# A NEURAL NETWORK APPROACH TO UNDERSTANDING EMPLOYEE RETENTION DYNAMICS: INSIGHTS FROM FEATURE IMPORTANCE ANALYSIS

Han-Chen Huang

Department of Tourism and MICE, Chung Hua University, Hsinchu, Taiwan.

### *ABSTRACT*

*In today's business environment, employee retention has become a significant challenge, as employee turnover can lead to decreased productivity, increased costs, and reduced morale. This study aims to leverage neural network technology to predict employees' retention intentions and conduct in-depth analyses based on employee data. A classification model was established considering various factors, including gender, marital status, number of children, education level, years of service, weekly working hours, career development opportunities, salary, and bonuses, to identify the potential risk of employee turnover. The results indicated that the model achieved an accurate classification rate of 95.12% on the test set, demonstrating high effectiveness in identifying employee retention intentions. Feature importance analysis revealed that education level (29.63%), bonuses (27.50%), and the number of children (21.81%) were the primary factors influencing retention decisions. Additionally, working hours, marital status, and career development opportunities also impacted employees' retention intentions. This research not only provides insights into employee mobility but also offers data support for enterprises to develop effective retention strategies, suggesting that businesses prioritize enhancing education and training opportunities, designing competitive bonus systems, and providing flexible work arrangements and benefits tailored to employees with children.*

### *KEYWORDS*

*Neural Network Model, Employee Retention, Feature Importance Analysis*

## 1. INTRODUCTION

Employee retention has emerged as a critical issue for modern organizations. High turnover rates disrupt operations, escalate costs, and undermine team cohesion, ultimately affecting overall productivity and employee morale [1-3]. As organizations strive to maintain a stable workforce, human resource departments must proactively identify employees at risk of leaving. Timely intervention is essential to implement effective retention strategies that can mitigate the negative impacts of turnover [2-4]. Predicting employee behavior is inherently complex, influenced by various factors including individual circumstances, organizational culture, and external economic conditions. To address these challenges, organizations are increasingly leveraging advanced technologies such as machine learning and artificial intelligence to enhance the accuracy of retention predictions [5,6]. These technologies provide powerful tools to analyze vast datasets, uncovering insights that traditional methods might overlook.

DOI: 10.5121/ijcsit.2024.16502

Neural networks, in particular, are gaining attention for their capacity to process large amounts of data and identify intricate patterns within it [7,8]. By analyzing historical employee data, neural networks can effectively model and predict employees' intentions regarding retention. This ability to discern subtle correlations makes neural networks an appealing choice for tackling the multifaceted issue of employee turnover [9-12].

The current study aims to develop a neural network model that accurately classifies employees' intentions to stay or leave, utilizing actual employee data for model training and testing. This approach will provide businesses with actionable insights to enhance employee retention rates, ultimately fostering greater organizational stability and competitiveness.

## 2. LITERATURE REVIEW

The issue of employee retention has been widely studied in the fields of organizational behavior and human resource management. According to Das (2013), effective employee retention strategies are crucial for organizations seeking to maintain a competitive advantage in today's dynamic business environment. The author reviews a wealth of literature that emphasizes how factors such as job satisfaction, work-life balance, and organizational culture significantly influence employees' decisions to stay or leave [1]. This foundational work underscores the necessity of addressing employee issues comprehensively to foster loyalty and reduce turnover.
Nair (2024) highlights the evolving nature of employee retention, suggesting that organizations should reconsider their notice period policies as a strategic tool for retention. This paper emphasizes that when employees feel that their departure would disrupt team dynamics, they may be more likely to reconsider their decision to leave [2]. Similarly, Bajaj (2022) examines the factors influencing employee retention among IT professionals in India, identifying key elements such as career advancement opportunities, professional development, and job security that significantly affect retention rates [3].

Shrestha and Prajapati (2023) provide insights into the relationship between human resource management practices and employee retention. Their study shows that positive HR practices, such as effective communication, recognition programs, and employee engagement initiatives, significantly enhance employee loyalty and retention [4]. Moreover, Malik et al. (2020) explore the role of perceived supervisor support in the relationship between HR practices and employee retention, concluding that a supportive supervisory environment fosters greater employee commitment and satisfaction [5].

The integration of advanced technologies, particularly AI-driven predictive analytics, is becoming an important strategy for improving employee retention. Basnet (2024) discusses how artificial intelligence can analyze vast amounts of employee data to identify potential turnover risks, enabling organizations to take proactive measures to retain key talent [6]. Huang (2014) and Liu and Lai (2021) emphasize the importance of artificial intelligence in demand forecasting and advanced machine learning algorithms, such as convolutional neural networks, in enhancing predictive capabilities across various domains, including employee retention [7,8].

Additionally, existing literature highlights the critical role of employee satisfaction in retention. Al Kurdi et al. (2020) establish a theoretical and empirical foundation for understanding how employee satisfaction directly correlates with customer satisfaction and organizational performance, reinforcing the notion that satisfied employees are less likely to leave [9]. Insights from Murtiningsih (2020) further support this idea, indicating that compensation, training, and organizational culture are crucial in shaping job satisfaction and, consequently, employee retention [10]. Ineson et al. (2013) and Jung and Yoon (2013) also contribute to this discussion, emphasizing the significance of employee loyalty, indicating that loyalty directly impacts customer satisfaction and repeat business [11,12].

The advancements in neural network models offer promising avenues for enhancing prediction accuracy in employee retention. Studies by Dai (2024) and Jiang et al. (2024) explore the implications of neural networks in forecasting employee behavior, suggesting that these models can capture complex patterns that traditional methods may overlook [13,14]. Goodfellow et al. (2016) and Hagan et al. (2014) provide foundational insights into the design and functionality of neural networks, emphasizing their capabilities in handling intricate datasets [15,16]. These advancements in deep learning technologies signify a shift towards more sophisticated, data-driven approaches in understanding and improving employee retention strategies.

In conclusion, the literature presents a comprehensive view of employee retention, identifying crucial factors and advanced methodologies that can enhance retention strategies. As organizations continue to navigate the complexities of employee behavior, integrating advanced predictive analytics and fostering a supportive work environment will be key to improving employee retention.

## 3. Research Methodology

This study aims to apply neural networks to analyze employee data to predict retention intentions and identify potential turnover risks. The research methodology includes dataset selection, data preprocessing, model development, training and testing processes, parameter tuning, and evaluation and analysis of results.

### 3.1. Dataset Description

The dataset comprises 198 employee records, with each record including 9 input features and 1 target variable. The input features include gender, marital status, number of children, education level, years of service, working hours, career development opportunities, salary, and bonuses. The target variable is the employee's retention intention (Stay or Leave) (Table 1). These features cover multiple dimensions influencing employee retention or turnover, and the model aims to predict employees' intentions based on these features:

Table 1. Input and Output Features

| **Feature** | | **Description** |
|---|---|---|
| Input | Gender | Employee's gender |
| | Marital Status | Employee's marital status |
| | Number of Children | Number of children of the employee |
| | Education Level | Employee's highest education level |
| | Years of Service | Number of years the employee has worked |
| | Working Hours | Average weekly working hours of the employee |
| | Career Development Opportunities | Number of career development opportunities within the company |
| | Salary | Average monthly salary of the employee |
| | Bonuses | Average monthly bonuses of the employee in the previous year |
| Output | Employee Intention | indicates the most probable employee intention: "Stay" or "Leave" |

### 3.2. Model Training and Testing

The dataset was split into three subsets: a training set (58.58%), a validation set (20.71%), and a test set (20.71%). This allocation ensures that the model learns on most of the data, is fine-tuned using the validation set, and is finally evaluated on unseen test data. A backpropagation neural network was employed to predict whether an employee would stay or leave. The network structure includes an input layer (with 9 neurons corresponding to the 9 input features), 1 or 2 hidden layers (to capture complex relationships between features), and an output layer, used to predict retention intentions (Stay or Leave).

The training process includes:

- Activation Functions: The hidden layers use an activation function, which is commonly used for its ability to capture non-linear relationships. The output layer uses an activation function, ideal for binary classification tasks, converting results to probabilities between 0 and 1.
- Loss Function: The binary cross-entropy loss function was used, which is a standard choice for binary classification problems.

## 4. MODEL DEVELOPMENT AND ANALYSIS

This study uses a neural network model to analyze employee retention intentions through precise data modeling, aiming to reveal the key factors influencing employee retention. The model's development process includes architecture design, parameter optimization, and performance evaluation, providing a solid foundation for subsequent predictions and analyses.

### 4.1. Neural Network Architecture

This study employs a neural network-based model with the following specific settings [16-21]:

- Input activation function: Logistic
- Output error function: Cross-entropy
- Output activation function: Logistic
- Classification model: Confidence limits
- Acceptance threshold: 0.5
- Rejection threshold: 0.5

The network architecture consists of three layers:

- Input Layer: Comprising 17 input neurons corresponding to the 9 input features in the dataset.
- Hidden Layer: Designed with one to two layers, with the number of neurons in each layer optimized to improve model performance.
- Output Layer: Contains one output neuron responsible for binary classification (predicting employee retention or turnover).

The optimal structure of the model was identified through exhaustive search methods, with the hidden layer range set to 1-2 layers and evaluated based on test error. After iterations, the best five network architectures were selected, as shown in (Table 2).

Table 2. Best Five Network Architectures

| Architecture | Akaike's Criteria | Fitness | Test Error |
|---|---|---|---|
| 17-5-1 | -0.000958 | 41 | 0.97561 |
| 17-7-1 | -0.001034 | 41 | 0.97561 |
| 17-16-1 | -0.001599 | 41 | 0.97561 |
| 17-20-1 | -0.002113 | 41 | 0.97561 |
| 17-21-1 | -0.002297 | 41 | 0.97561 |

Akaike's Criterion is a statistical criterion used for model selection. Proposed by Japanese statistician Hirotugu Akaike in 1974 [22], the primary purpose is to select the model that best explains the data among a set of models while considering model complexity. Among the architectures, the 17-5-1 structure performed the best (with 17 input neurons, 1 hidden layer with 5 neurons, and 1 output neuron), with a fitness of 41, test error of 0.97561, and the lowest Akaike's Criteria value, thus selected as the optimal structure.

### 4.2. Model Training and Validation

The neural network model was trained using batch backpropagation, with both the hidden and output layers employing logistic activation functions. The dataset was divided into training (58.58%), validation (20.71%), and test sets (20.71%). Prediction errors were minimized by adjusting model parameters, and classification accuracy was used to evaluate the model's performance. The correct classification rate during the training phase reached 99.17%, with a classification accuracy of 95% during the validation phase. Ultimately, the model achieved a correct classification rate of 95.12% on the test set, demonstrating robust classification capabilities. The following are the classification results for the test set (Table 3):

Table 3. Classification Results

<table>
<tr><th colspan="2" rowspan="2">Employee Retention Intention</th><th colspan="3">Predict Retention Intention</th></tr>
<tr><th>High</th><th>Low</th><th>Correct Predictions</th></tr>
<tr><td rowspan="2">Actual Retention Intention</td><td>High</td><td>30</td><td>1</td><td>30</td></tr>
<tr><td>Low</td><td>1</td><td>9</td><td>9</td></tr>
<tr><td colspan="5">Overall Correct Classification Rate: 95.12%</td></tr>
</table>

From the table, it can be observed that the model can accurately differentiate the majority of employees' retention intentions. Among a total of 41 samples, only 2 samples were misclassified (1 high-intention employee was classified as low intention, and 1 low-intention employee was classified as high intention), resulting in an overall correct classification rate of 95.12%(=39/41). These results indicate that the model can efficiently and accurately identify employees who may leave. Identifying employees at risk of leaving is crucial in organizational management, allowing businesses to take appropriate intervention measures in advance, such as providing better development opportunities or adjusting compensation and benefits.

Further analysis of the impact of the nine input features on the model's output was conducted using feature importance to measure each feature's contribution to the prediction results. The importance percentages of each feature are presented in Table 4:

Table 4. Importance Percentages of Each Feature

| Feature | Importance (%) |
|---|---|
| Education Level | 29.63 |
| Bonus | 27.50 |
| Number of Children | 21.81 |
| Working Hours | 7.97 |
| Marital Status | 5.63 |
| Career Development Opportunities | 4.66 |
| Salary | 2.15 |
| Gender | 0.03 |
| Years of Service | 0.63 |

From the table, it is evident that “**Education Level**” (29.63%), “**Bonus**” (27.50%), and “**Number of Children**” (21.81%) have the most significant impact on employee retention decisions. This suggests that enhancing employee education and creating competitive bonus systems could improve retention rates. Specifically, increasing education levels not only aids in employees' career advancement but also fosters greater loyalty to the organization. Companies should consider providing ongoing training and development opportunities to support professional growth.

Additionally, factors such as “**Working Hours**” (7.97%), “**Marital Status**” (5.63%), and “**Career Development Opportunities**” (4.66%) also influence retention intentions. Although these factors have a relatively smaller impact, they should not be overlooked. Reasonable working hours and a good work-life balance can enhance employee satisfaction, thereby influencing their retention decisions. Moreover, marital status may affect employees' expectations and needs regarding work, which should be considered in the formulation of human resource strategies.

Ultimately, these findings provide organizations with concrete suggestions for improving employee retention, emphasizing that effective employee management should focus on creating a supportive and rewarding work environment. Specifically, organizations should pay attention to employees' individual needs and career development and implement corresponding incentive measures to enhance retention rates and overall job satisfaction.

## 5. CONCLUSION

This study demonstrates the potential of neural networks in predicting employee retention intentions through a comprehensive analysis of employee data. By effectively classifying employees based on various features, the model provides valuable insights into retention risks. The achieved overall correct classification rate of 95.12% reflects the model's strong performance in accurately identifying employees' intentions to stay or leave.

Feature importance analysis reveals that education level, bonuses, and the number of children significantly influence employee retention decisions. Organizations can utilize these findings to develop targeted retention strategies, such as enhancing training opportunities, improving compensation packages, and tailoring benefits to employees with families.

In conclusion, neural networks are a powerful tool for predicting employee retention, offering businesses actionable insights that can lead to improved employee satisfaction, reduced turnover, and enhanced organizational performance. Future research could explore additional variables and advanced model architectures to further improve prediction accuracy and deepen understanding of employee retention dynamics.

## REFERENCES

[1] Das, Bidisha. (2013). Employee Retention: A Review of Literature. IOSR Journal of Business and Management. 14, 8-16. https://doi.org/10.9790/487X-1420816.
[2] Nair, Manju. (2024). Employee retention a top priority and what if it doesn't work: Time to look at an ideal notice period. Human Systems Management. 43(1), 1-15. https://doi.org/10.3233/HSM-230220.
[3] Bajaj, K.K. (2022). Investigation into Factors Influencing Employee Retention Among IT Professionals: A Perspective from India. Journal of Survey in Fisheries Sciences. 8(3), 457-467. https://doi.org/10.53555/sfs.v8i3.2441.
[4] Shrestha, Prakash and Prajapati, Dhan. (2023). Human Resource Management Practices and Employee Retention. Journal of Balkumari College. 12, 1-9. https://doi.org/10.3126/jbkc.v12i1.60415.
[5] Malik, E., Baig, S. A., and Manzoor, U. (2020). Effect of HR practices on employee retention: The role of perceived supervisor support. Journal of Public Value and Administrative Insight, 3(1), 1-7.
[6] Basnet, Sunil. (2024). The Impact of AI-Driven Predictive Analytics on Employee Retention Strategies. International Journal of Research and Review. 11, 50-65. https://doi.org/10.52403/ijrr.20240906.
[7] Huang, Han-Chen. (2014). A study on artificial intelligence forecasting of resort demand. Journal of Theoretical and Applied Information Technology, 70, 265-272.
[8] Liu, Kuang-Tai, and Lai, Yu-Ting. (2021, May 21). Defect Detection in Coffee Beans Using the YOLO Module and Convolutional Neural Networks. In The 2021 Conference on Enterprise Competitiveness & Management (pp. 410-416). Chung Hua University, Hsinchu, Taiwan.
[9] Al Kurdi, Barween and Alshurideh, Muhammad & Alnaser, Ahmad. (2020). The impact of employee satisfaction on customer satisfaction: Theoretical and empirical underpinning. Management Science Letters. 3561-3570. https://doi.org/10.5267/j.msl.2020.6.038.
[10] Murtiningsih, Retno. (2020). The Impact of Compensation, Training Development, and Organizational Culture on Job Satisfaction and employee Retention. Indonesian Management and Accounting Research. 19(1), 33-50. https://doi.org/10.25105/imar.v19i1.6969.
[11] Ineson, E., Benke, E., & László, J. (2013). Employee loyalty in Hungarian hotels. International Journal of Hospitality Management, 34(1), 31-39. https://doi.org/10.1016/j.ijhm.2012.04.001.
[12] Jung, Hyo and Yoon, Hye Hyun. (2013). Do employees' satisfied customers respond with a satisfactory relationship? The effects of employees' satisfaction on customers' satisfaction and loyalty in a family restaurant. International Journal of Hospitality Management. 34, 1-8. https://doi.org/10.1016/j.ijhm.2013.02.003.
[13] Dai, Ying. (2024). Predicting the Reasons of Employee Turnover Based on BP Neural Network Model. Highlights in Science, Engineering and Technology. 100, 119-127. https://doi.org/10.54097/x2vjgb22.
[14] Jiang, Chunheng and Huang, Zhenhan & Pedapati, Tejaswini & Chen, Pin-Yu & Sun, Yizhou & Gao, Jianxi. (2024). Network properties determine neural network performance. Nature Communications. 15, Article 48069. https://doi.org/10.1038/s41467-024-48069-8.
[15] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
[16] Hagan, M. T., Demuth, H. B., Beale, M. H., & De Jesús, O. (2014). Neural network design (2nd ed.). Martin Hagan.
[17] Ma, Heng, and Lin, Shih-Hung. (2021, May 21). A Study on Algorithm Strategies and Long-Term Holdings: A Case Study of E.SUN Financial Holding Company. In The 2021 Conference on Enterprise Competitiveness & Management (pp. 213-219). Chung Hua University, Hsinchu, Taiwan.
[18] Liu, Kuang-Tai, and Tseng, Tzu-Ling. (2021, May 21). Pineapple Leaf Image Recognition Using Convolutional Neural Networks. In The 2021 Conference on Enterprise Competitiveness & Management (pp. 231-236). Chung Hua University, Hsinchu, Taiwan.

[19] Liu, Kuang-Tai, and Hsu, Chia-Hsuan. (2021, May 21). Recognition of Healthy Banana Leaves Using Convolutional Neural Networks. In The 2021 Conference on Enterprise Competitiveness & Management (pp. 535-540). Chung Hua University, Hsinchu, Taiwan.

[20] Che, Hui-Chung, Lai, Yi-Hsuan, and Wang, Szu-Yi. (2010). Assessment of Patent Legal Value by Regression and Back-Propagation Neural Network. International Journal of Systematic Innovation, 1(1), 31-47.

[21] Lai, Yi-Hsuan, and Che, Hui-Chung. (2009). Modeling Patent Legal Value by Extension Neural Network. Expert Systems with Applications, 36(7), 10520-10528. https://doi.org/10.1016/j.eswa.2009.02.072.

[22] Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19(6), 716-723.