TOPIC

Artificial Neural Network based Prediction in SPM

Muneeb Ahmed M-78

Section B

Introduction

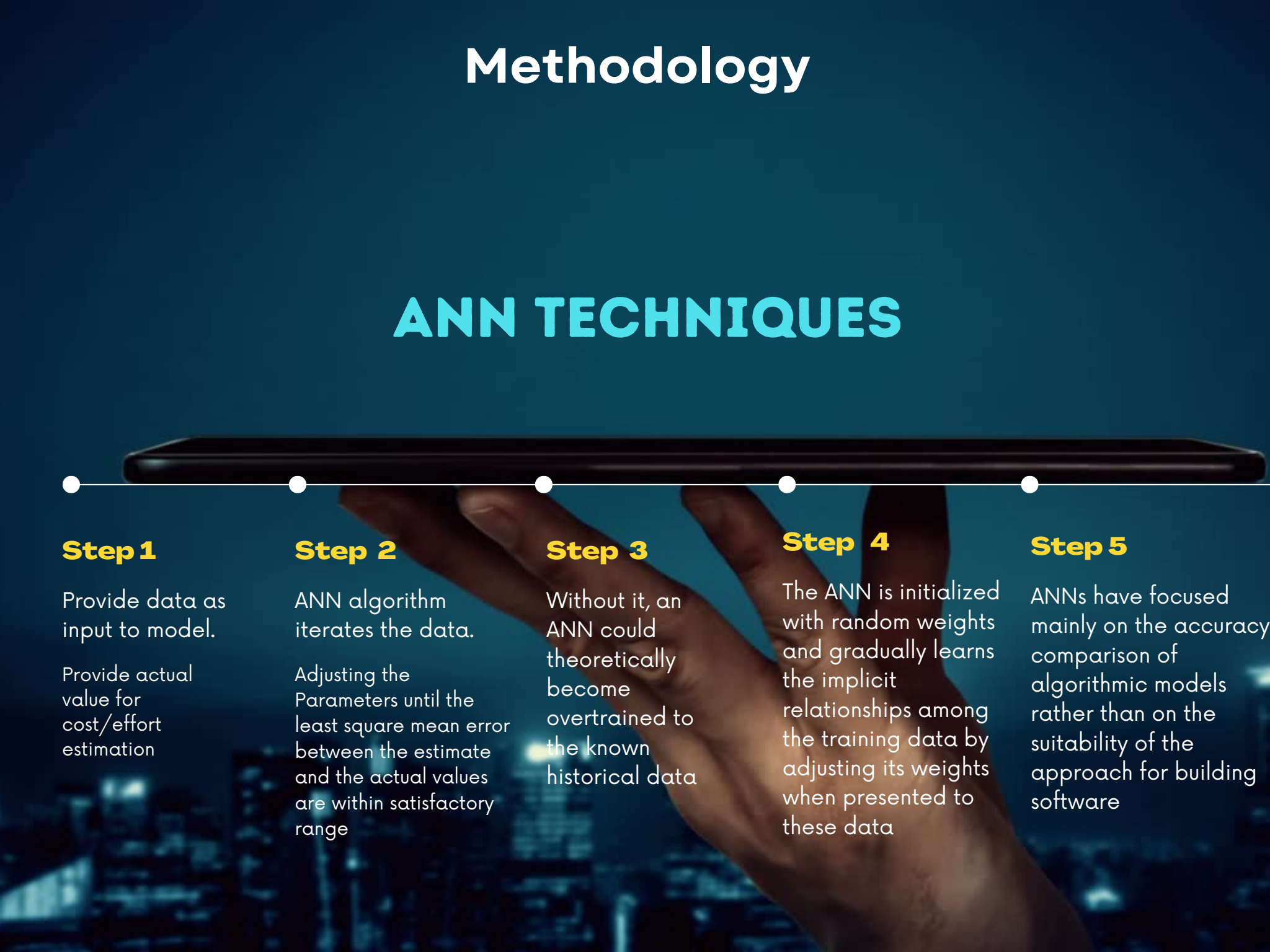
- An increase in software demand also increases the competition between companies to produce low-cost and effective software.
- A critical issue in software project management is the accurate estimation of size, effort, resources, cost, and time spent in the development process.
- Effective Techniques for Project Estimation
    1. Expert Judgment
    2. Algorithmic Model
    3. Machine Learning

Literature Review

Ref#	Method	Result	Parameter	Limitation
[1]	Clustering Method	Accuracy Above 20%	Tree-depth and pruning ratio used as parameter.	Cross-validation feature wasn't available owing to limitation of the respective tools used.
[2]	ESTOR	Error rate > 32%	Software size metrices, Productivity factors	ESTOR certainly fail to accurately estimate projects (e.g., embedded military systems)
[3]	Albus Perceptron	Accuracy improved as compared others	• Generalization width L = 2 • KDSZ values quantized into 63 categories.	The CMAC is not well supported by mature software products;
[4]	Function Point	Error rate 14.2%	Six hidden layers were developed to used as a parameter	The data set was limited for testing in the model
[5]	Ordinal Regression Model	Accuracy Above 30%	Maxwell dataset is used as parameter	The expected project productivity may be found to be lower than the nominal level in the organization

Methodology

ANN TECHNIQUES



Step 1

Provide data as input to model.

Provide actual value for cost/effort estimation

Step 2

ANN algorithm iterates the data.

Adjusting the Parameters until the least square mean error between the estimate and the actual values are within satisfactory range

Step 3

Without it, an ANN could theoretically become overtrained to the known historical data

Step 4

The ANN is initialized with random weights and gradually learns the implicit relationships among the training data by adjusting its weights when presented to these data

Step 5

ANNs have focused mainly on the accuracy comparison of algorithmic models rather than on the suitability of the approach for building software

Methodology

LINEAR AGGRESSION

Step 1

Linear Regression is used to predict effort estimation.

Step 2

Minimizing the mean square of the error across the range of observations in the data set

Step 3

The philosophy is essentially one of solving local prediction problems before attempting at constructing universal models

Step 4

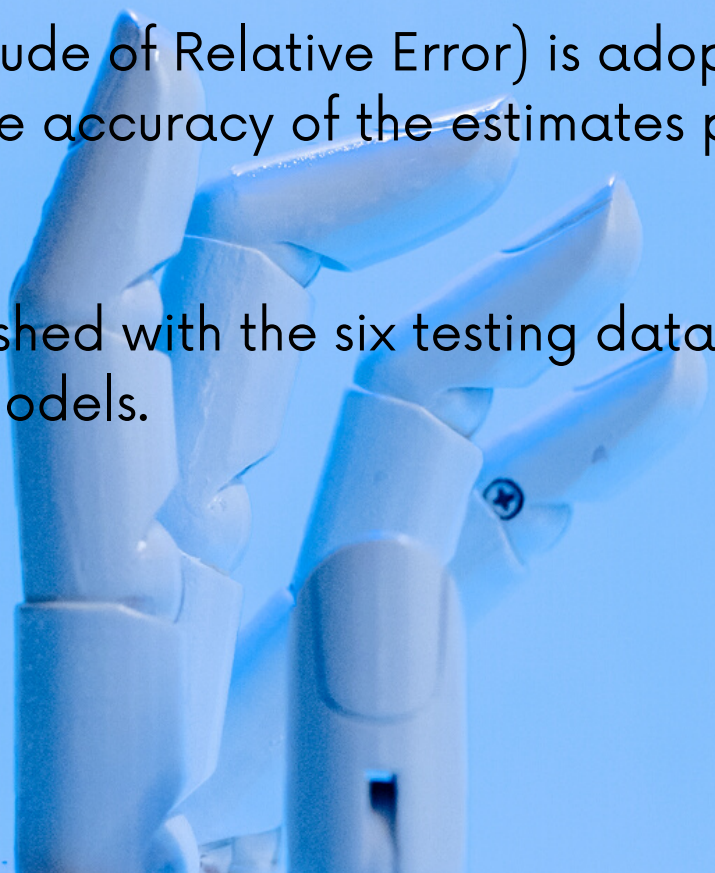
A disadvantage with this technique is its vulnerability to extreme outlier values although robust regression tech niques, that are less sensitivity to such problems, have been successfully used

Step 5

Another potential problem is the impact of co-linearity. The tendency of independent variables to be strongly correlated with one another

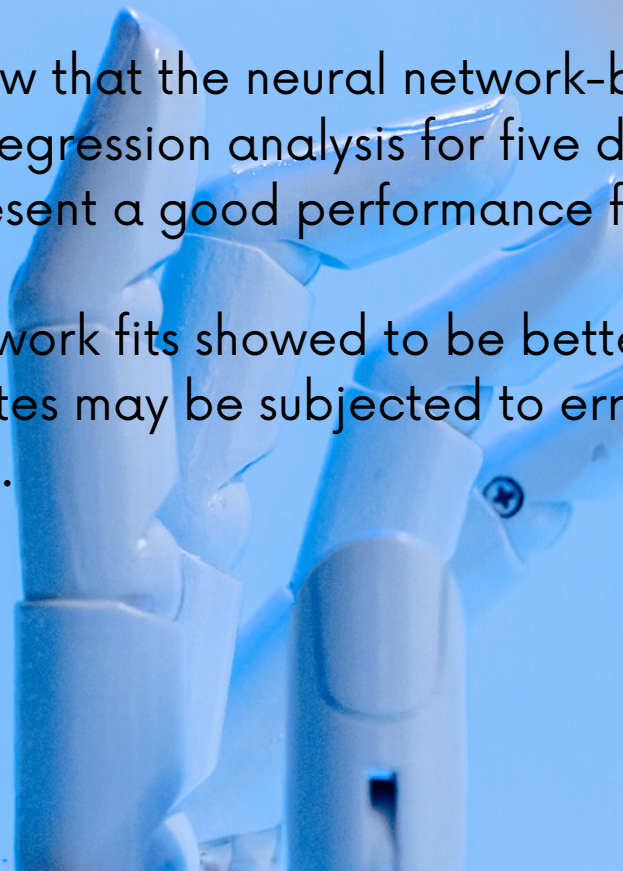
Discussion

- Two measures for estimating accuracy that is most popular in the cost estimation community are the mean magnitude relative error (MMRE) and the Pred(25) statistics.
- The MMRE (Mean Magnitude of Relative Error) is adopted as an indicator of the predictive accuracy of the estimates produced by the models.
- Estimates were accomplished with the six testing datasets using ANN and regression analysis models.



Discussion

- These results indicate that predictions from linear regression and ANN show a strong linear relationship with the actual development effort values for all test projects.
- The results achieved show that the neural network-based approach performed better than regression analysis for five datasets in terms of MMRE. But it did not present a good performance for the sixth dataset.
- Although the neural network fits showed to be better than the linear regression, these estimates may be subjected to error and should be interpreted with caution.

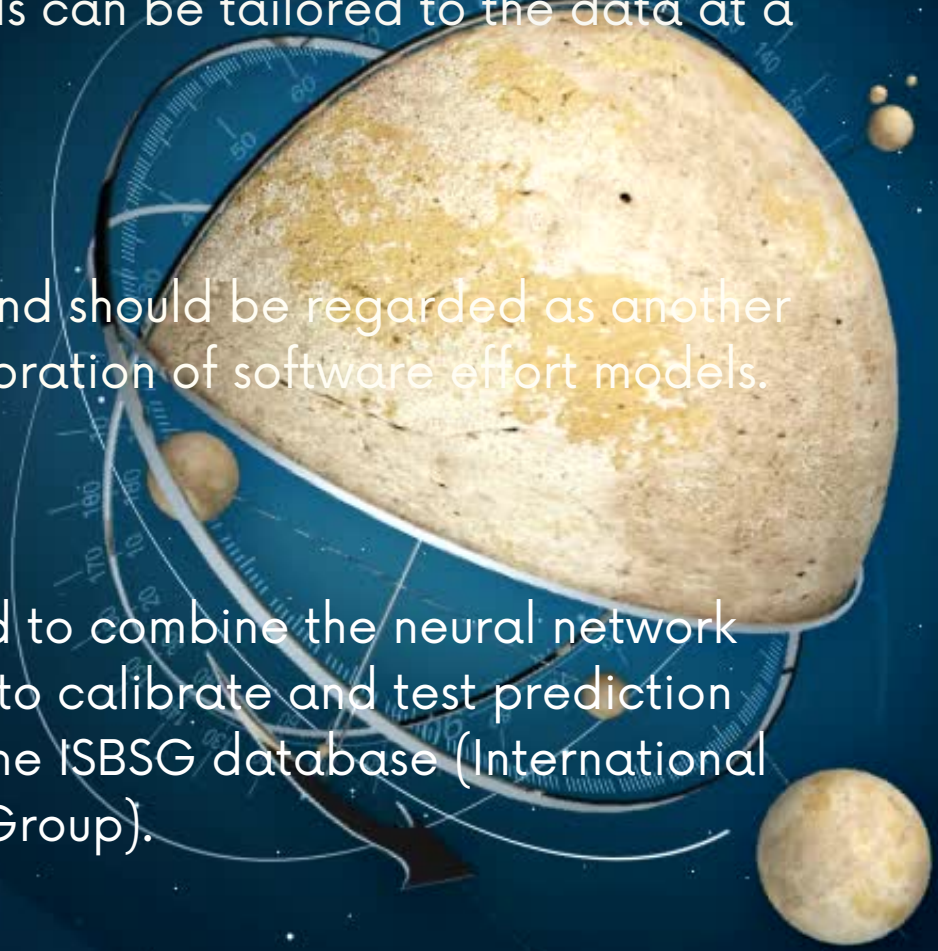


Conclusion

- A primary advantage of an ANN approach is that it is adaptable and nonparametric, thus predictive models can be tailored to the data at a particular site.

- ANN does not replace regression and should be regarded as another powerful tool to be used in the calibration of software effort models.

- New experiments will be conducted to combine the neural network and multiple regression techniques to calibrate and test prediction models on other datasets, such as the ISBSG database (International Software Benchmarking Standard Group).



Reference

1. Zhong, S., Khoshgoftaar, T. M., & Seliya, N. (2004). Analyzing software measurement data with clustering techniques. IEEE Intelligent Systems, 19(2), 20-27.

2. Vicinanza, S., Prietula, M. J., & Mukhopadhyay, T. (1990). CASE-BASED REASONING IN SOFTWARE EFFORT ESTIMATION.

3. Samson, B., Ellison, D., & Dugard, P. (1997). Software cost estimation using an Albus perceptron (CMAC). Information and Software Technology, 39(1), 55-60.

4. Wittig, G. E., & Finnie, G. R. (1994). Using artificial neural networks and function points to estimate 4GL software development effort. Australasian Journal of Information Systems, 1(2).

5. Sentas, P., Angelis, L., Stamelos, I., & Bleris, G. (2005). Software productivity and effort prediction with ordinal regression. Information and software technology, 47(1), 17-29.