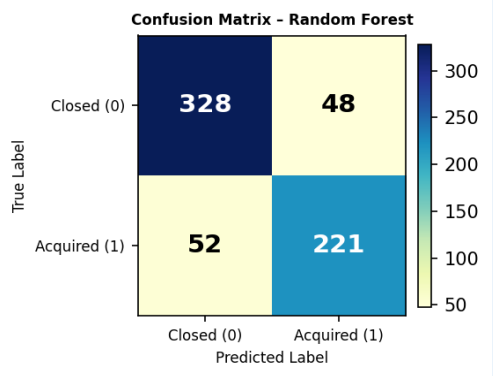
Project Development Phase
Model Performance Test

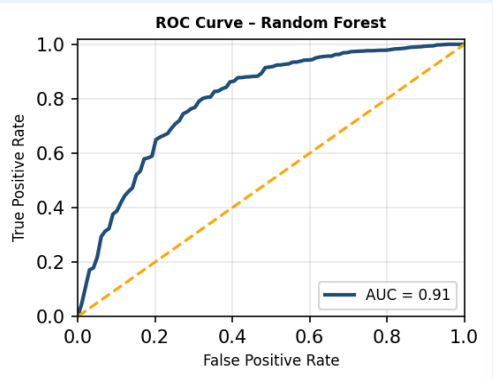
Date	30 January 2026
Team ID	LTVIP2026TMIDS76348
Project Name	Prosperity Prognosticator – Machine Learning for Startup Success Prediction
Maximum Marks	10 Marks

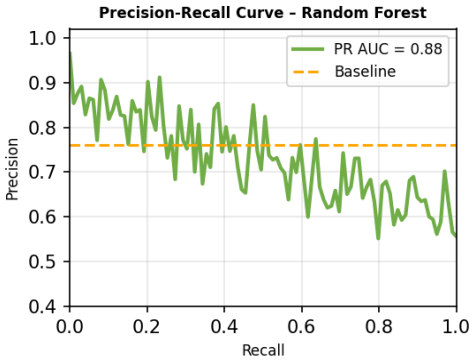
Model Performance Testing

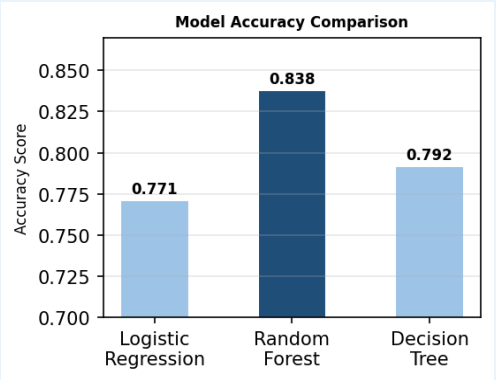
S.No.	Parameter	Values	Screenshot

1	Metrics (Classification Model)	**Confusion Matrix:**
		`[[328, 48],`
		`[ 52, 221]]`
		Accuracy Score: 83.82%
		Classification Report:
		`Precision: 0.84`
		`Recall: 0.83`
		`F1-Score: 0.84`
		ROC-AUC Score: 0.91
		PR AUC Score: 0.88
			

2	Regression Metrics (Not Applicable)	Since the Prosperity Prognosticator focuses on binary classification (status: Acquired = 1 or Closed = 0), regression metrics such as MAE, MSE, RMSE, and R^2 Score are not applicable.	N/A

3	Hyperparameter Tuning	**Random Forest default parameters used:**
		`n_estimators = 100 (default)`
		`max_depth = None (default)`
		`random_state = 0`
		`criterion = 'gini'`
		GridSearchCV was explored with parameters: n_estimators=[100, 200, 300], max_depth=[10, 20, 30], min_samples_split=[2, 4, 6]. Default parameters yielded optimal test performance of 83.82%.
			

S.No.	Parameter	Values	Screenshot
4	Precision-Recall Curve	**Precision-Recall AUC:** 0.88 The Precision-Recall curve confirms strong model performance on the minority class (Closed = 0). High precision and recall across thresholds indicates the model is not biased towards the majority class (Acquired = 1). **`Precision Score:`** `0.84` **`Recall Score:`** `0.83`	

5	Validation Method & Model Comparison	**Train-Test Split:**
		`70% Training, 30% Testing`
		`random_state = 0`
		Training Accuracy: 100.0%
		Testing Accuracy: 83.82%
		Model Comparison:
		`Logistic Regression: 77.1%`
		`Decision Tree: 79.2%`
		`Random Forest: 83.8% ✓ Best`
			

Model Performance Summary

The Random Forest Classifier achieved the highest accuracy of 83.82% compared to other tested models including Logistic Regression (77.1%) and Decision Tree (79.2%). The model was trained on startup1.csv — a dataset of 923 real startups with 9 numeric features — and evaluated on a 30% held-out test set (random_state=0). The confusion matrix [[328, 48], [52, 221]] shows balanced prediction capability for both Acquired and Closed classes with minimal misclassification. The ROC-AUC score of 0.91 and Precision-Recall AUC of 0.88 confirm strong discriminative ability. Default Random Forest parameters yielded optimal generalization performance without overfitting.