



Est. 1961

RISM YQS SUBCOMMITTEE
2022/2023

QS SYMPOSIUM 2022

SURVEYING AND BUILT ENVIRONMENT



ARTIFICIAL NEURAL NETWORK (ANN) APPLICATION FOR COST ESTIMATION OF CONSTRUCTION PROJECTS IN MALAYSIA: A STUDY ON QUALITY OF DATA

ARTIFICIAL NEURAL NETWORK (ANN):
An information processing system that is inspired by neural network in human brain

1. RESEARCH BACKGROUND

The Artificial Neural Network (ANN), is one of the Artificial Intelligence (AI) tools. It is a great technique that can be applied in the construction project cost estimation to solve classification, prediction, and regression problems (Juszczak, 2017). ANN is data-driven and is considered sensitive to input data (Tayefeh Hashemi et al., 2020). The ANN relies on the data input to execute tasks like prediction. Thus, to produce the best and most reliable cost estimation output, the best quality of data are required as input into ANN model.

2. PROBLEM STATEMENT

The Malaysian construction industry faces a **lack of access, accuracy, breadth, and depth of industry data**. Although open-data popularity has increased, the **problem with data quality remained unresolved** (Nikiforova, 2020).

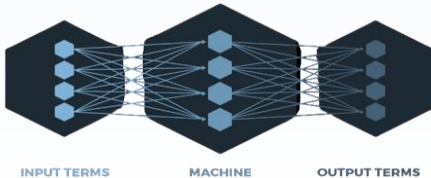
3. AIM

To investigate the quality of data for the implementation of Artificial Neural Networks (ANN) for cost estimation of construction projects in Malaysia

4. OBJECTIVES

- To identify the data required and the availability, accessibility, currency and reliability as input for cost estimation using ANN
- To identify data quality issues that can hinder the ANN practice for cost estimation of construction projects
- To identify appropriate strategies that can be proposed to improve the practice of ANN with specific reference to data input quality in the construction project

Example of ANN structure:



5. METHODOLOGY

A **qualitative research approach** was applied in this study to achieve the research objectives. This approach is employed because the research nature is subjective in nature that requires exploration of opinion, views and perceptions of respondents on the quality of data for the cost estimation using ANN technique. This research employed both data collection approaches to gather required information towards achieving the aims and objectives. Secondary data collection was done first following by the primary data collection.

Secondary data collection

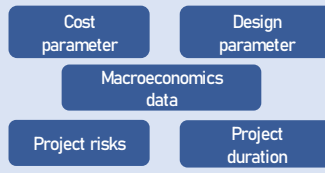
Literature review of journal articles, books, conference proceedings, online reading materials, dissertations and government publications.

Primary data collection

Semi-structured interview with nine (9) experts that have knowledge, skills or experience in Artificial Neural Network (ANN).

6. FINDINGS

THE DATA INPUT REQUIRED FOR COST ESTIMATION USING ANN



THE QUALITY OF ONLINE COST DATABASES IN MALAYSIA

Data Quality Attributes	Mean	Standard Deviation
Availability	2.96	1.13
Accessibility	2.82	1.11
Currency	2.82	1.18
Reliability	2.82	1.16

The standard deviation value achieved for each data quality attributes is higher than 1.00 which indicates fair level of consensus among respondents as shown in the following table:

Standard Deviation	Level of consensus achieved
$0 \leq X < 1$	High level of consensus
$1 \leq X < 1.5$	Reasonable/fair level of consensus
$1.5 \leq X < 2$	Low level of consensus
$2 \leq X$	No consensus

The respondents have assessed the quality of online cost databases in Malaysia which includes CIDB CONVINCE, JKR Rates Online, BCISM, JUBM and ARCADIS as well as Quantity Surveyors Online. Overall, the databases was considered poor quality and not ready to be used as input for cost estimation using ANN.

THE DATA QUALITY ISSUES THAT CAN HINDER ANN PRACTICE FOR COST ESTIMATION OF CONSTRUCTION PROJECTS AND THE MITIGATION STRATEGIES TO IMPROVE ANN PRACTICE FOR COST ESTIMATION OF CONSTRUCTION PROJECTS IN MALAYSIA

DATA QUALITY ISSUES THAT CAN HINDER ANN PRACTICE FOR COST ESTIMATION	MITIGATION STRATEGIES
Inaccurate data that involve many errors	Data quality assessment: Process of assessing data quality to identify errors and examine the suitability of data for intended use
Outdated data	Data quality monitoring: Continuously evaluate data quality and provide results via alerts
Access barriers to retrieve data from known source such as online database	Hire data entry personnel
Insufficient amount of data to be used as input for cost estimation	More fund in collecting data since the core of any ANN model or estimator depends on the data
Noise in training data	Data cleansing: Process of rectifying errors that are identified during data quality assessment

7. CONCLUSION AND SIGNIFICANCE

This research has presented a valuable findings where it has established that the Malaysia's cost database is not readily available, accessible, current and reliable to be used as input for cost estimation using the ANN approach. The data quality issues that can hinder the implementation of ANN for cost estimation of construction projects has been identified. Not only that, this study has also identified strategies to mitigate data quality issues to improve the practice of ANN for construction project cost estimation. All in all, this research has a pivotal role in the construction industry. It contributes to raise awareness among construction experts about cost estimating approaches such as the Artificial Neural Network (ANN), which is one of the critical integrated technologies in Industry Revolution 4.0. (IR4.0). Construction experts, particularly Quantity Surveyors, can assess the gap between their existing cost estimating approach and the use of the ANN technique. Furthermore, as Artificial Neural Networks (ANN) is one of Artificial Intelligence (AI) tools, this research outcomes also contributes to the body of knowledge related with AI.

Category: Undergraduate

Name of Contestants: Alya Farhani binti Mohd Zammari (Author), Mohd Fairullazi Ayob (Co-Author/Supervisor)

Contact Number: 601110990713

University: International Islamic University Malaysia (IIUM)

Correspondence Email: alya.farhani@live.iium.edu.my, fairullazi@iium.edu.my

