

An Evaluation of Software Cost Estimation Techniques: From Expert Analysis to Algorithmic Approaches

Veerpal Kaur¹, Dr. Sunny Arora²

¹Research Scholar, Computer Application, Guru Kashi University, Talwandi Sabo, Bathinda Punjab, sidhuveer94@gmail.com

²Professor, Department of CSE, Guru Kashi University, Talwandi Sabo, Bathinda, Punjab, sunnyarora@gku.ac.in

Abstract

The estimation of software costs is a critical element of project management, as it aids in the planning of budgets, the allocation of resources, and the scheduling of projects. This research work presents a comparative analysis of software cost estimation methodologies. The primary areas of investigation in the paper are expert judgment and algorithmic models. We provide a thorough examination of the numerous expert judgment methods, including the most common approaches and the associated advantages and disadvantages. Additionally, algorithmic models undergo analyses, which encompass sophisticated machine learning algorithms and regression-based methodologies. The associated algorithms' advantages and disadvantages are underscored in these analyses. A comparative study is conducted using predetermined criteria to provide insight into the relative efficacy of algorithmic models in comparison to expert judgment.

Keywords: Software, Cost, Estimation, Techniques, Algorithmic, Models

INTRODUCTION

Software development projects are inherently challenging endeavors that necessitate meticulous resource management, budgeting, and planning. This process is fundamentally centered around the software cost estimation procedure. It is a critical activity that is instrumental in the decision-making process throughout the project's entire lifecycle. In order to ensure that projects are completed on time and within their allocated budgets, it is imperative to have a precise cost estimate, which ultimately contributes to the success of the projects. Throughout its history, software cost estimation has consistently prioritized the application of expert judgment, relying on the perspectives and experiences of project participants. Although expert judgment has its benefits, it is often subjective and susceptible to biases, which can lead to estimates that are inconsistent or inaccurate. In recent years, there has been an increasing emphasis on the use of algorithmic models to improve the precision and dependability of cost estimation. In order to predict the costs of software development based on historical data, project parameters, and other relevant factors, algorithmic models employ mathematical formulas, statistical approaches, and machine learning algorithms. These models are expected to offer estimates that are more consistent and reliable than those that are exclusively based on the judgment of experts due to their impartiality, scalability, and the potential for automation.

SOFTWARE EFFORT ESTIMATION IS REQUIRED

Approximation is a straightforward procedure for smaller undertakings, and precision is not a significant concern. Nevertheless, the critical accuracy that is required diminishes as the project's scope broadens. Nevertheless, the necessity for precision becomes more critical as the project's scope increases, and it is exceedingly difficult to estimate. It is crucial for a reasonable estimate to possess a specific degree of granularity in order to be elucidated. due to the fact that the quantity of effort invested in an undertaking is one of the most significant and extensively researched factors. Consequently, the estimation of this value at the outset of the software initiatives is advantageous in terms of effectively planning any subsequent activities. There is an issue that has not yet been resolved: the estimation of the effort while maintaining a high level of reliability.

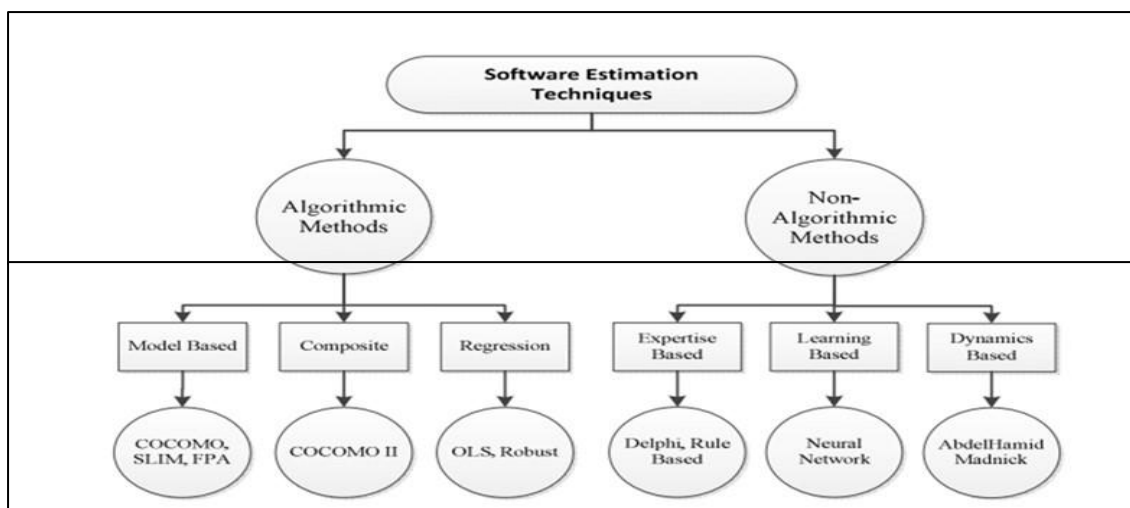


Figure 1. Software Estimation Techniques

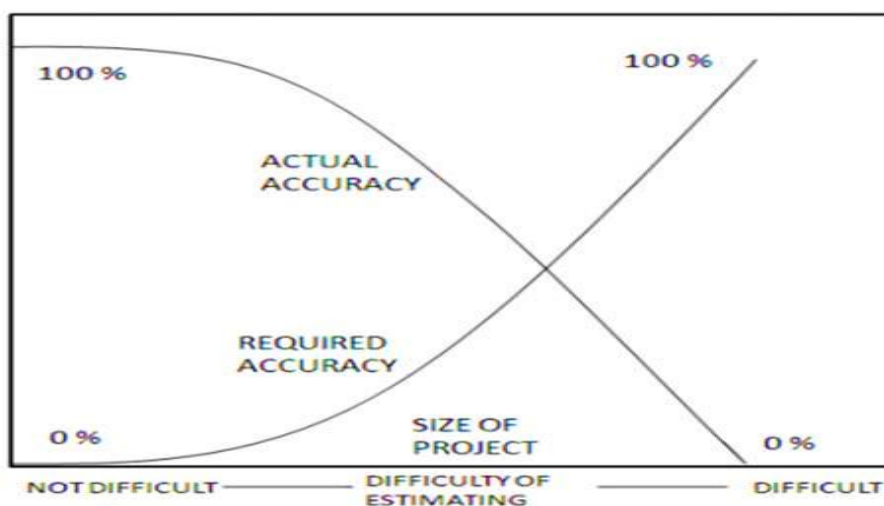


Figure 2. Accuracy of Estimating

EXPERT JUDGMENT TECHNIQUES

Expert Judgment Technique	Description	Advantages	Limitations
Delphi Method	An iterative consensus building technique where experts provide estimates anonymously and revise them based on feedback from other experts.	<ul style="list-style-type: none"> - Aggregates diverse perspectives - Reduces biases and conflicts among experts - Allows for complex and uncertain project scenarios 	<ul style="list-style-type: none"> - Time-consuming process - Requires skilled facilitation to prevent groupthink - May face challenges in maintaining participant engagement

Analogous Estimation	Uses historical data from similar past projects to estimate the cost of a current project, adjusting for differences in size, complexity, and other factors.	<ul style="list-style-type: none"> - Quick and straightforward method - Provides a basis for estimation even with limited project information - Encourages the use of past lessons learned 	<ul style="list-style-type: none"> - Accuracy heavily depends on the similarity between past and current projects - May overlook unique project characteristics - Susceptible to bias in selecting and adjusting historical data
Expert Opinion Polling	Involves soliciting estimates from individual experts or small groups of experts based on their knowledge and experience.	<ul style="list-style-type: none"> - Simple and easy to implement - Incorporates diverse perspectives and expert judgment - Can be tailored to specific project domains 	<ul style="list-style-type: none"> - Vulnerable to individual biases and subjective interpretations - Lack of consensus or validation mechanism may lead to unreliable estimates - Limited scalability and consensus building
Wideband Delphi	An adaptation of the traditional Delphi method that seeks to streamline the estimation process by reducing the number of rounds and increasing collaboration among experts.	<ul style="list-style-type: none"> - Combines the benefits of expert judgment with the efficiency of group collaboration - Allows for rapid iteration and convergence towards a consensus estimate - Facilitates knowledge sharing and consensus building among experts 	<ul style="list-style-type: none"> - Requires skilled facilitation to manage group dynamics - May still suffer from biases inherent in expert judgment - Relies on the availability and willingness of experts to participate actively

The objective of this table is to offer a concise summary of each expert judgement technique, emphasizing the benefits and cons of each technique in the context of software cost estimation.

ALGORITHMIC MODELS

Algorithmic Model	Description	Advantages	Limitations
Regression-Based Models	Utilizes regression analysis to model the relationship between project attributes (e.g., size, complexity) and cost.	<ul style="list-style-type: none"> - Provides interpretable relationship between input variables and cost estimates - Handles continuous and categorical variables 	<ul style="list-style-type: none"> - Assumes linear relationship between input variables and cost, which may not hold true in all cases - Limited flexibility to capture nonlinear relationships in the data
Machine Learning Models	Applies machine learning algorithms (e.g., decision trees, random forests, neural networks) to predict project costs based on historical data and project attributes.	<ul style="list-style-type: none"> - Captures complex, non-linear relationships in the data - Flexible to handle diverse data types and feature interactions 	<ul style="list-style-type: none"> - Requires significant computational resources for training and inference - Susceptible to overfitting with limited training data - Lack of interpretability compared to regression-based models

Parametric Models	Uses mathematical equations (e.g., COCOMO) to represent the relationship between project attributes and cost.	- Provides standardized approach based on empirical data and industry best practices - Tailorable to specific project domains or methodologies	- Assumes fixed relationships between input variables and cost, which may not accurately reflect project variability - Limited flexibility to accommodate projectspecific characteristics or changes in technology
--------------------------	---	--	--

A comparative overview of the various computational models that are employed in the software cost estimation process is provided in the following table. It emphasizes the benefits and disadvantages of each model. Organizations have the capacity to determine the most appropriate model for the purpose of generating cost estimates that are both accurate and reliable by analyzing the specific project context and the available data.

COMPARATIVE ANALYSIS

Aspect	Expert Judgment Techniques	Algorithmic Models
Accuracy	Subject to expertise and biases of individuals	Objectivity and consistency based on data and algorithms
Reliability	Relies on consensus-building among experts	Consistent predictions based on predefined algorithms
Scalability	Can involve multiple experts or panels for larger projects	Can handle large volumes of data efficiently
Adaptability	Tailorable to diverse project domains and contexts	Can be adapted to different project characteristics
Ease of Use	Relatively simple to implement, requiring minimal resources	May require technical expertise for development and deployment

This table provides a concise comparison of algorithmic models and expert judgement methodologies that are relevant to the estimation of software costs. The comparison is based on a variety of variables. By considering the specific requirements and constraints of a project, organizations can determine the most suitable approach for generating cost estimates that are both accurate and reliable.

MACHINE LEARNING METHODS

The majority of cost estimation methods employ statistical methodologies, which are incapable of producing results that are both reasonable and reliable. Given their capacity to enhance estimation accuracy through the training of estimation principles and the repetition of run cycles, this approach may be appropriate. In particular, it can be divided into two primary approaches: neural networks and fuzzy methods, which will be examined in the subsequent subsections.

NEURAL NETWORKS

Neurons comprise each stratum of a neural network. Neural networks are composed of numerous layers. Neurons generate the outputs by analyzing the weights assigned to the inputs. The primary objective is to obtain the actual effort as the outputs when it comes to estimation. The back propagation neural network is the optimal solution for software estimation problems due to its ability to modify the weights by comparing the network outputs with actual results. This is due to its exceptional adaptability. The training is also conducted in a manner that is both efficient and effective. The Cocomo approach has been the focus of the overwhelming majority of studies that have been conducted on the application of neural networks to estimate the cost of software. For example, the accompanying figure illustrates a neural network that has been suggested for the purpose of estimating the cost of software. The figure presented illustrates the layers, inputs, and transfer function of the neural network that was previously mentioned. The neural network utilizes scale factors (SF) and effort multipliers (EM) as inputs. The symbols π and q_j are used to represent the weights of SFs and Ems, respectively.

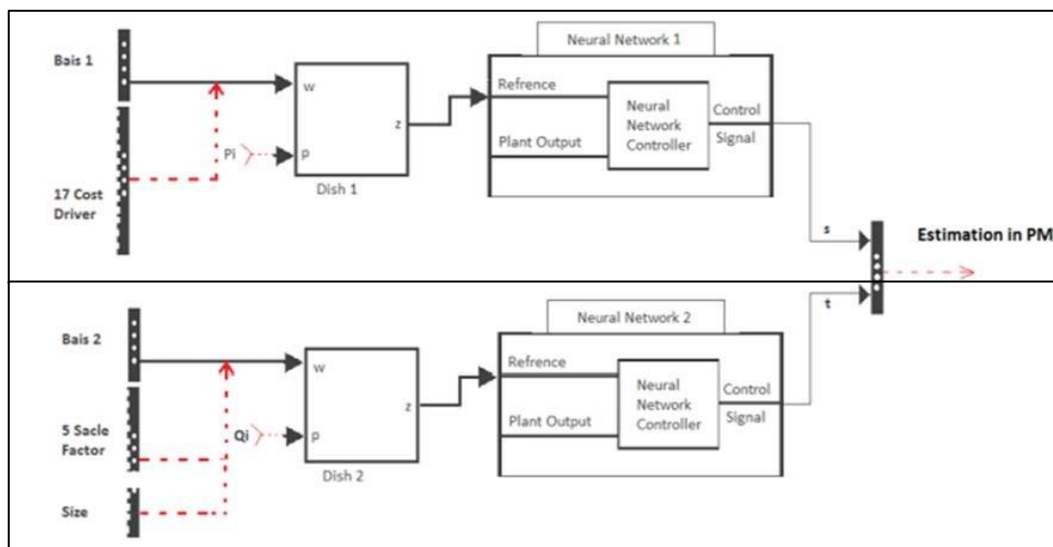


Figure 3. Layers, input & transfer function of neural network

FUZZY METHOD

In an effort to fulfill their functions, systems that operate on fuzzy logic attempt to replicate human behavior and thought. Fuzzy systems are a highly effective instrument that can be employed in situations where decision-making is extremely difficult and the conditions are ambiguous. In every instance, the imprecise method substantiates the facts that can be disregarded.

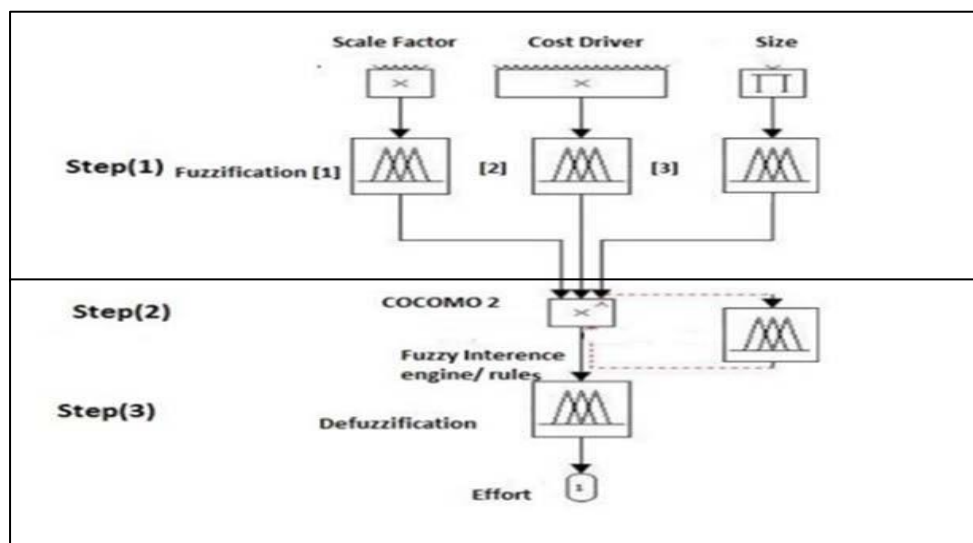


Figure 4. An example of using Fuzzy method

CONCLUSION

In our comparative analysis illuminates the complementary attributes that algorithmic models and individual expert judgment techniques contribute to the evaluation of software costs. Expert judgment offers context and intuition; however, it is subject to subjectivity. Conversely, computational models provide scalability and objectivity; however, they necessitate meticulous validation. By employing these methodologies, the precision and dependability of cost estimates can be enhanced, thereby facilitating the successful administration of projects in the constantly evolving software development environment. In order to enhance and advance software cost estimation techniques, it is imperative to engage in ongoing research and collaboration. In contrast, algorithmic models employ mathematical formulas, statistical methodologies, and machine learning algorithms to provide a data-driven approach to cost estimation. This method represents a substantial departure from conventional methods. These models, which incorporate regression-based techniques and machine learning algorithms, provide objectivity, scalability, and the potential for accuracy by analyzing historical data and project characteristics. However, in order to ensure that forecasts are precise, they must be meticulously selected, feature-engineered, and validated.

Furthermore, the fundamental components that are responsible for cost estimations may be challenging to understand due to the incapacity of certain computational models to be interpreted.

REFERENCES

- [1] Suri, Ranjan, (2012) "Comparative Analysis of Software Effort Estimation Techniques". International Journal of Computer Applications (0975 - 8887) Volume 48- No.21.
- [2] Patil, Waghmode, Joshi & Khanna, (2014) "Generic Model of Software Cost Estimation: A Hybrid Approach". IEEE International Advance Computing Conference (IACC).
- [3] Waghmode, Patil & Joshi (2013). "A Collective Study of PCA and Neural Network based on COCOMO for Software Cost Estimation". IJCA (0975 - 8887) Volume 74- No.16.
- [4] Madheswaran, M., & Sivakumar, D. (2014). "Enhancement of prediction accuracy in COCOMO model for software project using a neural network". In Computing, Communication and Networking Technologies (ICCCNT), International Conference on (pp. 1-5). IEEE.
- [5] Saroha, Sahu, (2015). "Tools & Methods for Software Effort Estimation Using Use Case Points Model - A Review", IEEE, ISBN:978-1-4799-8890-7/15.
- [6] Musilek, Pedrycz & Sun, (2002), "On the Sensitivity of COCOMO II Software Cost Estimation Model". Proceedings of the Eighth IEEE Symposium on Software Metrics (METRICS.02) 0-7695- 1339-5/02
- [7] Kuashik, Chauhan, Mittal, Gupta,(2012) ."COCOMO Estimates Using Neural Networks", MECS.
- [8] Gupta, A., Mishra, N., & Kushwaha, D. S. (2014). "Rule based test case reduction technique using decision table". In Advance Computing Conference (IACC), IEEE International, pp. 1398-1405.
- [9] S. Singh and G. Jagdev, "Execution of Big Data Analytics in Automotive Industry using Hortonworks Sandbox," 2020 Indo - Taiwan 2nd International Conference on Computing, Analytics and Networks (Indo-Taiwan ICAN), Rajpura, India, 2020, pp. 158-163, doi: 10.1109/Indo-TaiwanICAN48429.2020.9181314.
- [10] Singh, S., Jagdev, G. Execution of Structured and Unstructured Mining in Automotive Industry Using Hortonworks Sandbox. SN COMPUT. SCI. 2, 298 (2021).
- [11] Brar, T. P. S. (2021). Digital Marketing Performance: Understanding the Challenges and Measuring the Outcomes. In Big Data Analytics for Improved Accuracy, Efficiency, and Decision Making in Digital Marketing (pp. 51-63). IGI Global.
- [12] Mehta, S., Kukreja, V., Bhattacharjee, A., & Brar, T. P. S. (2023, April). Predicting Rice Leaf Disease Outbreaks using CNN-SVM Models: A Machine Learning Approach. In 2023 IEEE International Conference on Contemporary Computing and Communications (InC4) (Vol. 1, pp. 1-5). IEEE.
- [13] Banerjee, D., Kukreja, V., Gupta, A., Singh, V., & Brar, T. P. S. (2023, August). CNN and SVM-based Model for Effective Watermelon Disease Classification. In 2023 3rd Asian Conference on Innovation in Technology (ASIANCON) (pp. 1-6). IEEE.
- [14] Sharma, R. K., Brar, T. P. S., & Gandhi, P. (2021). Defense and Isolation in the Internet of Things. Internet of Things in Business Transformation: Developing an Engineering and Business Strategy for Industry 5.0, 141-168.
- [15] Brar, T. P. S. (2018). Business intelligence in banking: a study of bi technology implementation and challenges. CGC International Journal of Contemporary Technology and Research, 1(1).
- [16] Lata, S., & Singh, D. (2022, April). A Hybrid Approach for Cloud Load Balancing. In 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE) (pp. 548-552). IEEE.
- [17] Lata, S., & Singh, D. (2022). Intrusion detection system in cloud environment: Literature survey & future research directions. International Journal of Information Management Data Insights, 2(2), 100134.
- [18] Lata, S., & Singh, D. (2022, October). Cloud simulation tools: A survey. In AIP Conference Proceedings (Vol. 2555, No. 1). AIP Publishing.