**Learning about risk: Discriminating risky software project using neural networks**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Nguyen M.Thanh  MSIS3033.N22.CTTT  University of Information Technology  20521920@gm.uit.edu.vn | Nguyen V. Tan  MSIS3033.N22.CTTT  University of Information Technology  20521880@gm.uit.edu.vn | Tong T. Thinh  MSIS3033.N22.CTTT  University of Information Technology  20521958@gm.uit.edu.vn | Quach V. Quang  MSIS3033.N22.CTTT  University of Information Technology  20521811@gm.uit.edu.vn | Le V. Sy  MSIS3033.N22.CTTT  University of Information Technology  20521854@gm.uit.edu.vn |

***Abstract*— Early and accurate discrimination of risky software projects is critical to project success. Researchers have proposed many predictive approaches based on traditional modeling techniques, but the high misclassification rate of risky projects is common. To overcome this problem, this study proposes a typical three-layered neural network (NN) architecture with a back propagation algorithm that can learn the complex patterns of the OMRON dataset. This study uses four accuracy evaluation criteria and two performance charts to objectively quantify and visually illustrate the performance of the proposed approach. Experimental results indicate that the NN approach is useful for predicting whether a project is risky. Specifically, this approach improves accuracy and sensitivity by more than 12.5% and 33.3%, respectively, compared to a logistic regression model developed from the same database. These results imply that the proposed approach can be used for early planning of limited project/organization resources and appropriate action for risky projects that are likely to cause schedule slippage and cost overload.**

# CONTEXT

The article "Discriminating risky software project using neural networks" explores the use of a neural network (NN) model with a backpropagation algorithm to predict the risk associated with software projects to overcome the lack of approaches based on traditional modeling techniques,

This study uses four accuracy evaluation criteria (Accuracy, Sensitivity, Specificity, and Precision), and two performance charts (ROC (Receiver Operating Characteristic) And Precision-Recall curve) The author listed 2 approaches:

(1) predicting the overall degree of risk of a project and (2) predicting whether a project is risky.

In this article, the latter approach was considered- “predicting whether a project is risky". The author suggests that although neural networks (NNs) are recognized as a powerful and reliable technique for building models, there has been limited research on using them for software project risk management. Previous studies have focused on risk analysis and control, not on classifying projects as risky. This highlights an opportunity for further research and the potential benefits of using NNs for project risk classification. When experimenting, the authors approach improves accuracy and sensitivity by more than 12.5% and 33.3%, respectively, compared to a logistic regression model developed from the same database.

Overall, the author had realized that software projects have become more complex and diverse, project managers require more accurate predictive approaches to identify risky projects and the article highlights the need for further research in this area, particularly the use of NNs for classifying projects as risky or not, and the results indicate that the proposed NN model outperforms traditional modeling techniques such as logistic regression. By accurately predicting the degree of risk associated with a software project before implementation, limited resources can be effectively invested, ultimately improving project outcomes.

# DATA SET

Dataset wad retrieved through this website: https://link.springer.com/chapter/10.1007/978-3-540-24659-6\_19 The dataset indicates that the risk factors which could affect to the projects. This dataset comprises 40 valid projects completed in the past and 24 attributes (22 project risk factors and 2 project background features). These 22 project risk factors comprise 5 viewpoints: (1) requirements (Reqm), (2) estimations (Estm), (3) planning (Plan), (4) team organization (Team), and (5) project management activities (Prma). A risky project is defined as exceeding the budget and lacking control during project implementation; it is represented by the attribute Status to label whether the project is risky (confused) ornot risky (not confused)

# METHODS

The author proposes a risk assessment model using data from 40 valid projects completed in the past and 24 attributes (22 project risk factors and 2 project background features). These 22 project risk factors comprise five viewpoints: (1) requirements (Reqm), (2) estimations (Estm), (3) planning (Plan), (4) team organization (Team), and (5) project management activities (Prma).and compares the results with those obtained using other methods, including logistic regression and decision trees.

The model uses a feedforward neural network with one hidden layer, and the input variables are the risk factors. The network is trained using a backpropagation algorithm, and the output is a risk score indicating the risk level associated with the project.

Construction of the predictive model by neural network

Step 1: All 40 projects in the OMRON dataset were divided into training and validation sets containing 32 and 8 projects, respectively. The partition principle was adopted for consistency with the work of Takagi et al. to allow for a meaningful comparison.

Step 2: A typical three-layered NN architecture with a back prop- agitation algorithm was built using a training set; the set included 20 pro- jects in the notconfused group and 12 projects in the confused group. The network architecture included 23 attributes: 22 risk factors as input layer variables and the status as the output layer variable. The learning rate and momentum for network training were set to 0.4 and 0.9, respectively, and the training procedures were terminated once 10,000 training epochs were achieved. The software used to construct the NN was SPSS 18.0.

Step 3: Examined the predictive accuracy of our model on the validation set, which included five projects in the not-confused group and three projects in the confused group. In addition, used a cumulative gain chart and a lift chart to visualize the actual performance effectiveness of using the predictive model versus without the predictive model. The cumulative gain chart and the lift chart summarized the performance of the predictive model from the four accuracy evaluation measures and then compared our results.

# ADVANTAGE AND DISADVANTAGE

Advantage:

• Early identification of potential risk: by conducting risk management activities, potential risks can be identifies at an early stage, which allows project managers to develop a plan mitigate these risks before they can become actual problems

• Improved decision-making: with risk management strategies in place, project managers are better equipped to make informed decisions, which can help reduce the chances of project failure. • Resource allocation: risk management also involves identifying the resources required to manage the risk, which helps project managers allocate resources more effectively.

• Increased stakeholder confidence: hen project managers can demonstrate that they have take steps to mitigate potential risks, stakeholders are more confident that the project will be successful.

Disadvantage:

• Cost: risk management activities require time, effort, and resources, which can add to the project cost

• Over-reliance on risk management tools: over-reliance on risk management tool and techniques can lead to complacency, as project managers may assume that the tools will identify and mitigate all risk

• False sense of security: over-reliance on risk management can lead to a false sense of security, as some risks may still go undetected or may not be effectively mitigated

# IMPROVEMENT/RECOMMENDATITION

Advantage:

• Early identification of potential risk: by conducting risk management activities, potential risks can be identifies at an early stage, which allows project managers to develop a plan mitigate these risks before they can become actual problems

• Improved decision-making: with risk management strategies in place, project managers are better equipped to make informed decisions, which can help reduce the chances of project failure.

• Resource allocation: risk management also involves identifying the resources required to manage the risk, which helps project managers allocate resources more effectively.

• Increased stakeholder confidence: hen project managers can demonstrate that they have take steps to mitigate potential risks, stakeholders are more confident that the project will be successful.

Disadvantage:

• Cost: risk management activities require time, effort, and resources, which can add to the project cost

• Over-reliance on risk management tools: over-reliance on risk management tool and techniques can lead to complacency, as project managers may assume that the tools will identify and mitigate all risk

• False sense of security: over-reliance on risk management can lead to a false sense of security, as some risks may still go undetected or may not be effectively mitigated

# REFERENCES