Interpretable Multi-Class Classification for Client Potential Prediction in B2B IT Sales

***Abstract*—The modern B2B IT sales environment is characterized by long, complex sales cycles and multi-stakeholder decision-making, which presents a significant challenge for sales teams in prioritizing leads and maximizing conversion rates. This research addresses this problem by proposing a supervised machine learning approach to predict a multi-class "Client Potential Label" (High, Medium, Low, or No potential). The study's core objectives are to develop a classification model and to employ Explainable AI (XAI) techniques to provide a transparent, interpretable rationale for its predictions. The methodology involved a multi-class classification task on a private dataset of 700 historical B2B IT sales records during the period 2020-2022. After a comprehensive data preprocessing, modeling and evaluation phase, which included 10-fold cross-validation, hyperparameter tuning and a comparison of various algorithms, the research found that Gradient Boosting was the top-performing classifier with the accuracy, precision, recall and F1 score of 76%, 76.3%, 75.5%, and 76.2% respectively. Crucially, the application of the Feature Importance method, a XAI technique, revealed that Budget in USD, Engagement Level, and Required Service Mobile Application were consistently the most influential features in predicting client potential. These insights provide valuable and data-driven business intelligence for sales strategies.**

Keywords—Classification, Machine Learning, B2B IT sales, Explainable AI

I. Introduction

In today's highly competitive B2B technology sales environment, companies must make informed decisions about how they allocate time and resources to engage potential clients. Since B2B IT sales are characterized by long sales cycles, multiple stakeholders, and complex procurement processes, many sales teams still rely heavily on manual judgment and experience when qualifying leads as demonstrated in Fig. 1. Sales teams often spend time on low-potential leads, missing opportunities to engage with high-value prospects. This results in wasted resources, misaligned sales and marketing efforts, and reduced overall conversion rates.

Industry statistics show that only 10–20% of sales-qualified leads (SQLs) ultimately result in closed deals [1]. Moreover, 79% of marketing leads never convert into sales due to a lack of effective lead nurturing and prioritization [2]. Studies showed that companies find it difficult to analyse data opportunities that have the potential to be lost or won in the early stages of the sales pipeline. If potential label opportunities are not detected in early, it can lead to a risk of many opportunities being lost, and the company may be unable to estimate resource requirements needed to serve a new client [3].

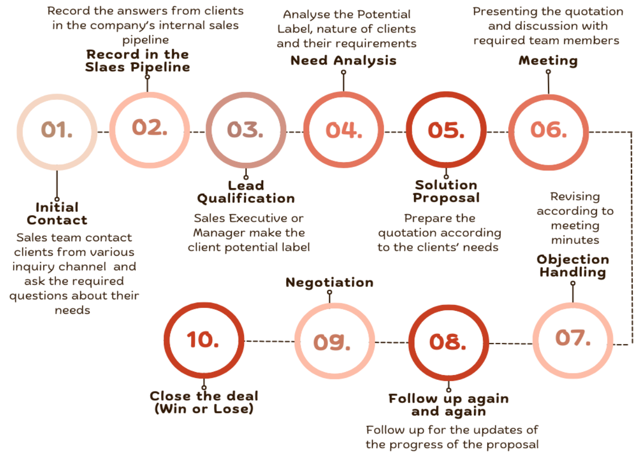

Fig. 1. Long and Complex Process of B2B IT sales

To tackle this challenge, machine learning (ML) is a promising solution since it offers an ability to learn patterns from historical sales data and generate predictive insights. By classifying potential clients into categories such as high, medium, low, or no potential, ML can support sales teams in prioritizing opportunities and allocating resources more effectively.

While prediction accuracy is valuable, relying solely on a classification model presents significant limitations in business decision-making. Predictions without explanation often fail to gain trust from managers and domain experts, as decision-makers need to understand why the model produces a particular output. Moreover, even when predictive models are applied, the lack of transparency in their decision-making limits trust and adoption by sales professionals. To solve this issue, Explainable AI (XAI) can highlight whether factors such as client budget or decision-making authority are key drivers of conversion. Such interpretability is crucial in B2B sales, as it not only supports better decisions but also bridges the gap between technical outputs and actionable business strategies.

The interpretable classification motivates the current research to develop a machine learning classification model enhanced with XAI techniques, aiming to improve customer prioritization and conversion outcomes in B2B IT sales. By combining predictive accuracy with explainability, the study aims to enhance both the effectiveness and usability of AI in real-world B2B sales environments.

The objectives of this research are 1) to develop and compare multi class classification models for client potential prediction 2) to propose a XAI method to interprete which features influence the prediction.

In this research, literature review is discussed in Section II. In Section III, methodology is proposed which includes the details of the dataset, research method, step by step process of using machine learning algorithms and XAI method. The IV section focuses on the results of the research and conclusion is presented in Section V.

II. Literature Review

A. Theoretical Background

1) Classification

Machine Learning (ML), a subfield of Artificial Intelligence enables algorithms to learn from empirical data and improve performance automatically with experience, similar to human learning [4].

In supervised learning, models are trained on labeled data to learn input–output relationships and evaluated on test data for prediction. Classification algorithms use labeled examples to identify patterns and assign classes to unseen data points [5]. Popular classification algorithms such as Gradient Boosting, LightGBM, XGBoost, Random Forest, Decision Tree, Logistic Regression, K-Nearest Neighbors(KNN), Support Vector Machine(SVM), and Naïve Bayes have been widely applied in prior researches.

Gradient Boosting is an ensemble method that sequentially combines weak decision trees, with each new model correcting errors of the previous ones, thereby improving accuracy and capturing complex data patterns. XGBoost is an optimized and scalable gradient boosting algorithm that builds decision trees sequentially to correct previous errors, while LightGBM is a highly efficient and fast gradient boosting framework that grows trees "leaf-wise" for quicker training on large datasets [6].

The Random Forest algorithm is an ensemble method that combines multiple decision trees to improve model performance. It works by building a random subset of features for each tree and uses a majority vote to determine the final prediction [5].

Decision Trees composed of a root, internal nodes, and leaves, uses the concept of entropy to manage complex datasets in fields like finance and healthcare [7].

Logistic regression is a linear model that analyzes the relationship between one or more predictor (independent) variables and a target (dependent) variable [8].

The KNN algorithm's performance depends on the number of neighbors, k, and the distance metric used, such as Euclidean distance. This metric is based on the core k-NN assumption that similar data points are close to each other [5].

SVM algorithm finds the optimal hyperplane to separate data points of different classes, aiming to maximize the margin between the hyperplane and the nearest data points. The hyperplane's dimension is determined by the number of features in the dataset [5].

The Naïve Bayes algorithm is a probabilistic classifier that applies Bayes' Theorem to classify data. Despite the "naïve" assumption that all features are independent, it is highly effective for tasks like document and text classification, particularly with large datasets [9].

Hyperparameter tuning is the process of identifying the optimal configuration parameters that govern the learning behavior of machine learning models; plays a pivotal role in ensuring model performance and generalization. Common strategies for hyperparameter tuning include Grid Search and Random Search. Grid Search exhaustively tries every possible combination of hyperparameters within a defined grid, while Random Search samples a fixed number of combinations from a specified distribution [10].

2) Explainable AI (XAI)

XAI is a set of processes and methods that allows human users to comprehend and trust the results and output created by machine learning algorithms [11]. The complexity and "black box" nature of modern models limit human interpretability, and high accuracy alone is insufficient for trust. XAI addresses this by providing methods to interpret and explain model outputs, enhancing transparency, fairness, and accountability [11].

Feature importance in XAI refers to the process of identifying which features contribute most significantly to a model's predictions. It is also referred to as feature attribution, feature detection, or model interpretability, and is conceptually related to the statistical notions of estimation and attribution. The outcome of feature importance analysis is typically a score or metric assigned to each feature, which enables the ranking of features according to their relative contribution to predictive performance [12].

Methods for calculating feature importance include Gini importance, which is often used in tree-based models, and Permutation Feature Importance, a model-agnostic method that measures the increase in a model's error when the values of a specific feature are shuffled.

Gini importance is calculated using Gini Impurity. For a dataset with C classes, the Gini Impurity is a measure of misclassification and is calculated as in (1). Importance values are normalized (e.g., scaled to sum to 1 or within [0,1]) to allow comparisons across features. Features are ranked according to their relative contribution. The ranked features are analyzed to provide insights into which factors most affect the model [13].

$$\text{Gini Impurity} = 1 - \sum_{i=1}^{C} (p_i)^2 \tag{1}$$

Where:

- C is the total number of classes.
- p_i is the probability of a randomly selected item belonging to class i.

3) Evaluation Methods

K-Folds cross-validation divides a dataset into k equal-sized folds. In each of the k iterations, one fold is used for validation and the remaining k-1 folds are used for training. This process is repeated until each fold has been used for validation once, and the results are then averaged to provide a more reliable estimate of the model's performance [14].

This research paper uses four key evaluation methods. Accuracy measures the overall percentage of correct predictions. Precision indicates the proportion of positive predictions that are truly positive. Recall determines the proportion of actual positive cases that are correctly identified. Finally, the F1-Score is the harmonic mean of precision and recall, providing a balanced metric for model performance [7].

B. Related Works

This section reviews prior research, academic studies, and industry cases that explore the application of machine learning across diverse business domains.

The application of machine learning to the challenges of B2B sales has been a growing area of research and industry innovation, largely due to the limitations of traditional lead scoring methods and the complexity of managing long sales cycles with multiple stakeholders. Effective lead qualification remains critical, as it directly impacts conversion rates and improves the efficiency of sales and marketing alignment [15].

Academic and industrial research confirms the viability of using machine learning classification models to predict B2B sales outcomes. In [15], a B2B software company's lead scoring model, developed on real CRM data between 2020 and 2024, evaluated fifteen different classification algorithms. The research found that the Gradient Boosting Classifier achieved superior performance in terms of accuracy and ROC AUC, significantly improving the company's ability to identify high-quality leads compared to its traditional methods.

[3] addressed the challenge faced by B2B sales teams in identifying which opportunities are likely to be lost or won in the early stages of a sales pipeline. The objective was to develop a machine learning model to improve pipeline management and enable earlier, data-driven decision-making.

Another study focused on the inefficiency of processing a high volume of RFQs in B2B spare parts sales, where only about 17% convert to sales. The objective was to use supervised machine learning to predict the sales potential of RFQs and help prioritize sales efforts [14].

Comparative analyses of different ML models have also been conducted. For example, one study tested eight classifiers, including Logistic Regression, Random Forest, Decision Tree, SVM, Naïve Bayes, KNN, XGBoost, and AdaBoost, using an 80:20 split, with Naïve Bayes achieving the highest accuracy [9]. Another study applied SMOTE to handle class imbalance in a survey dataset of 164 respondents, comparing Naïve Bayes, Decision Tree, and Random Forest, with Random Forest demonstrating the best performance under 10-fold cross-validation [7].

In various studies, researchers have applied XAI methods such as SHAP, XGB Explainer for Feature Importance [16], LIME, PDP, ICE [17] and Exploratory Data Analysis [17] to interpret machine learning models and generate business insights. These approaches helped identify both positive and negative drivers of conversion, uncover behavioral patterns, and provide actionable recommendations, such as targeting new visitors or leveraging seasonality effects. Among these techniques, Feature Importance has emerged as the most suitable for business decision-making, as it is straightforward, easy to use, and highly interpretable for non-technical stakeholders [16].

While prior studies have demonstrated the value of ML in lead scoring, pipeline management, and RFQ prioritization, most have focused on binary outcomes (win/loss or convert/not convert) or single-model evaluations. Moreover, the integration of XAI into multi-class prediction tasks in the B2B IT sales domain remains underexplored. Specifically, limited research has examined how ML can classify clients into multiple potential categories (e.g., High, Medium, Low, or None) while simultaneously providing interpretable insights for decision-makers.

This research seeks to address this gap by developing a multi-class classification framework to predict client potential labels from early sales interaction data and employing XAI technique to ensure model transparency and actionable business insights in the B2B IT sales environment.

III. Methodology

This section outlines the step-by-step methodology used in this research, from data collection to model evaluation and interpretation. The proposed methodology for this research is illustrated in Fig. 2. The first phase involves collecting private data from a B2B IT solution company. Following data collection, preprocessing procedures are applied, after which the dataset is used to train nine machine learning classification algorithms. The models are then evaluated, and the Feature Importance technique—an XAI method—is employed to enhance model interpretability.

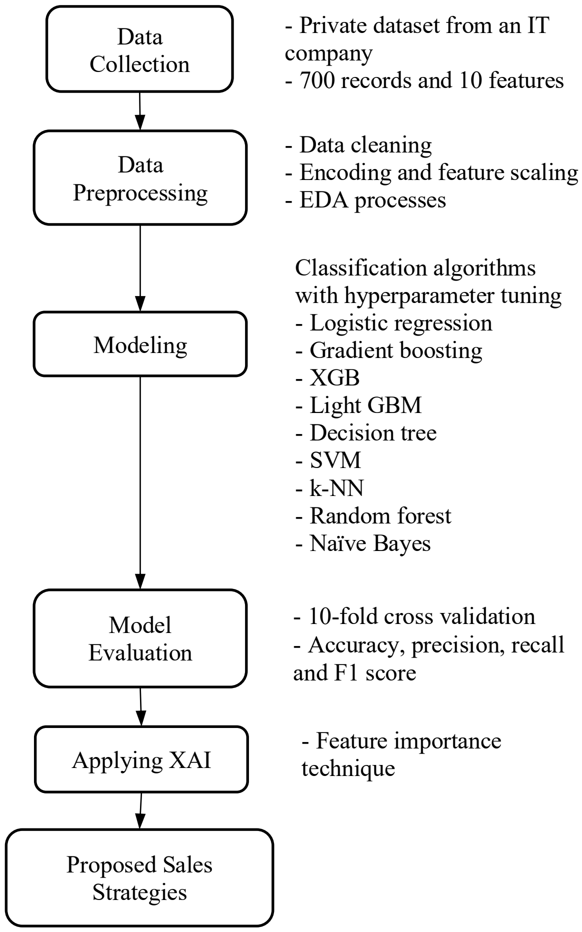

Fig. 2. Proposed Methodology

A. Data Collection

Data were obtained through structured questionnaires administered by B2B IT sales teams from the company's internal Excel-based sales pipeline, which records project inquiries managed by the sales team during the period 2020–2022. The dataset used in this study is a private dataset of a well-known IT company from Yangon, Myanmar comprising approximately 700 records, each containing 10 features per client. The included features are Client ID, Company Size, Required Service, Quotation Type, Meeting Type, Contact Person Role, Urgency Level, Budget in USD, Engagement Level and Potential Label which is target variable.

The description of features in the dataset are presented in Table I.

TABLE I. FEATURES OF THE DATASET

Name	Description	Example
Client ID	A unique identifier for each client	C001, C002
Company Size	The type of the client's company, it is not the ranking of the company.	SME, NGO, Corporate, Other
Required Service	The type of service the client is requesting	Mobile Application, Website, Software, Other IT Services
Quotation Type	The type of quotation that clients requested	Detailed, Standard, Just Pricing
Meeting Type	The type of meeting type that clients requested	Online Platform, In Person, No Meeting
Contact Person Role	The role of the contact person at the client's company.	CEO, IT Team, Sale Staff, Business Development Team, Digital Marketing Staff
Urgency Level	The urgency of the project	High, Medium, Low
Budget in USD	The client's budget in USD	2000, 3000
Engagement Level	The level of engagement level with clients	High, Medium, Low
Potential Label	The target variable for classification	High, Medium, Low and No

B. Data Preprocessing

To prepare the raw data for machine learning, a series of preprocessing steps were performed.

1) Data Cleaning: This step began with correcting a column name by changing "Conatct Person Role" to "Contact Person Role" and removing any extra spaces from all column names. The "Budget in USD" column was then converted to a numeric format (e.g., from 5000 to 5000.0). Missing values in the "Budget in USD" column were handled using Mean Imputation, as only 16 entries were missing. The "Client ID" column was subsequently dropped. An outlier detection analysis was conducted, no outliers were identified, so no data was removed.

2) Encoding and Feature Scaling : The categorical and numerical features were transformed. One-Hot Encoding was applied to categorical features, including Company Size, Required Service, Quotation Type, Meeting Type, and Contact Person Role, converting them into a numerical format. Ordinal Encoding was used for ordinal features such as Engagement Level, Urgency Level, and the target variable Potential Label to retain their natural ordering, with the Potential Label mapped as No = 0, Low = 1, Medium = 2, and High = 3. Finally, Min-Max Scaling was applied to scale all features into a common range.

3) Exploratory Data Analysis (EDA) Insights : The EDA process provided several key insights from the dataset. A Feature-to-Feature correlation matrix conducted in Fig. 3. revealed that there is no significant positive relationship between the features; therefore, no features were removed from the dataset. In terms of correlation between features, the highest correlation was found between 'Quotation Type' and 'Contact Person Role', which were moderately related (0.52). This indicates that certain roles are more likely to request specific quotation types. There is a moderate correlation between the Potential Label and three key features: Budget in USD (0.53), Engagement Level (0.50), and Urgency Level (0.44). This suggests these features are strong predictors of client potential.

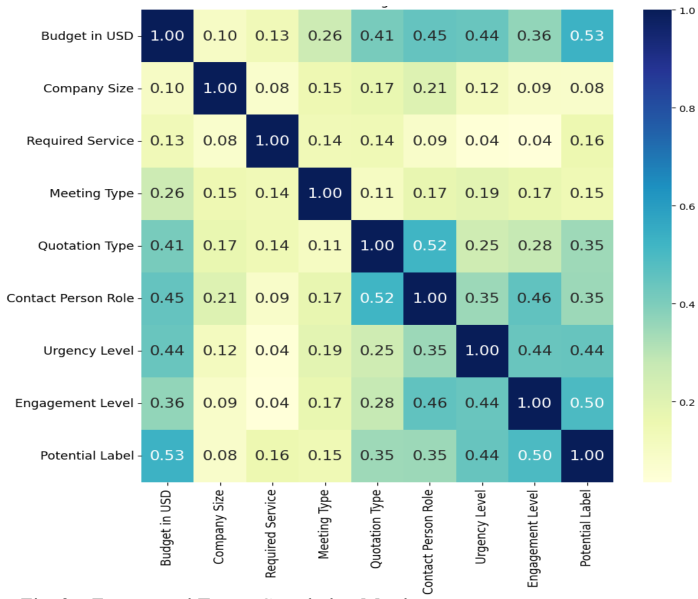

Fig. 3. Feature and Target Correlation Matrix

To better understand the overall data structure and visualize class separability, a PCA plot was generated. A PCA visualization as shown in Fig. 4, showed a significant overlap between the different potential labels, particularly between High(green) and Medium(orange), while Low(blue) and No (pink) potential labels appeared more clearly separated.

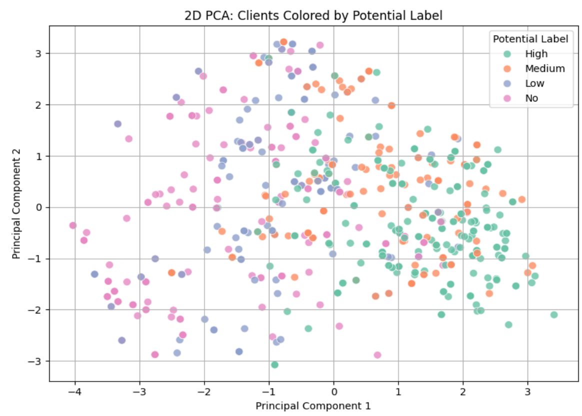

Fig. 4. 2D PCA Plot

C. Modeling

A range of supervised machine learning algorithms were used to predict the multi-class "Potential Label". The algorithms tested included: Gradient Boosting Classifier, Light GBM, XGB, Random Forest, Decision Tree, Logistic Regression, KNN, SVM, Naïve Byes. Each model was trained using the training dataset, and hyperparameter tuning was conducted to optimize the performance of the classification algorithms using Grid Search method. Each model has a set of hyperparameters that control its learning process, complexity, and generalization capability. The optimal parameter configurations identified through this tuning process are presented Table II.

D. Model Evaluation

The performance of the models was evaluated using a comprehensive set of metrics suitable for multi-class classification: Accuracy, Precision, Recall, and F1 Score. To obtain a more reliable estimate of the model performance, a 10-fold cross-validation technique was applied to the entire dataset. In this procedure, the dataset was partitioned into 10 folds, with each fold used once as the testing set while the remaining folds were used for training. This process was repeated until every fold had been used for testing, and the results were averaged to provide robust performance estimates. The mean scores across folds were reported as the final evaluation results.

E. Applying XAI

The Feature Importance method was selected as the primary XAI technique for this research. This method is global explanation technique that reveals the overall contribution of each feature to the model's predictive across the entire dataset. The built-in feature importance functionality of the top two best-performing models was used to compare their results and identify the most consistently influential features.

IV. Results

This section presents the results of the research, beginning with a comparison of classification models' performance using 10 fold cross validation method, optimal parameters and the performance of the models' after optimization. A detailed analysis of the multi-class classification model's performance and interpretability is provided using Feature Importance method.

TABLE II. OPTIMAL PARAMETERS FOR CLASSIFICATOIN ALGORITHMS

Model	Parameters Set
Gradient Boosting	learning_rate= 0.01, max_depth= 3, n_estimators= 300, subsample= 0.8
XGB	colsample_bytree= 1.0, learning_rate=0.01, max_depth=5, n_estimators=200, subsample=1.0
Light GBM	learning_rate= 0.01, max_depth=5, n_estimators= 300, num_leaves=31, subsample= 0.8
LR	C=100, penalty= l2, solver=lbfgs
Decision Tree	criterion= gini, max_depth=10, min_samples_leaf=4, min_samples_split= 2
SVM	C=100, gamma=scale, kernel=linear
KNN	Metric=euclidean, n_neighbors=7, weights=distance
Random Forest	max_depth=None, min_samples_leaf=1, min_samples_split=10, n_estimators=200}
Naïve Bayes	var_smoothing=3e-06

Model training was performed on the normalized dataset using 10 fold cross validation method with optimal parameters and the results showed that the Gradient Boosting model achieved the highest accuracy, precision, recall and F1 score of 76%, 76.3%, 75.5% and 76.2% followed by Light GBM. The performance of the models after optimization is summarized in Table III. This optimization enhances a model's performance by fine-tuning its parameters to find the best possible configuration.

TABLE III. COMPARISON OF CLASSIFICATION MODELS' PERFORMANCE AFTER OPTIMIZATION

Model	Accuracy	Precision	Recall	F1 Score
Gradient Boosting	**0.760**	**0.763**	**0.755**	**0.762**
XGB	0.750	0.750	0.747	0.747
Light GBM	**0.761**	**0.751**	0.733	**0.760**
LR	0.751	0.751	0.745	0.711
Decision Tree	0.751	0.756	0.745	0.755
SVM	0.705	0.701	0.710	0.710
KNN	0.677	0.677	0.686	0.677
Random Forest	0.758	0.733	0.735	0.735
Naïve Bayes	0.627	0.516	0.538	0.519

For interpretability of the model, the built in Feature Importance functionality of the top two best-performing models (Gradient Boosting and Light GBM) is used. As shown in Fig. 5, which presents the Feature Importance for the Gradient Boosting algorithm, Budget in USD is the top ranked feature followed by Required Service_Other IT services, Mobile Application and Engagement Level_Encoded and the lowest rank is Required Service_Website.

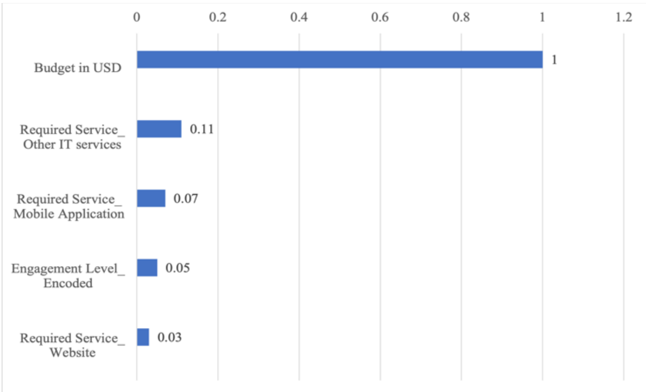

Fig. 5. Top 5 Feature Importance from Gradient Boosting Algorithm

Feature importance analysis was conducted to identify which variables contributed most to client potential prediction. Based on Fig. 6, Both Gradient Boosting and LightGBM ranked "Budget in USD" as the most important feature, confirming that financial capacity strongly influences client potential. Beyond this, the two algorithms differed in how they weighted other features. LightGBM emphasized Required Service (Mobile Application and Software), Urgency Level and Engagement Level, while Gradient Boosting highlighted Other IT Services and Website Requirements. Notably, some features considered important in one model were negligible in the other—for example, Urgency Level was influential in LightGBM but almost irrelevant in Gradient Boosting. This shows that although both models agree on budget as the main driver, they capture different aspects of client engagement and service needs.

In summary, the results show that budget is the strongest predictor, while service type, engagement, and urgency provide additional insights with varying importance across models.

Feature Importance analysis provides the transparency needed to build trust in the model's output. For instance, the sales team can now understand that a high-potential prediction is likely due to a combination of a high budget, high engagement, and specific required service, allowing them to tailor their strategies accordingly.

According to this, the sales team should prioritize leads with high budgets, as this is the most significant predictor of success. They should also focus on clients requesting Mobile Application services and actively work to increase the Engagement Level of all prospects. By focusing on these key factors, they can strategically allocate resources and tailor their approach to close more high-value deals.

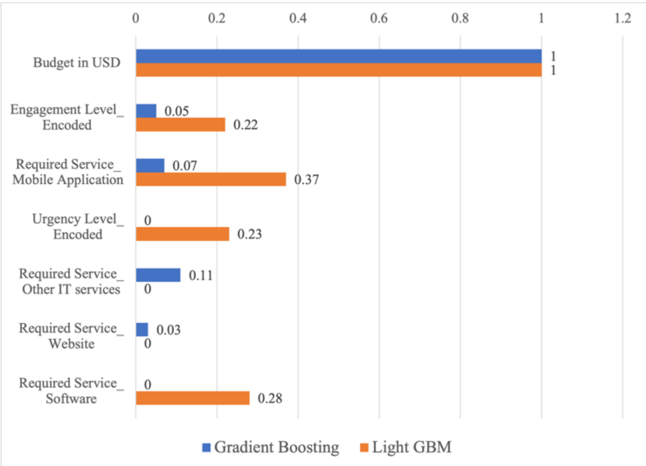

Fig. 6. Top Feature Importance from the Best Performing Algorithms

V. Conclusion

This research was undertaken to develop a multi-class classification model to predict a "Client Potential Label" for prospects and employed XAI to ensure the model's transparency. The research successfully achieved its two primary objectives: developing a predictive model and delivering interpretability for decision support.

The study developed a multi-class classification model to predict client potential in B2B IT sales, with Gradient Boosting and Light GBM achieving the highest performance. Feature Importance analysis identified "Budget in USD", "Engagement Level" and "Required Service_Mobile Application" as the most influential features, providing actionable insights for sales decision-making. These insights provide valuable and data-driven business intelligence for sales strategies

The model's accuracy was limited by class overlap in the feature space, as revealed by PCA. Future work will focus on data increasement, exploring other XAI methods, the application of deep learning techniques and validating the XAI results with domain expert in B2B sales context whether the results of XAI align with their domain knowledge.

Acknowledgment

I am deeply grateful for the financial support from Mae Fah Luang University and the essential dataset provided by a private company from Yangon, Myanmar, both of which were essential in making this research possible.

References

[1] L. Povolna, Marketing Communications on B2B Markets. 2017. doi: 10.7441/dokbat.2017.29.

[2] N. Tanska, "Lead Prioritization in B2B Sales: Creating Simple Rules with Machine Learning".

[3] A. Maulana and T. A. Napitupulu, "Lost Won Opportunity Prediction in Sales Pipeline B2B CRM using Machine Learning," . Vol., no. 10, 2022.

[4] D. J. C. Sihombing, J. U. Dexius, J. Manurung, M. Aritonang, and H. S. Adinata, "Design and Analysis of Automated Machine Learning (AutoML) in PowerBI Application Using PyCaret," in 2022 International Conference of Science and Information Technology in Smart Administration (ICSINTESA), Nov. 2022, pp. 89–94. doi: 10.1109/ICSINTESA56431.2022.10041543.

[5] H. T. Wai and S. Uttama, "Classification of Unsafe Children's Toys Based on NLP-Generated Corpus," in 2023 7th International Conference on Information Technology (InCIT), Nov. 2023, pp. 526–531. doi: 10.1109/InCIT60207.2023.10412857.

[6] J. Dutta, Y. W. Kim, and D. Dominic, "Comparison of Gradient Boosting and Extreme Boosting Ensemble Methods for Webpage Classification," in 2020 Fifth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), Nov. 2020, pp. 77–82. doi: 10.1109/ICRCICN50933.2020.9296176.

[7] P. Panwong, N. Kurubanjerdjit, P. Srikampha, J. Wattanasirisak, P. Chanchi, and W. Chongdarakul, "A Classification Model for Prediction of the Cafe Prospective Customer," in 2024 8th International Conference on Information Technology (InCIT), Nov. 2024, pp. 566–571. doi: 10.1109/InCIT63192.2024.10810550.

[8] R. Bhuvana, S. Maheshwari, and S. Sasikala, "Predict the Heart Disease Using a Logistic Regression Classifier Algorithm," in 2023 12th International Conference on System Modeling & Advancement in Research Trends (SMART), Dec. 2023, pp. 649–652. doi: 10.1109/SMART59791.2023.10428486.

[9] N. Sharma, P. Tiwari, M. Tanwani, A. Pandey, and S. K. Patel, "Customer Relationship Management (CRM) Automation Using Machine Learning Algorithm," in 2023 2nd International Conference on Futuristic Technologies (INCOFT), Nov. 2023, pp. 1–6. doi: 10.1109/INCOFT60753.2023.10425786.

[10] Monica and P. Agrawal, "A Survey on Hyperparameter Optimization of Machine Learning Models," Mar. 2024, pp. 11–15. doi: 10.1109/ICDT61202.2024.10489732.

[11] R. Dwivedi et al., "Explainable AI (XAI): Core Ideas, Techniques, and Solutions," ACM Comput. Surv., vol. 55, no. 9, pp. 1–33, Sept. 2023, doi: 10.1145/3561048.

[12] A. M. Musolf, E. R. Holzinger, J. D. Malley, and J. E. Bailey-Wilson, "What makes a good prediction? Feature

importance and beginning to open the black box of machine learning in genetics," Hum. Genet., vol. 141, no. 9, pp. 1515–1528, Sept. 2022, doi: 10.1007/s00439-021-02402-z.
[13] S. M. Lundberg and S. I. Lee, "A unified approach to interpreting model predictions," in Adv. Neural Inf. Process. Syst., 2017, pp. 4765-4774.
[14] D. Rohaan, E. Topan, and C. G. M. Groothuis-Oudshoorn, "Using supervised machine learning for B2B sales forecasting: A case study of spare parts sales forecasting at an after-sales service provider," Expert Syst. Appl., vol. 188, p. 115925, Feb. 2022, doi: 10.1016/j.eswa.2021.115925.
[15] L. González-Flores, J. Rubiano-Moreno, and G. Sosa-Gómez, "The relevance of lead prioritization: a B2B lead scoring model based on machine learning," Front. Artif. Intell., vol. 8, Mar. 2025, doi: 10.3389/frai.2025.1554325.
[16] J. Lee, O. Jung, Y. Lee, O. Kim, and C. Park, "A Comparison and Interpretation of Machine Learning Algorithm for the Prediction of Online Purchase Conversion," J. Theor. Appl. Electron. Commer. Res., vol. 16, no. 5, Art. no. 5, Aug. 2021, doi: 10.3390/jtaer16050083.
[17] S. Sarafopoulos, "Comparison of multiple prediction models on customer conversion of financial services customers," International Journal of Computer Science and Network Security, vol. 20, no. 2, pp. 1-8, 2020.