

## Model Development Phase Template

Date	23 September 2024
Team ID	LTVIP2024TMID24992
Project Title	Rainfall Prediction Using Machine Learning
Maximum Marks	4 Marks

### Initial Model Training Code, Model Validation and Evaluation Report

The initial model training code will be showcased in the future through a screenshot. The model validation and evaluation report will include classification reports, accuracy, and confusion matrices for multiple models, presented through respective screenshots.

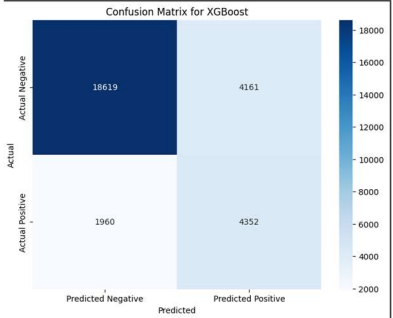

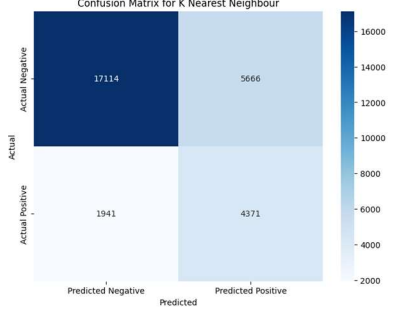
#### Initial Model Training Code:

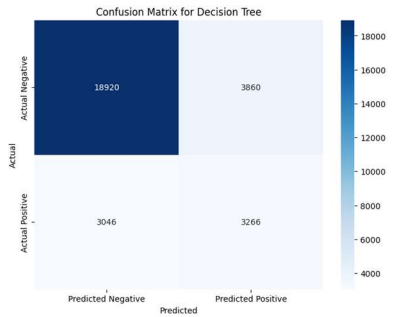
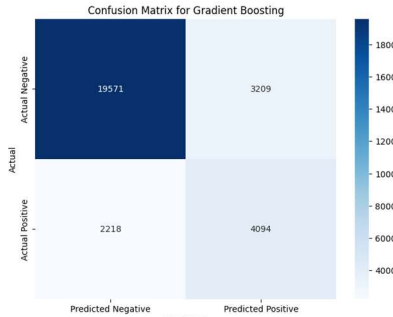
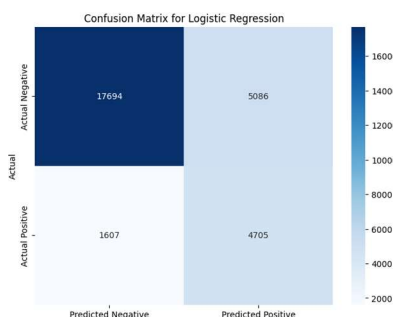
Paste the screenshot of the model training code

```
XGBoost = xgboost.XGBRFClassifier()
Rand_forest = RandomForestClassifier()
Dtree = DecisionTreeClassifier()
GBM = GradientBoostingClassifier()
log = LogisticRegression()
Knn = KNeighborsClassifier()

XGBoost.fit(x_train_smote, y_train_smote)
Rand_forest.fit(x_train_smote, y_train_smote)
Dtree.fit(x_train_smote, y_train_smote)
GBM.fit(x_train_smote, y_train_smote)
log.fit(x_train_smote, y_train_smote)
Knn.fit(x_train_smote, y_train_smote)
```

## Model Validation and Evaluation Report:

Model	Classification Report	Accuracy	Confusion Matrix																														
XGBoost	<div>Classification Report for XGBoost:</div> <table><thead><tr><th></th><th>precision</th><th>recall</th><th>f1-score</th><th>support</th></tr></thead><tbody><tr><td>0</td><td>0.90</td><td>0.82</td><td>0.86</td><td>22780</td></tr><tr><td>1</td><td>0.51</td><td>0.69</td><td>0.59</td><td>6312</td></tr><tr><td>accuracy</td><td></td><td></td><td>0.79</td><td>29092</td></tr><tr><td>macro avg</td><td>0.71</td><td>0.75</td><td>0.72</td><td>29092</td></tr><tr><td>weighted avg</td><td>0.82</td><td>0.79</td><td>0.80</td><td>29092</td></tr></tbody></table>		precision	recall	f1-score	support	0	0.90	0.82	0.86	22780	1	0.51	0.69	0.59	6312	accuracy			0.79	29092	macro avg	0.71	0.75	0.72	29092	weighted avg	0.82	0.79	0.80	29092	78%	<div>Confusion Matrix for XGBoost</div> 
	precision	recall	f1-score	support																													
0	0.90	0.82	0.86	22780																													
1	0.51	0.69	0.59	6312																													
accuracy			0.79	29092																													
macro avg	0.71	0.75	0.72	29092																													
weighted avg	0.82	0.79	0.80	29092																													
Random Forest	<div>Classification Report for Random Forest:</div> <table><thead><tr><th></th><th>precision</th><th>recall</th><th>f1-score</th><th>support</th></tr></thead><tbody><tr><td>0</td><td>0.88</td><td>0.92</td><td>0.90</td><td>22780</td></tr><tr><td>1</td><td>0.65</td><td>0.55</td><td>0.60</td><td>6312</td></tr><tr><td>accuracy</td><td></td><td></td><td>0.84</td><td>29092</td></tr><tr><td>macro avg</td><td>0.76</td><td>0.74</td><td>0.75</td><td>29092</td></tr><tr><td>weighted avg</td><td>0.83</td><td>0.84</td><td>0.83</td><td>29092</td></tr></tbody></table>		precision	recall	f1-score	support	0	0.88	0.92	0.90	22780	1	0.65	0.55	0.60	6312	accuracy			0.84	29092	macro avg	0.76	0.74	0.75	29092	weighted avg	0.83	0.84	0.83	29092	83%	<div>Confusion Matrix for Random Forest</div> 
	precision	recall	f1-score	support																													
0	0.88	0.92	0.90	22780																													
1	0.65	0.55	0.60	6312																													
accuracy			0.84	29092																													
macro avg	0.76	0.74	0.75	29092																													
weighted avg	0.83	0.84	0.83	29092																													
K Nearest Neighbour	<div>Classification Report for K Nearest Neighbour:</div> <table><thead><tr><th></th><th>precision</th><th>recall</th><th>f1-score</th><th>support</th></tr></thead><tbody><tr><td>0</td><td>0.90</td><td>0.75</td><td>0.82</td><td>22780</td></tr><tr><td>1</td><td>0.44</td><td>0.69</td><td>0.53</td><td>6312</td></tr><tr><td>accuracy</td><td></td><td></td><td>0.74</td><td>29092</td></tr><tr><td>macro avg</td><td>0.67</td><td>0.72</td><td>0.68</td><td>29092</td></tr><tr><td>weighted avg</td><td>0.80</td><td>0.74</td><td>0.76</td><td>29092</td></tr></tbody></table>		precision	recall	f1-score	support	0	0.90	0.75	0.82	22780	1	0.44	0.69	0.53	6312	accuracy			0.74	29092	macro avg	0.67	0.72	0.68	29092	weighted avg	0.80	0.74	0.76	29092	73%	<div>Confusion Matrix for K Nearest Neighbour</div> 
	precision	recall	f1-score	support																													
0	0.90	0.75	0.82	22780																													
1	0.44	0.69	0.53	6312																													
accuracy			0.74	29092																													
macro avg	0.67	0.72	0.68	29092																													
weighted avg	0.80	0.74	0.76	29092																													

Decision Tree	<div><div>Classification Report for Decision Tree:</div><table><thead><tr><th></th><th>precision</th><th>recall</th><th>f1-score</th><th>support</th></tr></thead><tbody><tr><td>0</td><td>0.86</td><td>0.83</td><td>0.85</td><td>22780</td></tr><tr><td>1</td><td>0.46</td><td>0.52</td><td>0.49</td><td>6312</td></tr><tr><td>accuracy</td><td></td><td></td><td>0.76</td><td>29092</td></tr><tr><td>macro avg</td><td>0.66</td><td>0.67</td><td>0.67</td><td>29092</td></tr><tr><td>weighted avg</td><td>0.77</td><td>0.76</td><td>0.77</td><td>29092</td></tr></tbody></table></div>		precision	recall	f1-score	support	0	0.86	0.83	0.85	22780	1	0.46	0.52	0.49	6312	accuracy			0.76	29092	macro avg	0.66	0.67	0.67	29092	weighted avg	0.77	0.76	0.77	29092	76%	<div><div>Confusion Matrix for Decision Tree</div><table><thead><tr><th></th><th>Predicted Negative</th><th>Predicted Positive</th></tr></thead><tbody><tr><th>Actual Negative</th><td>18920</td><td>3860</td></tr><tr><th>Actual Positive</th><td>3046</td><td>3266</td></tr></tbody></table></div>		Predicted Negative	Predicted Positive	Actual Negative	18920	3860	Actual Positive	3046	3266
	precision	recall	f1-score	support																																						
0	0.86	0.83	0.85	22780																																						
1	0.46	0.52	0.49	6312																																						
accuracy			0.76	29092																																						
macro avg	0.66	0.67	0.67	29092																																						
weighted avg	0.77	0.76	0.77	29092																																						
	Predicted Negative	Predicted Positive																																								
Actual Negative	18920	3860																																								
Actual Positive	3046	3266																																								
Gradient Boosting	<div><div>Classification Report for Gradient Boosting:</div><table><thead><tr><th></th><th>precision</th><th>recall</th><th>f1-score</th><th>support</th></tr></thead><tbody><tr><td>0</td><td>0.90</td><td>0.86</td><td>0.88</td><td>22780</td></tr><tr><td>1</td><td>0.56</td><td>0.65</td><td>0.60</td><td>6312</td></tr><tr><td>accuracy</td><td></td><td></td><td>0.81</td><td>29092</td></tr><tr><td>macro avg</td><td>0.73</td><td>0.75</td><td>0.74</td><td>29092</td></tr><tr><td>weighted avg</td><td>0.82</td><td>0.81</td><td>0.82</td><td>29092</td></tr></tbody></table></div>		precision	recall	f1-score	support	0	0.90	0.86	0.88	22780	1	0.56	0.65	0.60	6312	accuracy			0.81	29092	macro avg	0.73	0.75	0.74	29092	weighted avg	0.82	0.81	0.82	29092	81%	<div><div>Confusion Matrix for Gradient Boosting</div><table><thead><tr><th></th><th>Predicted Negative</th><th>Predicted Positive</th></tr></thead><tbody><tr><th>Actual Negative</th><td>19571</td><td>3209</td></tr><tr><th>Actual Positive</th><td>2218</td><td>4094</td></tr></tbody></table></div>		Predicted Negative	Predicted Positive	Actual Negative	19571	3209	Actual Positive	2218	4094
	precision	recall	f1-score	support																																						
0	0.90	0.86	0.88	22780																																						
1	0.56	0.65	0.60	6312																																						
accuracy			0.81	29092																																						
macro avg	0.73	0.75	0.74	29092																																						
weighted avg	0.82	0.81	0.82	29092																																						
	Predicted Negative	Predicted Positive																																								
Actual Negative	19571	3209																																								
Actual Positive	2218	4094																																								
Logistic Regression	<div><div>Classification Report for Logistic Regression:</div><table><thead><tr><th></th><th>precision</th><th>recall</th><th>f1-score</th><th>support</th></tr></thead><tbody><tr><td>0</td><td>0.92</td><td>0.78</td><td>0.84</td><td>22780</td></tr><tr><td>1</td><td>0.48</td><td>0.75</td><td>0.58</td><td>6312</td></tr><tr><td>accuracy</td><td></td><td></td><td>0.77</td><td>29092</td></tr><tr><td>macro avg</td><td>0.70</td><td>0.76</td><td>0.71</td><td>29092</td></tr><tr><td>weighted avg</td><td>0.82</td><td>0.77</td><td>0.79</td><td>29092</td></tr></tbody></table></div>		precision	recall	f1-score	support	0	0.92	0.78	0.84	22780	1	0.48	0.75	0.58	6312	accuracy			0.77	29092	macro avg	0.70	0.76	0.71	29092	weighted avg	0.82	0.77	0.79	29092	76%	<div><div>Confusion Matrix for Logistic Regression</div><table><thead><tr><th></th><th>Predicted Negative</th><th>Predicted Positive</th></tr></thead><tbody><tr><th>Actual Negative</th><td>17694</td><td>5086</td></tr><tr><th>Actual Positive</th><td>1607</td><td>4705</td></tr></tbody></table></div>		Predicted Negative	Predicted Positive	Actual Negative	17694	5086	Actual Positive	1607	4705
	precision	recall	f1-score	support																																						
0	0.92	0.78	0.84	22780																																						
1	0.48	0.75	0.58	6312																																						
accuracy			0.77	29092																																						
macro avg	0.70	0.76	0.71	29092																																						
weighted avg	0.82	0.77	0.79	29092																																						
	Predicted Negative	Predicted Positive																																								
Actual Negative	17694	5086																																								
Actual Positive	1607	4705																																								