

Evaluation and Analysis

Team: Anar Otgonbaatar, Adam Kaci

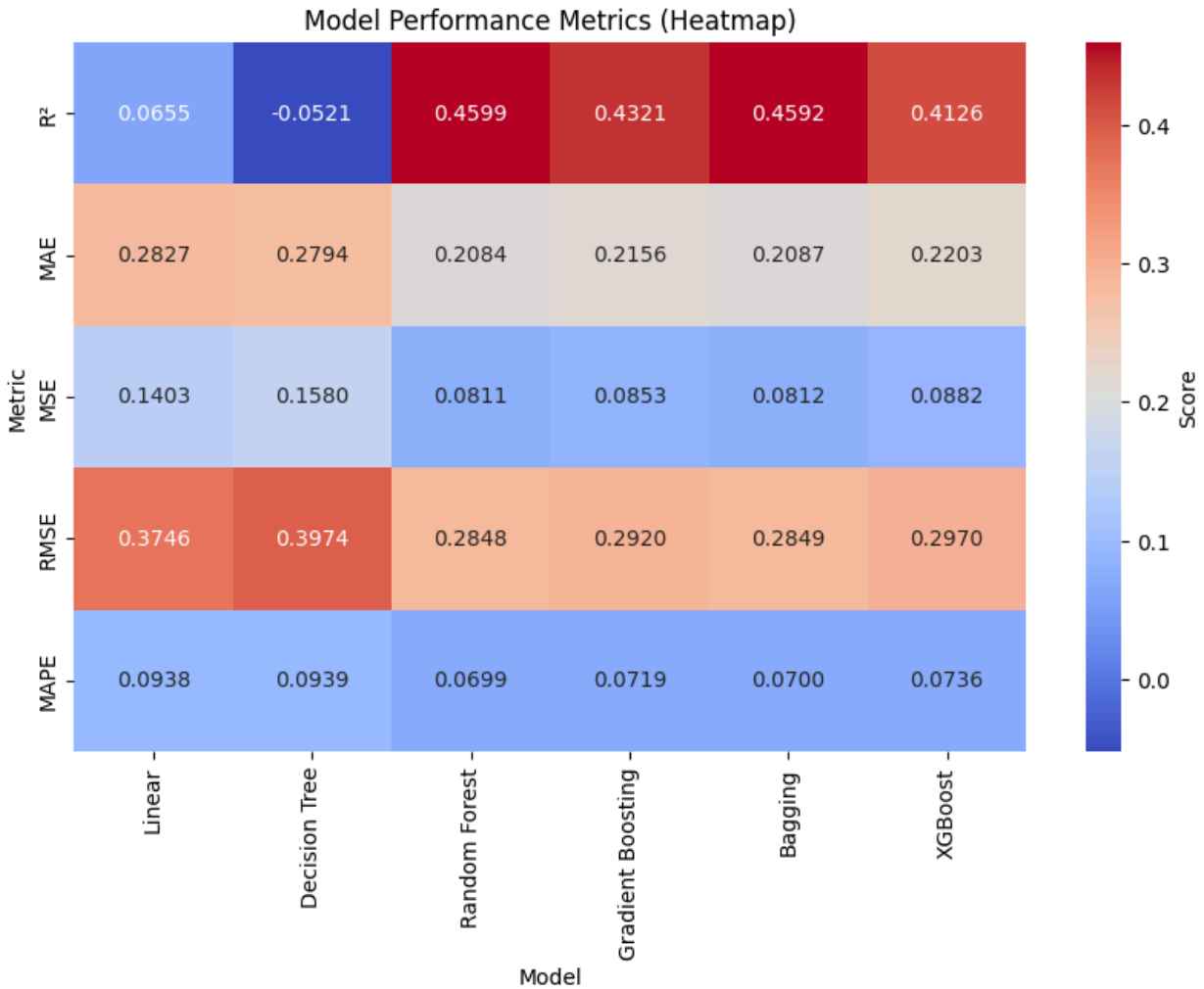
Performance Metrics

	R ²	MAE	MSE	RMSE	MAPE
Linear	0.0655	0.2827	0.1403	0.3746	0.0938
Decision Tree	-0.0521	0.2794	0.1580	0.3974	0.0939
Random Forest	0.4599	0.2084	0.0811	0.2848	0.0699
Gradient Boosting	0.4321	0.2156	0.0853	0.2920	0.0719
Bagging	0.4592	0.2087	0.0812	0.2849	0.0700
XGBoost	0.4126	0.2203	0.0882	0.2970	0.0736

Visualization Techniques

1. Heatmap of all metrics

```
plot.figure(figsize=(10, 6))
sns.heatmap(results_df.T, annot=True, cmap="coolwarm", fmt=".4f",
            cbar_kws={"label": "Score"})
plot.title("Model Performance Metrics (Heatmap)")
plot.ylabel("Metric")
plot.xlabel("Model")
plot.show()
```

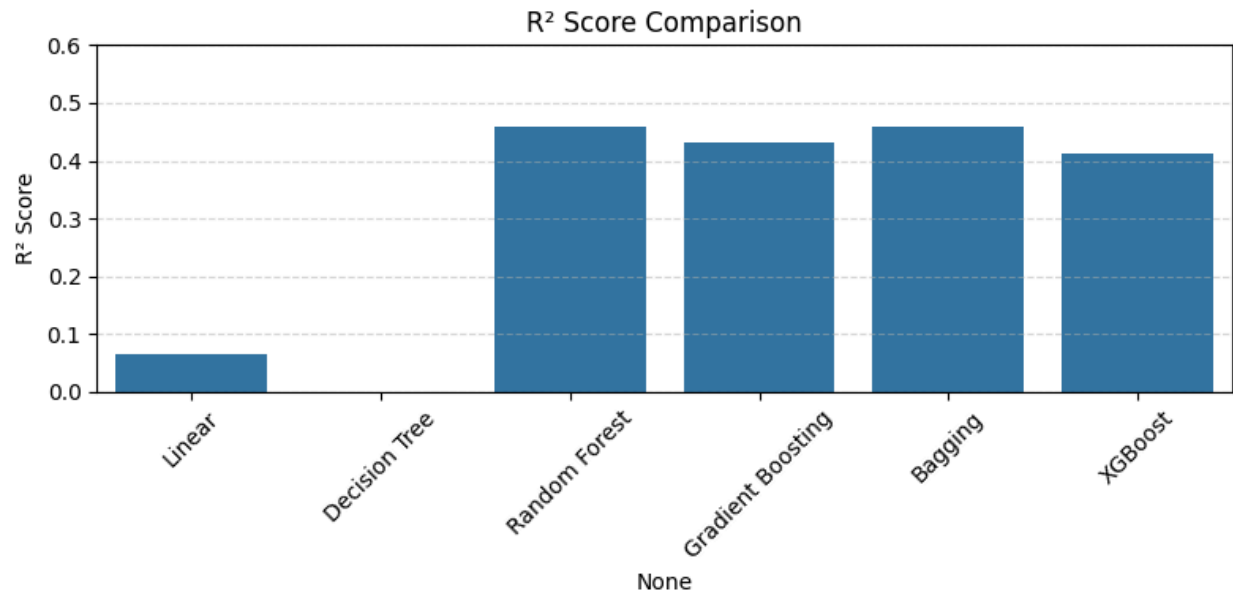


2. Bar plot of R² scores

```

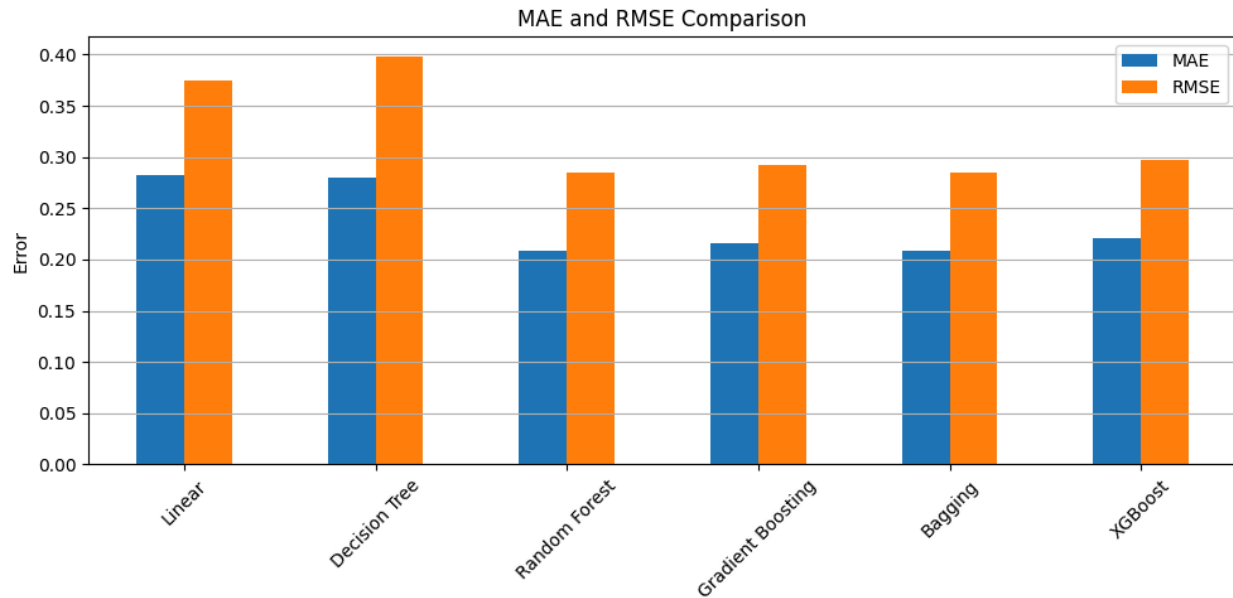
plot.figure(figsize=(8, 4))
sns.barplot(x=results_df.index, y=results_df["R2"])
plot.title("R2 Score Comparison")
plot.ylabel("R2 Score")
plot.ylim(0, 0.6)
plot.grid( axis = 'y', linestyle = '--', alpha = 0.5 )
plot.xticks(rotation=45)
plot.tight_layout()
plot.show()

```



3. Side-by-side bar plot for MAE and RMSE

```
mae_rmse_df = results_df[["MAE", "RMSE"]].copy()
mae_rmse_df.plot(kind="bar", figsize=(10, 5))
plot.title("MAE and RMSE Comparison")
plot.ylabel("Error")
plot.xticks(rotation=45)
plot.tight_layout()
plot.grid(axis='y')
plot.show()
```



Conclusion

Based on the performance across each evaluation metric, we can draw the following insights:

- **R²:** The best-performing model is **Random Forest**, while the **Decision Tree** model performs the worst.
- **MAE:** **Random Forest** and **Bagging** were nearly tied for the best performance, with only a 0.0003 difference in favor of **Random Forest**. **XGBoost** achieved the worst performance.
- **MSE:** **Random Forest** and **Bagging** again nearly tied for the best performance, differing by 0.0001 in favor of **Random Forest**. **Decision Tree** performed the worst here.
- **RMSE:** **Random Forest** and **Bagging** were nearly tied for the best, with only a 0.0001 difference in favor of **Random Forest**. The worst performance was from the **Decision Tree**.
- **MAPE:** **Random Forest** and **Bagging** were again closely matched for the best performance, with only a 0.0001 difference in favor of **Random Forest**. **Decision Tree** and **Linear Regression** were almost tied for the worst performance, with **Decision Tree** ahead by 0.0001.