Rank	Model	Train Score (%)	Test Score (%)	Accuracy (%) (Running code 3 times)	Model Type	Strength	Weakness	Remarks
High	XGBoost	100 %	97.9 %	1.97.9 % 2.97.9 % 3.97.9 %	Ensemble (Boosting)	Very high accuracy, can capture complex patterns	Train 100% , Overfitting may occur	Very good model, but overfitting needs to be checked.
Low	SVR	7.00%	8.20%	1.8.2 % 2.8.2 % 3.8.2 %	Supervised ML (Regression)	Works well on small datasets, Handles non-linear data using kernel, Less sensitive to outliers	Poor performance on large datasets, Requires hyperparameter tuning, Needs feature scaling	Model shows underfitting. Both train and test scores are very low, meaning it fails to capture data patterns effectively.
High	Ridge Regression	95.52%	94.34%	1.94.33 % 2.94.33 % 3.94.33 %	Supervised ML (Regression, Regularized Linear Model)	Prevents overfitting using L2 regularization, Stable performance, Works well with multicollinearity	Cannot capture complex non-linear patterns	Excellent performance. High accuracy with very small train-test gap. Model is well-fitted and generalizes well.
2nd	Random Forest	96.84%	88.47%	1.88.46 % 2.88.46 % 3.88.46 %	Supervised ML (Ensemble Regression)	Handles non-linear data well, High accuracy, Robust to outliers	Can overfit, Less interpretable, Computationally heavier	High train score but noticeable drop in test score (~8%). Slight overfitting observed.
3rd	MLP (Neural Network)	83.49%	86.28%	1.86.28 % 2.86.28 % 3.86.28 %	Supervised ML (Artificial Neural Network)	Captures complex non-linear patterns, Flexible model	Needs large data, Sensitive to hyperparameters, More Training time	Good generalization. Test score slightly higher than train score. No overfitting observed.
High	Linear Regression	100%	100% 100%	1.100 % 2.100 % 3.100 %	Supervised ML (Linear Model)	Simple, Fast, Highly interpretable	Cannot handle non-linear patterns	Perfect accuracy. Probably data leakage or perfectly linear relationship. Further validation recommended.
3rd	LightGBM	98%	91%	1.91 % 2.91 % 3.91 %	Gradient Boosting (Ensemble ML)	Very good predictive performance through gradient boosting	Overfitting Risk	High accuracy with slight overfitting (7% gap). Strong predictive performance.
2nd	Lasso Regression	96.25%	95.85%	1.95.85 % 2.95.85 % 3.95.85 %	Linear Model with L1 Regularization	Prevents overfitting, Performs feature selection automatically, High accuracy, Simple & interpretable	Cannot capture complex non-linear patterns, Sensitive to regularization parameter (alpha)	Excellent performance with minimal train-test gap. Well-generalized.

7th	K-Nearest Neighbors	58.27%	52.80%	1.52.80 % 2.52.80 % 3.52.80 %	Supervised ML (Instance-based Regression)	Simple, Easy to understand, No training phase needed	Low accuracy on large/complex datasets, Sensitive to feature scaling, Cannot extrapolate beyond training data, Slow prediction for large datasets	Poor performance. Underfitting observed, model fails to capture underlying patterns effectively.
4th	Gradient Boosting	99.30%	93.45%	1.93.45 % 2.93.45 % 3.93.45 %	Ensemble ML (Boosting)	High accuracy, Handles non-linear patterns well, Robust predictive power	Slight overfitting, Sensitive to hyperparameters, Computationally heavier	Strong performance with slight overfitting. Can be improved with learning rate tuning and early stopping.
3rd	Extra Trees	100%	95.37%	1.95.37 % 2.95.37 % 3.95.37 %	Ensemble ML (Randomized Decision Trees)	Handles non-linear patterns well, Robust to outliers, Fast training	Slight overfitting, Less interpretable, Sensitive to number of trees	Strong performance, very accurate, slight overfitting observed. Can tune n_estimators or max_depth to improve generalization.
5th	Decision Tree	93.07%	89.28%	1.89.28 % 2.89.28 % 3.89.28 %	Supervised ML (Tree-based Regression)	Easy to interpret, Handles non-linear patterns, Fast to train	Prone to overfitting, Sensitive to small data changes, Poor generalization if deep	Good performance with slight overfitting. Can improve with pruning or max_depth tuning.
6th	CatBoost	98.54%	88.26%	1.88.26 % 2.88.26 % 3.88.26 %	Gradient Boosting (Ensemble ML)	Handles categorical & numerical features well, High predictive accuracy, Reduces overfitting compared to other boosting methods	Slight overfitting observed, Sensitive to hyperparameters, Training time higher	Strong predictive power but slight overfitting. Can tune learning_rate, depth, or iterations for better generalization.
8th	AdaBoost	83.64%	70.03%	1.70.02 % 2.70.02 % 3.70.02 %	Ensemble ML (Boosting)	Can improve weak learners, Handles non-linear patterns moderately, Simple to implement	Overfitting risk, Sensitive to noisy data, Moderate accuracy	Moderate performance. Model overfits slightly and test accuracy is relatively low compared to other ensemble models.