

Part B –Supervised Learning

The goal of implementing the supervised learning algorithms (decision tree, gradient boosting and random forest algorithms) was to predict the rate of change of covid cases with respect to the temperature, the date and the proportion of fully vaccinated people for a given location.

The table below summarizes the results of the 3 algorithms.

Algorithm	MSE	RMSE	MAE	R2
Decision Tree Regressor	2599852	1619	1260	0.927
Gradient Boosting Regressor	3794678	1947	1416	0.909
Random Forest Regressor	657452	810	458	0.984

The table below shows the time taken to construct the models.

Algorithm	Time (in mins)
Decision Tree Regressor	15
Gradient Boosting Regressor	15
Random Forest Regressor	15

GitHub repository link – <https://github.com/bsewp045/CSI4142-Phase-4>

Actionable insights:

Since all of the 3 algorithms have a high R-squared score, it can be deduced that the temperature, the date and the proportion of fully vaccinated people are good predictors for predicting the rate of change of covid cases for a given location.