# **Title**

"Predictive Data Analytics for Banking Marketing Success using Term Deposit Subscription Classification"

# **Topic Area**

Predictive Data Analysis, Banking, Telemarketing, Finance

# **Research Objectives**

* To develop and compare predictive models (Random Forest, XGBoost, CatBoost, CNN, and TabNet) for classifying term deposit subscriptions using direct marketing campaign data.
* To evaluate the performance of these models based on accuracy, precision, recall, F1-score, and computational efficiency.
* To analyse the most important features on model prediction.
* To determine which model offers the best balance between interpretability and predictive power.

# **Statement of Hypotheses**

* **H1**: Deep learning-based models (CNN, TabNet) will outperform traditional ensemble methods (Random Forest, XGBoost, CatBoost) in predicting term deposit subscriptions, particularly in handling complex feature interactions.
* **H2**: Ensemble models (XGBoost, CatBoost) will provide higher interpretability and generalisability compared to deep learning models, making them more suitable for business use cases where transparency is important.

# **Literature Review**

## **Predictive Analytics in Banking**

Broby (2022) takes a closer look at the use of these techniques in the financial industry where classification, regression, clustering and time series are discussed. The study incorporates these techniques into Decision Support System (DSS), which assists the financial managers in their decisions making. More than 187 papers are inspected according to the proposed SPAR-4-SLR protocol; and despite the fact that the range includes many algorithms, it is dominated by Random Forests, Support Vector Machines (SVM), and Neural Networks. For instance, Huang et al. (2005) employed SVM to anticipate the overall direction of stock at 75%accuracy; Broby (2022). These methods are used in the determination of the financial future position and in the framework of risk management plans.

In the present study, Moinuddin et al. (2024) analyse the effect of the marketing analytics on the consumers and the campaign effectiveness using both discussed interviews and surveys. This research also finds out that checking of analytics tools is done weekly by 75% of the respondents, the mean score the respondents gave to the increase in the success of their campaigns as a result of this activity was 8.5 out of 10. Correlation analysis gave a value of 0.75 on the Pearson coefficient and this suggested that analytics usage had directly dependent on the campaign evaluation (Moinuddin et al., 2024/). This is particularly important in current society, yet to show why there is a need to shift towards data driven marketing.

Zaki et al, (2024) analyse predictive analytics and machine learning in term deposit subscriptions of the bank using models like SGD, KNN, and Random Forest. Feature engineering, cross tabulation, and heatmaps were used during data exploration conducted in the study. Random Forest model performed better than other models where accuracy was found to be 87 %, PPV = 87.83% and NPV = 92.99%. These findings offer relevant tips that can be used to improve advertising tactics in banking (Zaki et al., 2024).