study which suggest use of machine learning branch i.e, deep learning to analyse complex risk through a supervised learning technique. Once the model is trained with large amount of information it would have learnt to do risk categorisation. This model is then evaluated real time on the state of art monitored system e.g: an offshore oil and gas platform. The model used here is feed forward neural network because of its similarity with the hierarchical structure of indicator information and it's ability to extract useful features from the hidden layers of input and output data. It can be trained to minimise loss function.

$$R = f(s, p, c, k) \tag{1}$$

where R is the Risk, s is scenario, p is probability, c is consequence and k is the knowledge.

The author analysed 50 different indicator categories analysed over 30 years with total of 240 values per indicator categories. These indicators complex hierarchical structures allow operators to look into the trends whether the risk is improving or worsening with respect to time. Two data sets were created for training and testing using the DNN model. They used open-source library tensor flow for the DNN model. Multiple linear regression model was applied to the same data set to compare DNN model ability to predict risk increase.

3 Results

The Study shows that DNN model has higher precision and slightly higher accuracy as compared to MLR model. The MLR model slightly performs better in recall as compared to DNN model which has higher number of false negatives. It is to be noted that the derivative of risk is obtained for constant scenario s and consequence c. For a tolerance value equal to 0.0003 the DNN model reports 10% precision. Accuracy and recall also reaches 100% if the tolerance value is equal to 0.0005. Whereas for MLR model it reaches 100% accuracy and precision if the tolerance value is equal to 0.001. Recall also reaches 100% slightly faster if the tolerance value is equal to 0.0004.

4 Discussions

The trained machine learning model was able to identify specific risk factors through its computational power which otherwise could not be determined by human intelligence. They provided valuable indicators for risk assessment by updating on a regular basis. The DNN model used here gives 1.3% times more accuracy than the MLR model but I find that even though the model is capable of handling rare events due to their sensitivity to input data and capability to generalize it, there is still a need for further optimisation of this model.

The drawbacks I find in this paper are they require continuous improvement of machine learning approaches required to respond to potential unknown scenarios and for this data needs to be fed continuously to the model and the model needs to be constantly re-trained every time a new indicator arrives to gain knowledge. The process of retraining is computationally very expensive. Some data in these models are not completely accurate and they contain some errors. Since the model is highly sensitive, they do not tolerate inaccurate indicators that cause over fitting. This can be altered by random initialization of values before starting the training session.

5 Conclusion

In conclusion this study shows the main challenges faced by industrial risk assessment. It also notes the comparison between 2 machine learning models DNN and MLR and presents how the former is more accurate for dynamic risk assessment and has the flexibility to deal with unexpected events. It is important to have risk communication to raise general awareness, amplify information, share knowledge on common risk faced. It also includes their strong points and limitations in this study.

References

[1] Nicola Paltrinieri, Louise Comfort, and Genserik Reniers. "Learning about risk: Machine learning for risk assessment". In: *Safety science* 118 (2019), pp. 475–486.