Part D & E

Part D

I have hereby compared expected/correct results with three ensemble models. And also, as a testing method, I have used cross validation.

Comparison expected/correct results

In the dataset, there are 768 records. For testing I have splitted 20% of them. And for this comparison I have taken records.

Voting

	Actual Result	Predicted Result
653	0.0	0.0
331	0.0	0.0
568	0.0	1.0
196	0.0	0.0
76	0.0	0.0
64	1.0	1.0
671	0.0	0.0
52	0.0	0.0
310	0.0	0.0
416	0.0	0.0
476	1.0	0.0
482	0.0	0.0
230	1.0	0.0
527	0.0	0.0
380	0.0	0.0

Bagging

	Actual Result	Predicted Result
653	0.0	0.0
331	0.0	0.0
568	0.0	1.0
196	0.0	0.0
76	0.0	0.0
64	1.0	1.0
671	0.0	0.0
52	0.0	0.0
310	0.0	0.0
416	0.0	0.0

476	1.0	0.0
482	0.0	0.0
230	1.0	0.0
527	0.0	0.0
380	0.0	0.0

Boosting

	Actual Result	**Predicted Result**
653	0.0	0.0
331	0.0	0.0
568	0.0	1.0
196	0.0	0.0
76	0.0	0.0
64	1.0	1.0
671	0.0	0.0
52	0.0	0.0
310	0.0	0.0
416	0.0	0.0
476	1.0	0.0
482	0.0	0.0
230	1.0	0.0
527	0.0	0.0
380	0.0	0.0

Cross Validation

I have used cross validation as a testing method for this project. With this approach, the data is divided into several subsets (folds), the model is trained on various subsets, and its performance is assessed on the remaining subsets. Using this method, can get a good idea of the model's performance on unseen data.

	Ensemble Cross-Validated AUC	**Ensemble Cross-Validated Accuracy**
Voting	0.83	0.7698
Bagging	0.74	0.7812
Boosting	0.83	0.7683

Part E

After running the three aforementioned ensemble models, the following results are gained. Accuracy, precision, recall, F1-score, AUC are evaluation metrics of different machine learning algorithms. In this project, accuracy, precision, recall, F1-score, AUC of the different algorithms slightly differ from each other as their working mechanism is different from each other. Below table represents the accuracy, precision, recall, F1-score, AUC of predicting diabetes using ensemble learning models.

	Voting	**Bagging**	**Boosting**
Ensemble Accuracy	0.8247	0.8377	0.8377
Precision	0.7083	0.7619	0.7619
Recall	0.7234	0.6809	0.6809
F1-score	0.7158	0.7191	0.7191
AUC	0.85	0.79	0.86

Although an accuracy of higher than 80% has been acquired for each of the models, the dataset is still limited and requires further processing and larger population to draw a much stronger conclusion.