|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **k** | **TPR** | **TNR** | **PPV** | **NPV** | **FPR** | **FDR** | **FNR** | **F1** | **ACC** |
| 1 | 0.7193 | 1 | 1 | 0.8241 | 0 | 0 | 0.2807 | 0.8367 | 0.8788 |
| 2 | 0.9649 | 0.9153 | 0.9167 | 0.9643 | 0.0847 | 0.0833 | 0.03509 | 0.94017 | 0.93966 |
| 3 | 1 | 0.8372 | 0.7454 | 1 | 0.1628 | 0.2545 | 0 | 0.8541 | 0.8897 |
| 4 | 0.9642 | 0.9708 | 0.9 | 0.9900 | 0.0291 | 0.1 | 0.0357 | 0.9310 | 0.9694 |
| 5 | 0.9428 | 0.9484 | 0.8684 | 0.9787 | 0.0515 | 0.1315 | 0.0571 | 0.9041 | 0.9469 |
| AVG | 0.9182 | 0.9343 | 0.8861 | 0.9514 | 0.0656 | 0.1138 | 0.08172 | 0.8932 | 0.9249 |

2.a)

Cross validation is mainly used in the area where the goal is prediction and one wants to know how accurately the predictive model will perform in practice. The purpose of cross validation is to define the dataset to test the model in training phase so that it can limit some problems like overfitting. To reduce overfitting problem and reduce variablity we have used multiple round cross validation and taken their averaged value.

Yes the dataset justifies it correctly.

2.b)

For the given dataset, besides accuracy we can use false negative rate or miss rate in case of predicting. The rate in which cancer is actually present but the test does not classify it is called false negative rate or miss rate. For the given data set, it can play a great role in case of predicting cancer.