

```
In [13]: import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import accuracy_score, classification_report, confusion_matrix
```

```
In [14]: data = pd.read_csv('/content/fetal_health.csv')
```

```
In [15]: print("Shape of the original dataset:", data.shape)
```

Shape of the original dataset: (2126, 22)

```
In [16]: # Print the first few rows
print(data.head())
```

	baseline_value	accelerations	fetal_movement	uterine_contractions	\
0	120.0	0.000	0.0	0.000	
1	132.0	0.006	0.0	0.006	
2	133.0	0.003	0.0	0.008	
3	134.0	0.003	0.0	0.008	
4	132.0	0.007	0.0	0.008	

	light_decelerations	severe_decelerations	prolongued_decelerations	\
0	0.000	0.0	0.0	
1	0.003	0.0	0.0	
2	0.003	0.0	0.0	
3	0.003	0.0	0.0	
4	0.000	0.0	0.0	

	abnormal_short_term_variability	mean_value_of_short_term_variability	\
0	73.0	0.5	
1	17.0	2.1	
2	16.0	2.1	
3	16.0	2.4	
4	16.0	2.4	

	percentage_of_time_with_abnormal_long_term_variability	...	histogram_min	\
0	43.0	...	62.0	
1	0.0	...	68.0	
2	0.0	...	68.0	
3	0.0	...	53.0	
4	0.0	...	53.0	

	histogram_max	histogram_number_of_peaks	histogram_number_of_zeroes	\
0	126.0	2.0	0.0	
1	198.0	6.0	1.0	
2	198.0	5.0	1.0	
3	170.0	11.0	0.0	
4	170.0	9.0	0.0	

	histogram_mode	histogram_mean	histogram_median	histogram_variance	\
0	120.0	137.0	121.0	73.0	
1	141.0	136.0	140.0	12.0	
2	141.0	135.0	138.0	13.0	
3	137.0	134.0	137.0	13.0	
4	137.0	136.0	138.0	11.0	

	histogram_tendency	fetal_health
0	1.0	2.0
1	0.0	1.0
2	0.0	1.0
3	1.0	1.0
4	1.0	1.0

[5 rows x 22 columns]

```
In [17]: # Check number of null values in each coloumn
print(data.isnull().sum())
```

```
baseline value      0
accelerations       0
fetal_movement      0
uterine_contractions 0
light_decelerations 0
severe_decelerations 0
prolongued_decelerations 0
abnormal_short_term_variability 0
mean_value_of_short_term_variability 0
percentage_of_time_with_abnormal_long_term_variability 0
mean_value_of_long_term_variability 0
histogram_width     0
histogram_min       0
histogram_max       0
histogram_number_of_peaks 0
histogram_number_of_zeroes 0
histogram_mode      0
histogram_mean      0
histogram_median    0
histogram_variance  0
histogram_tendency  0
fetal_health        0
dtype: int64
```

```
In [18]: # Split the dataset into features and target variable
x = data.drop(columns=['fetal_health'])
y = data['fetal_health']
```

```
In [19]: # Split the data into training and testing sets
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2, random_state=42)
```

```
In [20]: # Create the decision tree classifier
clf = DecisionTreeClassifier()
```

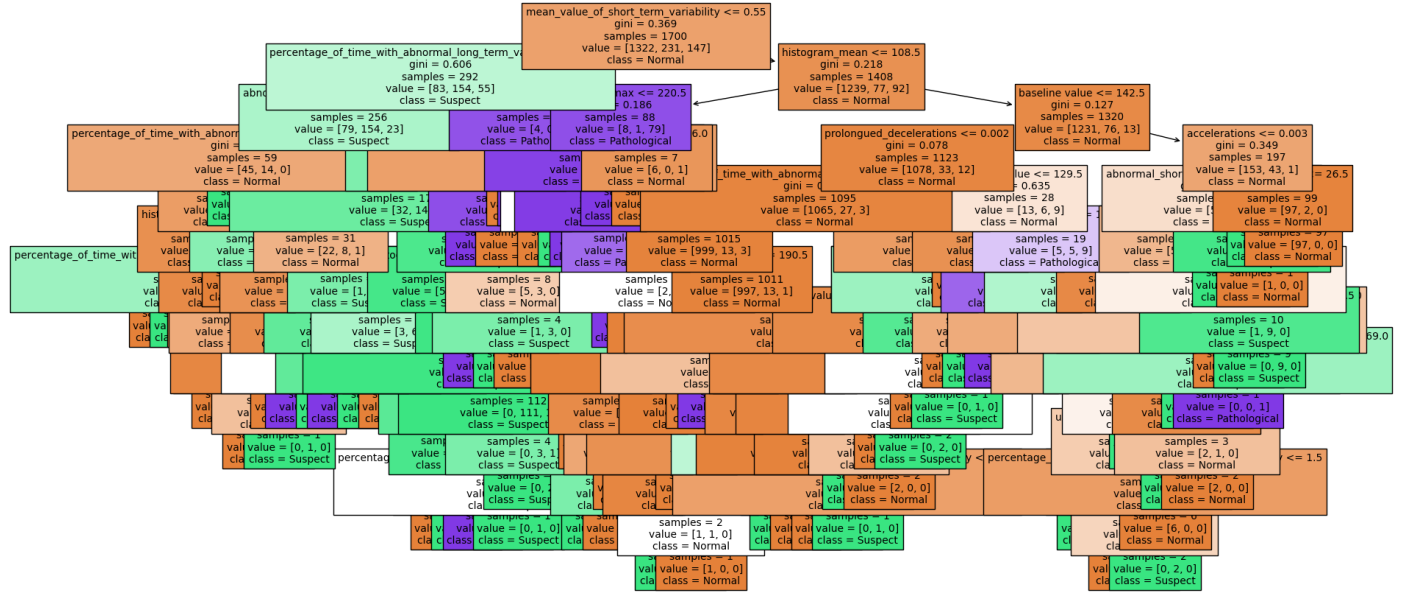
```
In [21]: # Train the classifier with training data
clf.fit(x_train, y_train)
```

```
Out[21]: □ DecisionTreeClassifier
DecisionTreeClassifier()
```

```
In [22]: # Predict the target variable
y_pred = clf.predict(x_test)
```

```
In [23]: import matplotlib.pyplot as plt
from sklearn.tree import plot_tree

plt.figure(figsize=(20,10))
# Plot and display the decision tree
plot_tree(clf, feature_names=x.columns, class_names=['Normal', 'Suspect', 'Pathological'])
plt.show()
```



```
In [24]: # Calculate the accuracy of DT classifier
accuracy = accuracy_score(y_test, y_pred)
print("Accuracy:", accuracy)

# Calculate precision, recall, and F1-score for DT
print("Decision Tree Classification Report:")
print(classification_report(y_test, y_pred, target_names=['Normal', 'Suspect', 'Patholog
```

Accuracy: 0.9272300469483568

Decision Tree Classification Report:

	precision	recall	f1-score	support
Normal	0.97	0.94	0.95	333
Suspect	0.76	0.83	0.79	64
Pathological	0.90	0.97	0.93	29
accuracy			0.93	426
macro avg	0.88	0.91	0.89	426
weighted avg	0.93	0.93	0.93	426

```
In [25]: # Confusion Matrix Display
from sklearn.metrics import confusion_matrix, ConfusionMatrixDisplay
cm = confusion_matrix(y_test, y_pred)
disp = ConfusionMatrixDisplay(confusion_matrix=cm, display_labels=['Normal', 'Suspect',
disp.plot(cmap=plt.cm.Blues)
plt.show()
```

