|  |
| --- |
| logo.jpg  **Pattern Recognition Course**  **Bioinformatics Department**  **Faculty of Computer and Information Sciences**  **Ain Shams University, Egypt** |
| **A Report of Final Project**  **By** |

|  |  |
| --- | --- |
| **Team no.: 1**  **Mohmed Soudy Mohmed – Sec(2)**  **Lumai Hesham Rashad – Sec(2)**  **Shourok Ashraf Khalil – Sec(1)**  **Abdelwahab Alaa – Sec(1)**  **Amr Qasem – Sec(1)** | |
| **Project Title** | |
| **"*Colon Cancer Diagnosis System*"** | |

**1st Semester 2017\2018**

# **Comparative Study**

Table 1.Results

**Number of Features 25**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Standards** | **Bayesian** | **mk-NN**  **with best k** | **SVM** | **Ensemble Classifier** |
| **Overall Accuracy (%)**  **From Confusion Matrix** | 66.6667 | 66.6667 – k = 7 | Sigmoid function  (cfitsvm) 83.3333 | 73.3333 |
| **Accuracy using 2-fold Cross Validation** | Mean  Varinace | Mean  Varinace | Mean  Varinace | Mean  Varinace |

Table 3.Results

**Number of Features 34**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Standards** | **Bayesian** | **mk-NN**  **with best k** | **SVM** | **Ensemble Classifier** |
| **Overall Accuracy (%)**  **From Confusion Matrix** | 66.6667  Bayes | 66.6667 – k = 7  MKNN | Sigmoid function  (cfitsvm) 80  C:\Users\Soudy\AppData\Local\Microsoft\Windows\INetCache\Content.Word\SVM_sigmoid34.png | 70  **C:\Users\Soudy\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Majority34.png** |
| **Accuracy using 2-fold Cross Validation** | Mean  Varinace | Mean  Varinace | Mean  Varinace | Mean  Varinace |

Table 4.Results

**Number of Features 25**

**TrainingSet 50(30 Cancer,20Normal)**

**TestingSet(12Samples)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Standards** | **Bayesian** | **mk-NN**  **with best k** | **SVM** | **Ensemble Classifier** |
| **Overall Accuracy (%)**  **From Confusion Matrix** | 41.6667 | 91.6667 – k = 41 | rbf function  (cfitsvm) 83.3333 | 91.6667  C:\Users\Soudy\AppData\Local\Microsoft\Windows\INetCache\Content.Word\MKNN_updated.png |
| **Accuracy using 2-fold Cross Validation** | Mean  Varinace | Mean  Varinace | Mean    Varinace | Mean      Varinace |

**Number of Features 34**

**TrainingSet 50(30 Cancer,20Normal)**

**TestingSet(12Samples)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Standards** | **Bayesian** | **mk-NN**  **with best k** | **SVM** | **Ensemble Classifier** |
| **Overall Accuracy (%)**  **From Confusion Matrix** | 50 | 91.6667 – k = 41  C:\Users\Soudy\AppData\Local\Microsoft\Windows\INetCache\Content.Word\MKNN_updated.png | rbf function  (cfitsvm) 83.3333  C:\Users\Soudy\AppData\Local\Microsoft\Windows\INetCache\Content.Word\SVM_sigmoid34.png | 91.6667 C:\Users\Soudy\AppData\Local\Microsoft\Windows\INetCache\Content.Word\MKNN_updated.png |
| **Accuracy using 2-fold Cross Validation** | Mean  Varinace | Mean  Varinace | Mean  Varinace | Mean  Varinace |

**Number of Features 61**

**TrainingSet 50(30 Cancer,20Normal)**

**TestingSet(12Samples)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Standards** | **Bayesian** | **mk-NN**  **with best k** | **SVM** | **Ensemble Classifier** |
| **Overall Accuracy (%)**  **From Confusion Matrix** | 75 | 91.6667 – k = 41  C:\Users\Soudy\AppData\Local\Microsoft\Windows\INetCache\Content.Word\MKNN_updated.png | rbf function  (cfitsvm) 83.3333  C:\Users\Soudy\AppData\Local\Microsoft\Windows\INetCache\Content.Word\SVM_sigmoid34.png | 91.6667 C:\Users\Soudy\AppData\Local\Microsoft\Windows\INetCache\Content.Word\MKNN_updated.png |
| **Accuracy using 2-fold Cross Validation** | Mean  Varinace | Mean  Varinace | Mean  Varinace | Mean  Varinace |

# **Conclusion**

With number of Features 25 out of 61 features extracted from principal component analysis it gives the best accuracies for the 3 classifiers, however we used 2 naïve classifiers (Bayesian, modified k nearest neighbor) with support vector machine which gives the highest accuracy as it was expected.