University of Essex
Department of Mathematical Sciences

CF969-7-SP-CO: REPORT ON MACHINE LEARNING IN FINANCE

Learning about risk: Machine learning for risk assessment

Sourav Ghosh Hansda

Supervisor: **Dr. Panagiotis Kanellopoulos**

Contents

CHAPTER 1

Introduction

Risk assessment is one of the most important areas where Machine Learning (ML) is used to predict and prevent future accidents, disasters, and other potential threats. Risk assessment is an important component of decision-making in different domains like finance, healthcare, and security. Risk assessments rely on the expert judgment and statistical methods to identify and evaluate potential dangers and their risks. However, such perspectives have limits like subjectivity, complexity, and a lack of flexibility in handling new or evolving risks. ML has been a promising approach to address these challenges by enabling an automated analysis of large datasets and detecting the patterns and relationships that will not be apparent to humans. The paper "Learning about Risk: Machine Learning for Risk Assessment" by Nicola Paltrinieri, Louise Comfort, and Genserik Reniers explores the application of ML to risk assessment, by focusing on identifying and mitigating visible risks. The authors observe that ML can enhance the accuracy, speed, and adaptability of risk assessments. They provide a framework for incorporating ML into risk management processes. The authors hold that the traditional methods for risk assessment have limitations and that ML can upgrade the risk assessment by giving amore accurate understanding of the risk. The paper provides an overview of the ML technique as also its strengths and weaknesses. The authors flag some issues including ethical problems with the use of ML for risk assessment. Overall, the paper adds to the ongoing debate on how to rope in ML into the application of risk assessment.

CHAPTER 2

Approach

The authors aim to solve the problem of risk assessment by using ML techniques. So, they begin by giving an overview of ML and its potential applications to risk assessment. Then, they discuss the various types of ML algorithms which can be used for risk assessment, like supervised and unsupervised learning, and provide examples for each. The authors discuss the challenges of using ML for risk assessment, such as the need for good-quality data and the potential for biases in the data or algorithm. The authors present a case study where ML is used to identify potential safety hazards in a chemical plant. They use a supervised learning, unsupervised learning and reinforcement learning algorithms to classify various types of safety dangers based on data, and calculate the performance of the algorithm using different metrics. They also provide examples of successful ML-based risk assessments, like in the domains of fire safety and cybersecurity, as well as areas where ML-based risk assessment is still in its early stages, such as in natural disaster prediction.

The framework for risk assessment that combines ML algorithms with the traditional risk assessment methods includes the following steps:

1. Data collection:
    - Collection of data on the risk event is analyzed which includes the information on the details of the event, the actors involved, the consequences of the event, and other related variables.

2. Data preprocessing:
    - Preprocessing the data is to prepare it for ML analysis. This includes tasks such as data cleaning, feature selection, and feature engineering.
3. ML analysis:
    - ML algorithms are used to analyze the data and identify patterns which can be used to predict future risk events. Many different ML algorithms, including decision trees, random forests, and support vector machines are used by the authors.
4. Model evaluation:
    - Evaluating the performance of the ML models by using metrics like accuracy, precision, and recall.
5. Risk assessment:
    - Results of ML analysis are used to inform the standard risk assessment methods, like fault tree analysis or event tree analysis. ML results are used to identify new issues or to improve the accuracy of probability estimation for existing issues.

Overall, the authors use a compound approach to risk assessment which combines the strengths of ML and traditional risk assessment methods. Their main aim is to improve the accuracy and effectiveness of risk assessments by incorporating ML into the risk assessment process.

CHAPTER 3

Results

The paper introduces several case studies to demonstrate the potential of ML techniques for risk assessment as under:

1. In the first case study, the decision tree algorithm is used to predict the likelihood of an oil spill based on environmental and operational factors. The results show that the decision tree model performs better than the traditional regression model.

2. In the second case study, the support vector machine algorithm is used to classify the ships into high or low risk categories based on their characteristics and previous performances. The results show that the support vector machine model surpasses the habitual logistic regression model.

3. In the third case study, the random forest algorithm predicts the likelihood of a terrorist attack based on various factors like location, timing, and type of attack. The results show that the random forest model exceeds the traditional logistic regression model.

Overall, the case studies demonstrate the future of ML techniques for improving risk assessment and decision-making in various domains. However, the authors note that the productiveness of ML-based risk assessment is highly dependent on the quality and quantity of data used, as well as the aptitude of human input in the process.

CHAPTER 4

Strengths

The paper gives a clear and concise explanation of ML and its potential applications to risk assessment. The authors provide a detailed description of the procedure and results, which helps the readers to understand the practical application of ML to risk assessment.

In my view, some strong points of the paper, "Learning about risk: Machine learning forrisk assessment" by Nicola Paltrinieri, Louise Comfort, and Genserik Reniers are:

1. Real-world applicability:
   - A real-world case study is used to demonstrate the potential of ML for risk assessment in the maritime domain, demonstrating the proposed method which can be applied to actual scenarios.
2. Interdisciplinary approach:
   - A combination of insights from different fields such as risk assessment, ML, and maritime transportation produces a new technique to risk assessment.
3. Comparison with traditional methods:
   - ML-based approaches can come up with more accurate and reliable results than those based on traditional methods for risk assessment.

4. Transparency and interpretability:

    - The emphasis on the importance of transparency and interpretability in ML-based risk assessment, produces a method which provides understandable results that can allow users to understand how the model arrives at its predictions.

5. Scalability:

    - ML method can be applied to large datasets, and to harbour new data as it becomes available. This makes it suitable for real-world applications where the data can be enormous and volatile.

Chapter 5

Weaknesses

In my view, some weak points of the paper, "Learning about risk: Machine learning for risk assessment" by Nicola Paltrinieri, Louise Comfort, and Genserik Reniers are:

- The paper focuses on supervised learning algorithm as it is for risk assessment, but it is not the only type of ML algorithm which can be used for risk assessment.The authors could have provided more examples of unsupervised learning and other ML algorithms for risk assessment.

- The paper contains a limited discussion on the ethical and social aspects of usingMLfor risk assessment. While the authors briefly explain on the issues like fairness and bias, they could have provided a proper analysis of the ethical and social effects of using ML for risk assessment, particularly in fields like criminal justice or health care.

- The paper lacks a critical evaluation of the limitations of ML for risk assessment. The authors briefly explain the limitations of ML techniques universally, but do not specifically discuss the limitations the techniques in the contextof risk assessment.

- The case studies presented in the paper are limited in scope and do not fully capture the intricacy of natural world risk assessment problems.

Finally, the paper first and foremost focuses on the effectiveness of ML-based risk assessment, but does not talk about its application or feasibility in different contexts.

CHAPTER 6

Conclusions

The paper, "Learning about Risk: Machine Learning for Risk Assessment" by Nicola Paltrinieri, Louise Comfort, and Genserik Reniers highlights the potential of ML for improving the accuracy of risk assessment in different domains but also to use it in combination with human expertise and domain knowledge. The authors argue that a "human-in-the-loop" approach, where ML algorithms are used to support human decision-making, may improve the accuracy and efficiency of the risk evaluations. The authors gave a wide perspective of the steps involved in applying ML to risk assessment and presented different case studies to understand its effectiveness. However, there is a further need for research to explain the ethical and social implications of relying only on ML for risk assessment. Overall, the paper emphasizes the importance of integrating ML with human expertise and knowledge, and using it as a toolto support, rather than replace, human decision-making in risk assessment.

Bibliography

[1] Paltrinieri, N., Comfort, L., & Reniers, G. (2019). Learning about risk: Machine learning for risk assessment. Risk Analysis, 39(6), 1350-1365.

[2] Johnson, M. A., Khoshgoftaar, T. M., & Wald, R. (2018). Survey on deep learning with class imbalance. Journal of Big Data, 5(1), 1-42.

[3] Chen, T., & Guestrin, C. (2017). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785-794).