Notes

Plot of dnr day and drn

#Given that we have reason to believe drn (do not rescussitate order) is not independant to DNRday I plot on a chart

#in order to ascertain indepandance since we see that they are clearly not independant I decide to remove DNRday in order to simplify the analysis

ax=sns.stripplot(x=X2['dnr'], y=X2['dnrday'],hue=X2['surv2m'])

sns.move\_legend(ax, "upper left", bbox\_to\_anchor=(1, 1))

A table of numbers and symbols

Description automatically generated

Ca

txt="Figure 1: Confusion Matrix and Metrics Model 1"

plt.figtext(0.5, 0.01, txt, wrap=True, horizontalalignment='center', fontsize=12)

plt.tight\_layout()

plt.show()

\*Figure 2: Correlation Matrix Independant Variables\*

dnr

Model 1:

Accuracy of the model: 0.85

Micro Precision: 0.8511806699615596

Micro Recall: 0.8511806699615596

r2: 0.2281363885195904

mse: 0.14881933003844042

bic: 4039.0100693185977

logloss: 5.363992349506177

Model 2:

Accuracy of the model: 0.89

Micro Precision: 0.8923668314113125

Micro Recall: 0.8923668314113125

r2: 0.4417517791507001

mse: 0.10763316858868753

bic: 4664.094168211044

logloss: 3.8794926217830654

Model 3:

Accuracy of the model: 0.89

Micro Precision: 0.8901702361339923

Micro Recall: 0.8901702361339923

r2: 0.43035895831704085

mse: 0.10982976386600769

bic: 4657.764559661429

logloss: 3.9586659405949645

model\_log\_loss = log\_loss(Y\_test, Y\_pred\_log)

print(f'logloss: {model\_log\_loss}')

Model 4:

Accuracy of the model: 0.89

Micro Precision: 0.8918176825919825

Micro Recall: 0.8918176825919825

r2: 0.4389035739422852

mse: 0.10818231740801758

bic: 4830.44702992389

logloss: 3.8992859514860405

A screenshot of a graph

Description automatically generated

For plotting many scatterplots

import matplotlib.pyplot as plt

#f, ax\_l = plt.subplots(1, 10, figsize=(14, 4))

f, ax\_l = plt.subplots(5,1, figsize=(6, 14))

#note that here I am only including the continous variables

for e, col\_name in enumerate(X\_correlation.loc[:,:'adlsc' ].columns):

ax\_l[e].scatter( X\_correlation[col\_name], X\_correlation.hospdead,alpha=0.025, color='b')

ax\_l[e].set\_xlabel(col\_name)

ax\_l[e].set\_ylabel('hospdead')

ax\_l[e].xaxis.set\_major\_locator(ticker.LinearLocator(2))

plt.tight\_layout()

///

ax = sns.stripplot(y=X\_correlation['hospdead'], x=X\_correlation['age'],hue=X\_correlation['hospdead'])

ax.xaxis.set\_major\_locator(ticker.LinearLocator(5))

ax.set(title="Plot of avtisst and sps Coma category")

////

race\_black

totcst the total ratio of costs to charges (RCC) cost

drn\_drn\_before\_sadm

drn\_drn\_before\_sadm

dzgroup\_ARF/MOSF w/Sepsis, dzgroup\_CHF, dzgroup\_COPD, dzgroup\_Cirrhosis, dzgroup\_Colon Cancer, dzgroup\_Coma, dzgroup\_Lung Cancer, dzgroup\_MOSF w/Malig

Totcst removed

Cirrhosis dz group removed

\*had earlier removed

dzgroup\_Coma

Hday,

Charges

Sps

aps

Model 1:

Mean Cross Validation Accuracy Score: 0.8628244532635312

//

Model 2

A screenshot of a computer

Description automatically generated

Model 3

A screenshot of a computer

Description automatically generated

Accuracy of the model: 0.89

Micro Precision: 0.8874244920373421

Micro Recall: 0.8874244920373421

r2: 0.41611793227496685

mse: 0.11257550796265788

bic: 4830.44702992389

logloss: 4.057632589109839

Bayesian Model Model 6