

## 1 CatBoost

```
Best hyperparameters: {'depth': 7, 'iterations': 1000,
'learning_rate': 0.05}
Expected RMSE error: 4.081210766971801
MAE: 2.5745905532392674
MSE: 16.656281324446557
R2: 0.935359818094655
95% confidence interval:
\left(\left[51.8243495 , 41.93664799, \dots, 42.298315...\right]\right)
```

## 2 Support Vector Machines

```
Best hyperparameters: {'C': 1000, 'coef0': 0, 'degree': 3,
'gamma': 0.1}
0.8907982261640797
SVC(C=1000, coef0=0, gamma=0.1, random_state=42)
[[96 6]
 [14 90]]
Expected RMSE error: 0.3115884764248779
MAE: 0.0970873786407767
MSE: 0.0970873786407767
R2: 0.6116138763197587
95% confidence interval: (0.397897447289056, 0.5341413876624003)
```

## 3 Random Forests

```
Best hyperparameters: {'max_depth': 10, 'min_samples_leaf': 1,
'min_samples_split': 2, 'n_estimators': 500}
-26.80270840004423
RandomForestRegressor(max_depth=10, n_estimators=500, random_state=42)
MAE: 3.736334329056867
MSE: 29.854417119395947
RMSE: 5.463919574755465
R2: 0.8841401081258502
Start: 6.579025407891052
End: 62.99933535147096
```

## 4 AdaBoost

```
Best hyperparameters: {'n_estimators': 200, 'loss': 'square',
'learning_rate': 1}
-55.76729644915056
AdaBoostRegressor(learning_rate=1, loss='square',
n_estimators=200,
random_state=42)
Mean Absolute Error: 6.270046832695762
Mean Squared Error: 57.745066862245814
Root Mean Squared Error: 7.599017493218832
R2 Score: 0.7759012619081165
95% confidence interval:
\left(\left[39.08921757, 37.82768758, \dots, 40.37628906...\right]\right)
```

## 5 Conclusion

The best model that was tested was the CatBoost model. The model was trained on 80 percent of the data and tested on the remaining 20 percent. The reason why the model performed the best was it produced, the Expected RMSE error: 4.081210766971801, MAE: 2.5745905532392674, MSE: 16.656281324446557, R2: 0.935359818094655. Since the model produced an RMSE that was low, it means that the predictions that the model produced were extremely close to the actual values. Furthermore, The R2 value was rather high, this means that the model was able to explain variance that occurred in the data. Finally when looking at the 95 percent confidence intervals. They are rather narrow which means that the model was able to predict the values with a high degree of accuracy.

Check out the mybest.py to see the speed of the Kfolds as well as permutation importance. Which will help your business focus on attributes that are important your business and the strength of concrete. This will allow you hone your formula to deliver the best product to your customers.