BREAST CANCER DETECTION

DATA SOURCE

https://www.kaggle.com/datasets/uciml/breast-cancer-wisconsin-data

Pycaret Module

In [1]:
```
!pip install pycaret
```

```
Collecting pycaret
  Obtaining dependency information for pycaret from https://files.pythonhosted.org/packa
ges/f5/4b/2002980b046ac396618dfc152d384b812a78182b776ca77fe0ae5f80deac/pycaret-3.0.4-py3
-none-any.whl.metadata
  Downloading pycaret-3.0.4-py3-none-any.whl.metadata (17 kB)
Requirement already satisfied: ipython>=5.5.0 in e:\anaconda3\lib\site-packages (from py
caret) (8.10.0)
Collecting ipywidgets>=7.6.5 (from pycaret)
  Obtaining dependency information for ipywidgets>=7.6.5 from https://files.pythonhoste
d.org/packages/b8/d4/ce436660098b2f456e2b8fdf76d4f33cbc3766c874c4aa2f772c7a5e943f/ipywid
gets-8.1.0-py3-none-any.whl.metadata
  Downloading ipywidgets-8.1.0-py3-none-any.whl.metadata (2.4 kB)
Requirement already satisfied: tqdm>=4.62.0 in e:\anaconda3\lib\site-packages (from pyca
ret) (4.65.0)
Requirement already satisfied: numpy<1.24,>=1.21 in e:\anaconda3\lib\site-packages (from
pycaret) (1.21.5)
Requirement already satisfied: pandas<2.0.0,>=1.3.0 in e:\anaconda3\lib\site-packages (f
rom pycaret) (1.4.2)
Requirement already satisfied: jinja2>=1.2 in e:\anaconda3\lib\site-packages (from pycar
et) (3.1.2)
Requirement already satisfied: scipy<2.0.0 in e:\anaconda3\lib\site-packages (from pycar
et) (1.7.3)
Collecting joblib>=1.2.0 (from pycaret)
  Obtaining dependency information for joblib>=1.2.0 from https://files.pythonhosted.or
g/packages/10/40/d551139c85db202f1f384ba8bcf96aca2f329440a844f924c8a0040b6d02/joblib-1.
3.2-py3-none-any.whl.metadata
  Downloading joblib-1.3.2-py3-none-any.whl.metadata (5.4 kB)
Requirement already satisfied: scikit-learn<1.3.0,>=1.0 in e:\anaconda3\lib\site-package
s (from pycaret) (1.0.2)
Collecting pyod>=1.0.8 (from pycaret)
  Downloading pyod-1.1.0.tar.gz (153 kB)
     ---------------------------------------- 153.4/153.4 kB 1.8 MB/s eta 0:00:00
  Preparing metadata (setup.py): started
  Preparing metadata (setup.py): finished with status 'done'
Collecting imbalanced-learn>=0.8.1 (from pycaret)
  Obtaining dependency information for imbalanced-learn>=0.8.1 from https://files.python
hosted.org/packages/a3/9e/fbe60a768502af54563dcb59ca7856f5a8833b3ad5ada658922e1ab09b7f/i
mbalanced_learn-0.11.0-py3-none-any.whl.metadata
  Downloading imbalanced_learn-0.11.0-py3-none-any.whl.metadata (8.3 kB)
Collecting category-encoders>=2.4.0 (from pycaret)
  Obtaining dependency information for category-encoders>=2.4.0 from https://files.pytho
nhosted.org/packages/1f/e2/495811f12b2e90753fff0e42a07adb0370a725de17cc23a579ac9d3ca67c/
category_encoders-2.6.2-py2.py3-none-any.whl.metadata
  Downloading category_encoders-2.6.2-py2.py3-none-any.whl.metadata (8.0 kB)
Collecting lightgbm>=3.0.0 (from pycaret)
  Obtaining dependency information for lightgbm>=3.0.0 from https://files.pythonhosted.o
rg/packages/87/0f/7630ee4fea60ebab5b0e3c35df570cb295c91ece537231a38105c0f243e8/lightgbm-
4.0.0-py3-none-win_amd64.whl.metadata
```

```
  Downloading lightgbm-4.0.0-py3-none-win_amd64.whl.metadata (19 kB)
Requirement already satisfied: numba>=0.55.0 in e:\anaconda3\lib\site-packages (from pyc
aret) (0.55.1)
Requirement already satisfied: requests>=2.27.1 in e:\anaconda3\lib\site-packages (from
pycaret) (2.31.0)
Requirement already satisfied: psutil>=5.9.0 in e:\anaconda3\lib\site-packages (from pyc
aret) (5.9.0)
Requirement already satisfied: markupsafe>=2.0.1 in e:\anaconda3\lib\site-packages (from
pycaret) (2.1.1)
Requirement already satisfied: importlib-metadata>=4.12.0 in e:\anaconda3\lib\site-packa
ges (from pycaret) (6.0.0)
Requirement already satisfied: nbformat>=4.2.0 in e:\anaconda3\lib\site-packages (from p
ycaret) (5.7.0)
Requirement already satisfied: cloudpickle in e:\anaconda3\lib\site-packages (from pycar
et) (2.2.1)
Collecting deprecation>=2.1.0 (from pycaret)
  Downloading deprecation-2.1.0-py2.py3-none-any.whl (11 kB)
Collecting xxhash (from pycaret)
  Obtaining dependency information for xxhash from https://files.pythonhosted.org/packag
es/30/25/ae5ede30b61e78d98cc0be2e94ad8a942ced11eafba0c2d2efe0db7647aa/xxhash-3.3.0-cp39-
cp39-win_amd64.whl.metadata
  Downloading xxhash-3.3.0-cp39-cp39-win_amd64.whl.metadata (12 kB)
Requirement already satisfied: matplotlib>=3.3.0 in e:\anaconda3\lib\site-packages (from
pycaret) (3.5.1)
Collecting scikit-plot>=0.3.7 (from pycaret)
  Downloading scikit_plot-0.3.7-py3-none-any.whl (33 kB)
Collecting yellowbrick>=1.4 (from pycaret)
  Downloading yellowbrick-1.5-py3-none-any.whl (282 kB)
     ---------------------------------------- 282.6/282.6 kB 2.2 MB/s eta 0:00:00
Requirement already satisfied: plotly>=5.0.0 in e:\anaconda3\lib\site-packages (from pyc
aret) (5.6.0)
Collecting kaleido>=0.2.1 (from pycaret)
  Downloading kaleido-0.2.1-py2.py3-none-win_amd64.whl (65.9 MB)
     ---------------------------------------- 65.9/65.9 MB 1.2 MB/s eta 0:00:00
Collecting schemdraw==0.15 (from pycaret)
  Downloading schemdraw-0.15-py3-none-any.whl (106 kB)
     ---------------------------------------- 106.8/106.8 kB 1.2 MB/s eta 0:00:00
Collecting plotly-resampler>=0.8.3.1 (from pycaret)
  Obtaining dependency information for plotly-resampler>=0.8.3.1 from https://files.pyth
onhosted.org/packages/08/1d/87d4ed45c26226630bcb0a205ff006c00645cc68977e22c0f6f16a7f5d2
b/plotly_resampler-0.9.1-py3-none-any.whl.metadata
  Downloading plotly_resampler-0.9.1-py3-none-any.whl.metadata (12 kB)
Requirement already satisfied: statsmodels>=0.12.1 in e:\anaconda3\lib\site-packages (fr
om pycaret) (0.13.2)
Collecting sktime!=0.17.1,!=0.17.2,!=0.18.0,>=0.16.1 (from pycaret)
  Obtaining dependency information for sktime!=0.17.1,!=0.17.2,!=0.18.0,>=0.16.1 from ht
tps://files.pythonhosted.org/packages/bc/6f/ae8d2e9779e517679b1cfa7542f7eaaa3b9c6582c9c1
3237b16378fff1e2/sktime-0.22.0-py3-none-any.whl.metadata
  Downloading sktime-0.22.0-py3-none-any.whl.metadata (25 kB)
Collecting tbats>=1.1.3 (from pycaret)
  Downloading tbats-1.1.3-py3-none-any.whl (44 kB)
     ---------------------------------------- 44.0/44.0 kB 2.1 MB/s eta 0:00:00
Collecting pmdarima!=1.8.1,<3.0.0,>=1.8.0 (from pycaret)
  Downloading pmdarima-2.0.3-cp39-cp39-win_amd64.whl (572 kB)
     ---------------------------------------- 572.7/572.7 kB 1.4 MB/s eta 0:00:00
Requirement already satisfied: patsy>=0.5.1 in e:\anaconda3\lib\site-packages (from cate
gory-encoders>=2.4.0->pycaret) (0.5.2)
Requirement already satisfied: packaging in e:\anaconda3\lib\site-packages (from depreca
tion>=2.1.0->pycaret) (23.0)
Requirement already satisfied: threadpoolctl>=2.0.0 in e:\anaconda3\lib\site-packages (f
rom imbalanced-learn>=0.8.1->pycaret) (2.2.0)
Requirement already satisfied: zipp>=0.5 in e:\anaconda3\lib\site-packages (from importl
ib-metadata>=4.12.0->pycaret) (3.11.0)
Requirement already satisfied: backcall in e:\anaconda3\lib\site-packages (from ipython>
=5.5.0->pycaret) (0.2.0)
Requirement already satisfied: decorator in e:\anaconda3\lib\site-packages (from ipython
```

```
>=5.5.0->pycaret) (5.1.1)
Requirement already satisfied: jedi>=0.16 in e:\anaconda3\lib\site-packages (from ipytho
n>=5.5.0->pycaret) (0.18.1)
Requirement already satisfied: matplotlib-inline in e:\anaconda3\lib\site-packages (from
ipython>=5.5.0->pycaret) (0.1.6)
Requirement already satisfied: pickleshare in e:\anaconda3\lib\site-packages (from ipyth
on>=5.5.0->pycaret) (0.7.5)
Requirement already satisfied: prompt-toolkit<3.1.0,>=3.0.30 in e:\anaconda3\lib\site-pa
ckages (from ipython>=5.5.0->pycaret) (3.0.37)
Requirement already satisfied: pygments>=2.4.0 in e:\anaconda3\lib\site-packages (from i
python>=5.5.0->pycaret) (2.14.0)
Requirement already satisfied: stack-data in e:\anaconda3\lib\site-packages (from ipytho
n>=5.5.0->pycaret) (0.2.0)
Requirement already satisfied: traitlets>=5 in e:\anaconda3\lib\site-packages (from ipyt
hon>=5.5.0->pycaret) (5.7.1)
Requirement already satisfied: colorama in e:\anaconda3\lib\site-packages (from ipython>
=5.5.0->pycaret) (0.4.6)
Collecting comm>=0.1.3 (from ipywidgets>=7.6.5->pycaret)
  Obtaining dependency information for comm>=0.1.3 from https://files.pythonhosted.org/p
ackages/fe/47/0133ac1b7dc476ed77710715e98077119b3d9bae56b13f6f9055e7da1c53/comm-0.1.4-py
3-none-any.whl.metadata
  Downloading comm-0.1.4-py3-none-any.whl.metadata (4.2 kB)
Collecting widgetsnbextension~=4.0.7 (from ipywidgets>=7.6.5->pycaret)
  Obtaining dependency information for widgetsnbextension~=4.0.7 from https://files.pyth
onhosted.org/packages/8e/d4/d31b12ac0b87e8cc9fdb6ea1eb6596de405eaaa2f25606aaa755d0eebbc
0/widgetsnbextension-4.0.8-py3-none-any.whl.metadata
  Downloading widgetsnbextension-4.0.8-py3-none-any.whl.metadata (1.6 kB)
Collecting jupyterlab-widgets~=3.0.7 (from ipywidgets>=7.6.5->pycaret)
  Obtaining dependency information for jupyterlab-widgets~=3.0.7 from https://files.pyth
onhosted.org/packages/74/5e/2475ac62faf2e342b2bf20b8d8e375f49400ecb38f52e4e0a7557eb1ced
b/jupyterlab_widgets-3.0.8-py3-none-any.whl.metadata
  Downloading jupyterlab_widgets-3.0.8-py3-none-any.whl.metadata (4.1 kB)
Requirement already satisfied: cycler>=0.10 in e:\anaconda3\lib\site-packages (from matp
lotlib>=3.3.0->pycaret) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in e:\anaconda3\lib\site-packages (from
matplotlib>=3.3.0->pycaret) (4.25.0)
Requirement already satisfied: kiwisolver>=1.0.1 in e:\anaconda3\lib\site-packages (from
matplotlib>=3.3.0->pycaret) (1.3.2)
Requirement already satisfied: pillow>=6.2.0 in e:\anaconda3\lib\site-packages (from mat
plotlib>=3.3.0->pycaret) (9.4.0)
Requirement already satisfied: pyparsing>=2.2.1 in e:\anaconda3\lib\site-packages (from
matplotlib>=3.3.0->pycaret) (3.0.4)
Requirement already satisfied: python-dateutil>=2.7 in e:\anaconda3\lib\site-packages (f
rom matplotlib>=3.3.0->pycaret) (2.8.2)
Requirement already satisfied: fastjsonschema in e:\anaconda3\lib\site-packages (from nb
format>=4.2.0->pycaret) (2.16.2)
Requirement already satisfied: jsonschema>=2.6 in e:\anaconda3\lib\site-packages (from n
bformat>=4.2.0->pycaret) (4.17.3)
Requirement already satisfied: jupyter-core in e:\anaconda3\lib\site-packages (from nbfo
rmat>=4.2.0->pycaret) (5.3.0)
Requirement already satisfied: llvmlite<0.39,>=0.38.0rc1 in e:\anaconda3\lib\site-packag
es (from numba>=0.55.0->pycaret) (0.38.0)
Requirement already satisfied: setuptools in e:\anaconda3\lib\site-packages (from numba>
=0.55.0->pycaret) (68.0.0)
Requirement already satisfied: pytz>=2020.1 in e:\anaconda3\lib\site-packages (from pand
as<2.0.0,>=1.3.0->pycaret) (2022.7)
Requirement already satisfied: tenacity>=6.2.0 in e:\anaconda3\lib\site-packages (from p
lotly>=5.0.0->pycaret) (8.0.1)
Requirement already satisfied: six in e:\anaconda3\lib\site-packages (from plotly>=5.0.0
->pycaret) (1.16.0)
Collecting dash<3.0.0,>=2.11.0 (from plotly-resampler>=0.8.3.1->pycaret)
  Obtaining dependency information for dash<3.0.0,>=2.11.0 from https://files.pythonhost
ed.org/packages/9b/b4/d522c16b41a8da013fd60a67f9618e57c504cd2c80e02a7a861413b93906/dash-
2.13.0-py3-none-any.whl.metadata
  Downloading dash-2.13.0-py3-none-any.whl.metadata (11 kB)
Collecting orjson<4.0.0,>=3.8.0 (from plotly-resampler>=0.8.3.1->pycaret)
```

```
  Obtaining dependency information for orjson<4.0.0,>=3.8.0 from https://files.pythonhos
ted.org/packages/bd/f3/38f25db1ddef6490c18ca6e2229097088947492f2a9e05b3047979df68c8/orjs
on-3.9.5-cp39-none-win_amd64.whl.metadata
  Downloading orjson-3.9.5-cp39-none-win_amd64.whl.metadata (50 kB)
     ---------------------------------------- 50.4/50.4 kB 2.5 MB/s eta 0:00:00
Collecting trace-updater>=0.0.8 (from plotly-resampler>=0.8.3.1->pycaret)
  Downloading trace_updater-0.0.9.1-py3-none-any.whl (185 kB)
     ---------------------------------------- 185.2/185.2 kB 1.9 MB/s eta 0:00:00
Collecting tsdownsample==0.1.2 (from plotly-resampler>=0.8.3.1->pycaret)
  Downloading tsdownsample-0.1.2-cp39-none-win_amd64.whl (1.0 MB)
     ---------------------------------------- 1.0/1.0 MB 1.1 MB/s eta 0:00:00
Requirement already satisfied: Cython!=0.29.18,!=0.29.31,>=0.29 in e:\anaconda3\lib\site
-packages (from pmdarima!=1.8.1,<3.0.0,>=1.8.0->pycaret) (0.29.28)
Requirement already satisfied: urllib3 in e:\anaconda3\lib\site-packages (from pmdarima!
=1.8.1,<3.0.0,>=1.8.0->pycaret) (1.26.16)
Requirement already satisfied: charset-normalizer<4,>=2 in e:\anaconda3\lib\site-package
s (from requests>=2.27.1->pycaret) (2.0.4)
Requirement already satisfied: idna<4,>=2.5 in e:\anaconda3\lib\site-packages (from requ
ests>=2.27.1->pycaret) (2.10)
Requirement already satisfied: certifi>=2017.4.17 in e:\anaconda3\lib\site-packages (fro
m requests>=2.27.1->pycaret) (2023.7.22)
Collecting scikit-base<0.6.0 (from sktime!=0.17.1,!=0.17.2,!=0.18.0,>=0.16.1->pycaret)
  Obtaining dependency information for scikit-base<0.6.0 from https://files.pythonhoste
d.org/packages/16/fe/57ecb5cdc28995dacebe785381341a8265acb4c32ccb07327dbfbddd5b4a/scikit
_base-0.5.1-py3-none-any.whl.metadata
  Downloading scikit_base-0.5.1-py3-none-any.whl.metadata (8.6 kB)
Requirement already satisfied: Flask<2.3.0,>=1.0.4 in e:\anaconda3\lib\site-packages (fr
om dash<3.0.0,>=2.11.0->plotly-resampler>=0.8.3.1->pycaret) (1.1.2)
Requirement already satisfied: Werkzeug<2.3.0 in e:\anaconda3\lib\site-packages (from da
sh<3.0.0,>=2.11.0->plotly-resampler>=0.8.3.1->pycaret) (2.0.3)
Collecting dash-html-components==2.0.0 (from dash<3.0.0,>=2.11.0->plotly-resampler>=0.8.
3.1->pycaret)
  Downloading dash_html_components-2.0.0-py3-none-any.whl (4.1 kB)
Collecting dash-core-components==2.0.0 (from dash<3.0.0,>=2.11.0->plotly-resampler>=0.8.
3.1->pycaret)
  Downloading dash_core_components-2.0.0-py3-none-any.whl (3.8 kB)
Collecting dash-table==5.0.0 (from dash<3.0.0,>=2.11.0->plotly-resampler>=0.8.3.1->pycar
et)
  Downloading dash_table-5.0.0-py3-none-any.whl (3.9 kB)
Requirement already satisfied: typing-extensions>=4.1.1 in e:\anaconda3\lib\site-package
s (from dash<3.0.0,>=2.11.0->plotly-resampler>=0.8.3.1->pycaret) (4.7.1)
Collecting retrying (from dash<3.0.0,>=2.11.0->plotly-resampler>=0.8.3.1->pycaret)
  Downloading retrying-1.3.4-py3-none-any.whl (11 kB)
Collecting ansi2html (from dash<3.0.0,>=2.11.0->plotly-resampler>=0.8.3.1->pycaret)
  Downloading ansi2html-1.8.0-py3-none-any.whl (16 kB)
Requirement already satisfied: nest-asyncio in e:\anaconda3\lib\site-packages (from dash
<3.0.0,>=2.11.0->plotly-resampler>=0.8.3.1->pycaret) (1.5.6)
Requirement already satisfied: parso<0.9.0,>=0.8.0 in e:\anaconda3\lib\site-packages (fr
om jedi>=0.16->ipython>=5.5.0->pycaret) (0.8.3)
Requirement already satisfied: attrs>=17.4.0 in e:\anaconda3\lib\site-packages (from jso
nschema>=2.6->nbformat>=4.2.0->pycaret) (22.1.0)
Requirement already satisfied: pyrsistent!=0.17.0,!=0.17.1,!=0.17.2,>=0.14.0 in e:\anaco
nda3\lib\site-packages (from jsonschema>=2.6->nbformat>=4.2.0->pycaret) (0.18.0)
Requirement already satisfied: wcwidth in e:\anaconda3\lib\site-packages (from prompt-to
olkit<3.1.0,>=3.0.30->ipython>=5.5.0->pycaret) (0.2.5)
Requirement already satisfied: platformdirs>=2.5 in e:\anaconda3\lib\site-packages (from
jupyter-core->nbformat>=4.2.0->pycaret) (3.10.0)
Requirement already satisfied: pywin32>=300 in e:\anaconda3\lib\site-packages (from jupy
ter-core->nbformat>=4.2.0->pycaret) (305.1)
Requirement already satisfied: executing in e:\anaconda3\lib\site-packages (from stack-d
ata->ipython>=5.5.0->pycaret) (0.8.3)
Requirement already satisfied: asttokens in e:\anaconda3\lib\site-packages (from stack-d
ata->ipython>=5.5.0->pycaret) (2.0.5)
Requirement already satisfied: pure-eval in e:\anaconda3\lib\site-packages (from stack-d
ata->ipython>=5.5.0->pycaret) (0.2.2)
Requirement already satisfied: itsdangerous>=0.24 in e:\anaconda3\lib\site-packages (fro
```

```
m Flask<2.3.0,>=1.0.4->dash<3.0.0,>=2.11.0->plotly-resampler>=0.8.3.1->pycaret) (2.0.1)
Requirement already satisfied: click>=5.1 in e:\anaconda3\lib\site-packages (from Flask<
2.3.0,>=1.0.4->dash<3.0.0,>=2.11.0->plotly-resampler>=0.8.3.1->pycaret) (8.0.4)
Downloading pycaret-3.0.4-py3-none-any.whl (484 kB)
   ---------------------------------------- 484.4/484.4 kB 843.6 kB/s eta 0:00:00
Downloading category_encoders-2.6.2-py2.py3-none-any.whl (81 kB)
   ---------------------------------------- 81.8/81.8 kB 1.2 MB/s eta 0:00:00
Downloading imbalanced_learn-0.11.0-py3-none-any.whl (235 kB)
   ---------------------------------------- 235.6/235.6 kB 1.2 MB/s eta 0:00:00
Downloading ipywidgets-8.1.0-py3-none-any.whl (139 kB)
   ---------------------------------------- 139.3/139.3 kB 921.9 kB/s eta 0:00:00
```

Import Modules

In [3]:
```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import warnings
from pycaret.classification import *
%matplotlib inline
warnings.filterwarnings('ignore')
```

In [4]:
```
# Load Dataset

df = pd.read_csv(r"E:\PORTFOLIO\1. DATA ANALYTICS\5. BREAST CANCER DETECTION\DATA\data.c
```

In [5]:
```
df.head()
```

Out[5]:

	id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_mean	compactness
0	842302	M	17.99	10.38	122.80	1001.0	0.11840	0
1	842517	M	20.57	17.77	132.90	1326.0	0.08474	0
2	84300903	M	19.69	21.25	130.00	1203.0	0.10960	0
3	84348301	M	11.42	20.38	77.58	386.1	0.14250	0
4	84358402	M	20.29	14.34	135.10	1297.0	0.10030	0

5 rows × 33 columns

In [7]:
```
# delete unnecessary columns

df = df.drop(columns=['id','Unnamed: 32'], axis=1)
```

In [8]:
```
# Statistical Information

df.describe()
```

Out[8]:

	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_mean	compactness_mean	concavity
count	569.000000	569.000000	569.000000	569.000000	569.000000	569.000000	569
mean	14.127292	19.289649	91.969033	654.889104	0.096360	0.104341	0
std	3.524049	4.301036	24.298981	351.914129	0.014064	0.052813	0
min	6.981000	9.710000	43.790000	143.500000	0.052630	0.019380	0

25%	11.700000	16.170000	75.170000	420.300000	0.086370	0.064920	0
50%	13.370000	18.840000	86.240000	551.100000	0.095870	0.092630	0
75%	15.780000	21.800000	104.100000	782.700000	0.105300	0.130400	0
max	28.110000	39.280000	188.500000	2501.000000	0.163400	0.345400	0

8 rows × 30 columns

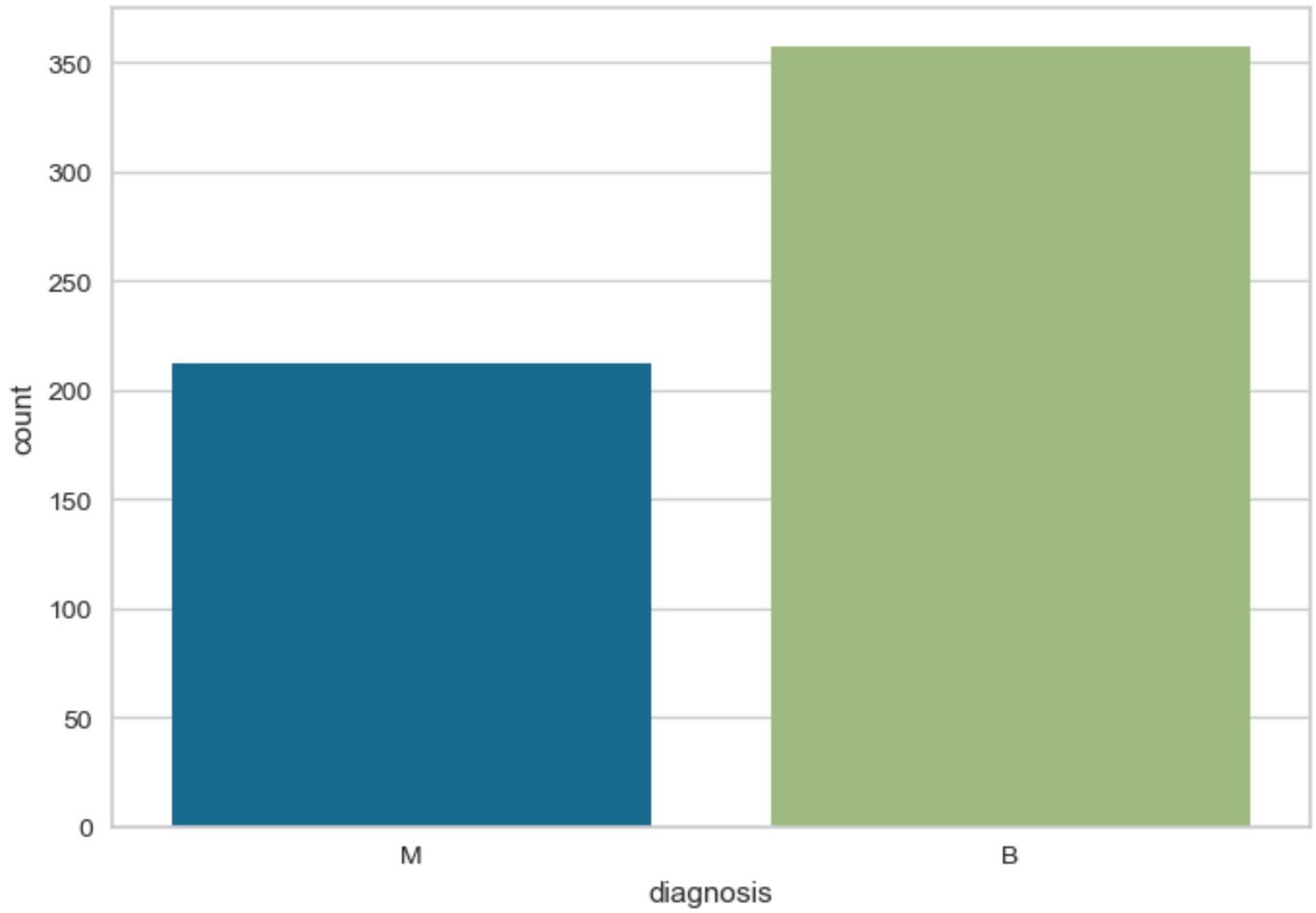
In [9]:
```
# datatype info

df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 569 entries, 0 to 568
Data columns (total 31 columns):
 #   Column                   Non-Null Count  Dtype  
---  ------                   --------------  -----  
 0   diagnosis                569 non-null    object 
 1   radius_mean              569 non-null    float64
 2   texture_mean             569 non-null    float64
 3   perimeter_mean           569 non-null    float64
 4   area_mean                569 non-null    float64
 5   smoothness_mean          569 non-null    float64
 6   compactness_mean         569 non-null    float64
 7   concavity_mean           569 non-null    float64
 8   concave points_mean      569 non-null    float64
 9   symmetry_mean            569 non-null    float64
 10  fractal_dimension_mean   569 non-null    float64
 11  radius_se                569 non-null    float64
 12  texture_se               569 non-null    float64
 13  perimeter_se             569 non-null    float64
 14  area_se                  569 non-null    float64
 15  smoothness_se            569 non-null    float64
 16  compactness_se           569 non-null    float64
 17  concavity_se             569 non-null    float64
 18  concave points_se        569 non-null    float64
 19  symmetry_se              569 non-null    float64
 20  fractal_dimension_se     569 non-null    float64
 21  radius_worst             569 non-null    float64
 22  texture_worst            569 non-null    float64
 23  perimeter_worst          569 non-null    float64
 24  area_worst               569 non-null    float64
 25  smoothness_worst         569 non-null    float64
 26  compactness_worst        569 non-null    float64
 27  concavity_worst          569 non-null    float64
 28  concave points_worst     569 non-null    float64
 29  symmetry_worst           569 non-null    float64
 30  fractal_dimension_worst  569 non-null    float64
dtypes: float64(30), object(1)
memory usage: 137.9+ KB
```

Exploratory Data

In [10]:
```
sns.countplot(df['diagnosis'])
```

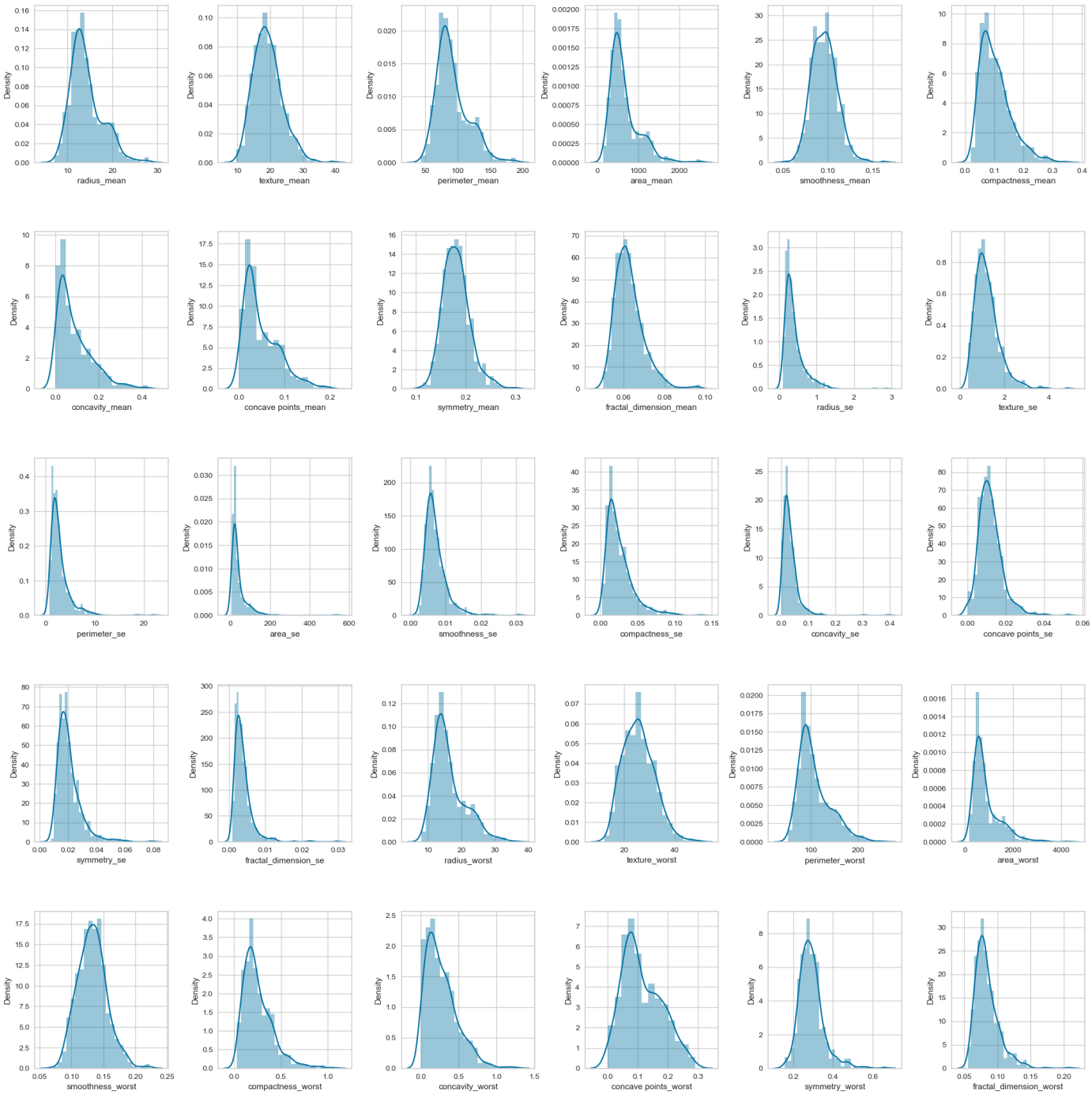
Out[10]:
```
<AxesSubplot:xlabel='diagnosis', ylabel='count'>
```



M = Malignant , B = Benign

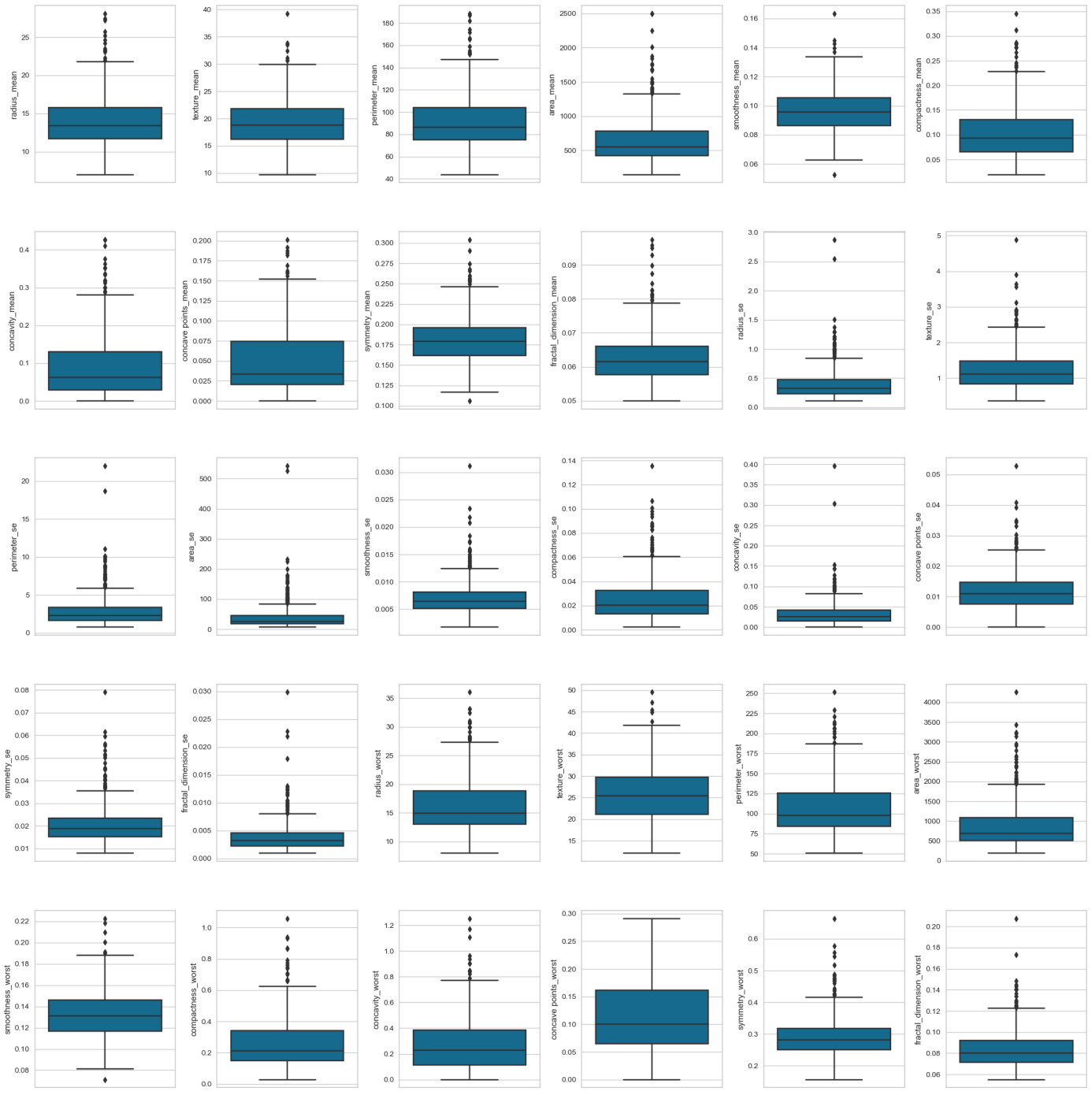
In [11]:
```
df_temp = df.drop(columns=['diagnosis'], axis=1)
```

In [13]:
```
# create dist plot
fig, ax = plt.subplots(ncols=6, nrows=5, figsize=(20, 20))
index = 0
ax = ax.flatten()

for col in df_temp.columns:
    sns.distplot(df[col], ax=ax[index])
    index+=1
plt.tight_layout(pad=0.5, w_pad=0.7, h_pad=5.0)
```



In [14]:

```
# create box plot

fig, ax = plt.subplots(ncols=6, nrows=5, figsize=(20, 20))
index = 0
ax = ax.flatten()

for col in df_temp.columns:
    sns.boxplot(y=col, data=df, ax=ax[index])
    index+=1
plt.tight_layout(pad=0.5, w_pad=0.7, h_pad=5.0)
```



Create and Train the Model

In [22]:
```
# Setup the Data

clf = setup(df, target='diagnosis')
```

	Description	Value
0	Session id	262
1	Target	diagnosis
2	Target type	Binary
3	Target mapping	B: 0, M: 1
4	Original data shape	(569, 31)
5	Transformed data shape	(569, 31)
6	Transformed train set shape	(398, 31)

7	Transformed test set shape	(171, 31)
8	Numeric features	30
9	Preprocess	True
10	Imputation type	simple
11	Numeric imputation	mean
12	Categorical imputation	mode
13	Fold Generator	StratifiedKFold
14	Fold Number	10
15	CPU Jobs	-1
16	Use GPU	False
17	Log Experiment	False
18	Experiment Name	clf-default-name
19	USI	ca23

In [23]:
```
# train and test the models

compare_models()
```

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
ada	Ada Boost Classifier	0.9724	0.9933	0.9529	0.9750	0.9629	0.9410	0.9422	0.3550
gbc	Gradient Boosting Classifier	0.9649	0.9952	0.9529	0.9585	0.9540	0.9257	0.9280	0.5100
lightgbm	Light Gradient Boosting Machine	0.9624	0.9938	0.9595	0.9464	0.9512	0.9208	0.9230	0.3960
et	Extra Trees Classifier	0.9623	0.9953	0.9324	0.9670	0.9477	0.9183	0.9206	0.4650
lda	Linear Discriminant Analysis	0.9574	0.9906	0.8929	0.9923	0.9378	0.9059	0.9110	0.1680
rf	Random Forest Classifier	0.9548	0.9933	0.9390	0.9467	0.9401	0.9039	0.9074	0.4810
qda	Quadratic Discriminant Analysis	0.9547	0.9887	0.9390	0.9440	0.9386	0.9029	0.9063	0.1390
ridge	Ridge Classifier	0.9524	0.0000	0.8857	0.9866	0.9305	0.8948	0.9007	0.1330
lr	Logistic Regression	0.9447	0.9908	0.9195	0.9335	0.9237	0.8806	0.8838	1.4840
dt	Decision Tree Classifier	0.9422	0.9373	0.9186	0.9279	0.9203	0.8751	0.8785	0.1440
nb	Naive Bayes	0.9347	0.9873	0.8857	0.9390	0.9081	0.8579	0.8624	0.1400
knn	K Neighbors Classifier	0.9072	0.9518	0.8314	0.9173	0.8668	0.7964	0.8042	0.1640
svm	SVM - Linear Kernel	0.9021	0.0000	0.7833	0.9513	0.8525	0.7811	0.7946	0.1320
dummy	Dummy Classifier	0.6282	0.5000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1760

Out[23]:
```
AdaBoostClassifier(algorithm='SAMME.R', base_estimator=None, learning_rate=1.0,
                   n_estimators=50, random_state=262)
```

In [28]:
```
# select the best model
model = create_model('ada')
```

Fold	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC

0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	0.9750	1.0000	0.9333	1.0000	0.9655	0.9459	0.9473
2	0.9500	0.9787	0.8667	1.0000	0.9286	0.8904	0.8958
3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	0.9250	0.9787	0.9333	0.8750	0.9032	0.8421	0.8433
5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6	0.9750	1.0000	0.9333	1.0000	0.9655	0.9459	0.9473
7	0.9250	0.9787	0.9333	0.8750	0.9032	0.8421	0.8433
8	0.9744	0.9971	0.9286	1.0000	0.9630	0.9434	0.9449
9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Mean	0.9724	0.9933	0.9529	0.9750	0.9629	0.9410	0.9422
Std	0.0284	0.0096	0.0429	0.0500	0.0372	0.0601	0.0593

In [29]:
```
# hyperparameter tuning

best_model = tune_model(model)
```

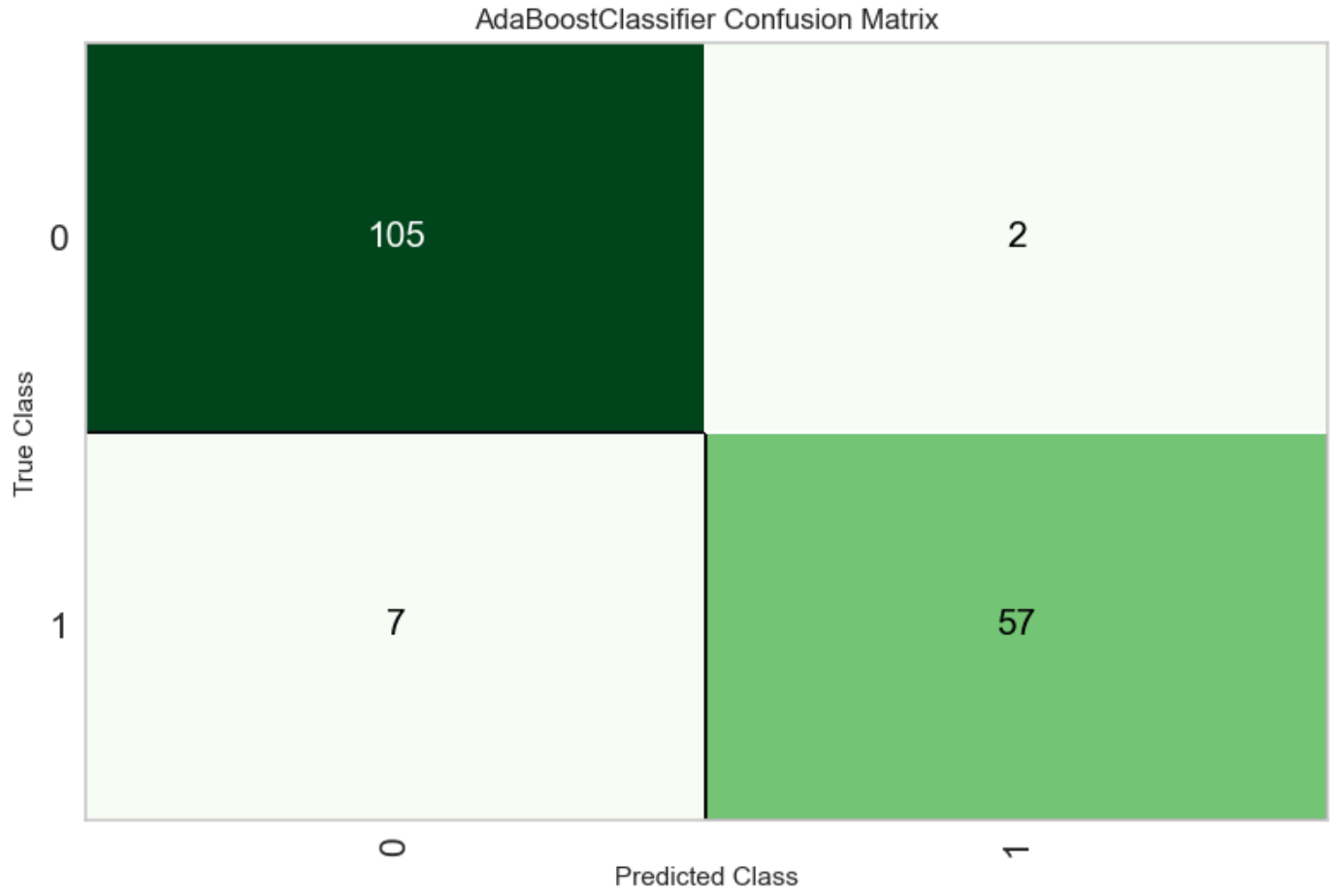
Fold	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0	0.9750	0.9973	1.0000	0.9375	0.9677	0.9474	0.9487
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	0.9500	0.9813	0.8667	1.0000	0.9286	0.8904	0.8958
3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	0.9500	0.9840	0.9333	0.9333	0.9333	0.8933	0.8933
5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6	0.9750	0.9973	1.0000	0.9375	0.9677	0.9474	0.9487
7	0.9500	0.9893	0.9333	0.9333	0.9333	0.8933	0.8933
8	0.9487	0.9943	0.9286	0.9286	0.9286	0.8886	0.8886
9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Mean	0.9749	0.9944	0.9662	0.9670	0.9659	0.9460	0.9468
Std	0.0225	0.0067	0.0451	0.0331	0.0309	0.0486	0.0480

```
Fitting 10 folds for each of 10 candidates, totalling 100 fits
```

In [30]:
```
evaluate_model(best_model)
```


```
interactive(children=(ToggleButtons(description='Plot Type:', icons=('',), options=(('Pi
peline Plot', 'pipelin…
```

In [31]:
```
# plot the result

plot_model(estimator=best_model, plot='confusion_matrix')
```


AdaBoostClassifier Confusion Matrix
0
1
True Class
105
2
7
57
0
1
Predicted Class

In [ ]: