

1 Paper Problems [40 points + [0/10 bonus]

Table 1: DC Data – MSE on ['id', 'align', 'eye', 'hair', 'sex', 'alive', 'year']

| Regressor | Train MSE | Test MSE |
|----------------|-----------|----------|
| Linear Reg. | 6305.183 | 8631.92 |
| Ridge Reg. | 8631.937 | 6305.183 |
| Logistic Reg. | 7287.173 | 9932.885 |
| Lasso Reg. | 8632.874 | 6305.22 |
| Random Forest | 5604.131 | 8350.651 |
| Gradient Boost | 3650.85 | 11788.79 |

Table 2: Marvel Data – MSE on ['id', 'align', 'eye', 'hair', 'sex', 'alive', 'year']

| Regressor | Train MSE | Test MSE |
|----------------|------------|----------|
| Linear Reg. | 10482.545 | 5716.500 |
| Ridge Reg. | 10482.545 | 5716.498 |
| Logistic Reg. | 11211.87 | 6210.674 |
| Lasso Reg. | 10482.5862 | 5715.585 |
| Random Forest | 9047.23 | 5844.28 |
| Gradient Boost | 7056.088 | 4909.267 |

Sklearn Linear Regression Train: 0.04414 Test: 0.04074 MSE TEST: 5716.5004302757925
MSE TRAIN: 10482.545429957441 Sklearn Ridge Regression Train: 0.04414 Test: 0.04074 MSE
TEST: 5716.498030488444 MSE TRAIN: 10482.545430045162 Sklearn Lasso Regression Train:
0.04430 Test: 0.04073 MSE TEST: 5715.585236948249 MSE TRAIN: 10482.586276824079

Table 3: DC Data – Regression on Eye, Hair color and Sex

| Regressor | KFold Test MSE |
|------------------|----------------|
| Linear Reg. | 7447.156 |
| Ridge Reg. | 7447.156 |
| Logistic Reg. | 8219.782 |
| Lasso Reg. | 7447.157 |
| Random Forest | 7323.726 |
| Gradient Boost | 7308.014 |
| K-Neighbors Reg. | 8921.632 |

Table 4: DC Data – Regression on Identity and Alignment

| Regressor | KFold Test MSE |
|------------------|----------------|
| Linear Reg. | 9245.027 |
| Ridge Reg. | 9253.201 |
| Logistic Reg. | 9257.364 |
| Lasso Reg. | 9397.039 |
| Random Forest | 9245.027 |
| Gradient Boost | 9257.364 |
| K-Neighbors Reg. | 14764.620 |

Table 5: Marvel Data – Regression on Identity and Alignment

| Regressor | KFold Test MSE |
|------------------|----------------|
| Linear Reg. | 9245.027 |
| Ridge Reg. | 9253.201 |
| Logistic Reg. | 9257.364 |
| Lasso Reg. | 9397.039 |
| Random Forest | 9245.027 |
| Gradient Boost | 9257.364 |
| K-Neighbors Reg. | 14764.620 |

Table 6: Marvel Data – Regression on ESH

| Regressor | KFold Test MSE |
|------------------|----------------|
| Linear Reg. | 9206.066 |
| Ridge Reg. | 9206.066 |
| Logistic Reg. | 9257.364 |
| Lasso Reg. | 9206.005 |
| Random Forest | 9171.763 |
| Gradient Boost | 9201.730 |
| K-Neighbors Reg. | 10307.132 |

Table 7: Marvel Data – Regression [id,align,eye,hair,sex,alive,year]

| Regressor | KFold Test MSE |
|------------------|----------------|
| Linear Reg. | 9062.278 |
| Ridge Reg. | 9062.278 |
| Logistic Reg. | 9712.405 |
| Lasso Reg. | 9062.277 |
| Random Forest | 8297.963 |
| Gradient Boost | 7766.886 |
| K-Neighbors Reg. | 8581.421 |

Table 8: DC Data – Regression [id,align,eye,hair,sex,alive,year]

| Regressor | KFold Test MSE |
|------------------|----------------|
| Linear Reg. | 7018.653 |
| Ridge Reg. | 7018.653 |
| Logistic Reg. | 7848.797 |
| Lasso Reg. | 7018.652 |
| Random Forest | 6791.093 |
| Gradient Boost | 7114.129 |
| K-Neighbors Reg. | 7542.965 |