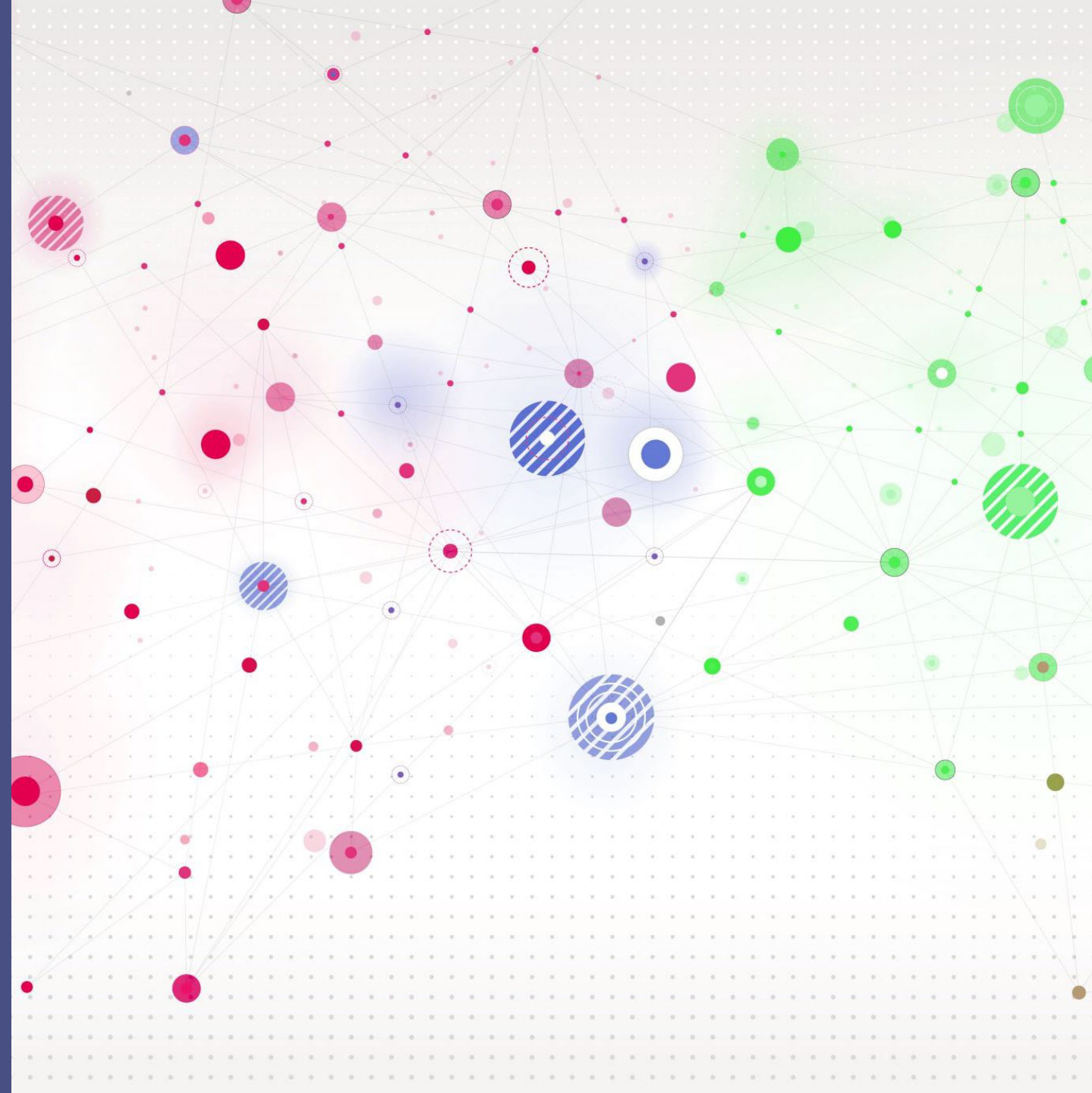


# K-NEAREST NEIGHBORS

*with scikit learn*  
*KNeighborsClassifier*



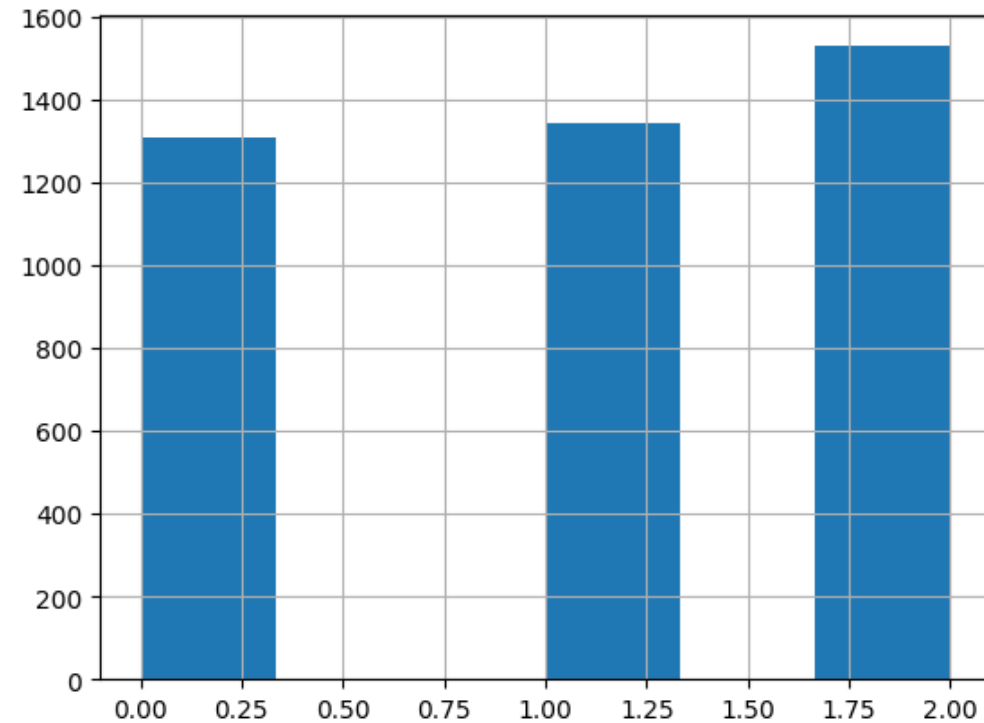
# DATA

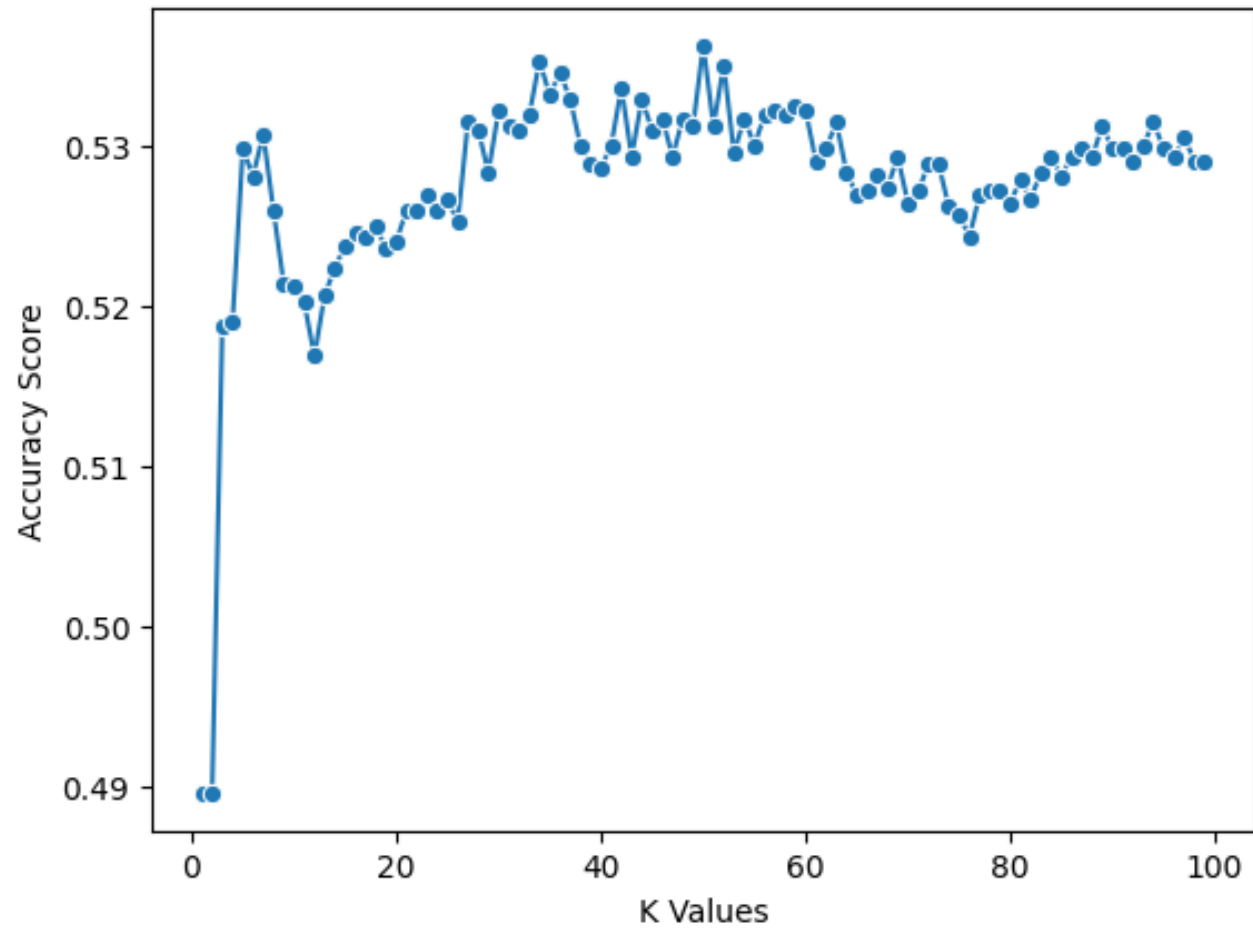
- 4177 samples
- 9 feature columns
- Task 1 : Sex
- Task 2 : Rings
- Splitted Data
- 80% train & 20% test

# TASK 1 : SEX

- Classes : F, I and M
- Distribution
- Correlation

<b>Length</b>	-0.036066
<b>Diameter</b>	-0.038874
<b>Height</b>	-0.042077
<b>Whole weight</b>	-0.021391
<b>Shucked weight</b>	-0.001373
<b>Viscera weight</b>	-0.032067
<b>Shell weight</b>	-0.034854
<b>Rings</b>	-0.034627

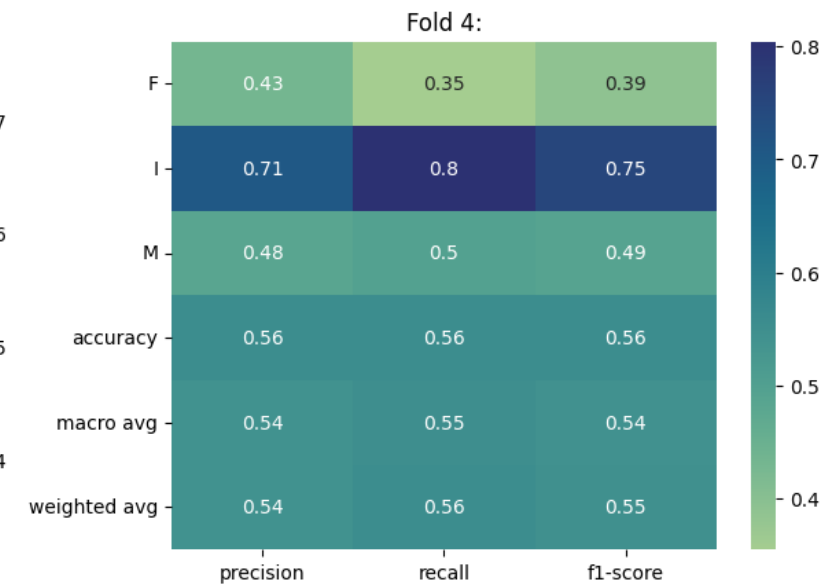
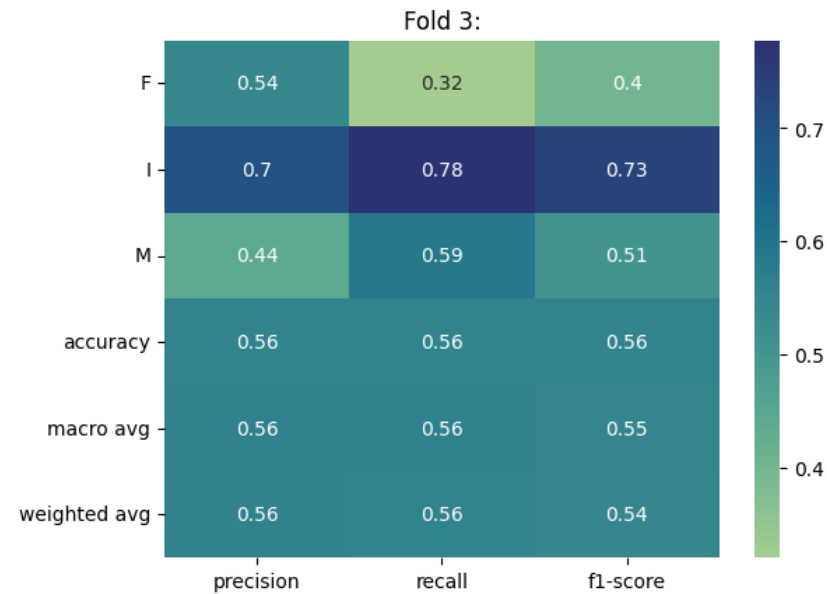
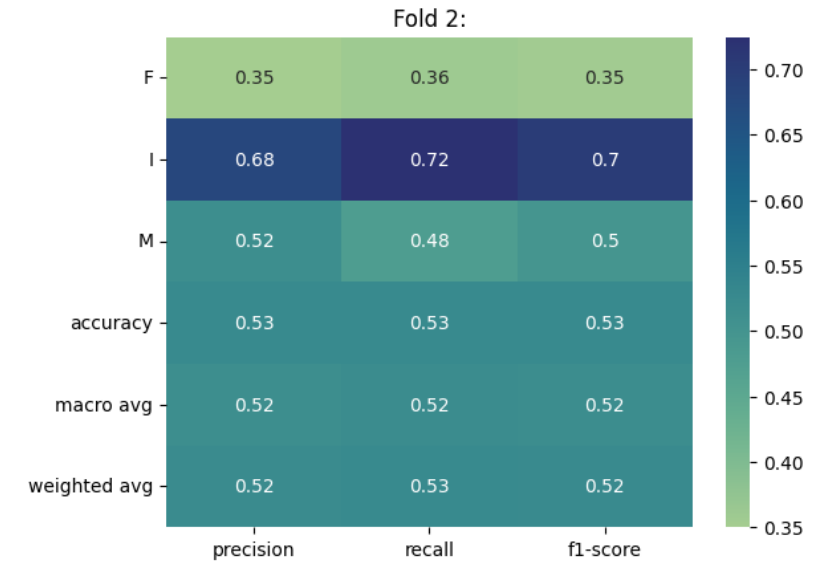
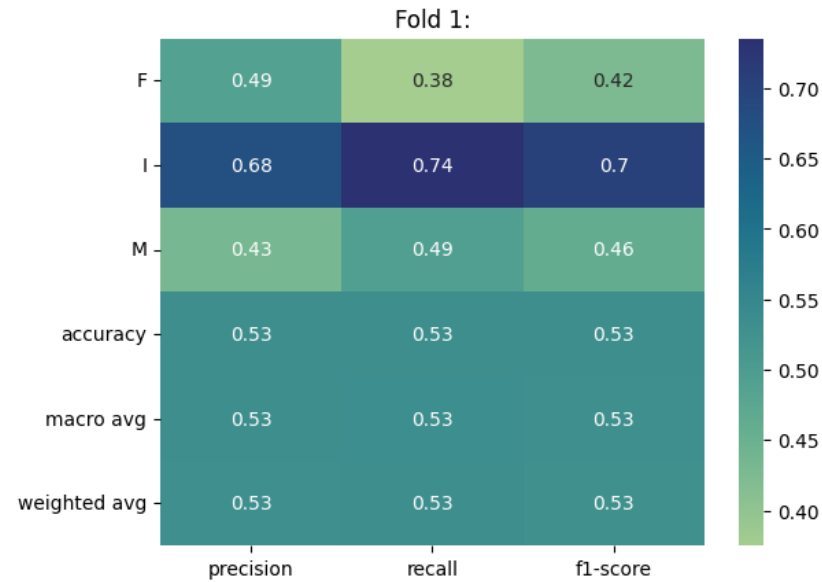




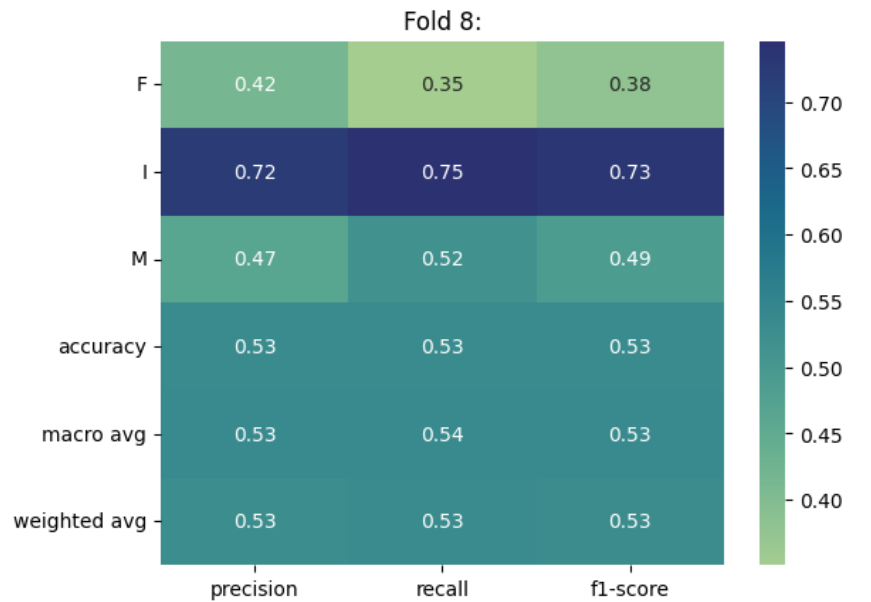
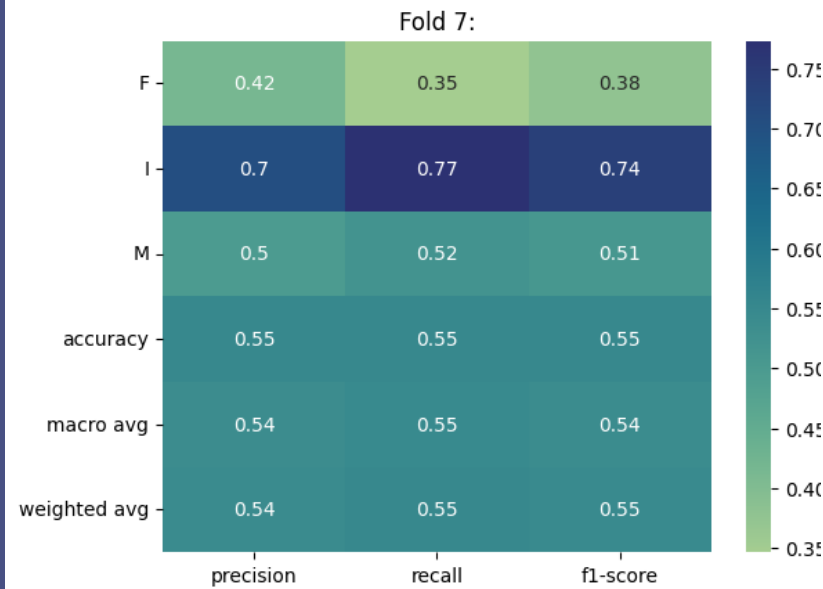
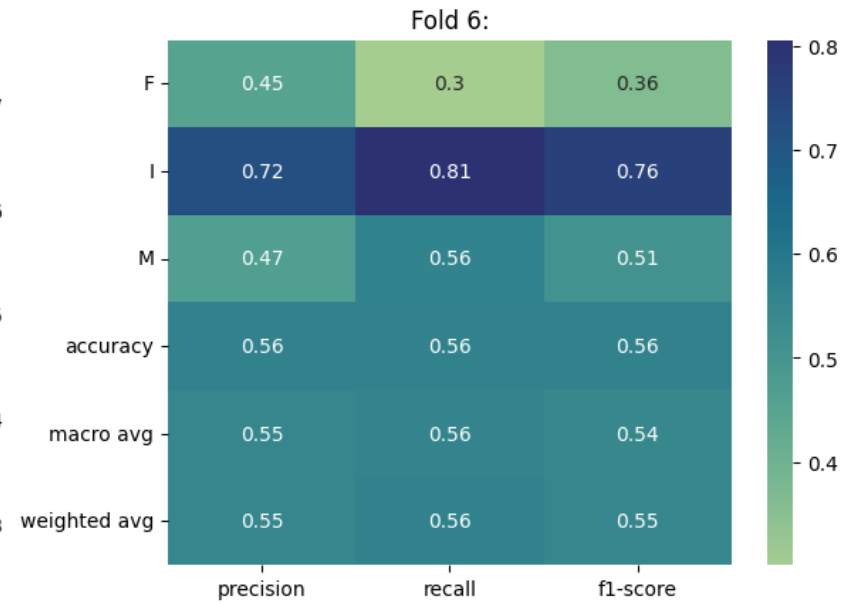
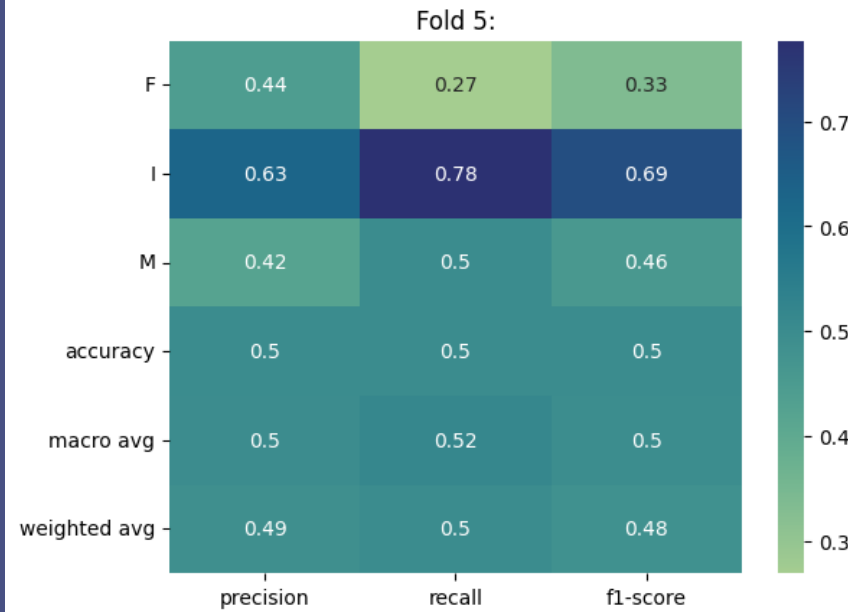
## BEST K

- Cross validation (10-fold)
- K = 50

# 10-FOLD CROSS VALIDATION

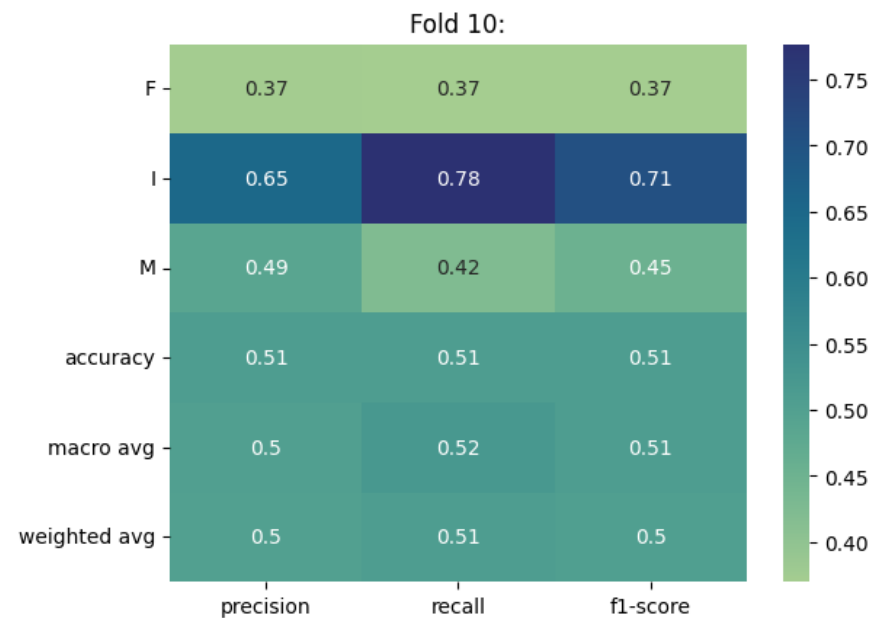
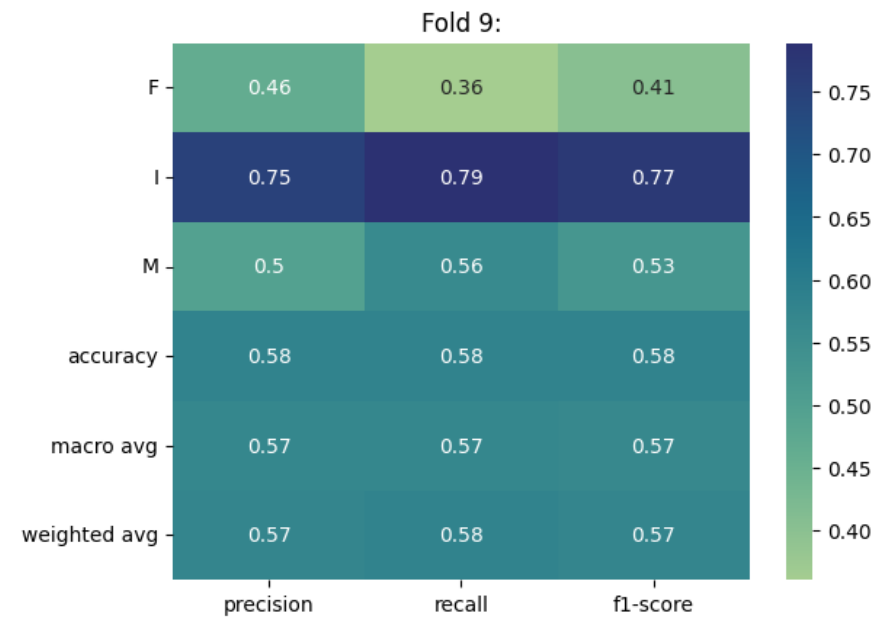


# 10-FOLD CROSS VALIDATION

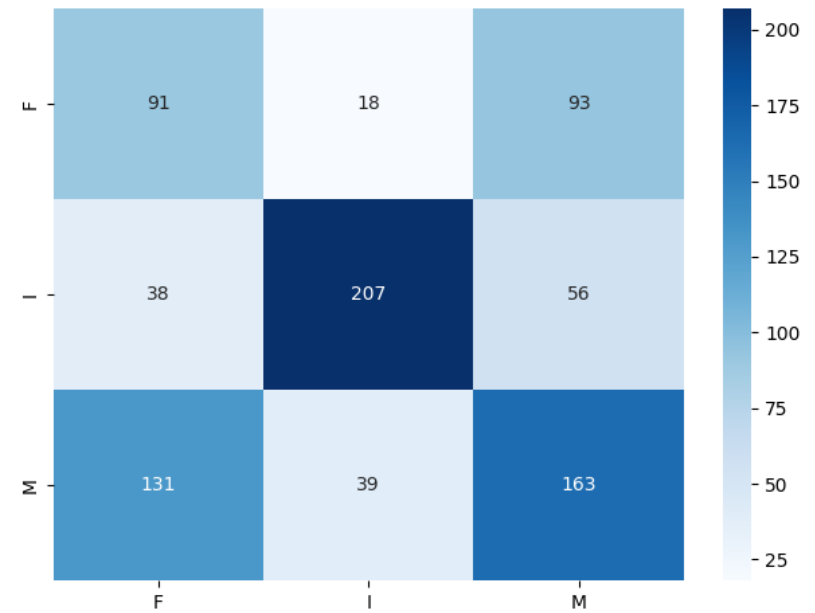
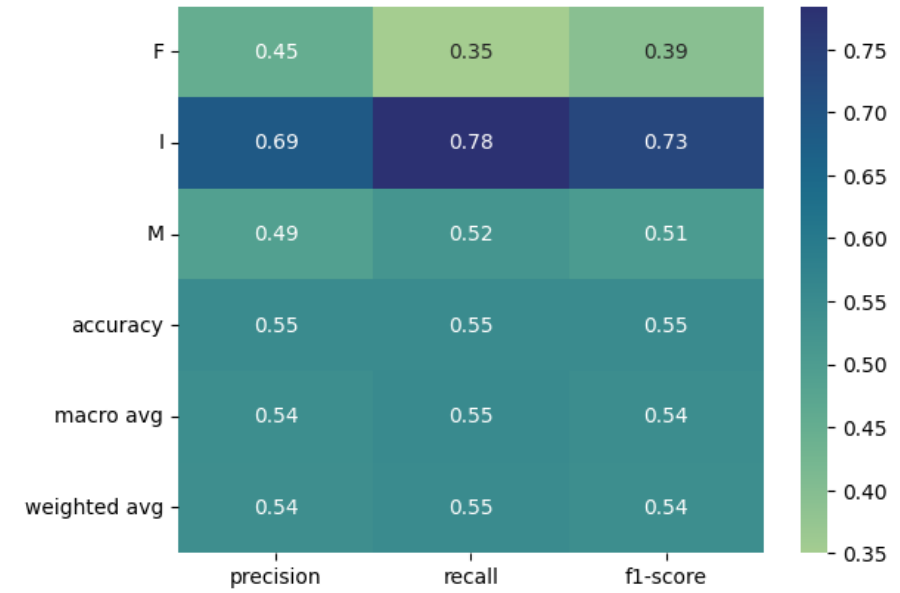


# 10-FOLD CROSS VALIDATION

- average accuracy = 0.5408
- average precision = 0.5339
- average recall = 0.5408
- average f1-score = 0.5326



# MODEL EVALUATION

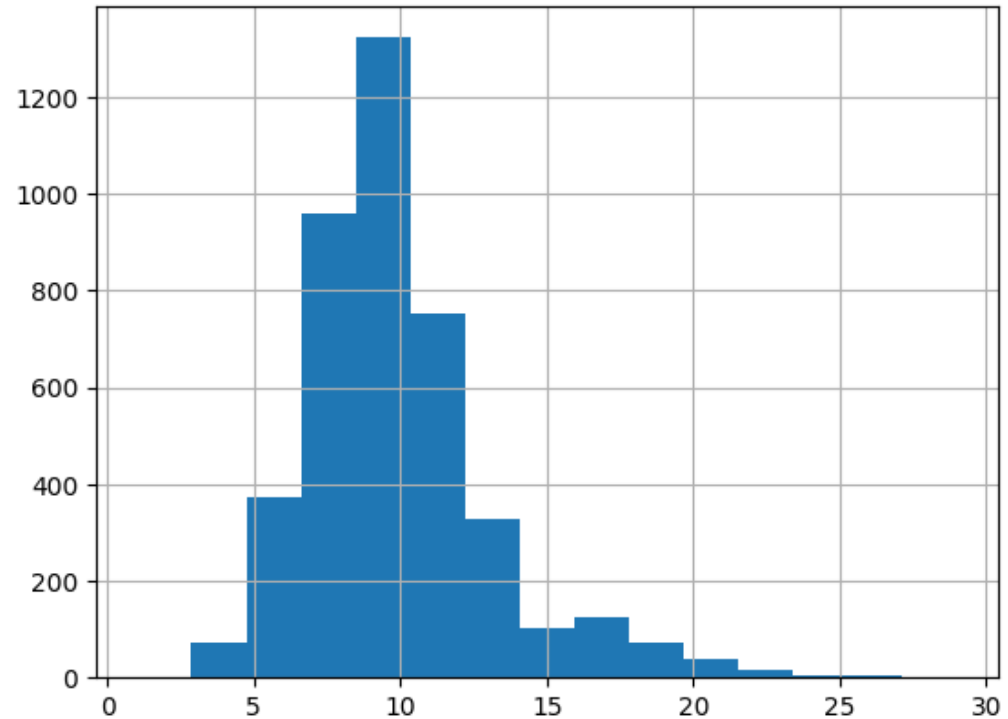


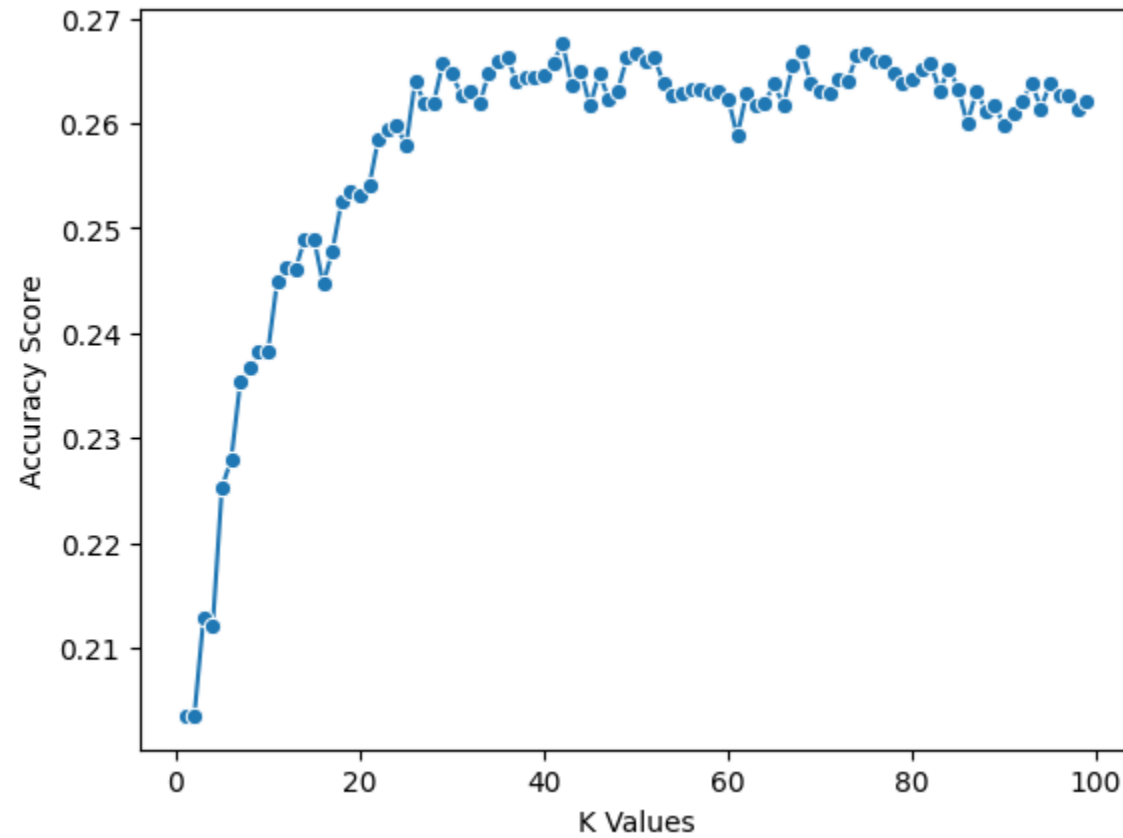


## TASK 2 : RINGS

- Classes : 1-29
- Distribution
- Correlation

<b>Sex</b>	-0.034627
<b>Length</b>	0.556720
<b>Diameter</b>	0.574660
<b>Height</b>	0.557467
<b>Whole weight</b>	0.540390
<b>Shucked weight</b>	0.420884
<b>Viscera weight</b>	0.503819
<b>Shell weight</b>	0.627574

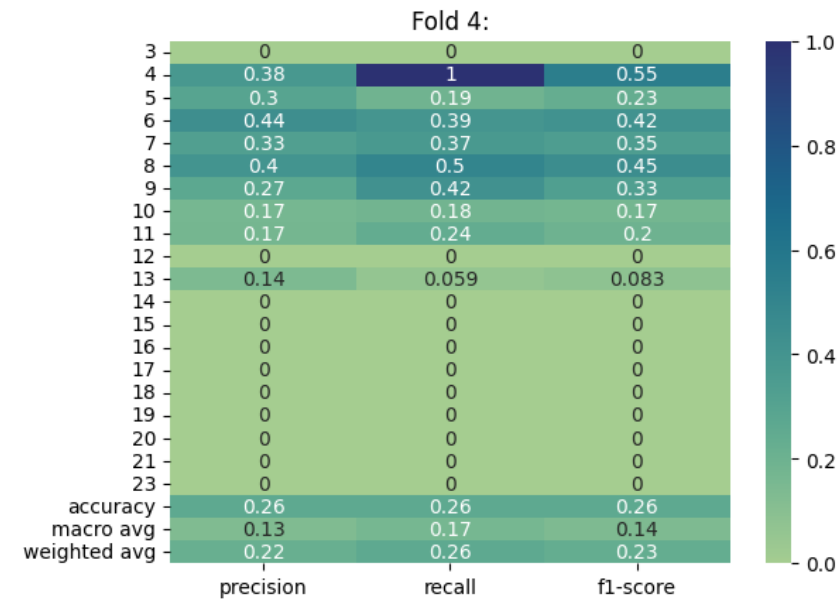
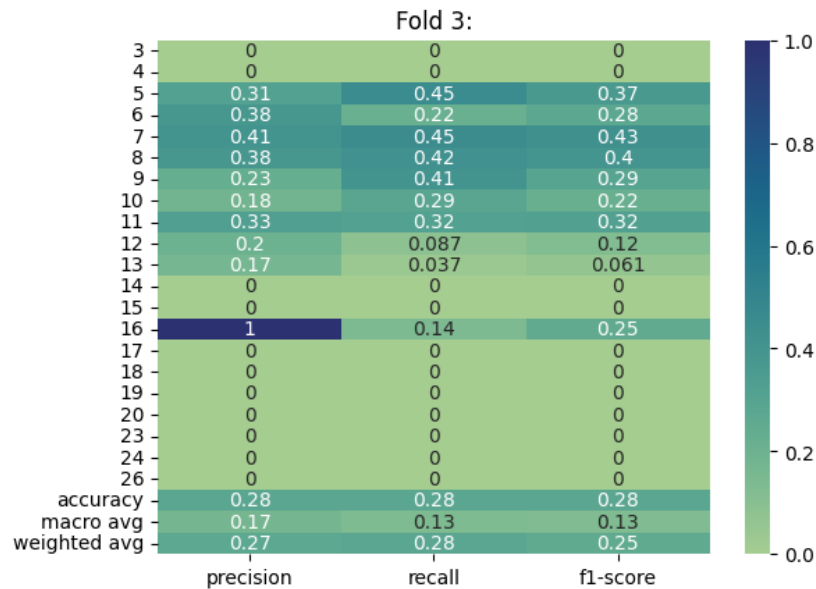
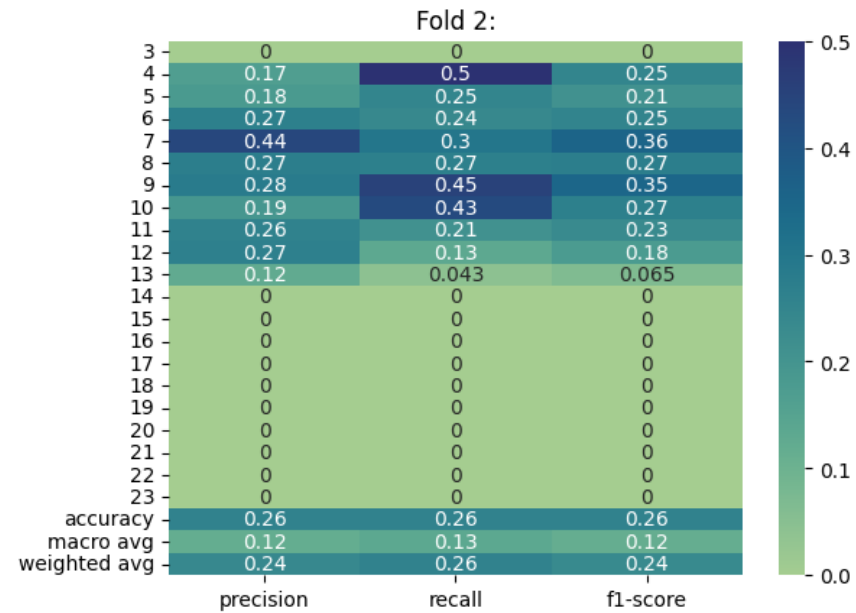
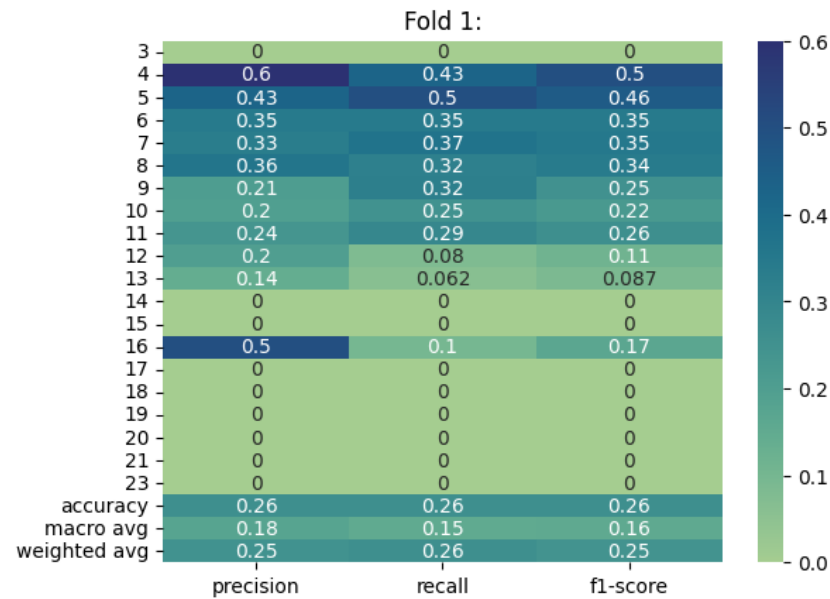




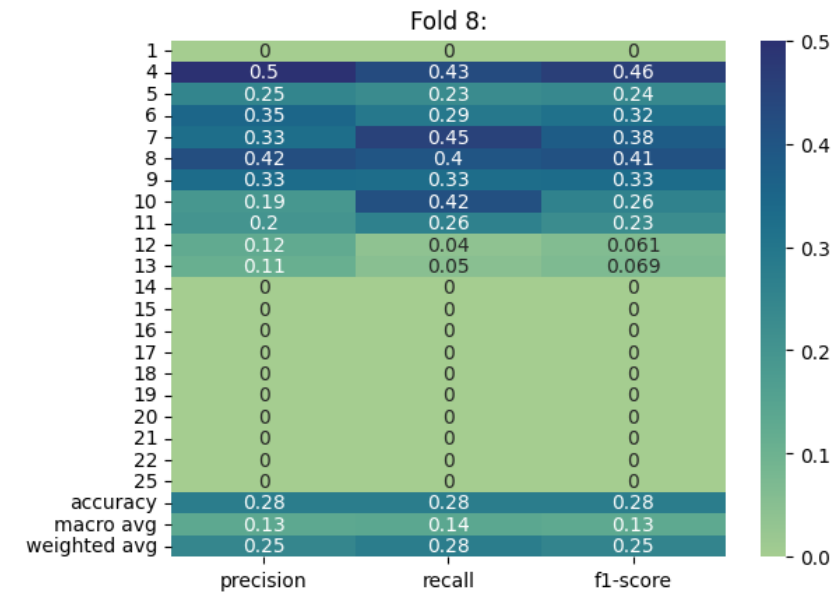
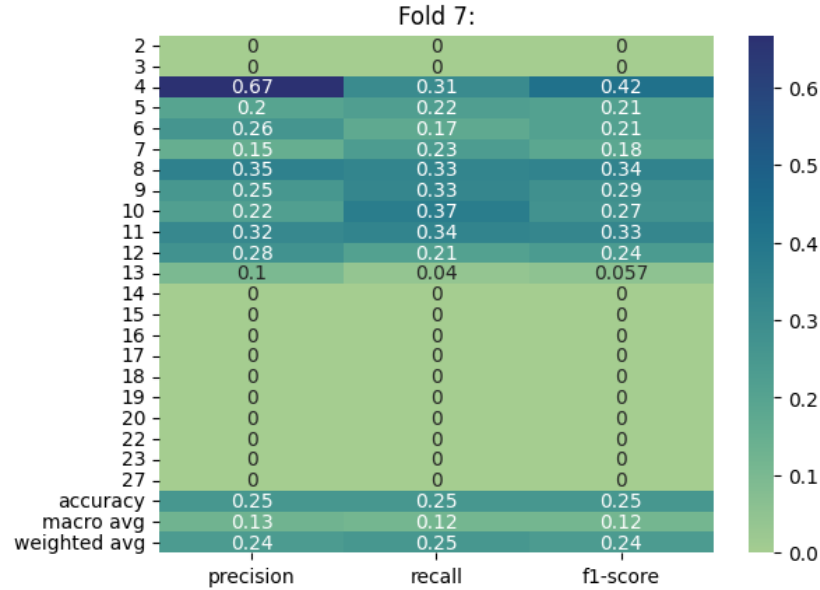
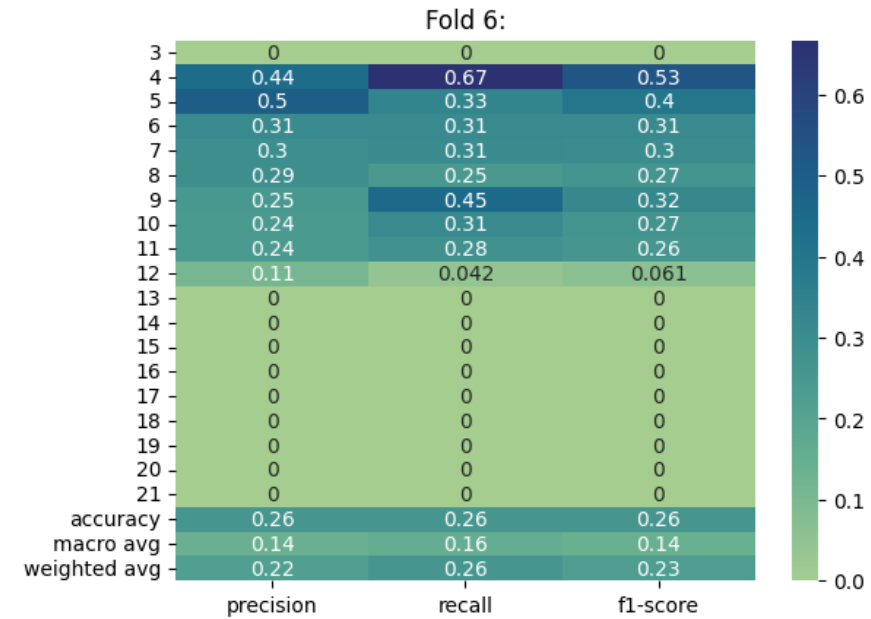
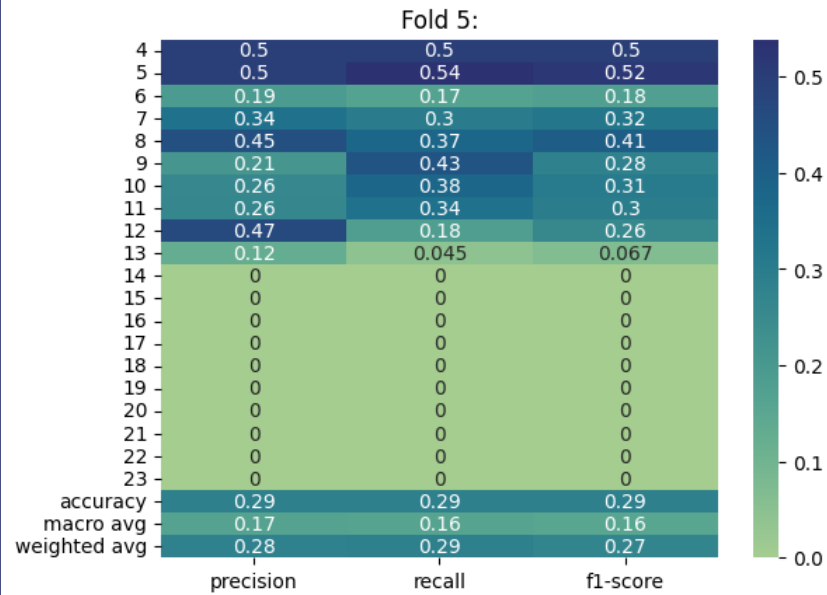
## BEST K

- Cross validation (10-fold)
- K = 42

# 10-FOLD CROSS VALIDATION

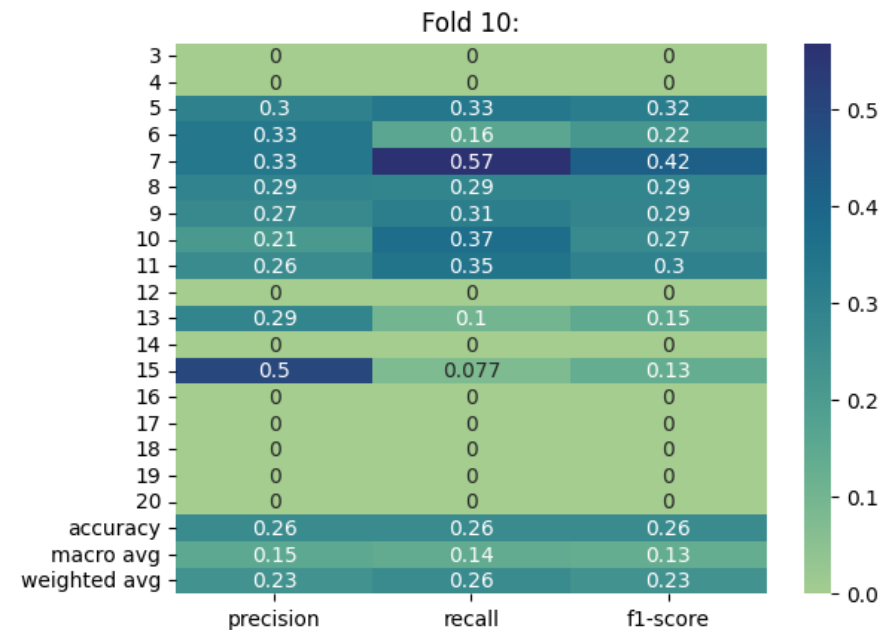
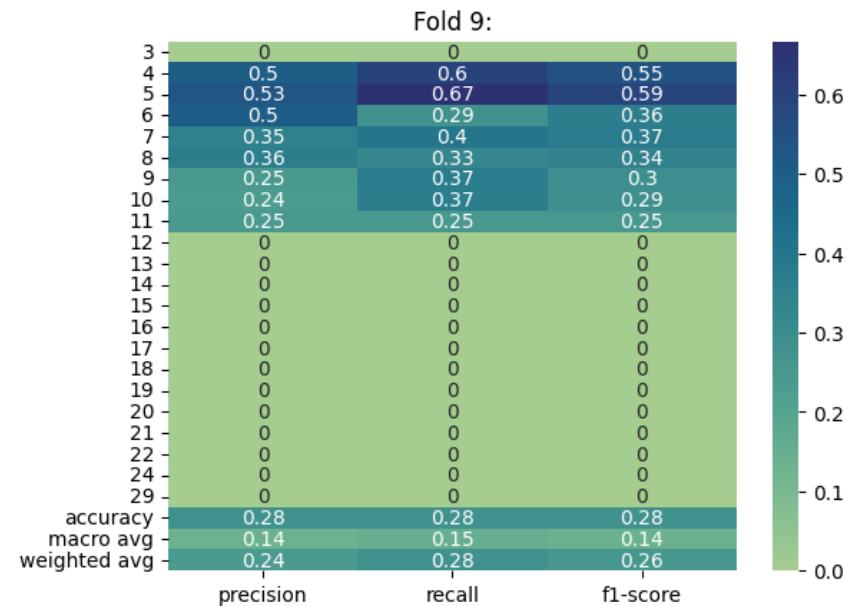


# 10-FOLD CROSS VALIDATION



# 10-FOLD CROSS VALIDATION

- average accuracy = 0.2678
- average precision = 0.2438
- average recall = 0.2678
- average f1-score = 0.24509



# MODEL EVALUATION

