Human Factors, Business Management and Society, Vol. 215, 2026, 43–50
https://doi.org/10.54941/ahfe1007584

Construction of a Model for Estimating Sales Thinking Processes by Learning Tacit Knowledge

Masato Saegusa[1], Masahiko Sakata[2], and Miwa Nakanishi[1]

[1]Keio University, Yokohama, Kanagawa 223-8522, Japan
[2]Mitsubishi Electric Corporation, Chiyoda-ku, Tokyo 110-8310, Japan

ABSTRACT

In BtoB sales, the tacit knowledge possessed by sales personnel is a source of competitive advantage; however, passing on tacit knowledge is difficult, making transfer to junior sales personnel a critical issue. Existing sales support systems often merely recommend the next action without showing the rationale behind the judgment. This study aims to construct a model that learns and estimates the thinking process based on tacit knowledge of experts to provide educational support. First, we extracted specific tacit knowledge through interviews with experts and defined the thinking process in sales activities as a four-stage structure consisting of Phase, Focal Information, Recalled Knowledge, and Decision Making. Second, we conducted a scenario-based questionnaire survey and collected 365 valid responses from experienced sales personnel. We constructed a machine learning model that sequentially estimates these four elements, using the output of the previous stage as the input for the next. In the model evaluation, we selected models prioritizing Recall for the estimation of Information and Knowledge from a training perspective, and Weighted Precision considering importance for Decision Making. Results showed that the integrated model can learn the context from previous stages and simulate the thinking process. Furthermore, we implemented this model into an AI agent. This agent structures and presents the recommended action along with its rationale (Phase, Information, and Knowledge) following the thinking process. This goes beyond merely presenting the correct answer; it is expected to contribute to the development of sales personnel who can make autonomous judgments and execute projects smoothly.

Keywords: Tacit knowledge, Thinking process, Machine learning, Decision support, Sales enablement

INTRODUCTION

In Business-to-Business (BtoB) sales activities targeting corporate customers, sales personnel play a crucial role in building long-term and complex relationships of trust with customers and making various strategic decisions. They not only elicit potential needs to offer optimal proposals but also, in some situations, make critical judgments, such as terminating relationships with customers (Mitrega et al., 2012). Experienced sales personnel utilize tacit knowledge cultivated through experience to anticipate potential troubles or complaints and make sophisticated judgments to avoid risks (Teece, 1998).

Received March 6, 2026; Revised April 4, 2026; Accepted April 18, 2026; Available online July 20, 2026

However, tacit knowledge is difficult to articulate and codify into manuals. Current training for inexperienced sales personnel relies heavily on an inefficient method where skills are acquired through time-consuming practical experience such as On-the-Job Training (OJT) (Naruko, 2006). Furthermore, it is reported to take an average of 18 to 24 months for inexperienced personnel to consistently produce results (Peasley & Hochstein, 2024). Meanwhile, the average tenure of sales personnel has remained under two years recently (The Bridge Group, 2025), making the efficiency of talent development a critical issue for companies. Therefore, organizational talent development requires a "Sales Enablement" perspective—improving and supporting capabilities leveraging the resources of the entire organization rather than relying solely on individual experience (Peterson et al., 2021).

In recent years, support systems such as Sales Force Automation (SFA) and Customer Relationship Management (CRM) have become widespread to improve sales activity efficiency, and research on AI-based support is progressing. For example, one study constructed a model that learns actions with high order probability from text data in sales reports and recommends the next action to take (Nakayama et al., 2022). Another study used machine learning to estimate appropriate rents for restaurant properties by quantifying tacit knowledge, such as the intuition and experience of experts through natural language processing (Arakawa et al., 2018). However, many of these existing studies focus primarily on operational efficiency or provide the correct answers. They do not reveal the thinking process or the rationale behind why an expert reached a specific judgment. Consequently, these approaches are insufficient for transferring tacit knowledge or providing educational support to inexperienced sales personnel. Developing sales talent requires not merely imitating correct answers but learning the thinking process itself to respond flexibly to varying situations (Schön, 1983).

Therefore, the purpose of this study is to construct a model that learns the tacit knowledge of experienced sales personnel and estimates their thinking process. The uniqueness of this study lies in defining judgment in sales activities not merely as a simple input-output relationship, but as a structure comprising specific actions and the thinking process leading up to them, and sequentially estimating these elements using machine learning. This study does not aim to automate operations by simply presenting optimal solutions; instead, by presenting the thinking process to the user, it is expected to contribute to the development of sales personnel capable of anticipating risks and autonomously executing projects smoothly.

TACIT KNOWLEDGE EXTRACTION AND COLLECTION

This study targets BtoB sales within an electronics manufacturing company. Data collection was conducted through interviews and a questionnaire survey.

First, we conducted interviews to understand specific sales activities and extract tacit knowledge possessed by experienced sales personnel. The participants were three experts with extensive sales experience. The interviews covered two main topics. First, participants were asked to recall specific past projects. We investigated the actual sales activities performed and the critical points where they made independent judgments. For each decision point, we elicited the rationale, the reasoning behind the action, and potential alternative actions. Second, participants were presented with case studies derived from the other participants' experiences and asked how they would judge the situation. Similar to the first topic, the content of the judgment, its rationale, and reasoning were explored.

The information obtained from the interviews was visualized using flowcharts. Finally, the individual flowcharts were consolidated into a single comprehensive flowchart covering the entire sales activity process (see Figure 1). The interviews revealed three types of sales activity flows: sales activities triggered by customer inquiries, sales activities triggered by introductions or active prospecting, and periodic sales activities.

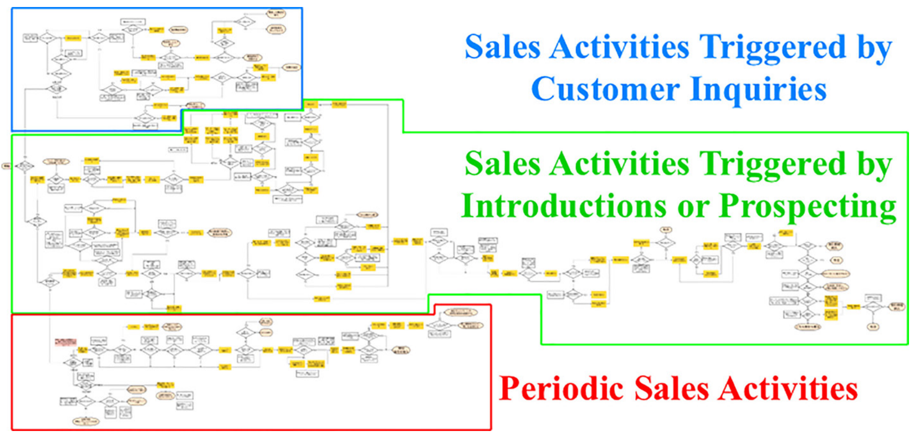

Figure 1: Integrated flowchart of sales activities visualized based on interviews.

Based on the analysis of the flowcharts, we defined the thinking process of sales personnel at each decision point (see Figure 2). This process involves recognizing the current phase and its goal, focusing on specific information, and recalling knowledge based on accumulated experience to reach a final decision. These four elements—phase, information, knowledge, and decision-making—are defined as the tacit knowledge cultivated through experience.

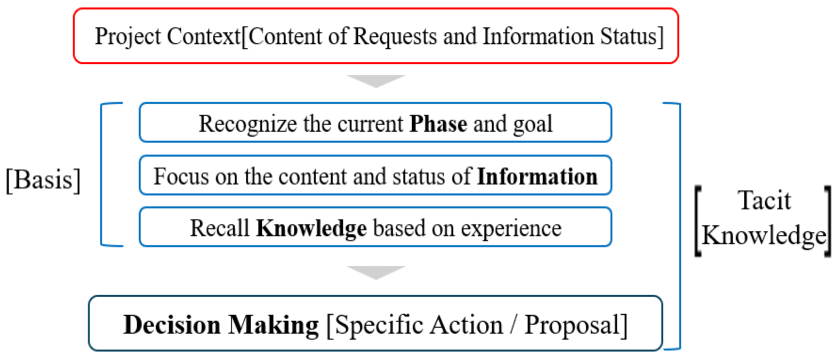

Figure 2: Thinking process of sales personnel at decision points.

Regarding the phases, we defined five phases that constitute a sales project based on the overall flow of the flowcharts. For the tacit knowledge of information, knowledge, and decision-making, elements were extracted from the branching points in the flowcharts. Specifically for the tacit knowledge related to information, we also extracted the detailed classifications for each element. After consolidating similar elements based on their meaning and purpose, we finally obtained 46 elements for information, 57 elements for knowledge, and 66 elements for decision-making.

Second, we conducted a questionnaire survey to collect tacit knowledge utilized in various situations. The participants were sales personnel with at least five years of experience, and the survey was distributed via a web-based platform. The project scenarios were systematically created by combining the elements and classifications of information extracted from the interviews, resulting in a total of 130 distinct scenarios. Each scenario consisted of a list of 15 to 20 information elements and their specific classifications. For each presented scenario, respondents were asked to select the project's current phase, focused information, recalled knowledge, and decision-making from the list of tacit knowledge elements. We instructed the respondents to make judgments specifically for the presented scene rather than considering possibilities for the entire project. A total of 365 valid responses were obtained. This dataset was used to construct the machine learning model.

CONSTRUCTION OF ESTIMATION MODEL

We constructed a four-stage machine learning model to simulate the thinking process of sales personnel. This model is structured to sequentially estimate (1) Phase, (2) Information, (3) Knowledge, and (4) Decision Making based on the input Project Scenario Elements.

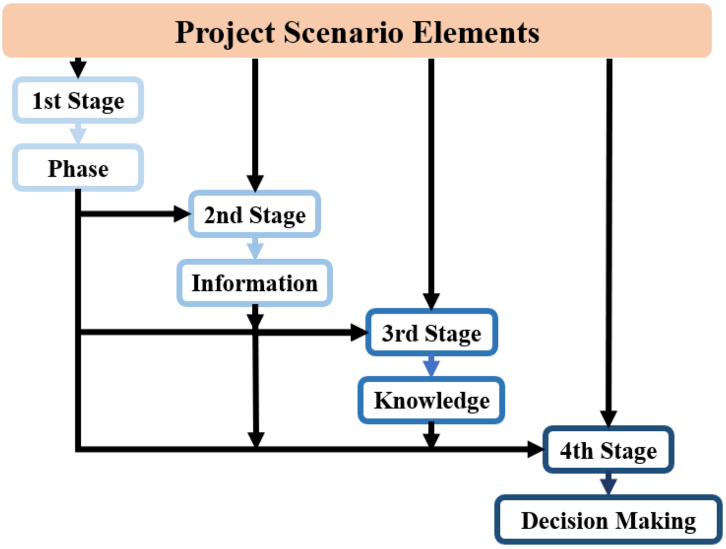

Figure 3: Structure of the four-stage sequential estimation model.

The estimation results output by the previous stage model are added to the input data for the next stage. Through this integrated model, we reproduced the series of thinking processes in which sales personnel recognize the situation, identify the current Phase, focus on critical Information, recall necessary Knowledge, and make a Decision. The detailed input and output variables for each stage are summarized in Table 1.

Table 1: Input and output variables for each stage of the machine learning model.

Stage	Input Variables	Output Variables
1st (Phase)	Project Scenario Elements	Probability of each Phase
2nd (Information)	Project Scenario Elements +Estimated Phase (0/1)	Selection probability of each Information element
3rd (Knowledge)	Project Scenario Elements +Estimated Phase (0/1) +Selection status of Information (0/1)	Selection probability of each Knowledge element
4th (Decision Making)	Project Scenario Elements +Estimated Phase (0/1) +Selection status of Information (0/1) + Selection status of Knowledge (0/1)	Selection probability of each Decision Making element

As preprocessing for the dataset, we unified the responses for Phase and Decision Making based on selection ratios to account for variations in judgment among respondents. Specifically, we calculated the selection ratio for each element within each project scenario and assigned these as the values. Furthermore, to prevent data leakage resulting from the presence of identical project scenarios in both training and test datasets, we performed an 8:2 split based on project scenario units.

To maximize the estimation performance at each stage, we compared and evaluated multiple machine learning algorithms (Random Forest, Extra Trees, Neural Network, LightGBM, XGBoost). We compared these five algorithms for the 1st and 4th stages. However, for the 2nd and 3rd stages, where multiple elements can be selected simultaneously, we also included algorithms applying the Classifier Chain method to account for dependencies between elements. Consequently, a total of eight algorithm types were compared for these stages. Furthermore, we employed Bayesian optimization for tuning the hyperparameters that significantly influence model performance.

In selecting the optimal model, we set distinct evaluation metrics for each stage. For the 2nd and 3rd stages, which estimate Information and Knowledge, we prioritized Recall. This is because our objective is to facilitate user learning by presenting tacit knowledge comprehensively. Conversely, for the 4th Stage (Decision Making), since the judgments of sales personnel significantly influence the smooth progression of the project, the certainty of the presented actions is crucial. Therefore, we defined and utilized Weighted Precision to evaluate whether the correct answers are included among the elements with high recommendation levels. This metric is calculated as the ratio of the sum of the selection ratios of the top five elements predicted by the model to the sum of the selection ratios of the top five elements in the ground truth data. A higher value indicates that the model accurately predicts elements with high selection ratios—that is, important elements—within the top ranks. Through this approach, we aimed to construct a model that satisfies both educational effectiveness and operational stability.

RESULTS AND SYSTEM IMPLEMENTATION

We selected the optimal model for each stage based on the evaluation metrics. Table 2 presents the adopted models, training scores, and integration scores for each stage.

Table 2: Adopted models, training scores, and integration scores for each stage.

Stage	Evaluation Metric	Adopted Model	Training Score	Integration Score
1st (Phase)	Accuracy	Neural Network	0.644	0.644
2nd (Information)	Recall	LightGBM-Chain	0.765	0.807
3rd (Knowledge)	Recall	LightGBM-Chain	0.544	0.800
4th (Decision Making)	Weighted Precision	XGBoost	0.371	0.452

The Accuracy for the 1st Stage (Phase) was 0.644, indicating that the model successfully learned the appropriate phases corresponding to the project scenarios.

For the 2nd and 3rd stages (Information and Knowledge), models employing the Classifier Chain method, which learns relationships between elements, were adopted. This suggests that the selected elements tend to form groups. In fact, since the Information elements are derived from the presented project scenarios, they inherently formed groups specific to each scenario. Similarly, regarding Knowledge, it is inferred that correlations exist between elements, as sales personnel likely recall multiple related knowledge items simultaneously to make decisions. Furthermore, in these stages, the Integration scores (Recall) were higher than the Training scores. This suggests that the model accurately captured the context—including discrepancies in the prediction results of the previous stages—thereby learning the variations in respondent judgments and the sales personnel's thinking process.

Conversely, for the final 4th Stage (Decision Making), the Weighted Precision remained low for both Training and Integration. This can be attributed to the vast number of input features. Given the large number of Decision Making elements (66), it is considered that XGBoost, which learns elements individually, achieved the highest score, rather than Random Forest or Extra Trees, which learn all elements simultaneously. Additionally, since the 4th Stage uses all elements of tacit knowledge regarding Knowledge as input, the learning rules likely became complex, making the model susceptible to overfitting specific data. However, similar to the 2nd and 3rd stages, the Integration score was higher than the Training score, suggesting that the errors from the previous stages had a positive influence.

Based on the constructed integrated model, we implemented an AI agent that simulates the thinking process of sales personnel (see Figure 2). When a user inputs the project scenario elements, the agent presents not only the recommended Decision Making (action) but also the supporting rationale—specifically, the Phase, Information, and Knowledge. By visualizing the thinking process leading to a judgment rather than merely providing the correct answer, the system aims to facilitate user learning in utilizing appropriate tacit knowledge adaptable to specific situations. Regarding the display format, the predicted Phase and its goal are shown. For Information, Knowledge, and Decision Making, the top five elements with the highest probabilities are displayed by default, with lower-ranked elements accessible by expanding the view. Furthermore, to promote autonomous decision-making, a disclaimer is clearly stated, instructing users to make the final judgment themselves.

DISCUSSION

The machine learning model and AI agent constructed in this study can support the learning of inexperienced sales personnel by presenting not only the final action but also the underlying rationale (Phase, Information, and Knowledge) as a thinking process. However, prioritizing Recall during model training to enhance educational effectiveness resulted in a large number of presented tacit knowledge elements. This presents a challenge where the user's cognitive load may increase in situations with time constraints, such as immediately before a business negotiation. Considering that experienced sales personnel make decisions by taking various risks into account to ensure the smooth progression of the project and successful order receipt, refining the AI agent's interface is required to prioritize elements. Specifically, highlighting important elements directly linked to the smooth progression of the project according to the situation is necessary. To achieve this, it is necessary to conduct interviews regarding the essential purposes of sales activities and specific actions directly linked to risks. We believe that based on these findings, the AI agent can be improved to a more effective and safe system.

CONCLUSION

In this study, focusing on BtoB sales at an electronics manufacturer, we collected the tacit knowledge of experienced sales personnel through interviews and a questionnaire survey. We defined the thinking process as a four-stage structure consisting of Phase, Information, Knowledge, and Decision Making, and constructed a machine learning model to estimate these sequentially. Using this model, we implemented an AI agent. Future challenges include the following two points. The first is the improvement of the presentation format. It is necessary to investigate the essential purposes of sales activities more deeply and explore more effective presentation formats. The second is the quantitative verification of educational effects. It is necessary to introduce this system to inexperienced sales personnel and verify its contribution to changes in their thinking processes and improvements in their autonomous judgment capabilities.

ACKNOWLEDGMENT

We would like to thank Mitsubishi Electric Corporation in the interviews and a questionnaire survey.

REFERENCES

Arakawa, S., Suwa, H., Ogawa, Y., Arakawa, Y., Yasumoto, K., & Ohta, T. (2018). Proposal of rent estimation model for restaurant properties based on tacit knowledges. Journal of Information Processing, 59(1), 33–42.

Mitrega, M., Forkmann, S., Ramos, C., & Henneberg, S. C. (2012). Networking capability in business relationships—Concept and scale development. Industrial Marketing Management, IMP 2011 The Impact of Globalization on Networks, 41(5), 739–751.

Nakayama, Y., Mori, M., Saito, S., Narusue, Y., & Morikawa, H. (2022). Construction of Decision-making Support System in Sales Activities by Deep Reinforcement Learning. Journal of Information Processing, 63(4), 1008–1018.

Naruko, Y. (2006). Emerging synthesis of expert knowledge for design and manufacturing. Journal of Information Processing and Management, 49(8), 439–448.

Peasley, M. C., & Hochstein, B. (2024). The early-tenure salesperson: Sales effort and sales growth during the ramp-up period. Journal of Personal Selling & Sales Management, 44(3), 219–236.

Peterson, R. M., Malshe, A., Friend, S. B., & Dover, H. (2021). Sales enablement: Conceptualizing and developing a dynamic capability. Journal of the Academy of Marketing Science, 49(3), 542–565.

Schön, D. A. (1983). The reflective practitioner: How professionals think in action. New York: Basic Books.

Teece, D. J. (1998). Capturing value from knowledge assets: The new economy, markets for know-how, and intangible assets. California Management Review, 40(3), 55–79.

The Bridge Group. (2025). TBG 2025 sales development report. The Bridge Group. https://blog.bridgegroupinc.com/hubfs/resources/TBG%202025%20Sales%20Development%20Report.pdf