

```

In[1]:= dataPath = FileNameJoin[
    {ParentDirectory[NotebookDirectory[]], "data/dataOccupancyPreprocessed.csv"}];

In[2]:= data = SemanticImport[dataPath, {"DateTime", "Real", "Real",
    "Real", "Real", "Real", "Boolean"}, "Dataset", HeaderLines → 1];

In[3]:= trainData = Values@
    Normal[data[[All, {"Temperature", "Humidity", "Light", "CO2", "HumidityRatio"}]]];

In[15]:= cluster[method_ : "GaussianMixture", numberClusters_ : 2,
    distanceFunction_ : EuclideanDistance] := Module[{clusters, rules, classifier},
    clusters = ClusteringComponents[trainData, numberClusters, 1, Method → method,
        DistanceFunction → distanceFunction, PerformanceGoal → "Quality"];
    rules = Map[clusters[[#]] → data[[#, "Occupancy"]] &, Range[Length[data]]];
    classifier = Classify[rules, Method → Automatic];
    Return[ClassifierMeasurements[classifier, rules]];
]

In[56]:= cm = cluster["KMeans", 10, ManhattanDistance];

In[57]:= cm["ClassMeanCrossEntropy"]

Out[57]= <| False → 0.082856, True → 0.24876 |>

In[58]:= cm["Accuracy"]

Out[58]= 0.960274

In[59]:= cm["FScore"]

Out[59]= <| False → 0.974663, True → 0.908064 |>

In[60]:= cm["ClassMeanCrossEntropy"]

Out[60]= <| False → 0.082856, True → 0.24876 |>

In[61]:= cm["Specificity"]

Out[61]= <| False → 0.982992, True → 0.954609 |>

In[62]:= cm["Perplexity"]

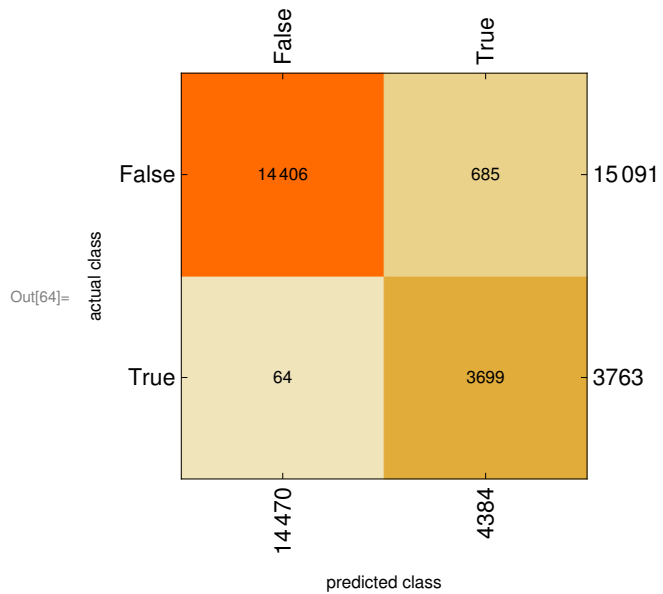
Out[62]= 1.12296

In[63]:= cm["Precision"]

Out[63]= <| False → 0.995577, True → 0.84375 |>

In[64]:= cm["ConfusionMatrixPlot"]

```



In[65]:= `cm["ROCCurve"]`

