Mohd Talha Patrawala
CMPN-B
23102B0025

ML : Assignment - 6

Q.2.

① The almost useless models become strong in AdaBoost because:
(a) Each decision stump is slightly better than random.
(b) AdaBoost trains them sequentially, increasing weight on misclassified samples.
(c) Final prediction is the weighted vote of all weak learners.
(d) As long as each learner performs slightly better than random, the combined model becomes a strong classifier.

② This can fail when:
(a) Algorithm keeps focusing on noise leading to overfitting.
(b) If weak learners are not better than random, boosting guarantee breaks.
(c) Sensitive to outliers due to exponential loss.

Q.2.

1] ① Regularization:
ⓐ Controls tree complexity by penalizing large leaf weights.
ⓑ Prevents overfitting and improves generalization.

② Second-order Gradient Optimization:
ⓐ Uses both first-order (gradient) and second-order (Hessian) information.
ⓑ makes more accurate split decisions.
ⓒ Leads to faster convergence and more stable learning.

2] Trade-offs:
① It will have more hyperparameters which requires proper tuning.
② Higher computational cost and complexity, ~~compared to Scikit~~
③ Less interpretable due to deeper trees and many boosting rounds.