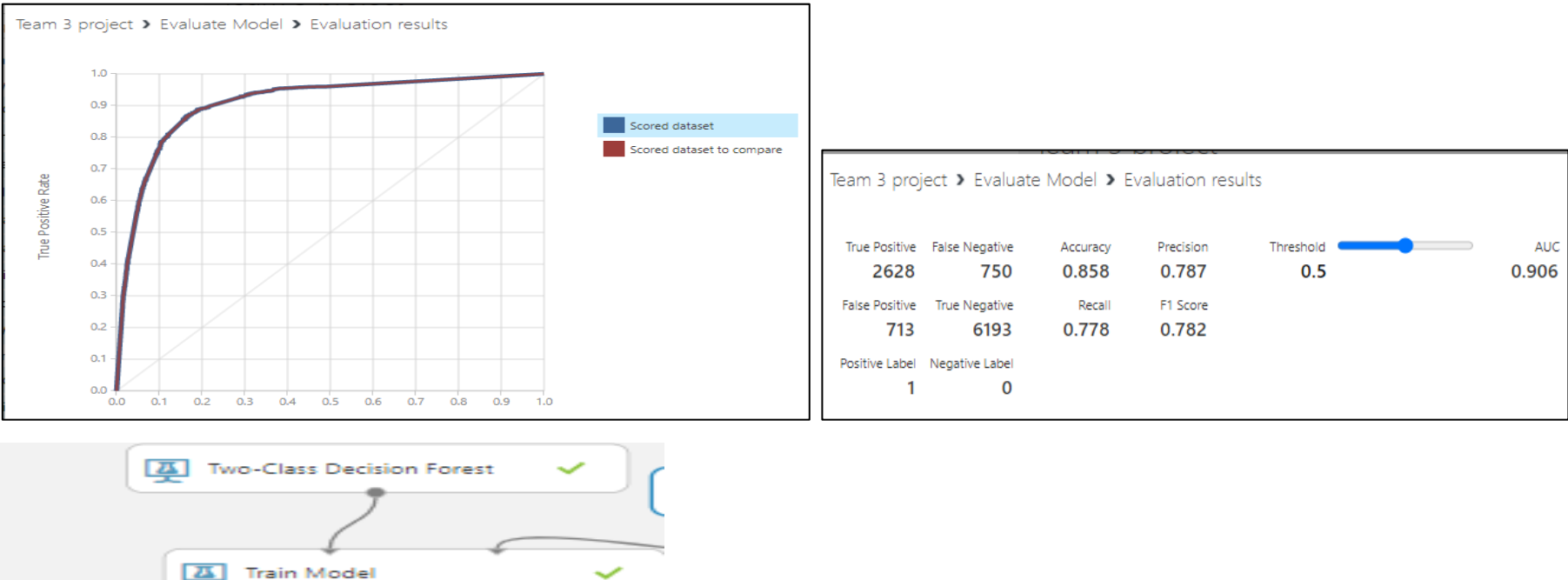


Model Comparison :

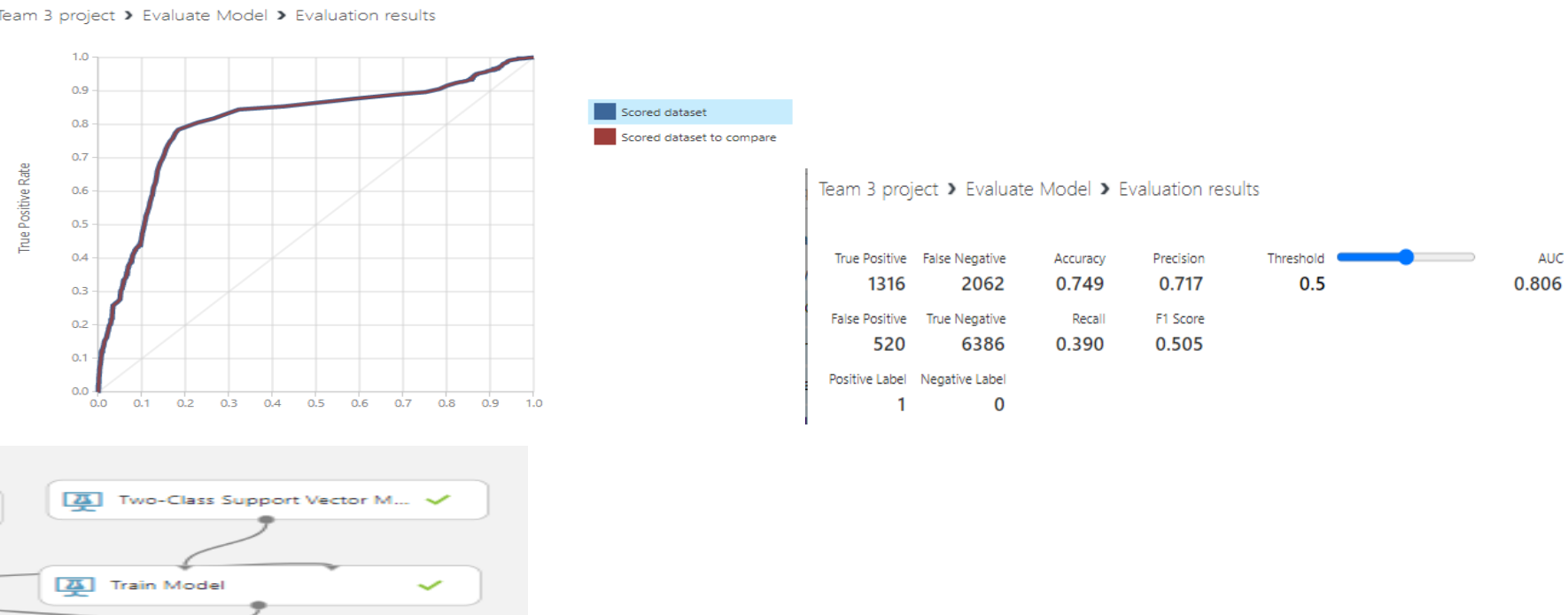
1) Two class Decision Forests model :

- Decision forests are fast, supervised ensemble models.
- This module is a good choice if you want to predict a target with a maximum of two outcomes.
- Ensemble methods are based on the general principle that rather than relying on a single model, you can get better results and a more generalized model by creating multiple related models and combining them in some way.
- After using this algorithm to create model.The preform of model is pretty good.
- As we see most of the area is under the curve i.e AUC more than 90 %.
- After doing hyperparameter tuning,we achieve **85% accuracy**



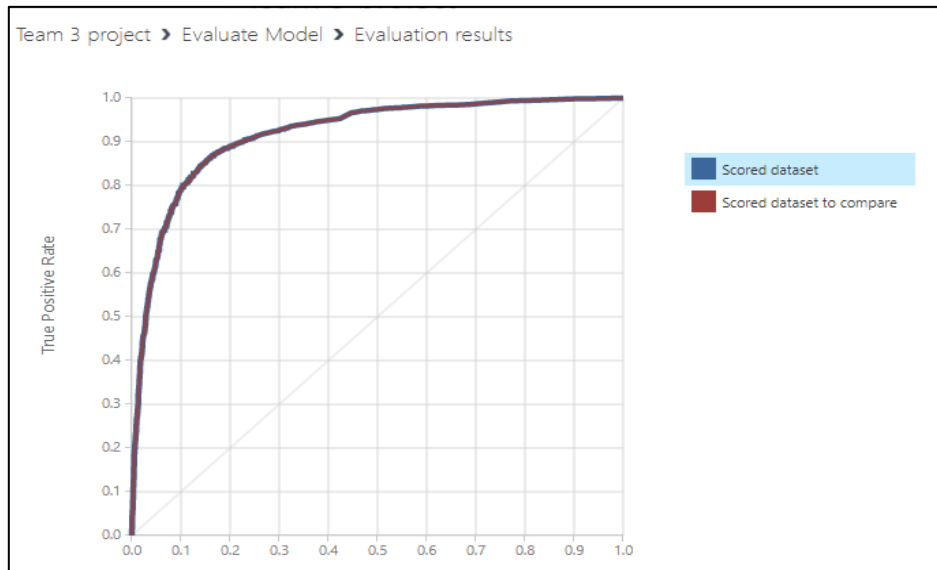
2) Two class Support Vector Machine model :

- Support vector machines are among the earliest of machine learning algorithms, and SVM models have been used in many applications, from information retrieval to text and image classification.
- SVMs can be used for both classification and regression tasks.
- It requires labeled data because it is supervised learning model.
- In the training process, the algorithm analyzes input data and recognizes patterns in a multi-dimensional feature space called the hyperplane.
- As we see, after implement of this model is good but not compare to Two class Decision Forests.
- ROC –AUC is less it is 80% And also Accuracy 74% which less than other model.
- We can improve it with using hyper meter tuning.



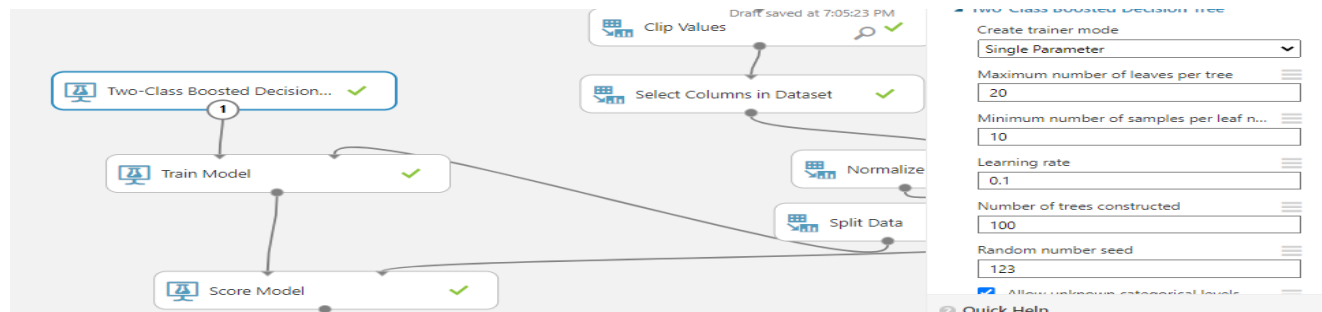
3) Two-Class Boosted Decision Tree :

- A boosted decision tree is an ensemble learning method in which the second tree corrects for the errors of the first tree, the third tree corrects for the errors of the first and second trees, and so forth.
- Predictions are based on the entire ensemble of trees together that makes the prediction.
- Boosted decision trees are the easiest methods with which to get top performance on a wide variety of machine learning tasks.
- After implement, we get excellent perform with 86% accuracy & AUC with 91% with help of hyperparameter tuning.
- This model can explain 86% of data from dataset.



Team 3 project > Evaluate Model > Evaluation results

True Positive	False Negative	Accuracy	Precision	Threshold	AUC
2626	752	0.863	0.801	0.5	0.919
False Positive	True Negative	Recall	F1 Score		
652	6254	0.777	0.789		
Positive Label	Negative Label				
1	0				



4) Conclusion :

- After see all model perform we can decide to use Two class Boosted Decision Forests.
- Cause it give best accuracy compare to other model.
- We can improve it with performing more hyperparameter tuning & learning rate.