Please implement the **Decision Tree**, **Random Forest**, **Bagging,** and **Boosting** classification approaches on the *Customer Churn* (*Telecom Churn)* dataset in *R Studio*. You should try the best setting for hyperparameters to fetch the best result. For detailed information about the implementation, please refer to the lectures up to the second session, ended September 11, 2024.

This dataset is a binary-labeled dataset with two classes: 0 for customers who have not churned and 1 for those who have churned. You can download the dataset from the following link:

https://www.kaggle.com/datasets/barun2104/telecom-churn?resource=download

The deliverable for each group is a PDF file as the report with members' name. Your report should include the following tasks, along with screenshots of the codes and results and the explanation for each model respectively.

- ✓ **Partitioning**: Splitting the dataset into training and test sets.
- ✓ **Modeling**: Building the models requested above using the training set.
- ✓ **Prediction**: Making predictions using the test set.
- ✓ **Evaluation**:
    - Calculating *Accuracy*, *F1 Score*, and *Classification Error* (*CE*).
    - Reporting the results using the *Confusion matrix*.
- ✓ **Feature Importance Calculation**: Calculation of variable importance via *Decision Tree*, *Random Forest* and *XGBoost* models and plotting the results for the second and the third approach.

	Excellent	**Good**	**Satisfactory**	**Needs Improvement**
Partitioning	• Splitting the dataset into train and test sets • Paying attention to the partitioning type • Paying attention to the partitioning percentage.	• The partitioning type is not properly selected. • The partitioning percentage is not properly selected.	• The explanation for the selection of partitioning type and percentage shows a moderate understanding of the concepts.	• No partitioning (the model is built based on the whole dataset) • Not enough or readable screenshots and explanations.
(1 Points)	(1 Point)	(0.75 Point)	(0.5 Point)	(0 – 0.25 Point)
Modeling	Making classification models: • *Decision Tree* • *Random Forest* • *Bagging* • *Boosting* • Proper settings of hyperparameters and providing the reasons and explanations	Some of the hyperparameter setting needs modification and they are not set properly.	• All requested models have not been represented. • The hyperparameter setting is inaccurate. • The provided explanation for the settings of hyperparameters shows a moderate understanding of the concepts.	Not enough or readable screenshots and explanations.
(1 Points)	(1 Point)	(0.75 Point)	(0.5 Point)	(0 – 0.25 Point)
Prediction	• Prediction of target values for all of the models above. • Proper selection of columns to be compared.	Prediction is done either based on the whole dataset or the trainset.	• The selection of columns for prediction phase is not correct. • The prediction phase is done for some models not for all of the constructed models.	Not enough or readable screenshots and explanations.
(1 Points)	(1 Point)	(0.75 Point)	(0.5 Point)	(0 – 0.25 Point)
Evaluation	Comparing all of the models constructed above based on • The *Confusion matrix* • Reporting the *Accuracy*, *CE* and *F1*.	• No usage of *Metrics* package. • No comparison among models.	• All of the requested metrics have not been represented. • All of the constructed models have not been evaluated.	Not enough or readable screenshots and explanations.
(1 Points)	(1 Point)	(0.75 Point)	(0.5 Point)	(0 – 0.25 Point)
Feature Importance Calculation	Calculation of feature importance and plotting the results based on: • *Decision Tree* • *Random Forest* • *Boosting*	The plotting needs improvement.	• Plotting of the results is not provided. • The importance calculation has been done on some of the models not all the requested models.	Not enough or readable screenshots and explanations.
(1 Points)	(1 Point)	(0.75 Point)	(0.5 Point)	(0 – 0.25 Point)