

Name: Rajdeep Savani
UID : U79032196

Question 1:

Classification Results:

Classifier/ Data Type	Original			Translated			Rotated X/Y/Z		
Metric	Accuracy	Precision	Recall	Accuracy	Precision	Recall	Accuracy	Precision	Recall
Random Forest	0.9999	0.9999	0.9999	0.9988	0.9988	0.9988	0.9999 0.9999 0.9999	0.9999 0.9999 0.9999	0.9999 0.9999 0.9999
SVM	0.7940	0.7938	0.7941	0.7894	0.7892	0.7895	0.7935 0.7936	0.7933 0.7934	0.7936 0.7937
Decision Tree	0.9848	0.9848	0.9847	0.9580	0.9581	0.9580	0.9845 0.9842 0.9848	0.9845 0.9842 0.9848	0.9845 0.9842 0.9848

Question 2:

For Original Datatype – Random Forest worked best

For Translated Original Datatype – Random Forest worked best

For Rotated Original Datatype – Random Forest worked best

Random Forest (RF) excel with the provided dataset because they can effectively handle three-dimensional data, resulting in more accurate emotion prediction due to the dataset's limited features and well-separated classes.

The similar performance observed between Support Vector Machines (SVM) and Decision Tree models likely stems from their shared approach of constructing decision trees based on feature selection.

When we refer to a model performing "best," we mean it achieves higher accuracy according to the specified performance metrics.

Accuracy proves more beneficial for evaluating model performance as it represents the proportion of correct predictions to the total predictions made.

In contrast, precision identifies false positive predictions, while recall indicates the ratio of true positive predictions to the total actual positive examples. However, these metrics don't encompass all predictions like accuracy does, reinforcing the superiority of accuracy as the primary evaluation metric.

Question 3:

Given the very high accuracy, precision, and recall scores for the Random Forest classifier across all data types, misclassifications were minimal. The specific instances of misclassification aren't detailed in the provided data, but generally, misclassifications in facial expression recognition can occur due to similarities in facial landmarks between certain expressions, e.g., "sad" and "happy" might be confused if the changes in facial landmarks are subtle.

Question 4:

The Random Forest classifier's strong performance is likely attributed to its ensemble approach, where multiple decision trees are combined to enhance accuracy and robustness. This method reduces the risk of overfitting compared to a single decision tree and effectively handles the complexity of facial landmarks.

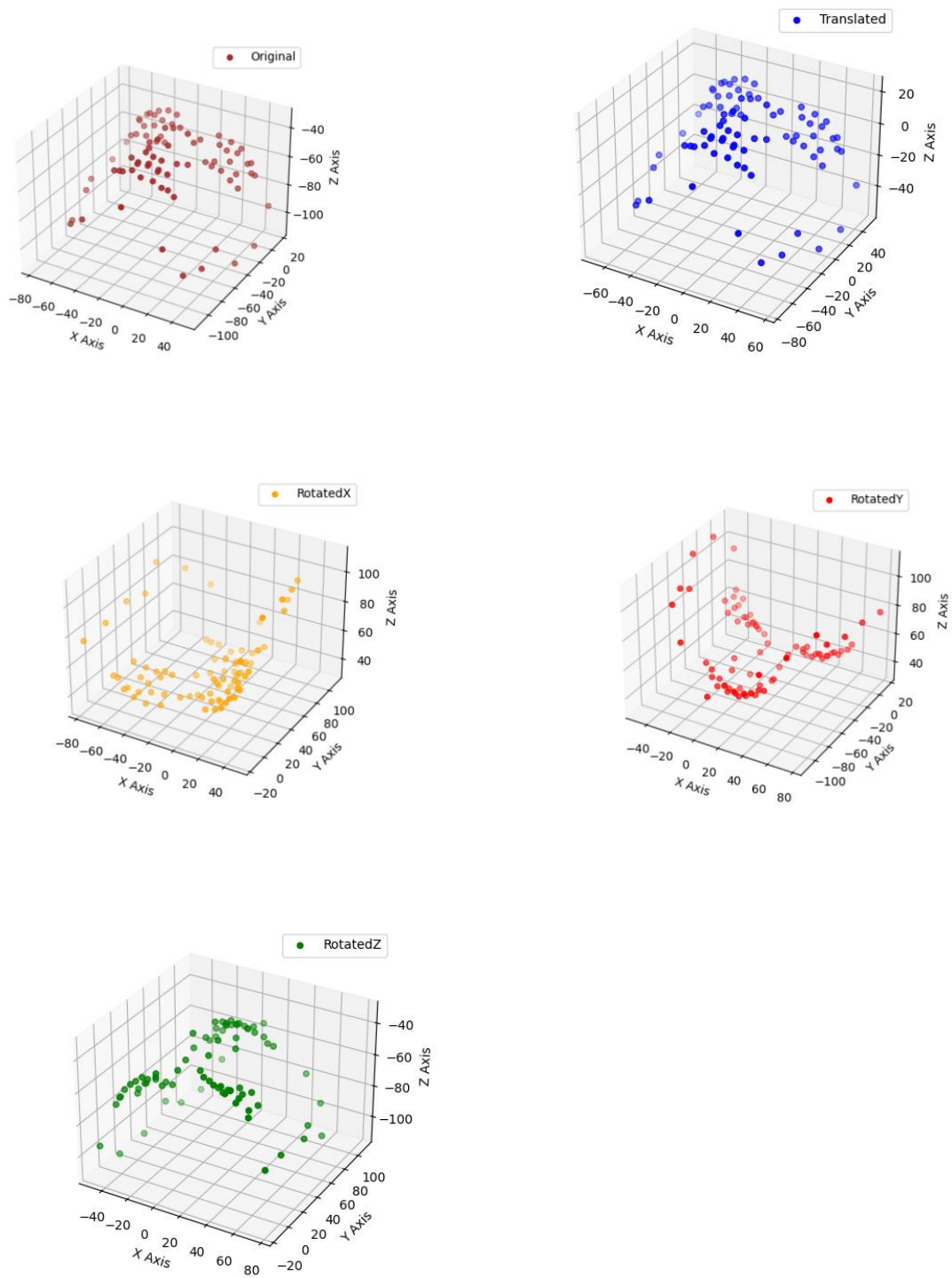
On the other hand, SVM exhibited lower performance, potentially because of the high-dimensional nature of the data and the challenge of identifying the optimal hyperplane for diverse facial expressions.

Decision Trees performed well, particularly with original and rotated data types, thanks to their ability to establish clear decision boundaries based on landmarks. However, they are more susceptible to overfitting compared to Random Forest, which aggregates results from multiple trees to counteract this issue.

The translated data type, despite being centered around the origin, may introduce new variance that classifiers must adjust to, resulting in a slight performance decrease across all classifiers. Nonetheless, Random Forest maintains high accuracy by averaging decisions from multiple trees, effectively managing the variance introduced by translation.

Question 5:

sample of each data type – original, translated, rotated



Answer 6:

Random Forest: Random Forest is an ensemble learning method that operates by constructing a multitude of decision trees during training and outputting the mode of the

classes (classification) or mean prediction (regression) of the individual trees. Each tree in the forest is built using a random selection of features from the training set, which helps to decorrelate the trees and reduce overfitting. During prediction, each tree "votes" for the most popular class, and the class with the most votes is chosen as the final prediction. Random Forest is robust, handles high-dimensional data well, and is less prone to overfitting compared to individual decision trees.

Support Vector Machine (SVM): SVM is a supervised machine learning algorithm used for classification and regression tasks. It works by finding the hyperplane that best separates different classes in the feature space. SVM aims to maximize the margin between the hyperplane and the nearest data points of each class, making it robust to outliers and effective in high-dimensional spaces. In cases where the data is not linearly separable, SVM uses the kernel trick to transform the input space into a higher-dimensional space, where a separating hyperplane exists. SVM is effective for both linear and non-linear classification tasks, although it can be computationally expensive for large datasets.

Decision Tree: Decision Tree is a simple and intuitive supervised learning algorithm used for both classification and regression tasks. It operates by partitioning the feature space into segments, based on the features' values, in a hierarchical manner. Each internal node of the tree represents a decision based on a feature, and each leaf node represents a class label or a numerical value. Decision Trees are easy to interpret and visualize, making them suitable for understanding the underlying decision-making process. However, they are prone to overfitting, especially with complex data, and can create biased trees if not properly pruned or constrained. Various techniques like pruning, limiting tree depth, and using ensembles help mitigate these issues.

DETAILED RESULTS:

1. Random Forest With Original Data

Precision: 0.999900360555471

Recall: 0.9999004131750542

Accuracy: 0.999900670473618

	Angry	Disgust	Fear	Happy	Sad	Surprise
Angry	1012.2	0	0	0	0	0
Disgust	0.2	1017.1	0	0	0	0
Fear	0	0	1004.2	0	0	0
Happy	0	0	0	997.3	0	0
Sad	0	0	0	0	1014.2	0.1
Surprise	0	0	0.2	0	0	994.7

2. Random Forest With Translated Data

Precision: 0.9988257685511014

Recall: 0.9988218471631074

Accuracy: 0.9988245416804157

	Angry	Disgust	Fear	Happy	Sad	Surprise
Angry	1011.5	0.5	0.2	0	0.1	0
Disgust	0.7	1015.4	0.2	0.2	0	0.6
Fear	0	0.5	1003	0.7	0	0.7
Happy	0	0	0.6	996.3	0	0.1
Sad	0.2	0.6	0.1	0.1	1013.9	0.3
Surprise	0	0.1	0.3	0	0.2	993.1

3. Random Forest With Rotated on X-axis Data

Precision: 0.9999170253641042

Recall: 0.999916572125847

Accuracy: 0.9999172240243548

	Angry	Disgust	Fear	Happy	Sad	Surprise
Angry	1012.2	0	0	0	0	0
Disgust	0.2	1017.1	0	0	0	0
Fear	0	0	1004.2	0	0	0
Happy	0	0	0	997.3	0	0
Sad	0	0	0	0	1014.2	0.1
Surprise	0	0	0.2	0	0	994.7

4. Random Forest With Rotated on Y-axis Data

Precision: 0.999916845883529

Recall: 0.9999172822708708

Accuracy: 0.9999172267650087

	Angry	Disgust	Fear	Happy	Sad	Surprise
Angry	1012.2	0	0	0	0	0
Disgust	0.2	1017.1	0	0	0	0
Fear	0	0	1004.2	0	0	0
Happy	0	0	0	997.3	0	0
Sad	0	0	0	0	1014.2	0.1
Surprise	0	0	0.2	0	0	994.7

5. Random Forest With Rotated on Z-axis Data

Precision: 0.9999168949143625

Recall: 0.9999171299921723

Accuracy: 0.9999172240243546

	Angry	Disgust	Fear	Happy	Sad	Surprise
Angry	1012.3	0	0	0	0	0
Disgust	0.1	1017.1	0	0	0	0
Fear	0	0	1004.2	0	0	0.1
Happy	0	0	0	997.3	0	0
Sad	0	0	0	0	1014.2	0.1
Surprise	0	0	0.2	0	0	994.6

6. Decision Tree With Original Data

Precision: 0.9845248416792532

Recall: 0.9845095642324534

Accuracy: 0.9845203416828274

	Angry	Disgust	Fear	Happy	Sad	Surprise
Angry	994.4	4.8	2.6	2.5	1.5	2.4
Disgust	3.9	999.4	5.2	2.3	1.5	4.2
Fear	2.6	5.0	985.3	3.7	1.2	5.5
Happy	2.3	2.6	4.6	983.2	2.4	3.3
Sad	1.9	1.7	2.5	2.8	1005.4	3.7
Surprise	2.3	3.6	4.2	2.8	2.2	975.7

7. Decision Tree With Translated Data

Precision: 0.9584947050832857

Recall: 0.9584414001997494

Accuracy: 0.9584782024817171

	Angry	Disgust	Fear	Happy	Sad	Surprise
Angry	974.8	11.9	6.5	5.8	7.1	6.1

Disgust	11.1	974.	13.1	8.4	4.1	9.6
Fear	6.7	11.2	953.1	13.1	7.3	15.7
Happy	5.9	8.1	12.1	957.4	6.3	9.1
Sad	8.1	5.4	8.4	5.8	981.6	8.1
Surprise	5.8	6.6	11.4	6.9	8.7	946.2

8. Decision Tree With Rotated on X-axis Data

Precision: 0.9855005804514001

Recall: 0.9854870187600507

Accuracy: 0.9854971601342264

	Angry	Disgust	Fear	Happy	Sad	Surprise
Angry	1000.6	4.4	1.8	2.1	2.1	2.6
Disgust	3.7	999.1	5.5	3.1	1.4	4.1
Fear	2.4	5.1	986.2	4.2	1.6	5.9
Happy	1.5	2.5	3.8	981.8	2.1	4.6
Sad	2.3	2.3	2.2	3.1	1004.8	3.3
Surprise	1.9	3.8	4.9	3.2	2.4	974.4

9. Decision Tree With Rotated on Y-axis Data

Precision: 0.9844199127989371

Recall: 0.9844082291216465

Accuracy: 0.984421061488219

	Angry	Disgust	Fear	Happy	Sad	Surprise
Angry	999.8	3.6	2.9	2.3	1.9	2.2
Disgust	3.7	997.7	5.7	2.6	1.6	4.8
Fear	2.8	5.5	984.1	3.6	1.9	5.1
Happy	1.8	3.4	4.4	983.3	1.6	3.7
Sad	2.1	2.1	2.3	2.7	1005.1	3.6
Surprise	2.2	4.8	5.1	2.8	2.2	975.4

10. Decision Tree With Rotated on Z-axis Data

Precision: 0.9846443310041091

Recall: 0.9846376188656587

Accuracy: 0.9846528249018023

	Angry	Disgust	Fear	Happy	Sad	Surprise
Angry	1001.1	4.1	2.3	1.8	2.1	2.4
Disgust	3.5	998.6	4.9	2.4	1.8	3.8
Fear	2.5	4.5	985.5	4.1	1.6	5.6
Happy	2.1	3.2	4.6	983.5	1.6	4.1
Sad	1.5	1.9	2.4	3.1	1005.1	3.8
Surprise	1.8	3.9	4.7	2.4	2.2	975.1

11. SVM With Original Data

Precision: 0.7938306153391884

Recall: 0.7941618243284989

Accuracy: 0.7940631512479295

	Angry	Disgust	Fear	Happy	Sad	Surprise
Angry	845.3	100.5	49.6	37.8	55.1	32.2
Disgust	44.4	751.1	84.5	25.6	18.6	17.8
Fear	25.8	74.6	653.3	44.6	23.7	32.9
Happy	20.8	37.8	75.5	835.1	23.5	30.3
Sad	56.9	23.6	53.2	28.1	870.9	40.9
Surprise	19.2	29.5	88.3	26.3	22.4	840.7

12. SVM With Translated Data

Precision: 0.7892539785508116

Recall: 0.789508983157966

Accuracy: 0.7894275129879598

	Angry	Disgust	Fear	Happy	Sad	Surprise
Angry	841.8	98.3	50.1	40.1	59.6	29.7
Disgust	45.7	749.5	87.1	25.3	18.2	19.6
Fear	23.4	72.3	645.2	42.9	28.2	34.1
Happy	20.5	38.1	81.7	836.3	20.7	34.3
Sad	60.7	27.1	53.8	25.8	867.2	48.8
Surprise	23.3	31.8	86.6	27.1	20.3	828.3

13. SVM With Rotated on X-axis Data

Precision: 0.7937509014980593

Recall: 0.7940602179556135

Accuracy: 0.7939638299435096

	Angry	Disgust	Fear	Happy	Sad	Surprise
Angry	845.8	100.4	49.5	37.9	55.9	32.5
Disgust	44.1	749.6	84.4	25.5	18.1	17.8
Fear	25.6	75.4	652.9	44.5	23.6	32.6
Happy	20.9	38.5	75.6	834.9	23.8	30.5
Sad	57.3	23.3	53.9	27.6	870.3	41.6
Surprise	18.7	29.9	88.1	26.9	22.6	839.8

14. SVM With Rotated on Y-axis Data

Precision: 0.7936850208291734

Recall: 0.7940126782581383

Accuracy: 0.7939141555880294

	Angry	Disgust	Fear	Happy	Sad	Surprise
Angry	845.8	100.1	49.1	38.1	55.4	32.2
Disgust	44.1	750.7	84.4	25.8	18.3	18.1
Fear	25.5	75.1	652.7	43.9	23.6	32.9
Happy	20.8	38.5	76.3	834.2	24.3	29.9
Sad	57.1	23.1	53.3	28.4	870.3	41.4
Surprise	19.2	29.8	88.6	26.9	22.3	840.4

15. SVM With Rotated on Z-axis Data

Precision: 0.793629399944264

Recall: 0.7939625611841148

Accuracy: 0.7938644675292783

	Angry	Disgust	Fear	Happy	Sad	Surprise
Angry	845.7	100.1	49.4	37.6	54.9	32.3
Disgust	44.1	750.5	84.5	25.6	18.3	17.8
Fear	26.1	75.2	652.9	44.5	23.5	33.1
Happy	21.1	38.3	76.3	834.1	23.5	30.1
Sad	56.8	22.8	53.2	28.9	871.9	41.4
Surprise	18.7	30.3	88.1	26.6	22.1	840.2

