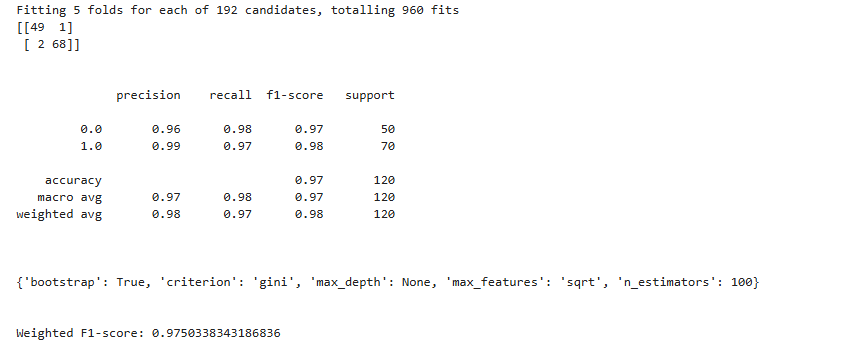
**ASSIGNMENT**

* Identify your **problem statement** - **Predict the Chronic Kidney Disease** **(CKD)** based on the several parameters given
* **Basic info about the dataset** – Total number of **rows is 399**, – Total number of **columns is 25**
* **Pre-Processing Method –** Added **dummies to avoid nominal values** and also done some **standardization with sklearn preprocessing library**

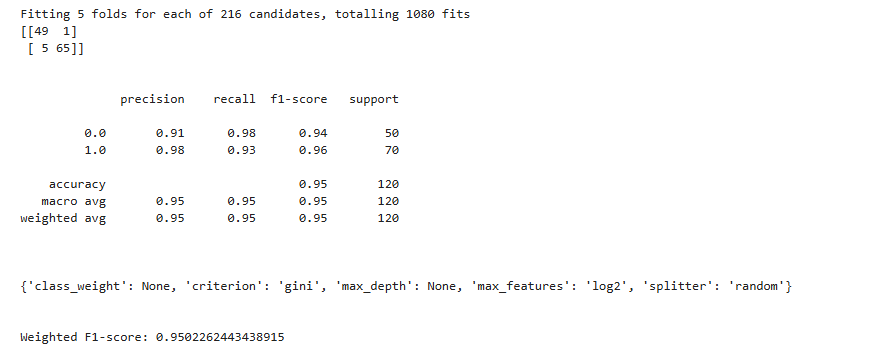
Models are created with many algorithms with many **hyper tuning parameters**  
With **RandomForestClassifier**, I got following result and **Weighted F1-score is 0.975** with hyper tuning parameters as

**bootstrap is True, criterion is gini, max\_depth is None, max\_features is sqrt, n\_estimators is 100**

****

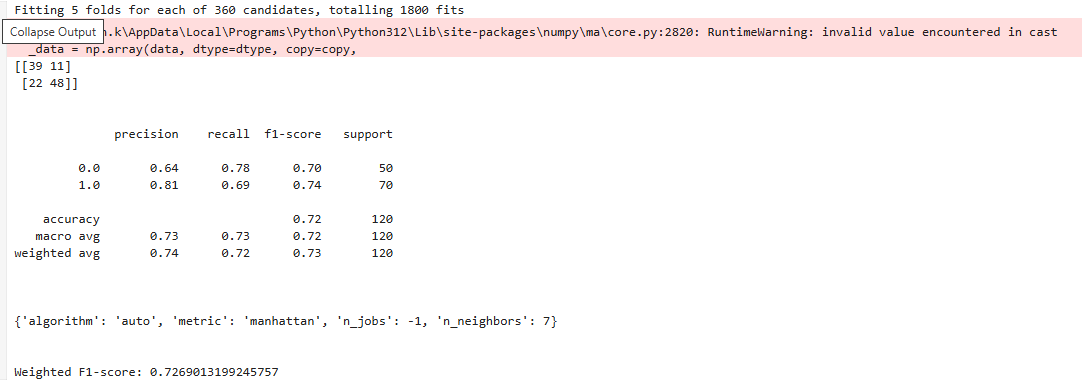
With **DecisionTreeClassifier**, I got following result and **Weighted F1-score is 0.95** with hyper tuning parameters as

**class\_weight is None, criterion is gini,max\_depth is None, max\_features is log2, splitter is random**



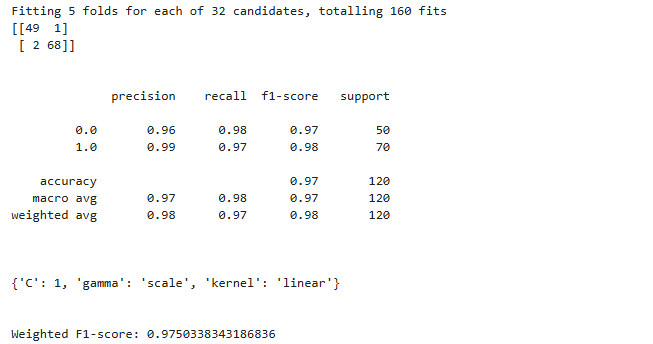
With **KNeighborsClassifier**, I got following result and **Weighted F1-score is 0.73** with hyper tuning parameters as

**Algorithm is auto, metric is manhattan, n\_jobs is -1, n\_neighbors is 7**



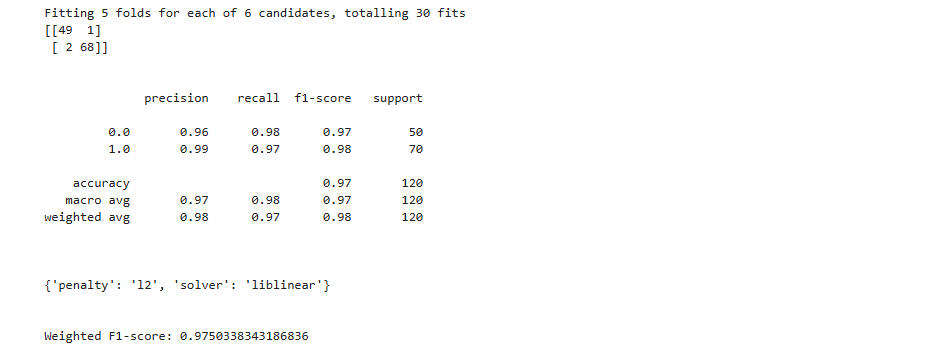
With **Support Vector Machine**, I got following result and **Weighted F1-score is 0.975** with hyper tuning parameters as

**C is 1, gamma is scale,kernel is linear**



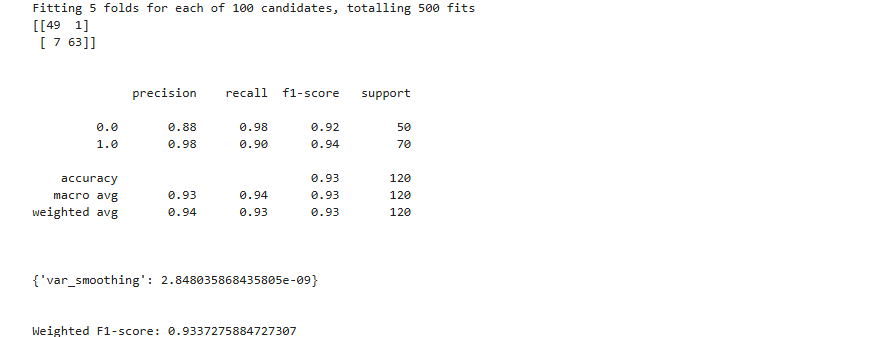
With **Logistic Regression**, I got following result and **Weighted F1-score is 0.975** with hyper tuning parameters as

**Penalty is l2, solver is liblinear**



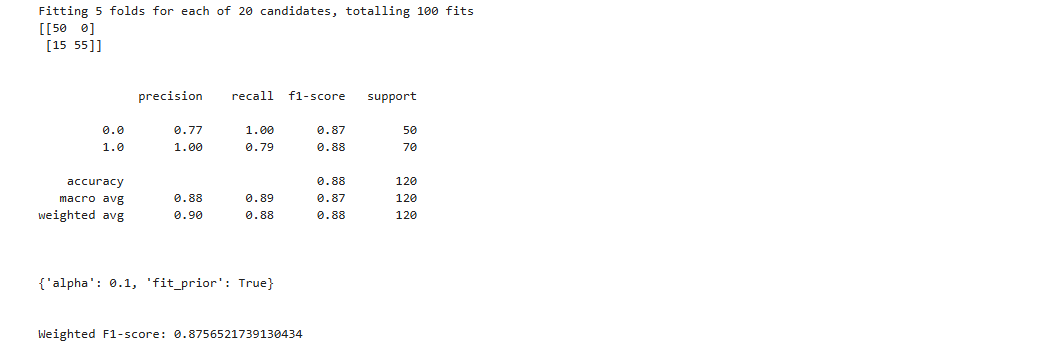
With **Gaussian Naive Bayes**, I got following result and **Weighted F1-score is 0.93** with hyper tuning parameters as

**var\_smoothing is 2.848035868435805e-09**



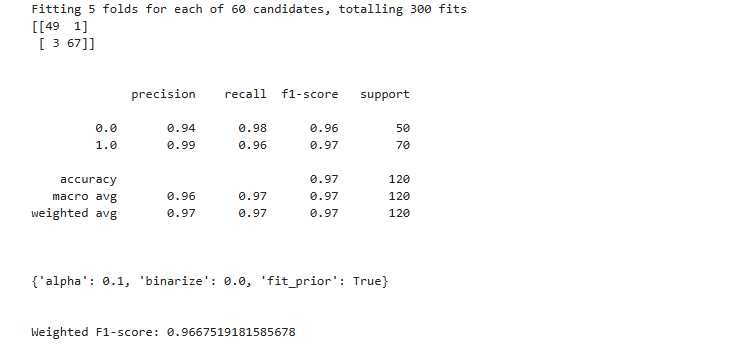
With **Multinomial Naive Bayes**, I got following result and **Weighted F1-score is 0.93** with hyper tuning parameters as

**Alpha is 0.1, fit\_prior is True**



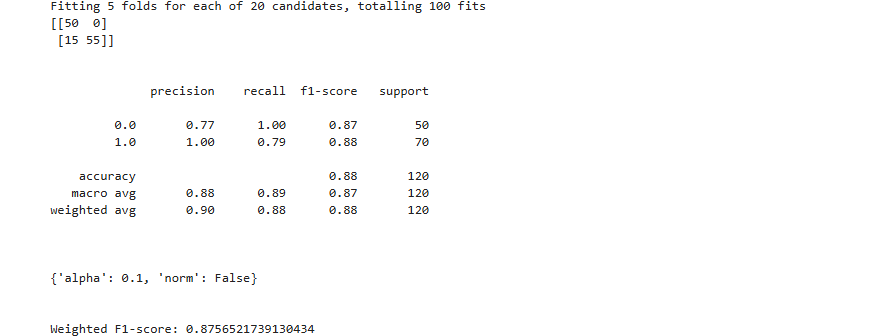
With **Bernoulli Naive Bayes**, I got following result and **Weighted F1-score is 0.97** with hyper tuning parameters as

**Alpha is 0.1, fit\_prior is True, binarize is 0.0**



With **Complement Naive Bayes**, I got following result and **Weighted F1-score is 0.875** with hyper tuning parameters as

**Alpha is 0.1, fit\_ norm is False**



According to this Random Forest Classifier have most Weighted F1 score with most fits

So we can choose the model with **Random Forest Classifier** with hyper parameters as **bootstrap is True, criterion is gini, max\_depth is None, max\_features is sqrt, n\_estimators is 100**