modelling_results_1_0-interp_data-rem_feat

May 14, 2023

[20]:
```
import pandas as pd
import seaborn as sns
```

[21]:
```
data = pd.read_pickle("df_results_1_0")
```

[22]:
```
data
```

[22]:

	Classifier	Precision	Recall	F1-score	Timestamp
1	Support Vector Mashines	0.72	0.57	0.64	10052023_2137
3	Decision Tree	0.42	0.38	0.39	10052023_2137
5	Random Forest	0.81	0.55	0.66	10052023_2137
7	Neural Network	0.33	0.62	0.43	10052023_2137
1	Support Vector Mashines	0.79	0.48	0.59	10052023_2156
3	Decision Tree	0.62	0.20	0.30	10052023_2156
5	Random Forest	0.86	0.60	0.71	10052023_2156
7	Neural Network	0.46	0.32	0.38	10052023_2156
1	Support Vector Mashines	0.63	0.42	0.51	10052023_2210
3	Decision Tree	0.54	0.38	0.44	10052023_2210
5	Random Forest	0.70	0.65	0.68	10052023_2210
7	Neural Network	0.47	0.42	0.45	10052023_2210

[31]:
```
data.groupby(by="Classifier").mean().round(2)
```

[31]:

Classifier	Precision	Recall	F1-score
Decision Tree	0.53	0.32	0.38
Neural Network	0.42	0.45	0.42
Random Forest	0.79	0.60	0.68
Support Vector Mashines	0.71	0.49	0.58

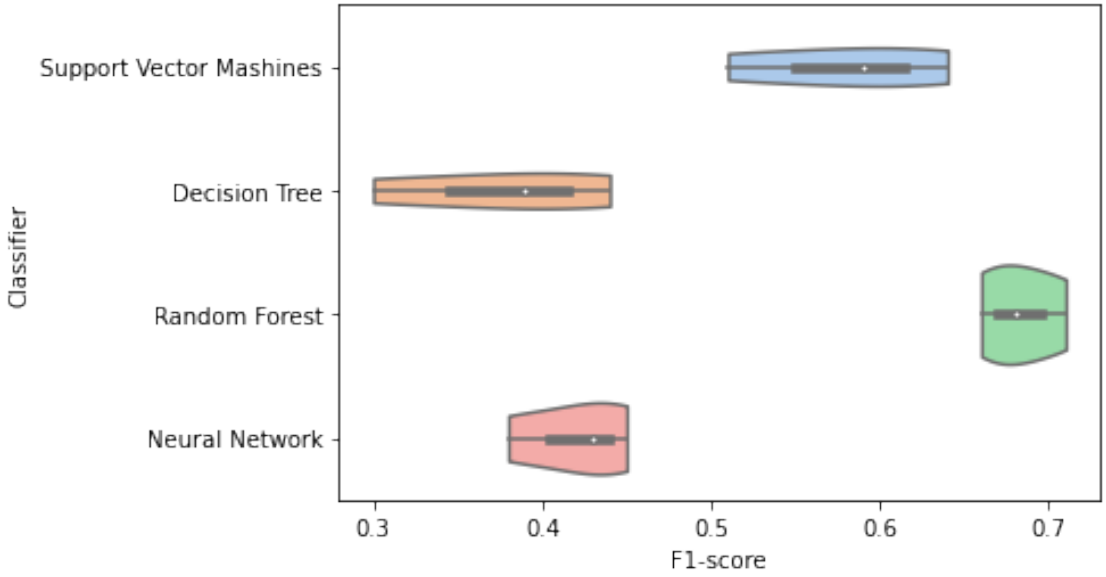
[29]:
```
data.describe()
```

[29]:

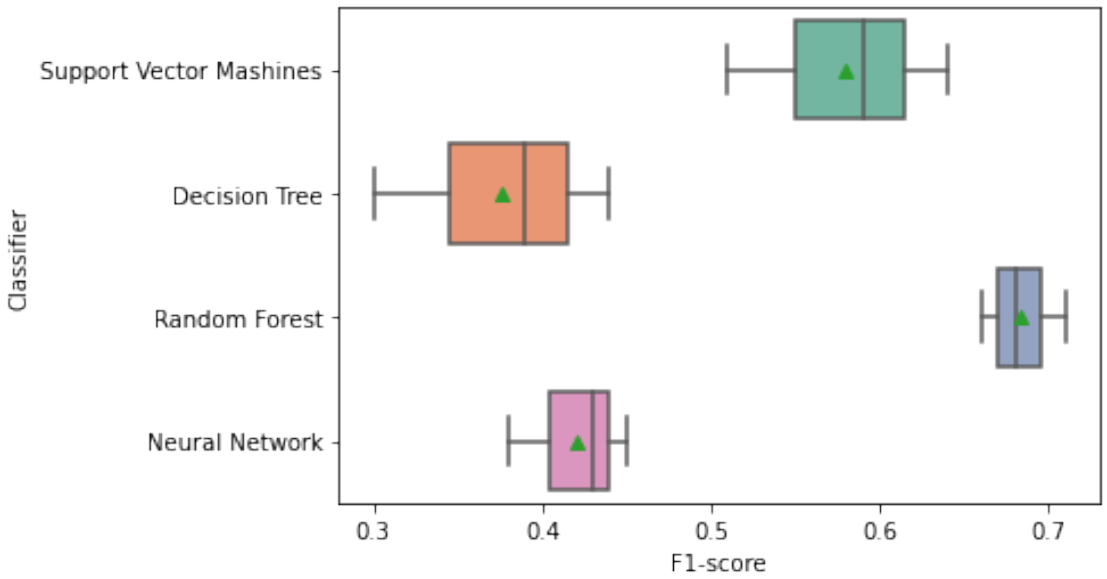
	Precision	Recall	F1-score
count	12.000000	12.000000	12.000000
mean	0.612500	0.465833	0.515000
std	0.169766	0.136412	0.136415
min	0.330000	0.200000	0.300000
25%	0.467500	0.380000	0.420000

```
50%     0.625000   0.450000   0.480000
75%     0.737500   0.577500   0.645000
max     0.860000   0.650000   0.710000
```

[23]:
```
ax = sns.violinplot(data=data, y="Classifier", x="F1-score", orient="h",␣
↪palette="pastel", showmeans=True, cut=0)
```

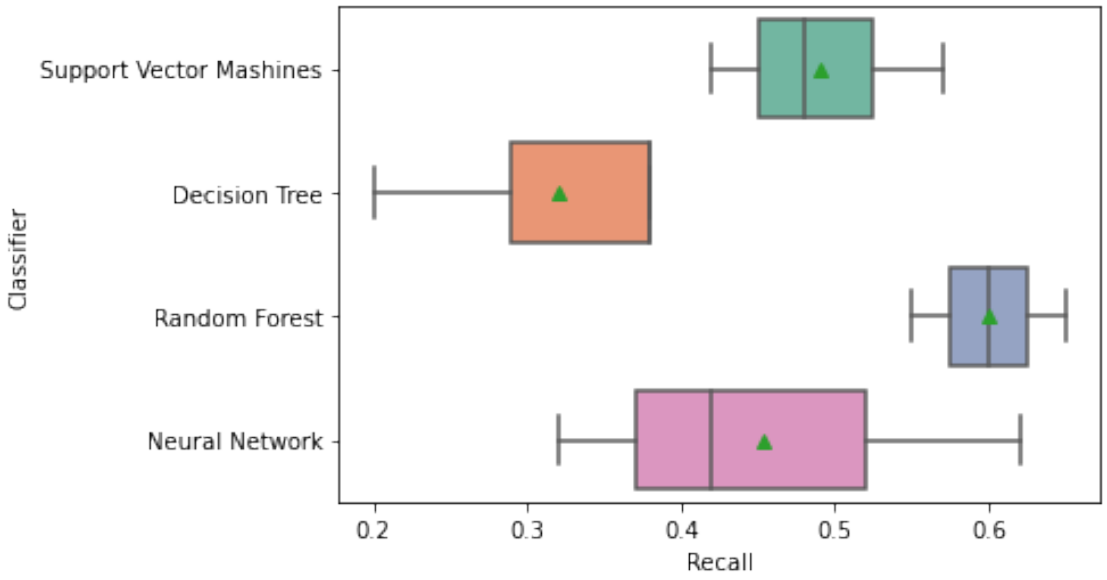


[24]:
```
ax = sns.boxplot(data=data, y="Classifier", x="F1-score", orient="h",␣
↪palette="Set2", showmeans=True)
# sns.boxplot(data=data, y="Classifier", x="Recall", orient="h", color="white",␣
↪showmeans=True, ax=ax)
```



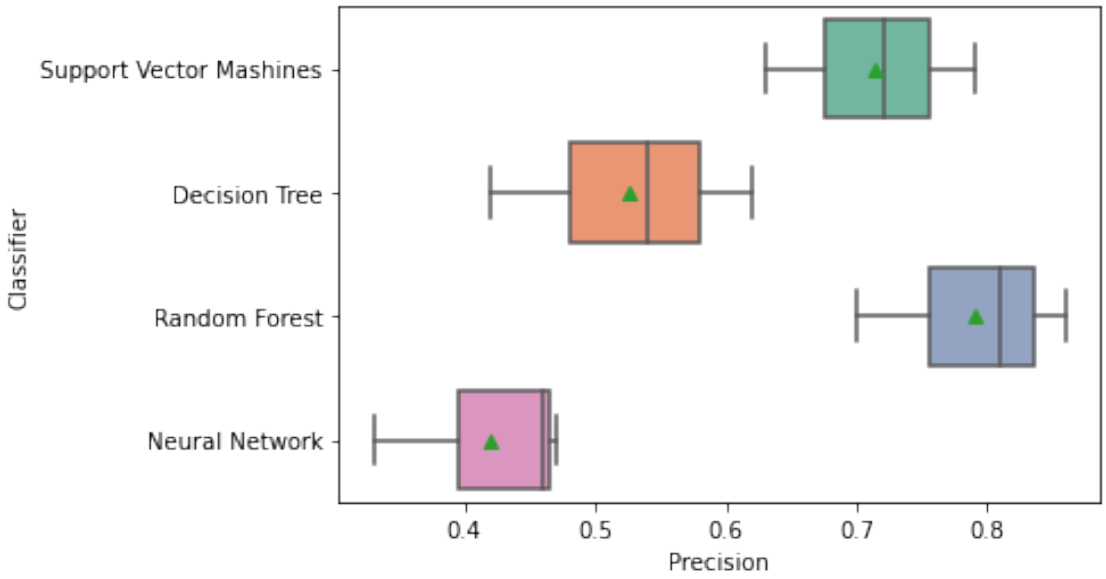
[25]:
```
sns.boxplot(data=data, y="Classifier", x="Recall", orient="h", palette="Set2",␣
↪showmeans=True)
```

[25]: <AxesSubplot:xlabel='Recall', ylabel='Classifier'>

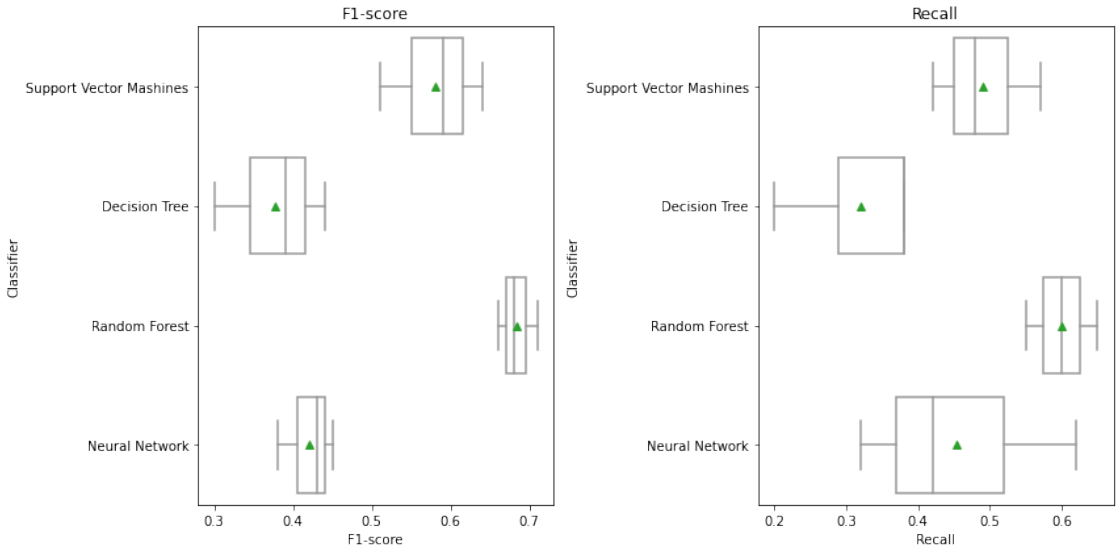


[26]:
```
sns.boxplot(data=data, y="Classifier", x="Precision", orient="h",␣
↪palette="Set2", showmeans=True)
```

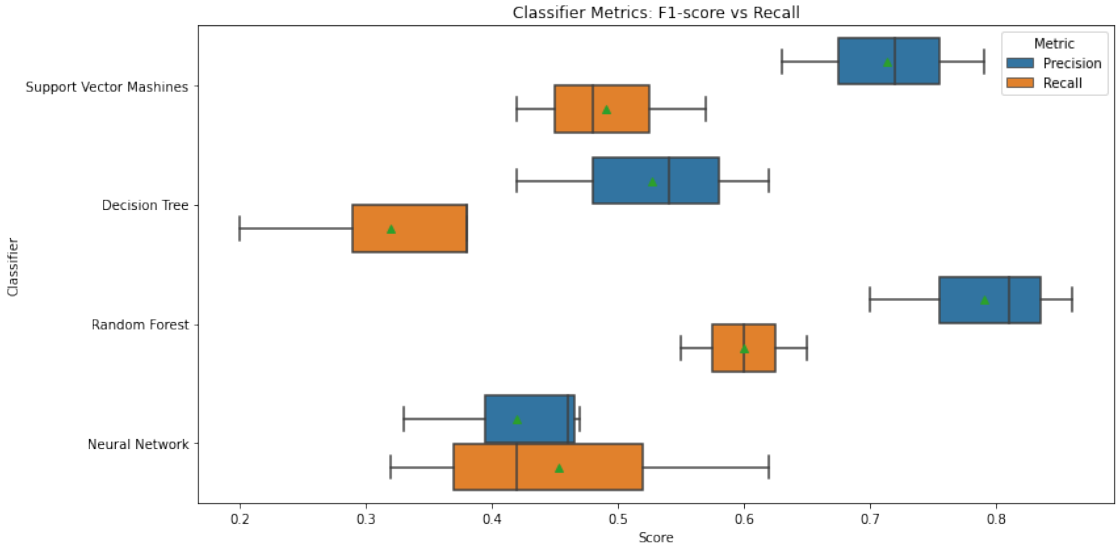
[26]: <AxesSubplot:xlabel='Precision', ylabel='Classifier'>



[27]:
```
from matplotlib import pyplot as plt

fig, axs = plt.subplots(ncols=2, figsize=(12, 6))

ax1 = sns.boxplot(data=data, y="Classifier", x="F1-score", orient="h",␣
 ↪color="white", showmeans=True, ax=axs[0])
ax1.set_title("F1-score")

ax2 = sns.boxplot(data=data, y="Classifier", x="Recall", orient="h",␣
 ↪color="white", showmeans=True, ax=axs[1])
ax2.set_title("Recall")

plt.tight_layout()
plt.show()
```



[28]:
```
# Reshape the data
data_melted = pd.melt(data, id_vars='Classifier', value_vars=['Precision',␣
 ↪'Recall'], var_name='Metric', value_name='Score')

# Create the boxplot
plt.figure(figsize=(12, 6))
ax = sns.boxplot(data=data_melted, y="Classifier", x="Score", orient="h",␣
 ↪hue="Metric", showmeans=True)
ax.set_title("Classifier Metrics: F1-score vs Recall")
plt.tight_layout()
plt.show()
```



[ ]: