

Final Model: Random Forest Classifier

Features

Star rating is based solely on patient health statistics. Each dialysis center is graded on nine separate health statistics. These include:

- **Mortality Ratios**
 - Mortality Rate Standardized
- **Hospitalizations**
 - Hospitalization Rate Standardized
- **Blood Transfusions**
 - Transfusion Rate Standardized
- **Incidents of Hypercalcemia**
 - Percentage Of Adult Patients With Hypercalcemia (Serum Calcium Greater Than 10.2 mg/dL)
- **Percentage of AV fistulas**
 - Fistula Rate Standardized
- **Percentage of Catheters In Use Over 90 Days**
 - Percentage Of Adult Patients With Long Term Catheter In Use
- **Percentage of Waste Removed During Hemodialysis in Adults and Children & Percentage of Waste Removed in Adults During Peritoneal Dialysis**
 - Percent Of Adult HD Patients With Kt/V ≥ 1.2
 - Percentage Of Adult Patients With Hypercalcemia (Serum Calcium Greater Than 10.2 mg/dL)
 - *Children data was too sparse to be use*

Percentage of Waste Removed During Hemodialysis in Adults and Children & Percentage of Waste Removed in Adults During Peritoneal Dialysis are broad terms encompassing multiple features. In order to avoid overfitting, use the features that are at least moderately correlated (≥ 0.3) with the dependent variable (star rating).

Parameters

- min_samples_split = 2
- min_samples_leaf = 1
- min_weight_fraction_leaf = 0.0
- max_features = 'auto'
- max_leaf_nodes = None
- min_impurity_decrease = 0.0
- min_impurity_split = None
- _weight = None
- bootstrap = True
- oob_score = False
- n_jobs = None
- random_state = None
- verbose = 0
- warm_start = False
- class

Hyperparameters

- `n_estimators = 500`
- `criterion = 'entropy'`
- `max_depth = 12`

Performance Metrics

F1-Score

- Accuracy = 0.80
- Macro average = 0.72
- Weighted average = 0.80

Cross validation

- Mean cross validation train score: 0.7657427660059238
- Standard deviation in train cv scores: 0.019451199874457762
- Mean cross validation test score: 0.7657427660059238
- Standard deviation in test cv scores: 0.019451199874457762

Runtime

- 0.5211247709994495