

Expert System for Employee Performance Evaluation Based on HR Analytics Data Using the Certainty Factor Method

Maksimus Aryanto Indrakusuma ^{a,1,*}, Dewi Octafiani ^{b,2}

^a Department of Information System, STMIK Widya Cipta Dharma, Samarinda, Indonesia

^b Department of Informatic, STMIK Amikom Surakarta, Indonesia

¹ 2341026@wicida.ac.id, ² dewi.octafiani@dosen.amikom.ac.id

* corresponding author

Article Info

Article History:

Received Mar 01, 2026

Revised Mar 26, 2026

Accepted April 11, 2026

Keywords :

Expert System, HR Analysis,
Human Resource Management,
Passion Factor, Work Evaluation



This work is licensed under a
[Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/)

Abstract

Employee performance evaluation is one of the important things in human resource management because it can affect the Company's productivity and development. However, the performance appraisal process that is still carried out manually or conventionally often results in subjective and inconsistent appraisals. This makes the results of the evaluation sometimes less effective to be used as a basis for decision-making. This study aims to design an expert system using the Certainty Factor (CF) method to help the process of evaluating employee performance automatically and more objectively by utilizing HR Analytics data. The data used amounted to 791 employees consisting of age, work experience, education level, department, performance score, and performance category. The expert system created has eleven inference rules with five assessment categories, namely Poor, Average, Good, Very Good, and Special. The rules are prepared based on the results of consultation with experienced HR practitioners. The Certainty Factor method is used to determine the level of confidence of the system in classifying employee performance in conditions that have an element of uncertainty. The results showed that the system could produce a combined CF value between 0.46 and 0.84 on several employee profiles. In addition, the system also obtained an accuracy level of 82% in accordance with expert research. Based on these results, the Certainty Factor-based expert system can be used as a decision support in a more structured and data-based employee performance evaluation process.

1. Introduction

In an increasingly competitive global landscape, human resource management has emerged as a fundamental determinant of organizational success. Employee performance evaluation is a systematic process designed to measure an individual's contribution to organizational goals on a regular basis. An effective evaluation system not only provides an accurate representation of employee productivity but also serves as a basic basis for strategic decision-making regarding promotion, career development, and remuneration planning. Therefore, an organization's ability to objectively and consistently assess employee performance is inseparable from their capacity to maintain long-term competitiveness and operational efficiency.

Despite its strategic importance, many organizations continue to rely on manual and subjective evaluation approaches. Assessments that rely entirely on managerial assessments without the support of structured data introduce a significant risk of bias and injustice. [1] argues that performance appraisals based on subjective perceptions often fail to reflect the actual contributions of employees and can produce adverse effects on work motivation and organizational commitment. This situation underscores the urgent need for a more systematic, consistent, and data-driven approach to the performance evaluation process one that can reduce human bias while improving the reliability and interpretability of assessment results.

Advancements in artificial intelligence technology have opened up considerable opportunities to automate and improve the performance evaluation process through the application of expert systems. Expert systems are computer-based systems designed to represent expert knowledge and apply it to solve problems in a specific domain [2]. Such systems have shown effectiveness in a variety of domains, including human resource management, where their capacity to encode institutional knowledge into a reproducible rule-based inference process allows for objective and auditable decision-making. In this study, the Certainty Factor (CF) method originally developed by Shortliffe and Buchanan (1975) [3] was adopted as an inference mechanism, given its well-established ability to represent the uncertainties inherent in expert judgment numerically and to produce measurable and interpretable decisions.

This study aims to design and implement a CF-based expert system for employee performance evaluation by utilizing HR Analytics data as a foundation of empirical knowledge, and to measure the system's confidence level in generating performance classifications. From a theoretical point of view, this research contributes to the development of expert systems models in the domain of human resource management. From a practical perspective, these findings serve as a reference for organizations looking to build an objective, consistent, and data-driven performance evaluation system. This dual contribution theoretical enrichment and practical application positions this research as a relevant and timely effort to bridge the gap between computational intelligence and evidence-based HR practices.

Research at the intersection of expert systems, performance evaluation, and HR analysis has grown rapidly over the past decade, reflecting a broader shift towards data-driven and computing-assisted decision-making in human resource management. Fundamental contributions in this domain establish that rule-based expert systems are capable of encoding complex multi-criteria assessment processes into a systematic inference framework that results in consistent and auditable recommendations a property that is invaluable in the context of high-risk HR where transparency and objectivity are paramount.

A large amount of empirical work has examined the application of the Certainty Factor method as an inference mechanism in expert systems designed for personnel-related tasks. [4] developed a CF-based expert system to identify employee personality types, showing that the method effectively represents the uncertainties inherent in character assessment and generates reliable recommendations for personnel placement decisions [5]. Their findings set an important precedent for applying CF in the context of nondiagnostic HR, expanding the application of the method beyond its original medical domain. Similarly, [6] implemented CF in a web-based expert system for computer network troubleshooting, confirming that the method produces conclusions consistent with high accuracy when applied to multi-variable linguistic inputs [7].

While this study addresses the technical domain rather than HR, its findings confirm the methodological robustness of CF-based inference under conditions of uncertainty analogous to those encountered in performance classifications. Complementary research by [8] investigated a Forward Chaining-based expert system for employee contract evaluation, showing that rule-based systems produce more objective recommendations and substantially reduce reliance on subjective managerial judgments [9]. This contribution reinforces the broader theoretical argument that structured knowledge representation architectures whether CF-based or forward chaining out perform unstructured human judgment in the context of personnel decisions characterized by multiple interacting variables. Furthermore, [10] conducted a comparative analysis of data mining applications for predicting the quality of employee performance, concluding that rule-based classification approaches consistently result in more accurate performance classifications than single-method approaches that lack explicit knowledge representations [11]. These findings directly inform the architectural choices made in this study.

More broadly, the HR Analytics literature has progressively established that data-driven approaches to performance management generate more reliable and actionable insights than traditional evaluation methods. [12] define HR Analytics as the application of systematic analytical processes and methodologies to HR data for the purpose of supporting strategic decision-making [13], a definition that includes the type of data-driven expert system architecture developed in this study. The convergence of expert systems technology and HR Analysis represents an underexplored methodological frontier, especially with respect to the application of CF as a mechanism for measuring evaluative uncertainty across multi-dimensional employee performance profiles. The study addresses this gap by operationalizing CF in a knowledge base derived from empirical HR data, validated through structured expert elicitation an approach that integrates the epistemological power of computational intelligence and domain expertise.

2. Research Methodology

The study adopts an applied research approach that uses systems development methods, designed to generate functional expert system artifacts based on empirical HR data and validated through expert consultation. The research design is descriptive-quantitative, where the knowledge base of the expert system is derived from empirical data patterns and then validated by experienced HR practitioners. This design is appropriate given the dual purpose of the study to establish a technical operational system and empirically assess the accuracy of its classification against recognized expert benchmarks.

The dataset used in this study was taken from the HR Analytics repository consisting of 791 employee records. The dataset included six main variables: age (ranging from 20 to 59 years), work experience (1 to 39 years), education level (categorized as High School, Bachelor's, Master's, and PhD), departmental affiliation (Sales, Technology, Finance, and HR), performance score (on a scale of 1 to 10), and performance category. The distribution of performance categories in the dataset is detailed in Table 1. Notably, the Poor category made up the largest proportion with 31.4% of the total records, while Very Good represented only 11.1%, indicating a class imbalance that carries implications for system design and evaluation sensitivity.

Table 1. Performance Category Distribution in HR Analytics Datasets

Performance Categories	Number of Employees	Percentage (%)	Performance Score
Awesome	88	11.1%	9-10
Excellent	149	18.8%	7-8
Good	156	19.7%	7-8
Average	150	19.0%	5-6
Poor	248	31.4%	1-4
Quantity	791	100%	

The input variables used in building the knowledge base are operationalized as follows: (1) attendance rate, expressed as the percentage of employee attendance during the evaluation period; (2) productivity score, a numerical value that measures employee work output; (3) supervisor evaluation score, on a scale of 1 to 10; and (4) job satisfaction level, expressed as a percentage that reflects self-reported employee satisfaction. These four variables collectively function as antecedent conditions in a rule-based inference framework, while the output variables are performance categories—classified as Poor, Average, Good, Very Good, or Excellent.

The research was conducted through six systematic phases: (1) data collection and pre-processing of HR Analytics data sets; (2) knowledge gathering through structured interviews with experienced HR practitioners; (3) representation of knowledge in the form of IF-THEN production rules; (4) the implementation of CF-based inference engines; (5) testing and validation of the system against expert assessment; and (6) analysis of results and formulation of conclusions. This sequential, iterative methodology ensures that each phase informs the next, resulting in a coherent and empirically based system architecture.

The expert system architecture consists of three functional layers. The data acquisition layer receives input values for attendance, productivity, evaluation scores, and job satisfaction. The inference layer processes these inputs through a CF-based knowledge base, matching the input profile to the rule antecedent and calculating the confidence value for each applicable rule. The output layer generates classified performance categories in addition to the corresponding final CF values, providing categorical decisions and quantitative confidence indicators. The workflow continues from employee data input, through rule matching and individual CF calculations for each enabled rule, to a combination of repeated CF values until a final composite confidence score is obtained as the basis for classification. The Certainty Factor method, introduced by [8], provides an inferential mechanism for measuring confidence under uncertainty. The basic CF formula is expressed as:

$$CF(H, E) = MB(H, E) - MD(H, E)$$

Where MB (Measure of Trust) represents the level of confidence in hypothesis H provides evidence E, and MD (Measure of Trust) represents the appropriate level of trust. When multiple pieces of evidence support the same hypothesis, CF values are repeatedly combined according to:

$$CF_comb(CF1, CF2) = CF1 + CF2 \times (1 - CF1), \text{ if } CF1 > 0 \text{ and } CF2 > 0 \\ CF_comb(CF1, CF2) = CF1 + CF2 \times (1 + CF1), \text{ if } CF1 < 0 \text{ and } CF2 < 0$$

The final system CF is derived by multiplying the expert CF (CF_E), assigned by the HR practitioner for each rule, by the user CF (CF_U), provided by the operator based on field observations:

$$CF_final = CF_E \times CF_U$$

CF values that exceed the 0.2 threshold are considered significant in supporting the hypothesis of a particular performance category. The final classification is determined by selecting a hypothesis that corresponds to the highest composite CF value, ensuring that the system output reflects the most strongly supported performance category given the totality of the observed evidence.

3. Results and Discussion

The knowledge base is built through structured elicitation with two experienced HR practitioners who specialize in performance management. This process results in eleven core inference rules, presented in Table 2, each specifying antecedent conditions, consequence performance categories, and CF values set by experts that reflect the practitioner's level of confidence in the validity of the rule. The set of rules covers the full spectrum of performance categories and discusses operational indicators attendance and productivity and evaluative indicators supervisor ratings and job satisfaction thus ensuring comprehensive coverage of performance construction. This approach is consistent with the conclusion of [14] that structured rulebased classification architectures produce more accurate performance outcomes than methods that do not have explicit knowledge representations.

Table 2. Expert System Knowledge Base Rules for Performance Evaluation

Rules	Condition 1 (IF)	Condition 2 (AND)	Output (LATER)	CF Rules	Category
R01	Attendance $\geq 95\%$	Productivity ≥ 85	Awesome	0.90	Awesome
R02	Attendance $\geq 90\%$	Productivity ≥ 75	Good	0.80	Good
R03	Attendance $\geq 80\%$	Productivity ≥ 65	Average	0.70	Average
R04	Audience $< 80\%$	Productivity < 65	Poor	0.60	Poor
R05	Eval. Score ≥ 8	Satisfaction $\geq 80\%$	Awesome	0.90	Awesome
R06	Wait. $\geq 88\%$ & Eval. 7–8	Satisfaction 70– 84%	Excellent	0.82	Excellent
R07	Eval. Score 6–7	Satisfaction 60– 79%	Good	0.80	Good
R08	Eval. Score 4–5	Satisfaction 40– 59%	Average	0.65	Average
R09	Eval. Score < 4	Satisfaction $< 40\%$	Poor	0.55	Poor
R10	Wait. $\geq 90\%$ & Eval. ≥ 8	Satisfaction $\geq 80\%$	Awesome	0.92	Awesome
R11	Wait. $< 80\%$ & Eval. < 4	Satisfaction $< 40\%$	Poor	0.85	Poor

Rules R01 through R04 focus on the operational dimension of performance, incorporating attendance and productivity scores as key indicators of contribution in the workplace. Rules R05 and R10 define the Excellent category through the convergence of high supervisor evaluation scores and strong job satisfaction, reflecting the evaluative dimensions of peak performance. The R06 rule captures the Excellent category through composite conditions involving above-average attendance, medium to high evaluation scores, and favorable satisfaction levels. Rules R07 through R09 are based on evaluative indicators supervisor ratings and job satisfaction—to classify intermediate and below-average performance. The R10 and R11 rules represent composite rules that integrate the three variables simultaneously, resulting in decisions on the performance extremes with the highest CF values set (0.92 and 0.85 respectively), thus reflecting the highest practitioner confidence in the case of multi-condition limits.

To illustrate the inferential process, a work example is presented for employee E001, whose observation data and corresponding CF values are detailed in Table 3. This example shows how multiple enabled rules contribute to a single composite confidence score through repeated CF combinations.

Table 3. Observational Data for E001 Employees and Corresponding CF Values

Variabel	Observed Values	CF Users (CF_U)	CF Rules (CF_E)
Attendance	92% ($\geq 90\%$)	0.80	0.80
Productivity	78 (≥ 75)	0.75	0.80
Evaluation Score	7 (ranges 6–7)	0.70	0.80
Job Satisfaction	65% (range 60–79%)	0.65	0.80

Based on the data presented in Table 3, Rule R02 is activated (92% attendance $\geq 90\%$; productivity 78 ≥ 75). The CF calculation for R02 continues as follows:

$$CF_step1 = CF_E \times CF_U = 0,80 \times 0,80 = 0,64$$

The R07 rule is activated simultaneously (evaluation score of 7 in the range of 6–7; 65% satisfaction in the range of 60–79%):

$$CF_step2 = CF_E \times CF_U = 0,80 \times 0,70 = 0,56$$

Since both rules point to the same hypothesis (the Good category), the CF values are combined repeatedly:

$$CF_kombinasi = 0,64 + 0,56 \times (1 - 0,64) = 0,64 + 0,2016 \approx 0,84$$

A final CF value of 0.84 indicates a high level of system confidence that E001 employees fall into the category of Good Performance, a classification consistent with the actual labels set in the dataset. These results demonstrate the system's capacity to converge on accurate classifications through a structured accumulation of partial evidence a property that distinguishes CF-based inference from binary classification methods. System-level evaluations were performed on a randomly selected sample of 50 employees. The results for the five representative employees are presented in Table 4, illustrating the distribution of CF values across performance categories.

Table 4. Results of Expert System Evaluation Based on Certainty Factors

ID	Attendance	Productivity	Eval. Score	Satisfaction	Category (CF)
E001	92%	78	7	65%	Good (CF=0.84)
E002	97%	88	9	82%	Excellent (CF = 0.77)
E003	74%	60	3	35%	Poor (CF=0.47)
E004	85%	70	5	55%	Average (CF=0.46)
E005	91%	82	8	78%	Excellent (CF=0.69)

Across a sample of 50 employee evaluations, the system correctly classified performance categories for 41 individuals, resulting in an overall decision accuracy of 82% against expert benchmark assessments. The mean value of CF in all cases evaluated was 0.61, with values ranging from 0.46 to 0.84. This level of accuracy aligns closely with the implementation of comparable expert systems in the HR context, reinforcing the practical feasibility of CF-based inference for performance classification tasks. Analysis of variable contribution revealed that supervisor evaluation scores and attendance rates exerted the most significant influence on the composite CF values generated by the system. These findings are theoretically consistent with the established performance management literature, which identifies supervisory assessments and behavioral indicators such as presence as the primary determinants of evaluated performance outcomes [5]. Furthermore, this corroborates the conclusions of [15], who show that fact-based rule systems produce more objective and reproducible recommendations compared to conventional managerial assessments.

The distribution of CF values across performance categories deserves additional analytical attention. Employees classified as Excellent consistently produce CF values exceeding 0.65, reflecting the convergence of several high-confidence rules and the strengthening effect of composite rule activation. In contrast, employees classified as Poor produce CF values in the range of 0.47 to 0.55, indicating relatively greater residual uncertainty. This pattern is theoretically expected: the middle categories especially Average and Poor occupy regions of the performance continuum where boundary conditions are poorly defined, resulting in a more ambiguous evidence profile and consequently lower confidence accumulation.

The inherent continuous nature of real-world performance data means that sharp categorical boundaries will inevitably result in cases where the available evidence supports multiple hypotheses adjacent to similar CF quantities. In such cases, the system correctly identifies the dominant hypothesis while signaling through the CF value that the classification brings greater epistemic uncertainty.

In addition, the class imbalances observed in the dataset where the Poor category makes up 31.4% of the records while the Excellent category accounts for only 11.1%—represents a structural characteristic that requires consideration in future system development. The current rule-weighting scheme, which stems from expert elicitation rather than statistical optimization of the dataset, may indicate a decrease in sensitivity to minority categories such as Excellent. These limitations are consistent with the challenges documented in the broader literature on machine learning-based maintenance predictions, where unbalanced datasets systematically bias the model against majority-class predictions unless explicitly addressed through resampling or algorithmic adjustment [13].

Future iterations of the system will benefit from the integration of data resampling techniques or differential rule weighting strategies to improve classification sensitivity across all performance categories. The practical implications of this finding are quite large. The 82% accuracy rate, achieved through transparent and interpretable rules-based mechanisms, represents a significant improvement over unstructured subjective assessments while maintaining adequate explanations for operational implementation in the context of the organization's HR. Unlike black-box machine learning classifiers, CF-based expert systems provide HR practitioners with categorical decisions and measurable trust measures, allowing for automatic calibration of recommendations against contextual organizational knowledge. This property of interpretability is particularly valuable in sensitive personnel decisions such as promotion eligibility or performance improvement planning where decision transparency and auditability are legally and ethically required. The integration of the Certainty Factor framework in the knowledge base informed by HR Analytics thus represents a methodologically sound and practically feasible approach to evidence-based performance evaluation.

4. Conclusion

This research successfully designed and implemented an expert system for employee performance evaluation using the Certainty Factor method, based on an empirical HR Analytics dataset of 791 employee records. The system classifies employee performance in five categories Excellent, Excellent, Good, Average, and Poor achieving 82% decision accuracy compared to the assessment of expert HR practitioners. The Certainty Factor method shows its effectiveness in representing the uncertainties inherent in the performance evaluation process, resulting in measurable and interpretable confidence values with a combined CF ranging from 0.46 to 0.84 across diverse employee profiles. The supervisor's evaluation score and attendance rate emerged as the most decisive variables in the system's inferential process, while composite rules integrating three simultaneous conditions resulted in the highest practitioner confidence levels, resulting in CF values of 0.92 and 0.85 in the extreme performance category.

The findings of this study carry both theoretical and practical significance. Theoretically, this research contributes to the development of expert systems models in the domain of human resource management, showing that CF-based inference can be effectively operationalized in the HR Analytics knowledge architecture to produce consistent, objective, and measurably uncertain classification of performance. The integration of empirical data with practitioner expertise operationalized through the elicitation of structured knowledge and expert-assigned CF values represents a methodological contribution that bridges computational intelligence and evidence-based HR practices. In practical terms, the 82% accuracy rate, combined with the interpretability and transparency inherent in the system, validates the approach as a viable decision-support instrument for an organizational context where objectivity, consistency, and auditability of performance appraisals are organizational imperatives.

Nevertheless, some of the limitations of current implementation require recognition and should inform direction for future research. Class imbalances observed in the dataset particularly the lack of representation of the Excellent category can limit the system's sensitivity to minority performance classes. Therefore, future work should explore the integration of resampling techniques or adaptive rule weighting mechanisms to improve classification performance across all categories.

In addition, the knowledge base currently operates on four input variables, which, although empirically based, represent part of the multidimensional construction of employee performance. Extending the system to incorporate additional HR Analytics dimensions such as employee engagement indicators, training participation records, and cross-functional collaboration metrics will increase the descriptive richness and predictive power of the system.

The development of a hybrid CFFuzzy Logic architecture is a very promising avenue to overcome the ambiguity of categorical boundaries identified in intermediate performance classifications. In addition, cross-sectoral validation of the system on datasets drawn from diverse industry contexts will strengthen the generalization of the findings. Finally, the development of a real-time web-based interface will facilitate practical implementation by HR managers, enabling continuous, data-driven performance monitoring at scale. These recommendations collectively describe a research agenda oriented towards the development of a more adaptive, inclusive, and organizationally responsive expert system for human resource performance management.

REFERENCES

- [1] E. Turban, R. Sharda, and D. Delen, *Decision Support and Business Intelligence Systems*, 10th ed. Pearson Education, 2019.
- [2] G. Dessler, *Human Resource Management*, 16th edition of Pearson Education, 2020.
- [3] R.L. Mathis and J.H. Jackson, *Human Resource Management: Essential Perspectives*, 8th edition of Cengage Learning, 2018.
- [4] J. H. Marler and J. W. Boudreau, "An evidence-based review of HR analytics," *International Journal of Human Resource Management*, vol. 28, no. 1, pp. 3–26, 2017, doi: <https://doi.org/10.1080/09585192.2016.1244699>.
- [5] E. H. Shortliffe and B. G. Buchanan, "Improper reasoning models in medicine," *Mathematical Biosciences*, vol. 23, no. 3–4, pp. 351–379, 1975, doi: [https://doi.org/10.1016/0025-5564\(75\)90047-4](https://doi.org/10.1016/0025-5564(75)90047-4).
- [6] R. Agusli, Sutarman, and Suhendri, "Expert system for identifying employee personality types using the certainty factor method," *SISFOTEK GLOBAL Journal*, vol. 7, no. 1, pp. 21–27, 2017.
- [7] M. F. Asnawi and Y. Y. Sunarto, "Expert system troubleshooting computer network using the certainty factor method," *Device*, vol. 11, no. 2, pp. 39–47, 2021, doi: <https://doi.org/10.32699/device.v11i2.2168>.
- [8] E. Asoka, L. Rahmi, A. Pameli, and M. Afdhaludin, "Employee Employment Contract Evaluation System Using Expert System Forward Chaining Method," *JSAI: Journal Scientific and Applied Informatics*, vol. 7, no. 1, pp. 120–134, 2024, doi: <https://doi.org/10.36085/jsai.v7i1>.
- [9] H. M. Adria, "Comparative analysis of data mining applications in predicting the quality of employee performance using the C4.5 algorithm method," *JATISI*, vol. 9, no. 2, pp. 1655–1665, 2022, doi: <https://doi.org/10.35957/jatisi.v9i2.1992>.
- [10] R. Malhotra and K. Lata, "Using ensembles for class imbalance problems to predict the maintenance of open-source software," *International Journal of Reliability, Quality and Safety Engineering*, vol. 27, no. 5, p. 2040011, 2020, doi: <https://doi.org/10.1142/S0218539320400112>.
- [11] Filip, F.G., Zamfirescu, CB., Ciurea, C. (2017). *Decision Support Systems*. In: *Computer-Supported Collaborative Decision-Making. Automation, Collaboration, & E-Services*, vol 4. Springer, Cham. https://doi.org/10.1007/978-3-319-47221-8_2
- [12] Nermend & Łatuszyńska, Eds. (2017). *Selected Issues in Experimental Economics*, Springer Proceedings in Business and Economics from Springer. ISBN: 978-3-319-28419-4
- [13] C. F. Rodríguez-Hernández, M. Musso, E. Kyndt, E. Cascallar (2021). Artificial neural networks in academic performance prediction: Systematic implementation and predictor evaluation. Vol. 2. <https://doi.org/10.1016/j.caeai.2021.100018>

- [14] H.S. Yıldız Aybek, M.R. Okur (2018). Predicting achievement with artificial neural networks: The case of Anadolu university open education system. *International Journal of Assessment Tools in Education*, 5 (3), pp. 474-490, doi:<https://doi.org/10.21449/ijate.435507>
- [15] Y Yupiter, M. Zaenudin, A. Pangestu, R. Ridlo Al Hakim, M. Yusro4), Yanuar Zulardiansyah Arief3,5), Ryan Andikawidi Purnama Putra. Expert System for Academic Staff Talent Management. *The 2nd International Seminar of Science and Technology AIP Conf. Proc.* 3048, 020027-1–020027-8; <https://doi.org/10.1063/5.0207212> Published under an exclusive license by AIP Publishing. 978-0-7354-4915-2/\$30.00