

30 Python Libraries **to (Hugely) Boost** **Your Data Science** **Productivity**

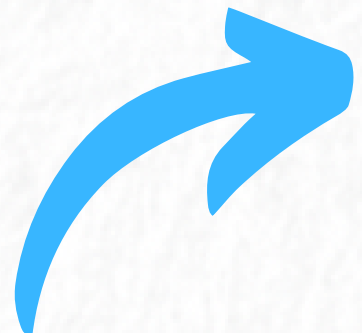
Avi Chawla



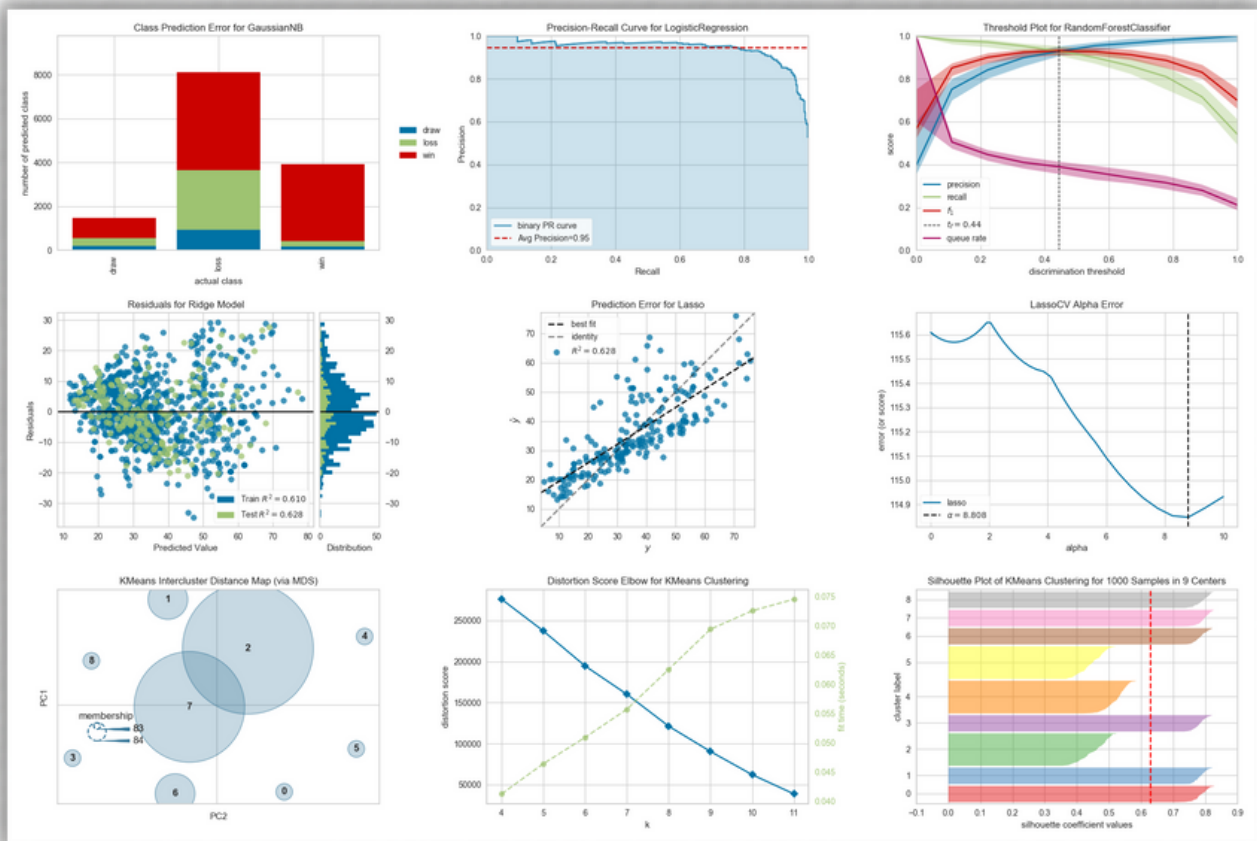
avichawla.substack.com

**Data Science is much
more than Pandas,
NumPy and
Sklearn.**

Here are **30 open-source
libraries** to upgrade
your data game.



1. YellowBrick

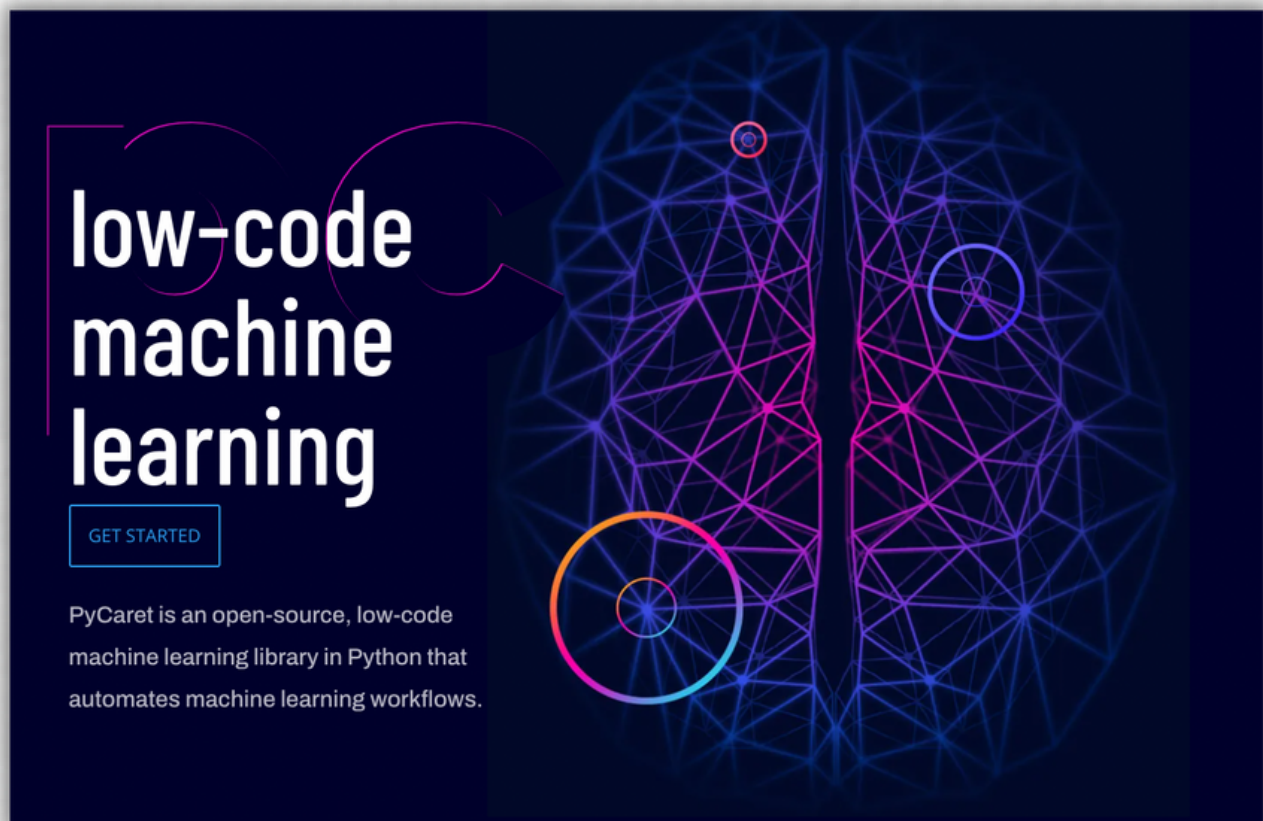


A suite of visualization and diagnostic tools for **faster model selection.**

Matplotlib Sklearn



2. PyCaret



Automate ML workflows with this **low-code library**.



3. imbalanced-learn

```
target.value_counts()
1      548
0      352
```

```
from imblearn.over_sampling import SMOTE

os = SMOTE()

X_new, y_new = os.fit_resample(data, target)

y_new.value_counts()
1      548
0      548
```



A variety of methods to handle **class imbalance**.