CS7005 - BIG DATA ANALYTICS
(MINI PROJECT)

TOPIC : FAVOURITE ACTOR

Submitted by,

Nikhil Kumar A P - 2017103560

Gokkul A – 2017103525

Objective :

To locally collect data from the students, to build machine learning models using the data and to visualize the model in the forms of accuracy score, confusion matrix etc .

Dataset :

Sourced from the students of CEG using Google forms .

Explanation :

The Trained 5 machine learning models are :

- Naive Bayes
- Random Forest
- MLP
- Logistic
- Multi Class

The following pictures are in order of :

Train Test Split - 50% / Train Test Split - 66% / Train Test Split - 75% / Cross Validation With 10 Folds

i) Naive Bayes :

=== Summary ===

Correctly Classified Instances	61	92.4242 %
Incorrectly Classified Instances	5	7.5758 %
Kappa statistic	0.9004	
Mean absolute error	0.018	
Root mean squared error	0.1002	
Relative absolute error	11.1067 %	
Root relative squared error	35.6846 %	
Total Number of Instances	66	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.900	0.000	1.000	0.900	0.947	0.940	0.998	0.991	Surya
	0.960	0.024	0.960	0.960	0.960	0.936	0.998	0.997	Vijay
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Ajith
	1.000	0.016	0.800	1.000	0.889	0.887	1.000	1.000	Rajinikanth
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Vijay Sethupathi
	1.000	0.015	0.500	1.000	0.667	0.702	1.000	1.000	Vikram
	1.000	0.015	0.500	1.000	0.667	0.702	1.000	1.000	Dhanush
	1.000	0.015	0.500	1.000	0.667	0.702	1.000	1.000	Simbu
	0.333	0.000	1.000	0.333	0.500	0.568	1.000	1.000	Kamal
	0.000	0.000	?	0.000	?	?	1.000	1.000	Sivakarthikeyan
Weighted Avg.	0.924	0.011	?	0.924	?	?	0.999	0.997	

=== Confusion Matrix ===

```
 a  b  c  d  e  f  g  h  i  j   <-- classified as
 9  1  0  0  0  0  0  0  0  0 |  a = Surya
 0 24  0  1  0  0  0  0  0  0 |  b = Vijay
 0  0 17  0  0  0  0  0  0  0 |  c = Ajith
 0  0  0  4  0  0  0  0  0  0 |  d = Rajinikanth
 0  0  0  0  3  0  0  0  0  0 |  e = Vijay Sethupathi
 0  0  0  0  0  1  0  0  0  0 |  f = Vikram
 0  0  0  0  0  0  1  0  0  0 |  g = Dhanush
 0  0  0  0  0  0  0  1  0  0 |  h = Simbu
 0  0  0  0  0  0  1  1  1  0 |  i = Kamal
 0  0  0  0  0  1  0  0  0  0 |  j = Sivakarthikeyan
```

=== Summary ===

Correctly Classified Instances	44	97.7778 %
Incorrectly Classified Instances	1	2.2222 %
Kappa statistic	0.9698	
Mean absolute error	0.0087	
Root mean squared error	0.0706	
Relative absolute error	5.4593 %	
Root relative squared error	25.4446 %	
Total Number of Instances	45	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.833	0.000	1.000	0.833	0.909	0.901	0.996	0.976	Surya
	1.000	0.037	0.947	1.000	0.973	0.955	0.998	0.997	Vijay
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Ajith
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Rajinikanth
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Vijay Sethupathi
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Vikram
	?	0.000	?	?	?	?	?	?	Dhanush
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Simbu
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Kamal
	?	0.000	?	?	?	?	?	?	Sivakarthikeyan
Weighted Avg.	0.978	0.015	0.979	0.978	0.977	0.969	0.999	0.996	

=== Confusion Matrix ===

```
 a  b  c  d  e  f  g  h  i  j   <-- classified as
 5  1  0  0  0  0  0  0  0  0 |  a = Surya
 0 18  0  0  0  0  0  0  0  0 |  b = Vijay
 0  0 12  0  0  0  0  0  0  0 |  c = Ajith
 0  0  0  3  0  0  0  0  0  0 |  d = Rajinikanth
 0  0  0  0  2  0  0  0  0  0 |  e = Vijay Sethupathi
 0  0  0  0  0  1  0  0  0  0 |  f = Vikram
 0  0  0  0  0  0  0  0  0  0 |  g = Dhanush
 0  0  0  0  0  0  0  1  0  0 |  h = Simbu
 0  0  0  0  0  0  0  0  2  0 |  i = Kamal
 0  0  0  0  0  0  0  0  0  0 |  j = Sivakarthikeyan
```

=== Summary ===

```
Correctly Classified Instances        32               96.9697 %
Incorrectly Classified Instances       1                3.0303 %
Kappa statistic                        0.9582
Mean absolute error                    0.0091
Root mean squared error                0.0792
Relative absolute error                5.7362 %
Root relative squared error           28.5915 %
Total Number of Instances             33
```

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.750	0.000	1.000	0.750	0.857	0.851	1.000	1.000	Surya
	1.000	0.053	0.933	1.000	0.966	0.940	1.000	1.000	Vijay
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Ajith
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Rajinikanth
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Vijay Sethupathi
	?	0.000	?	?	?	?	?	?	Vikram
	?	0.000	?	?	?	?	?	?	Dhanush
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Simbu
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Kamal
	?	0.000	?	?	?	?	?	?	Sivakarthikeyan
Weighted Avg.	0.970	0.022	0.972	0.970	0.968	0.957	1.000	1.000	

=== Confusion Matrix ===

```
 a  b  c  d  e  f  g  h  i  j   <-- classified as
 3  1  0  0  0  0  0  0  0  0 |  a = Surya
 0 14  0  0  0  0  0  0  0  0 |  b = Vijay
 0  0  8  0  0  0  0  0  0  0 |  c = Ajith
 0  0  0  2  0  0  0  0  0  0 |  d = Rajinikanth
 0  0  0  0  2  0  0  0  0  0 |  e = Vijay Sethupathi
 0  0  0  0  0  0  0  0  0  0 |  f = Vikram
 0  0  0  0  0  0  0  0  0  0 |  g = Dhanush
 0  0  0  0  0  0  0  1  0  0 |  h = Simbu
 0  0  0  0  0  0  0  0  2  0 |  i = Kamal
 0  0  0  0  0  0  0  0  0  0 |  j = Sivakarthikeyan
```

=== Summary ===

```
Correctly Classified Instances       127               96.2121 %
Incorrectly Classified Instances       5                3.7879 %
Kappa statistic                        0.952
Mean absolute error                    0.0116
Root mean squared error                0.0795
Relative absolute error                7.2126 %
Root relative squared error           28.1978 %
Total Number of Instances            132
```

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.944	0.009	0.944	0.944	0.944	0.936	0.997	0.985	Surya
	1.000	0.023	0.957	1.000	0.978	0.967	0.998	0.995	Vijay
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Ajith
	0.800	0.000	1.000	0.800	0.889	0.887	0.957	0.916	Rajinikanth
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Vijay Sethupathi
	0.750	0.000	1.000	0.750	0.857	0.863	1.000	1.000	Vikram
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Dhanush
	1.000	0.008	0.750	1.000	0.857	0.863	1.000	1.000	Simbu
	1.000	0.008	0.800	1.000	0.889	0.891	1.000	1.000	Kamal
	0.000	0.000	?	0.000	?	?	0.992	0.500	Sivakarthikeyan
Weighted Avg.	0.962	0.009	?	0.962	?	?	0.996	0.986	

=== Confusion Matrix ===

```
 a  b  c  d  e  f  g  h  i  j   <-- classified as
17  1  0  0  0  0  0  0  0  0 |  a = Surya
 0 44  0  0  0  0  0  0  0  0 |  b = Vijay
 0  0 33  0  0  0  0  0  0  0 |  c = Ajith
 0  1  0  8  0  0  0  0  1  0 |  d = Rajinikanth
 0  0  0  0  6  0  0  0  0  0 |  e = Vijay Sethupathi
 1  0  0  0  0  3  0  0  0  0 |  f = Vikram
 0  0  0  0  0  0  9  0  0  0 |  g = Dhanush
 0  0  0  0  0  0  0  3  0  0 |  h = Simbu
 0  0  0  0  0  0  0  0  4  0 |  i = Kamal
 0  0  0  0  0  0  0  1  0  0 |  j = Sivakarthikeyan
```

ii) Random Forest :

```
Correctly Classified Instances        53              80.303 %
Incorrectly Classified Instances      13              19.697 %
Kappa statistic                       0.7255
Mean absolute error                   0.0774
Root mean squared error               0.1711
Relative absolute error               47.6462 %
Root relative squared error           60.9437 %
Total Number of Instances             66
```

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.700	0.018	0.875	0.700	0.778	0.749	0.982	0.892	Surya
	1.000	0.195	0.758	1.000	0.862	0.781	0.998	0.997	Vijay
	1.000	0.020	0.944	1.000	0.971	0.962	1.000	1.000	Ajith
	0.750	0.016	0.750	0.750	0.750	0.734	0.976	0.850	Rajinikanth
	0.000	0.000	?	0.000	?	?	0.995	0.917	Vijay Sethupathi
	0.000	0.015	0.000	0.000	0.000	−0.015	0.892	0.125	Vikram
	1.000	0.015	0.500	1.000	0.667	0.702	1.000	1.000	Dhanush
	0.000	0.000	?	0.000	?	?	1.000	1.000	Simbu
	0.000	0.000	?	0.000	?	?	0.979	0.810	Kamal
	0.000	0.000	?	0.000	?	?	0.500	0.015	Sivakarthikeyan
Weighted Avg.	0.803	0.083	?	0.803	?	?	0.985	0.933	

=== Confusion Matrix ===

```
 a  b  c  d  e  f  g  h  i  j   <-- classified as
 7  2  0  0  0  1  0  0  0  0 |  a = Surya
 0 25  0  0  0  0  0  0  0  0 |  b = Vijay
 0  0 17  0  0  0  0  0  0  0 |  c = Ajith
 0  1  0  3  0  0  0  0  0  0 |  d = Rajinikanth
 0  1  1  0  0  0  1  0  0  0 |  e = Vijay Sethupathi
 1  0  0  0  0  0  0  0  0  0 |  f = Vikram
 0  0  0  0  0  0  1  0  0  0 |  g = Dhanush
 0  1  0  0  0  0  0  0  0  0 |  h = Simbu
 0  2  0  1  0  0  0  0  0  0 |  i = Kamal
 0  1  0  0  0  0  0  0  0  0 |  j = Sivakarthikeyan
```

```
Correctly Classified Instances        40              88.8889 %
Incorrectly Classified Instances       5              11.1111 %
Kappa statistic                       0.8442
Mean absolute error                   0.07
Root mean squared error               0.1578
Relative absolute error               43.7968 %
Root relative squared error           56.8818 %
Total Number of Instances             45
```

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.833	0.026	0.833	0.833	0.833	0.808	0.983	0.911	Surya
	1.000	0.111	0.857	1.000	0.923	0.873	1.000	1.000	Vijay
	1.000	0.030	0.923	1.000	0.960	0.946	1.000	1.000	Ajith
	0.333	0.000	1.000	0.333	0.500	0.564	0.984	0.867	Rajinikanth
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Vijay Sethupathi
	0.000	0.000	?	0.000	?	?	0.932	0.250	Vikram
	?	0.000	?	?	?	?	?	?	Dhanush
	0.000	0.000	?	0.000	?	?	1.000	1.000	Simbu
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Kamal
	?	0.000	?	?	?	?	?	?	Sivakarthikeyan
Weighted Avg.	0.889	0.056	?	0.889	?	?	0.995	0.963	

=== Confusion Matrix ===

```
 a  b  c  d  e  f  g  h  i  j   <-- classified as
 5  1  0  0  0  0  0  0  0  0 |  a = Surya
 0 18  0  0  0  0  0  0  0  0 |  b = Vijay
 0  0 12  0  0  0  0  0  0  0 |  c = Ajith
 0  1  1  1  0  0  0  0  0  0 |  d = Rajinikanth
 0  0  0  0  2  0  0  0  0  0 |  e = Vijay Sethupathi
 1  0  0  0  0  0  0  0  0  0 |  f = Vikram
 0  0  0  0  0  0  0  0  0  0 |  g = Dhanush
 0  1  0  0  0  0  0  0  0  0 |  h = Simbu
 0  0  0  0  0  0  0  0  2  0 |  i = Kamal
 0  0  0  0  0  0  0  0  0  0 |  j = Sivakarthikeyan
```

```
Correctly Classified Instances          28                84.8485 %
Incorrectly Classified Instances         5                15.1515 %
Kappa statistic                          0.782
Mean absolute error                      0.0705
Root mean squared error                  0.1587
Relative absolute error                 44.3874 %
Root relative squared error             57.3014 %
Total Number of Instances               33
```
running classification
```
=== Detailed Accuracy By Class ===
```

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.750	0.034	0.750	0.750	0.750	0.716	1.000	1.000	Surya
	1.000	0.158	0.824	1.000	0.903	0.833	1.000	1.000	Vijay
	1.000	0.040	0.889	1.000	0.941	0.924	1.000	1.000	Ajith
	0.500	0.000	1.000	0.500	0.667	0.696	0.935	0.667	Rajinikanth
	0.000	0.000	?	0.000	?	?	1.000	1.000	Vijay Sethupathi
	?	0.000	?	?	?	?	?	?	Vikram
	?	0.000	?	?	?	?	?	?	Dhanush
	0.000	0.000	?	0.000	?	?	0.969	0.500	Simbu
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Kamal
	?	0.000	?	?	?	?	?	?	Sivakarthikeyan
Weighted Avg.	0.848	0.081	?	0.848	?	?	0.995	0.965	

```
=== Confusion Matrix ===

 a  b  c  d  e  f  g  h  i  j   <-- classified as
 3  1  0  0  0  0  0  0  0  0 |  a = Surya
 0 14  0  0  0  0  0  0  0  0 |  b = Vijay
 0  0  8  0  0  0  0  0  0  0 |  c = Ajith
 0  0  1  1  0  0  0  0  0  0 |  d = Rajinikanth
 1  1  0  0  0  0  0  0  0  0 |  e = Vijay Sethupathi
 0  0  0  0  0  0  0  0  0  0 |  f = Vikram
 0  0  0  0  0  0  0  0  0  0 |  g = Dhanush
 0  1  0  0  0  0  0  0  0  0 |  h = Simbu
 0  0  0  0  0  0  0  0  2  0 |  i = Kamal
 0  0  0  0  0  0  0  0  0  0 |  j = Sivakarthikeyan
```

```
=== Summary ===

Correctly Classified Instances         116                87.8788 %
Incorrectly Classified Instances        16                12.1212 %
Kappa statistic                          0.8429
Mean absolute error                      0.0669
Root mean squared error                  0.1541
Relative absolute error                 41.6828 %
Root relative squared error             54.6418 %
Total Number of Instances              132

=== Detailed Accuracy By Class ===
```

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.778	0.018	0.875	0.778	0.824	0.799	0.979	0.918	Surya
	1.000	0.091	0.846	1.000	0.917	0.877	0.998	0.996	Vijay
	0.970	0.040	0.889	0.970	0.928	0.904	0.999	0.998	Ajith
	0.800	0.008	0.889	0.800	0.842	0.831	0.957	0.873	Rajinikanth
	0.667	0.008	0.800	0.667	0.727	0.719	0.988	0.775	Vijay Sethupathi
	0.000	0.000	?	0.000	?	?	0.941	0.391	Vikram
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Dhanush
	0.667	0.000	1.000	0.667	0.800	0.813	1.000	1.000	Simbu
	0.750	0.000	1.000	0.750	0.857	0.863	0.998	0.950	Kamal
	0.000	0.000	?	0.000	?	?	0.313	0.008	Sivakarthikeyan
Weighted Avg.	0.879	0.044	?	0.879	?	?	0.986	0.940	

```
=== Confusion Matrix ===

 a  b  c  d  e  f  g  h  i  j   <-- classified as
14  3  1  0  0  0  0  0  0  0 |  a = Surya
 0 44  0  0  0  0  0  0  0  0 |  b = Vijay
 0  1 32  0  0  0  0  0  0  0 |  c = Ajith
 0  1  1  8  0  0  0  0  0  0 |  d = Rajinikanth
 0  0  2  0  4  0  0  0  0  0 |  e = Vijay Sethupathi
 2  2  0  0  0  0  0  0  0  0 |  f = Vikram
 0  0  0  0  0  0  9  0  0  0 |  g = Dhanush
 0  1  0  0  0  0  0  2  0  0 |  h = Simbu
 0  0  0  1  0  0  0  0  3  0 |  i = Kamal
 0  0  0  0  1  0  0  0  0  0 |  j = Sivakarthikeyan
```

iii) Multi Layer Perceptron :

```
=== Summary ===

Correctly Classified Instances        59               89.3939 %
Incorrectly Classified Instances       7               10.6061 %
Kappa statistic                      0.8585
Mean absolute error                  0.0346
Root mean squared error              0.1301
Relative absolute error             21.2705 %
Root relative squared error         46.3333 %
Total Number of Instances             66
```

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.800	0.018	0.889	0.800	0.842	0.817	0.991	0.961	Surya
	0.960	0.073	0.889	0.960	0.923	0.875	0.988	0.983	Vijay
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Ajith
	1.000	0.016	0.800	1.000	0.889	0.887	1.000	1.000	Rajinikanth
	1.000	0.016	0.750	1.000	0.857	0.859	1.000	1.000	Vijay Sethupathi
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Vikram
	1.000	0.015	0.500	1.000	0.667	0.702	1.000	1.000	Dhanush
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Simbu
	0.000	0.000	?	0.000	?	?	0.995	0.917	Kamal
	0.000	0.000	?	0.000	?	?	0.969	0.333	Sivakarthikeyan
Weighted Avg.	0.894	0.032	?	0.894	?	?	0.994	0.974	

=== Confusion Matrix ===

```
 a  b  c  d  e  f  g  h  i  j   <-- classified as
 8  1  0  0  0  0  1  0  0  0 |  a = Surya
 0 24  0  1  0  0  0  0  0  0 |  b = Vijay
 0  0 17  0  0  0  0  0  0  0 |  c = Ajith
 0  0  0  4  0  0  0  0  0  0 |  d = Rajinikanth
 0  0  0  0  3  0  0  0  0  0 |  e = Vijay Sethupathi
 0  0  0  0  0  1  0  0  0  0 |  f = Vikram
 0  0  0  0  0  0  1  0  0  0 |  g = Dhanush
 0  0  0  0  0  0  0  1  0  0 |  h = Simbu
 1  2  0  0  0  0  0  0  0  0 |  i = Kamal
 0  0  0  0  1  0  0  0  0  0 |  j = Sivakarthikeyan
```

```
=== Summary ===

Correctly Classified Instances        44               97.7778 %
Incorrectly Classified Instances       1                2.2222 %
Kappa statistic                      0.9698
Mean absolute error                  0.0224
Root mean squared error              0.0916
Relative absolute error             14.0088 %
Root relative squared error         33.0151 %
Total Number of Instances             45
```

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.833	0.000	1.000	0.833	0.909	0.901	0.983	0.933	Surya
	1.000	0.037	0.947	1.000	0.973	0.955	0.994	0.991	Vijay
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Ajith
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Rajinikanth
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Vijay Sethupathi
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Vikram
	?	0.000	?	?	?	?	?	?	Dhanush
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Simbu
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Kamal
	?	0.000	?	?	?	?	?	?	Sivakarthikeyan
Weighted Avg.	0.978	0.015	0.979	0.978	0.977	0.969	0.995	0.987	

=== Confusion Matrix ===

```
 a  b  c  d  e  f  g  h  i  j   <-- classified as
 5  1  0  0  0  0  0  0  0  0 |  a = Surya
 0 18  0  0  0  0  0  0  0  0 |  b = Vijay
 0  0 12  0  0  0  0  0  0  0 |  c = Ajith
 0  0  0  3  0  0  0  0  0  0 |  d = Rajinikanth
 0  0  0  0  2  0  0  0  0  0 |  e = Vijay Sethupathi
 0  0  0  0  0  1  0  0  0  0 |  f = Vikram
 0  0  0  0  0  0  0  0  0  0 |  g = Dhanush
 0  0  0  0  0  0  0  1  0  0 |  h = Simbu
 0  0  0  0  0  0  0  0  2  0 |  i = Kamal
 0  0  0  0  0  0  0  0  0  0 |  j = Sivakarthikeyan
```

```
=== Summary ===

Correctly Classified Instances        32              96.9697 %
Incorrectly Classified Instances       1               3.0303 %
Kappa statistic                        0.9582
Mean absolute error                    0.025
Root mean squared error                0.1032
Relative absolute error               15.731  %
Root relative squared error           37.2454 %
Total Number of Instances             33

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
                 0.750    0.000    1.000      0.750   0.857      0.851  0.974     0.893     Surya
                 1.000    0.053    0.933      1.000   0.966      0.940  0.985     0.979     Vijay
                 1.000    0.000    1.000      1.000   1.000      1.000  1.000     1.000     Ajith
                 1.000    0.000    1.000      1.000   1.000      1.000  1.000     1.000     Rajinikanth
                 1.000    0.000    1.000      1.000   1.000      1.000  1.000     1.000     Vijay Sethupathi
                 ?        0.000    ?          ?       ?          ?      ?         ?         Vikram
                 ?        0.000    ?          ?       ?          ?      ?         ?         Dhanush
                 1.000    0.000    1.000      1.000   1.000      1.000  1.000     1.000     Simbu
                 1.000    0.000    1.000      1.000   1.000      1.000  1.000     1.000     Kamal
                 ?        0.000    ?          ?       ?          ?      ?         ?         Sivakarthikeyan
Weighted Avg.    0.970    0.022    0.972      0.970   0.968      0.957  0.990     0.978

=== Confusion Matrix ===

  a  b  c  d  e  f  g  h  i  j   <-- classified as
  3  1  0  0  0  0  0  0  0  0 |  a = Surya
  0 14  0  0  0  0  0  0  0  0 |  b = Vijay
  0  0  8  0  0  0  0  0  0  0 |  c = Ajith
  0  0  0  2  0  0  0  0  0  0 |  d = Rajinikanth
  0  0  0  0  2  0  0  0  0  0 |  e = Vijay Sethupathi
  0  0  0  0  0  0  0  0  0  0 |  f = Vikram
  0  0  0  0  0  0  0  0  0  0 |  g = Dhanush
  0  0  0  0  0  0  0  1  0  0 |  h = Simbu
  0  0  0  0  0  0  0  0  2  0 |  i = Kamal
  0  0  0  0  0  0  0  0  0  0 |  j = Sivakarthikeyan



=== Summary ===

Correctly Classified Instances       122              92.4242 %
Incorrectly Classified Instances      10               7.5758 %
Kappa statistic                        0.9036
Mean absolute error                    0.0249
Root mean squared error                0.1065
Relative absolute error               15.5387 %
Root relative squared error           37.7572 %
Total Number of Instances            132

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
                 0.889    0.009    0.941      0.889   0.914      0.902  0.991     0.961     Surya
                 0.955    0.045    0.913      0.955   0.933      0.899  0.992     0.982     Vijay
                 1.000    0.020    0.943      1.000   0.971      0.961  1.000     1.000     Ajith
                 0.900    0.008    0.900      0.900   0.900      0.892  0.915     0.899     Rajinikanth
                 0.833    0.000    1.000      0.833   0.909      0.909  0.999     0.976     Vijay Sethupathi
                 0.500    0.000    1.000      0.500   0.667      0.702  1.000     1.000     Vikram
                 0.889    0.016    0.800      0.889   0.842      0.831  0.998     0.980     Dhanush
                 1.000    0.000    1.000      1.000   1.000      1.000  1.000     1.000     Simbu
                 1.000    0.000    1.000      1.000   1.000      1.000  1.000     1.000     Kamal
                 0.000    0.000    ?          0.000   ?          ?      0.733     0.028     Sivakarthikeyan
Weighted Avg.    0.924    0.023    ?          0.924   ?          ?      0.987     0.971

=== Confusion Matrix ===

  a  b  c  d  e  f  g  h  i  j   <-- classified as
 16  1  0  0  0  0  1  0  0  0 |  a = Surya
  0 42  0  1  0  0  1  0  0  0 |  b = Vijay
  0  0 33  0  0  0  0  0  0  0 |  c = Ajith
  0  1  0  9  0  0  0  0  0  0 |  d = Rajinikanth
  0  0  1  0  5  0  0  0  0  0 |  e = Vijay Sethupathi
  1  0  1  0  0  2  0  0  0  0 |  f = Vikram
  0  1  0  0  0  0  8  0  0  0 |  g = Dhanush
  0  0  0  0  0  0  0  3  0  0 |  h = Simbu
  0  0  0  0  0  0  0  0  4  0 |  i = Kamal
  0  1  0  0  0  0  0  0  0  0 |  j = Sivakarthikeyan
```

iv) Multi Class Classifier :

```
Correctly Classified Instances          50              75.7576 %
Incorrectly Classified Instances        16              24.2424 %
Kappa statistic                          0.6714
Mean absolute error                      0.0565
Root mean squared error                  0.2138
Relative absolute error                 34.7888 %
Root relative squared error             76.1661 %
Total Number of Instances               66
```

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.100	0.000	1.000	0.100	0.182	0.294	0.873	0.515	Surya
	1.000	0.146	0.806	1.000	0.893	0.830	0.959	0.927	Vijay
	1.000	0.041	0.895	1.000	0.944	0.926	0.996	0.990	Ajith
	0.750	0.032	0.600	0.750	0.667	0.647	0.782	0.668	Rajinikanth
	0.667	0.000	1.000	0.667	0.800	0.810	0.995	0.917	Vijay Sethupathi
	0.000	0.046	0.000	0.000	0.000	−0.027	0.446	0.027	Vikram
	1.000	0.015	0.500	1.000	0.667	0.702	1.000	1.000	Dhanush
	1.000	0.031	0.333	1.000	0.500	0.568	0.969	0.333	Simbu
	0.000	0.000	?	0.000	?	?	0.958	0.587	Kamal
	0.000	0.000	?	0.000	?	?	0.585	0.036	Sivakarthikeyan
Weighted Avg.	0.758	0.069	?	0.758	?	?	0.934	0.814	

=== Confusion Matrix ===

```
 a  b  c  d  e  f  g  h  i  j   <-- classified as
 1  3  1  0  0  3  1  1  0  0 |  a = Surya
 0 25  0  0  0  0  0  0  0  0 |  b = Vijay
 0  0 17  0  0  0  0  0  0  0 |  c = Ajith
 0  1  0  3  0  0  0  0  0  0 |  d = Rajinikanth
 0  0  0  0  2  0  0  1  0  0 |  e = Vijay Sethupathi
 0  0  1  0  0  0  0  0  0  0 |  f = Vikram
 0  0  0  0  0  0  1  0  0  0 |  g = Dhanush
 0  0  0  0  0  0  0  1  0  0 |  h = Simbu
 0  1  0  2  0  0  0  0  0  0 |  i = Kamal
 0  1  0  0  0  0  0  0  0  0 |  j = Sivakarthikeyan
```

=== Summary ===

```
Correctly Classified Instances          37              82.2222 %
Incorrectly Classified Instances         8              17.7778 %
Kappa statistic                          0.7529
Mean absolute error                      0.0383
Root mean squared error                  0.1773
Relative absolute error                 23.9323 %
Root relative squared error             63.9009 %
Total Number of Instances               45
```

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.500	0.051	0.600	0.500	0.545	0.485	0.902	0.604	Surya
	1.000	0.074	0.900	1.000	0.947	0.913	0.990	0.984	Vijay
	1.000	0.061	0.857	1.000	0.923	0.897	1.000	1.000	Ajith
	0.667	0.000	1.000	0.667	0.800	0.807	0.746	0.695	Rajinikanth
	0.500	0.000	1.000	0.500	0.667	0.699	0.988	0.833	Vijay Sethupathi
	0.000	0.023	0.000	0.000	0.000	−0.023	0.977	0.500	Vikram
	?	0.022	0.000	?	?	?	?	?	Dhanush
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Simbu
	0.000	0.000	?	0.000	?	?	0.826	0.313	Kamal
	?	0.000	?	?	?	?	?	?	Sivakarthikeyan
Weighted Avg.	0.822	0.053	?	0.822	?	?	0.957	0.871	

=== Confusion Matrix ===

```
 a  b  c  d  e  f  g  h  i  j   <-- classified as
 3  2  0  0  0  0  1  0  0  0 |  a = Surya
 0 18  0  0  0  0  0  0  0  0 |  b = Vijay
 0  0 12  0  0  0  0  0  0  0 |  c = Ajith
 0  0  1  2  0  0  0  0  0  0 |  d = Rajinikanth
 1  0  0  0  1  0  0  0  0  0 |  e = Vijay Sethupathi
 0  0  1  0  0  0  0  0  0  0 |  f = Vikram
 0  0  0  0  0  0  0  0  0  0 |  g = Dhanush
 0  0  0  0  0  0  0  1  0  0 |  h = Simbu
 1  0  0  0  0  1  0  0  0  0 |  i = Kamal
 0  0  0  0  0  0  0  0  0  0 |  j = Sivakarthikeyan
```

=== Summary ===

Correctly Classified Instances	27	81.8182 %
Incorrectly Classified Instances	6	18.1818 %
Kappa statistic	0.7458	
Mean absolute error	0.0419	
Root mean squared error	0.1787	
Relative absolute error	26.3755 %	
Root relative squared error	64.5056 %	
Total Number of Instances	33	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.500	0.034	0.667	0.500	0.571	0.528	0.948	0.729	Surya
	1.000	0.053	0.933	1.000	0.966	0.940	0.985	0.979	Vijay
	1.000	0.080	0.800	1.000	0.889	0.858	0.985	0.953	Ajith
	0.500	0.000	1.000	0.500	0.667	0.696	0.758	0.559	Rajinikanth
	1.000	0.032	0.667	1.000	0.800	0.803	1.000	1.000	Vijay Sethupathi
	?	0.030	0.000	?	?	?	?	?	Vikram
	?	0.000	?	?	?	?	?	?	Dhanush
	0.000	0.000	?	0.000	?	?	0.969	0.500	Simbu
	0.000	0.000	?	0.000	?	?	0.500	0.087	Kamal
	?	0.000	?	?	?	?	?	?	Sivakarthikeyan
Weighted Avg.	0.818	0.048	?	0.818	?	?	0.938	0.849	

=== Confusion Matrix ===

```
 a  b  c  d  e  f  g  h  i  j   <-- classified as
 2  1  1  0  0  0  0  0  0  0 |  a = Surya
 0 14  0  0  0  0  0  0  0  0 |  b = Vijay
 0  0  8  0  0  0  0  0  0  0 |  c = Ajith
 0  0  1  1  0  0  0  0  0  0 |  d = Rajinikanth
 0  0  0  0  2  0  0  0  0  0 |  e = Vijay Sethupathi
 0  0  0  0  0  0  0  0  0  0 |  f = Vikram
 0  0  0  0  0  0  0  0  0  0 |  g = Dhanush
 0  0  0  0  1  0  0  0  0  0 |  h = Simbu
 1  0  0  0  0  1  0  0  0  0 |  i = Kamal
 0  0  0  0  0  0  0  0  0  0 |  j = Sivakarthikeyan
```

=== Summary ===

Correctly Classified Instances	111	84.0909 %
Incorrectly Classified Instances	21	15.9091 %
Kappa statistic	0.7975	
Mean absolute error	0.0396	
Root mean squared error	0.1706	
Relative absolute error	24.6926 %	
Root relative squared error	60.4702 %	
Total Number of Instances	132	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.722	0.026	0.813	0.722	0.765	0.732	0.959	0.858	Surya
	0.955	0.045	0.913	0.955	0.933	0.899	0.970	0.895	Vijay
	1.000	0.030	0.917	1.000	0.957	0.943	0.992	0.974	Ajith
	0.600	0.033	0.600	0.600	0.600	0.567	0.788	0.493	Rajinikanth
	0.833	0.008	0.833	0.833	0.833	0.825	0.857	0.795	Vijay Sethupathi
	0.750	0.023	0.500	0.750	0.600	0.598	0.986	0.841	Vikram
	0.667	0.000	1.000	0.667	0.800	0.807	0.906	0.872	Dhanush
	0.667	0.016	0.500	0.667	0.571	0.566	0.992	0.833	Simbu
	0.250	0.008	0.500	0.250	0.333	0.340	0.941	0.422	Kamal
	0.000	0.000	?	0.000	?	?	0.908	0.077	Sivakarthikeyan
Weighted Avg.	0.841	0.030	?	0.841	?	?	0.950	0.850	

=== Confusion Matrix ===

```
 a  b  c  d  e  f  g  h  i  j   <-- classified as
13  2  0  1  1  0  0  1  0  0 |  a = Surya
 0 42  1  1  0  0  0  0  0  0 |  b = Vijay
 0  0 33  0  0  0  0  0  0  0 |  c = Ajith
 1  1  1  6  0  0  0  0  1  0 |  d = Rajinikanth
 0  0  0  1  5  0  0  0  0  0 |  e = Vijay Sethupathi
 1  0  0  0  0  3  0  0  0  0 |  f = Vikram
 1  1  0  0  0  1  6  0  0  0 |  g = Dhanush
 0  0  0  0  0  1  0  2  0  0 |  h = Simbu
 0  0  1  1  0  1  0  0  1  0 |  i = Kamal
 0  0  0  0  0  0  0  1  0  0 |  j = Sivakarthikeyan
```

v) Logistics :

=== Summary ===

Correctly Classified Instances	57	86.3636 %
Incorrectly Classified Instances	9	13.6364 %
Kappa statistic	0.8179	
Mean absolute error	0.0275	
Root mean squared error	0.1627	
Relative absolute error	16.9007 %	
Root relative squared error	57.9698 %	
Total Number of Instances	66	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0.700	0.000	1.000	0.700	0.824	0.815	0.989	0.948	Surya
1.000	0.073	0.893	1.000	0.943	0.910	0.976	0.926	Vijay
1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Ajith
1.000	0.016	0.800	1.000	0.889	0.887	1.000	1.000	Rajinikanth
0.667	0.000	1.000	0.667	0.800	0.810	1.000	1.000	Vijay Sethupathi
1.000	0.015	0.500	1.000	0.667	0.702	0.985	0.500	Vikram
0.000	0.062	0.000	0.000	0.000	-0.032	0.923	0.167	Dhanush
0.000	0.000	?	0.000	?	?	1.000	1.000	Simbu
0.333	0.000	1.000	0.333	0.500	0.568	0.979	0.700	Kamal
0.000	0.000	?	0.000	?	?	0.723	0.053	Sivakarthikeyan
Weighted Avg. 0.864	0.030	?	0.864	?	?	0.983	0.916	

=== Confusion Matrix ===

```
 a  b  c  d  e  f  g  h  i  j   <-- classified as
 7  1  0  0  0  1  1  0  0  0 |  a = Surya
 0 25  0  0  0  0  0  0  0  0 |  b = Vijay
 0  0 17  0  0  0  0  0  0  0 |  c = Ajith
 0  0  0  4  0  0  0  0  0  0 |  d = Rajinikanth
 0  0  0  1  2  0  0  0  0  0 |  e = Vijay Sethupathi
 0  0  0  0  0  1  0  0  0  0 |  f = Vikram
 0  1  0  0  0  0  0  0  0  0 |  g = Dhanush
 0  0  0  0  0  0  1  0  0  0 |  h = Simbu
 0  0  0  0  0  0  2  0  1  0 |  i = Kamal
 0  1  0  0  0  0  0  0  0  0 |  j = Sivakarthikeyan
```

=== Summary ===

Correctly Classified Instances	38	84.4444 %
Incorrectly Classified Instances	7	15.5556 %
Kappa statistic	0.7859	
Mean absolute error	0.0311	
Root mean squared error	0.1739	
Relative absolute error	19.4254 %	
Root relative squared error	62.6543 %	
Total Number of Instances	45	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0.667	0.000	1.000	0.667	0.800	0.796	0.949	0.848	Surya
0.944	0.074	0.895	0.944	0.919	0.863	0.973	0.934	Vijay
1.000	0.061	0.857	1.000	0.923	0.897	1.000	1.000	Ajith
0.667	0.048	0.500	0.667	0.571	0.543	0.968	0.810	Rajinikanth
0.500	0.000	1.000	0.500	0.667	0.699	1.000	1.000	Vijay Sethupathi
0.000	0.000	?	0.000	?	?	1.000	1.000	Vikram
?	0.000	?	?	?	?	?	?	Dhanush
0.000	0.000	?	0.000	?	?	0.977	0.500	Simbu
1.000	0.023	0.667	1.000	0.800	0.807	1.000	1.000	Kamal
?	0.000	?	?	?	?	?	?	Sivakarthikeyan
Weighted Avg. 0.844	0.050	?	0.844	?	?	0.980	0.930	

=== Confusion Matrix ===

```
 a  b  c  d  e  f  g  h  i  j   <-- classified as
 4  2  0  0  0  0  0  0  0  0 |  a = Surya
 0 17  0  1  0  0  0  0  0  0 |  b = Vijay
 0  0 12  0  0  0  0  0  0  0 |  c = Ajith
 0  0  1  2  0  0  0  0  0  0 |  d = Rajinikanth
 0  0  0  1  1  0  0  0  0  0 |  e = Vijay Sethupathi
 0  0  1  0  0  0  0  0  0  0 |  f = Vikram
 0  0  0  0  0  0  0  0  0  0 |  g = Dhanush
 0  0  0  0  0  0  0  0  1  0 |  h = Simbu
 0  0  0  0  0  0  0  0  2  0 |  i = Kamal
 0  0  0  0  0  0  0  0  0  0 |  j = Sivakarthikeyan
```

=== Summary ===

```
Correctly Classified Instances          29               87.8788 %
Incorrectly Classified Instances         4               12.1212 %
Kappa statistic                          0.8323
Mean absolute error                      0.0242
Root mean squared error                  0.1557
Relative absolute error                 15.2506 %
Root relative squared error             56.1978 %
Total Number of Instances               33
```

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.750	0.034	0.750	0.750	0.750	0.716	0.966	0.875	Surya
	1.000	0.053	0.933	1.000	0.966	0.940	0.974	0.933	Vijay
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Ajith
	1.000	0.032	0.667	1.000	0.800	0.803	0.984	0.833	Rajinikanth
	0.000	0.000	?	0.000	?	?	1.000	1.000	Vijay Sethupathi
	?	0.000	?	?	?	?	?	?	Vikram
	?	0.000	?	?	?	?	?	?	Dhanush
	0.000	0.000	?	0.000	?	?	0.906	0.250	Simbu
	1.000	0.032	0.667	1.000	0.800	0.803	0.976	0.583	Kamal
	?	0.000	?	?	?	?	?	?	Sivakarthikeyan
Weighted Avg.	0.879	0.030	?	0.879	?	?	0.979	0.898	

=== Confusion Matrix ===

```
 a  b  c  d  e  f  g  h  i  j   <-- classified as
 3  1  0  0  0  0  0  0  0  0 |  a = Surya
 0 14  0  0  0  0  0  0  0  0 |  b = Vijay
 0  0  8  0  0  0  0  0  0  0 |  c = Ajith
 0  0  0  2  0  0  0  0  0  0 |  d = Rajinikanth
 1  0  0  1  0  0  0  0  0  0 |  e = Vijay Sethupathi
 0  0  0  0  0  0  0  0  0  0 |  f = Vikram
 0  0  0  0  0  0  0  0  0  0 |  g = Dhanush
 0  0  0  0  0  0  0  0  1  0 |  h = Simbu
 0  0  0  0  0  0  0  0  2  0 |  i = Kamal
 0  0  0  0  0  0  0  0  0  0 |  j = Sivakarthikeyan
```

=== Summary ===

```
Correctly Classified Instances         118               89.3939 %
Incorrectly Classified Instances        14               10.6061 %
Kappa statistic                          0.8648
Mean absolute error                      0.0216
Root mean squared error                  0.1456
Relative absolute error                 13.4819 %
Root relative squared error             51.5992 %
Total Number of Instances              132
```

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.833	0.009	0.938	0.833	0.882	0.867	0.978	0.942	Surya
	0.977	0.045	0.915	0.977	0.945	0.917	0.985	0.950	Vijay
	1.000	0.020	0.943	1.000	0.971	0.961	0.990	0.943	Ajith
	0.800	0.016	0.800	0.800	0.800	0.784	0.929	0.755	Rajinikanth
	0.500	0.016	0.600	0.500	0.545	0.528	0.848	0.523	Vijay Sethupathi
	0.750	0.000	1.000	0.750	0.857	0.863	0.975	0.809	Vikram
	0.889	0.008	0.889	0.889	0.889	0.881	0.994	0.951	Dhanush
	0.667	0.000	1.000	0.667	0.800	0.813	0.995	0.867	Simbu
	0.750	0.016	0.600	0.750	0.667	0.659	0.996	0.917	Kamal
	0.000	0.000	?	0.000	?	?	0.977	0.250	Sivakarthikeyan
Weighted Avg.	0.894	0.024	?	0.894	?	?	0.976	0.901	

=== Confusion Matrix ===

```
 a  b  c  d  e  f  g  h  i  j   <-- classified as
15  2  0  0  1  0  0  0  0  0 |  a = Surya
 1 43  0  0  0  0  0  0  0  0 |  b = Vijay
 0  0 33  0  0  0  0  0  0  0 |  c = Ajith
 0  1  0  8  0  1  0  0  0  0 |  d = Rajinikanth
 0  0  1  1  3  0  0  0  1  0 |  e = Vijay Sethupathi
 0  0  1  0  0  3  0  0  0  0 |  f = Vikram
 0  1  0  0  0  0  8  0  0  0 |  g = Dhanush
 0  0  0  0  0  0  0  2  1  0 |  h = Simbu
 0  0  0  1  0  0  0  0  3  0 |  i = Kamal
 0  0  0  0  1  0  0  0  0  0 |  j = Sivakarthikeyan
```

Conclusion :

From the above models, it is clear that the **Naive Bayes Model** with a **Train Test Split of 66 %** has the highest accuracy of **97.778 %**

So the best model from the above date analysis for our dataset is, **Naive Bayes Model** with a **Train Test Split of 66 %**