Airline Survey Analysis

- Syeda Mariam Danish
- Anugraha Antoo Kanjookaran
- Fiyifoluwa Onafeso
- Ajay Kumar Paidimarri
- Hariharakumar Rathianar

R1: Project Introduction

Airline Dataset

Primary goal:
To predict customer Loyalty based on factors such as age, type of travel, flight distance, seat comfort, and more.

Why?
Customer satisfaction is crucial in the airline industry for improving services, retaining customers, and enhancing overall experiences. By exploring the key drivers behind satisfaction, we can offer insights to help airlines make data-driven decisions for service improvements.

Image

Primary Goal:
To predict the satisfaction level of a passenger by analyzing their facial expressions from survey images.

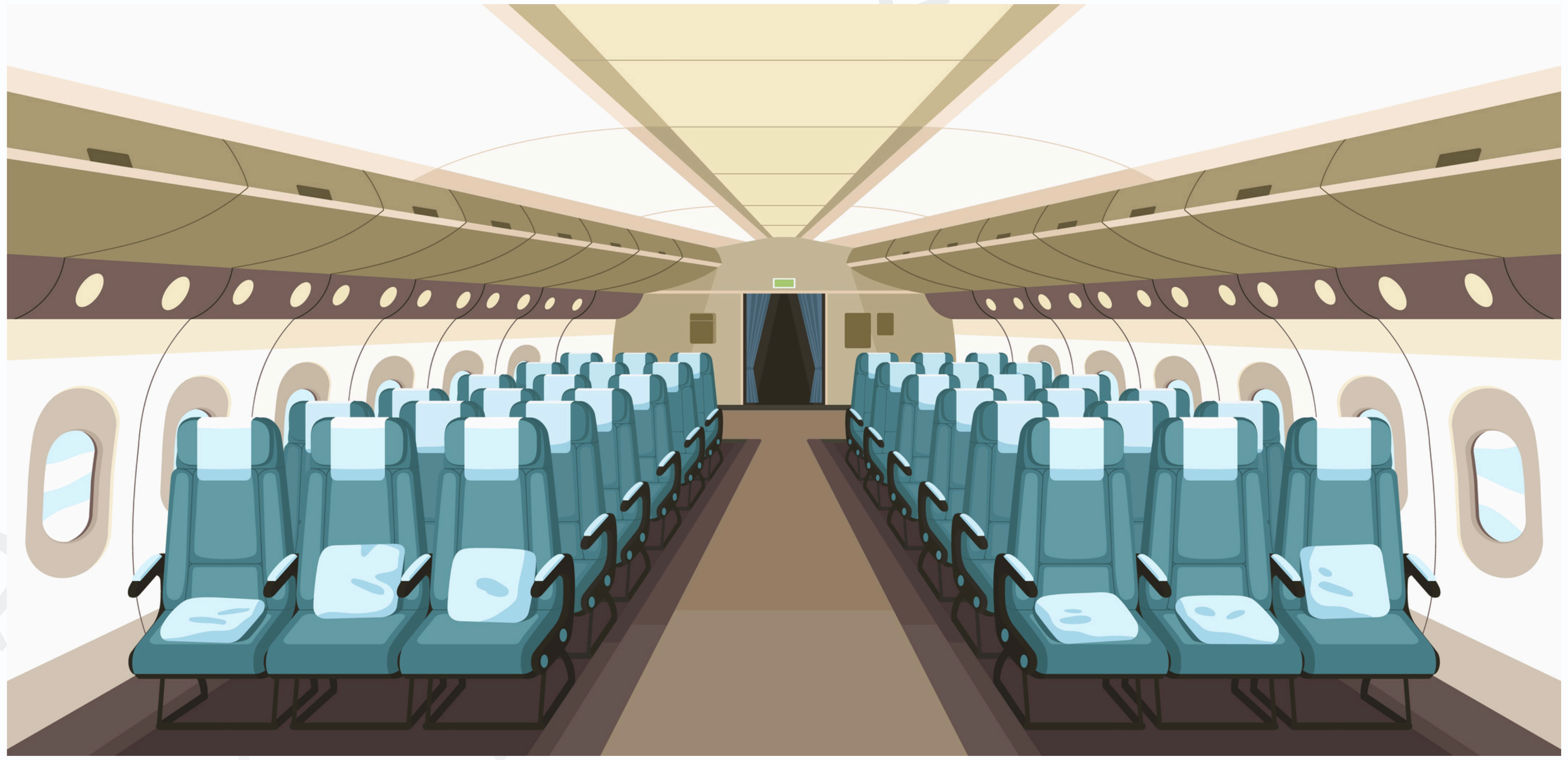
Why?
- Facial expressions often reveal underlying emotions, which can provide insights into a passenger's satisfaction level.
- This approach allows airlines to complement traditional survey data with visual emotion recognition for a holistic understanding of customer sentiment.

R2: Feature Selection

Airline Dataset

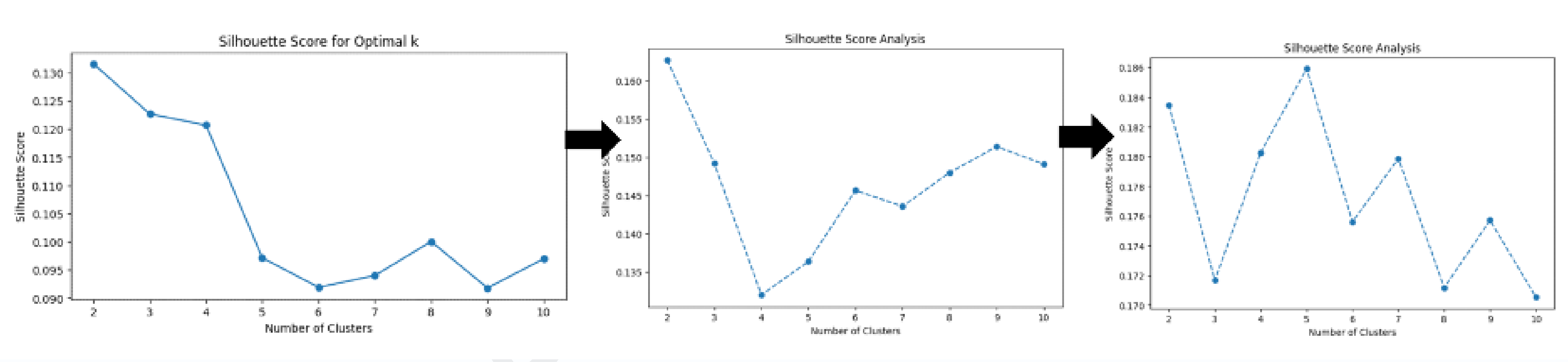
Features are selected using combined computation of different models Correlation Rank, KBest, LASSO Coefficient (Lasso from sklearn.linear_model), Chi-Square, Recursive Feature Elimination (RFE from from sklearn.feature_selection) & ExtraTreesClassifier from sklearn.ensemble

Feature	Importance	TC Rank	Correlation Rank	Kmean Rank	Chi2Score	Lasso	RFE	Overall rank
Type of Travel	0.157687	1	7	1	3	2	2	3
Online Boarding	0.114978	2	4	2	7	1	6	4
In-flight Wifi Service	0.109038	3	17	3	6	14	1	7
Ease of Online Booking	0.086492	4	18	4	8	13	4	9
Age	0.047013	7	3	8	2	4	18	7
In-flight Entertainment	0.046097	8	6	9	11	5	7	8
Flight Distance	0.04094	9	2	7	1	3	21	7
Departure and Arrival Time Convenience	0.052641	6	16	6	10	12	9	10
Seat Comfort	0.038821	11	5	10	12	7	15	10
Class	0.039614	10	15	5	9	22	3	11
Cleanliness	0.025059	15	8	14	14	17	12	13
On-board Service	0.029742	13	11	11	13	9	11	11
Leg Room Service	0.026262	14	10	13	15	8	14	12
In-flight Service	0.030306	12	19	12	18	16	8	14
Gate Location	0.057054	5	22	19	21	11	5	14
Baggage Handling	0.02316	16	13	16	19	10	13	15
Check-in Service	0.021555	17	14	15	16	22	10	16
Food and Drink	0.021267	18	12	18	20	15	16	17
Gender	0.011835	19	9	17	17	6	17	14
Arrival Delay	0.010298	20	20	20	4	22	19	18
Departure Delay	0.010139	21	21		5	22	20	18

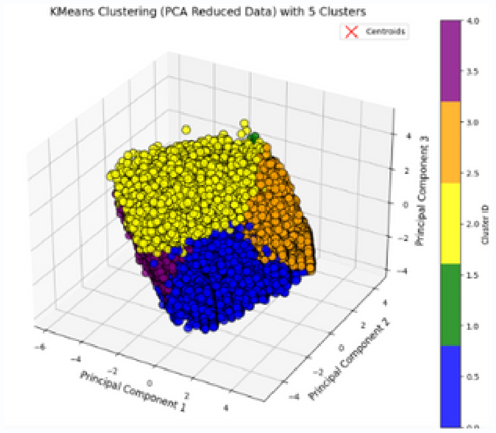


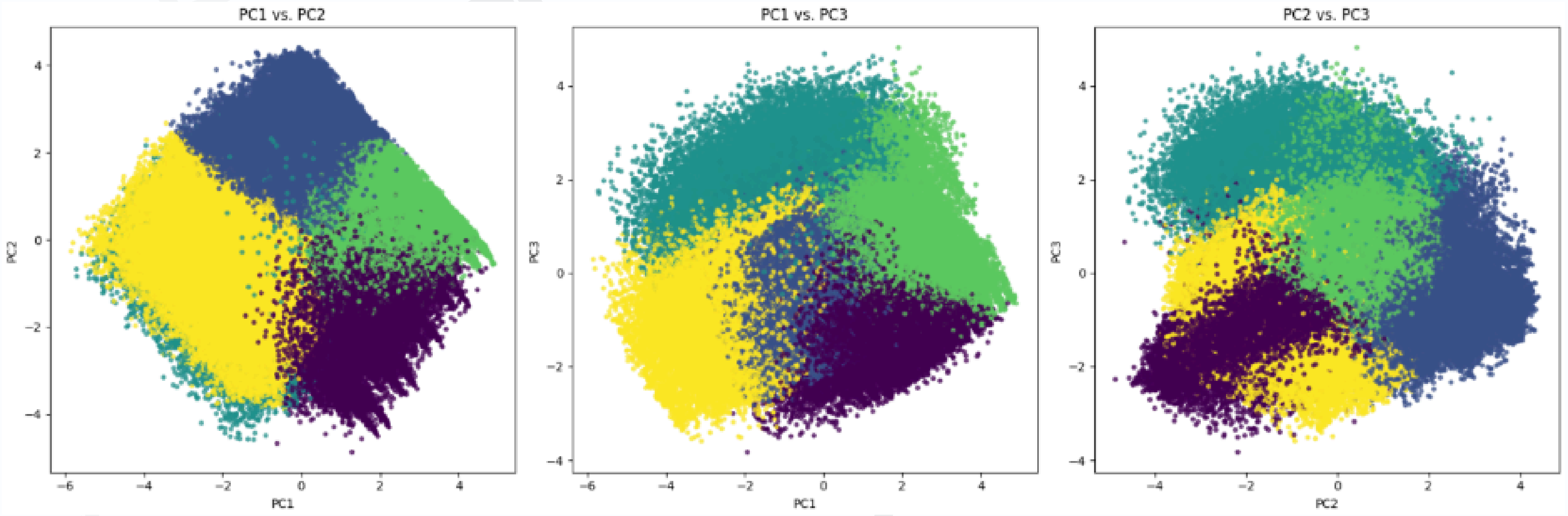
R3: Clustering

- Feature selected data with 8 features had the best kmean Silhouette peak

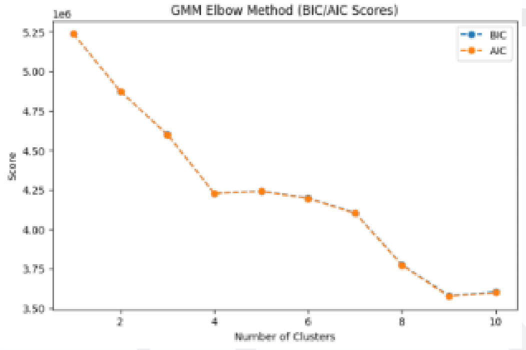


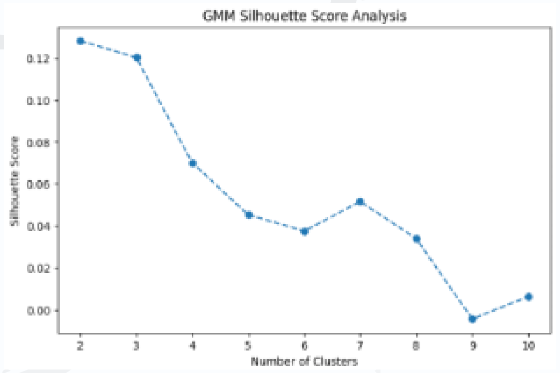
- Kmean with 5 clusters

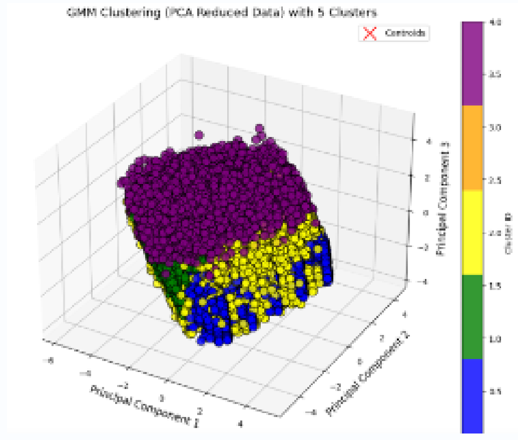




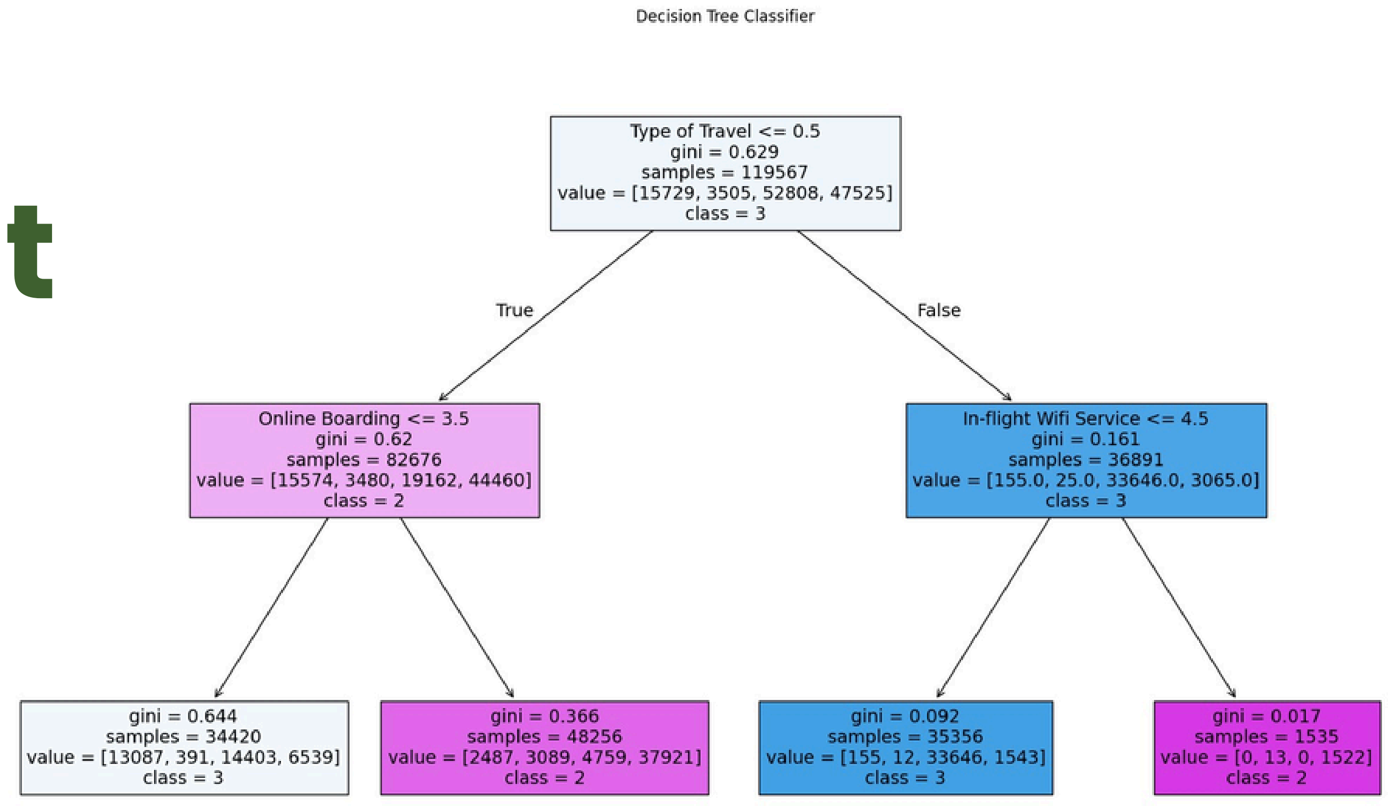
- GMM and Hierarchy Clustering werent effective







R4: Baseline Training Dataset



Naive Bayes Algorithm	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**	**Precision (%)**	**Recall (%)**	**ROC AUC**
Multinomial Naive Bayes	49.42	49.34	83.13	53.64	49.42	71.17
Gaussian Naive Bayes	78.72	78.71	92.91	80.22	78.72	94.59
Complement Naive Bayes	46.68	46.53	82.20	56.81	46.68	72.10
Bernoulli Naive Bayes	50.38	50.33	83.43	45.12	50.38	73.23
Categorical Naive Bayes	**82.37**	**82.37**	**94.13**	**82.86**	**82.37**	**95.47**

Solver	Max Iter	Batch Size	C	Accuracy	Precision	Recall	F1-Score	Observation
lbfgs	300	16	Default	81%	81.38%	81.18%	81.17%	Stable performance with default regularization and iterations; converges quickly.
lbfgs	1000	32	Default	81%	81.38%	81.18%	81.17%	No significant change in performance with increased iterations.
lbfgs	2000	64	Default	81%	81.38%	81.18%	81.17%	Larger batch size and more iterations did not significantly impact accuracy.
saga	300	16	0.1	81%	81.37%	81.18%	81.16%	Reduced regularization (C=0.1) yielded stable performance; effective for large datasets.
saga	1000	32	1	81%	81.37%	81.18%	81.16%	Default regularization (C=1) and more iterations maintained stability.
saga	2000	64	10	81%	81.37%	81.18%	81.16%	Increased regularization (C=10) did not negatively impact performance.
newton-cg	500	32	Default	81%	81.37%	81.18%	81.16%	Performs well with larger datasets; requires slightly more iterations to converge.
newton-cg	1500	64	1	81%	81.37%	81.18%	81.16%	Stable performance across increased batch sizes and regularization.
newton-cg	2000	16	0.5	81%	81.37%	81.18%	81.16%	Lower regularization (C=0.5) did not yield any noticeable improvement in accuracy.

Findings:
- Accuracy:
  - Decision Tree: Accuracy : 76%
  - Logistic Regression : Accuracy : 81%
  - Naive Bayes : Accuracy : 82.3% (Categorical Naive Bayes)

R4: Baseline Training Image

Establish baseline accuracy using simpler models to identify dataset trends and challenges.

Steps Taken

Preprocessing:

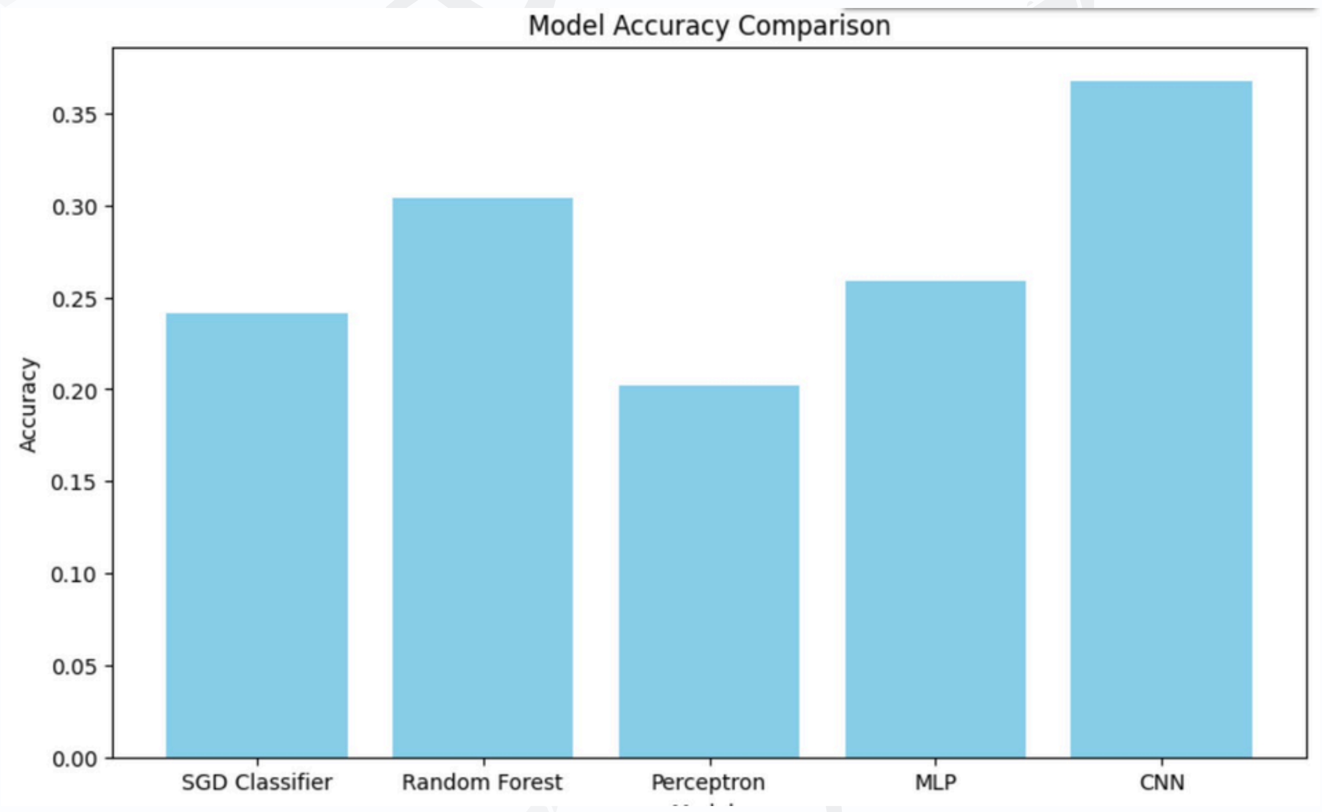
- Images resized to 56x56 pixels and converted to grayscale for simplicity.
- Normalized pixel values between 0 and 1.

Visualization:

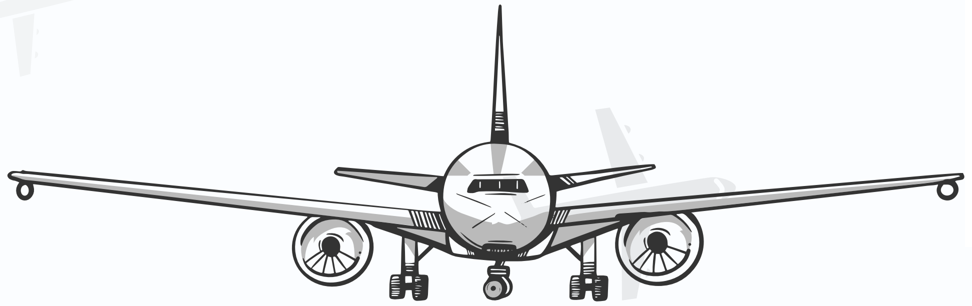
- Displayed sample images for each emotion class.
- Checked class distribution and identified imbalance issues (e.g., "Disgust" had fewer samples).

Findings:

- Accuracy: Baseline models achieved ~24%–30% accuracy.
- Challenges:
  - Struggled to capture spatial features in images.
  - Poor generalization due to limited complexity.
  - Imbalance in data affected minority classes.



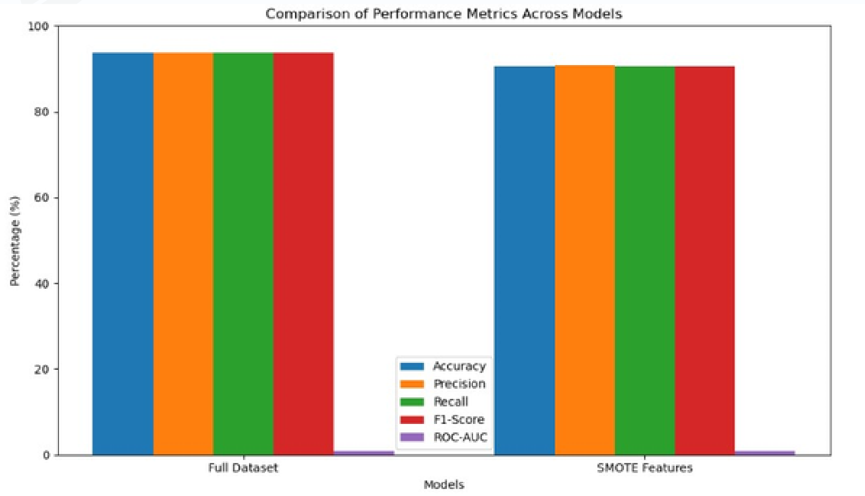
Model	Accuracy	Precision	Recall	F1 Score	ROC Score
Stochastic Gradient Descent	24%	20%	24%	19%	n/a
Random Forest Classifier	30%	28%	30%	26%	n/a
Perceptron	20%	21%	20%	13%	54%
Multi Layer Perceptron	26%	17%	26%	11%	50%
Convolutional Neural Network	37%	34%	37%	34%	70%



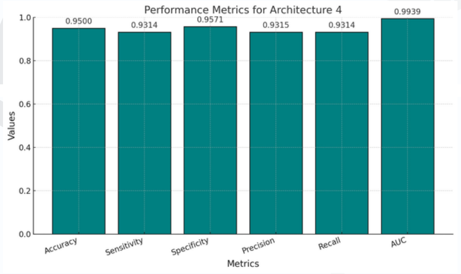
R5: Neural Networks Dataset

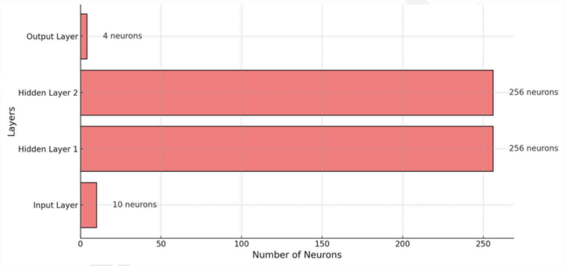
Airline Dataset

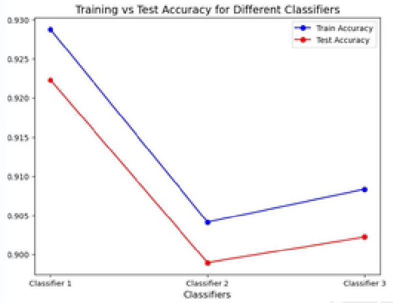
- 10 Folds Cross Validation
  Metric performance for Entire data and selected features.

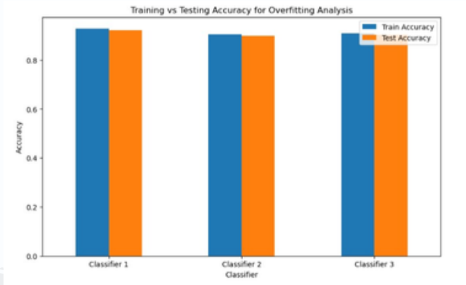


- Hyperparameter tuning
- Choosing Best Architecture
- Classifiers
- Overfitting Analysis
- Conclusion the data generalizes well as we can observe smaller gaps between train and test accuracies.

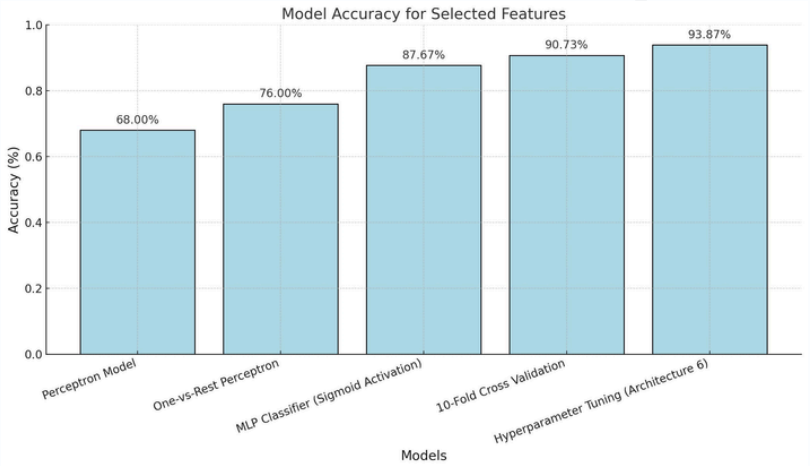


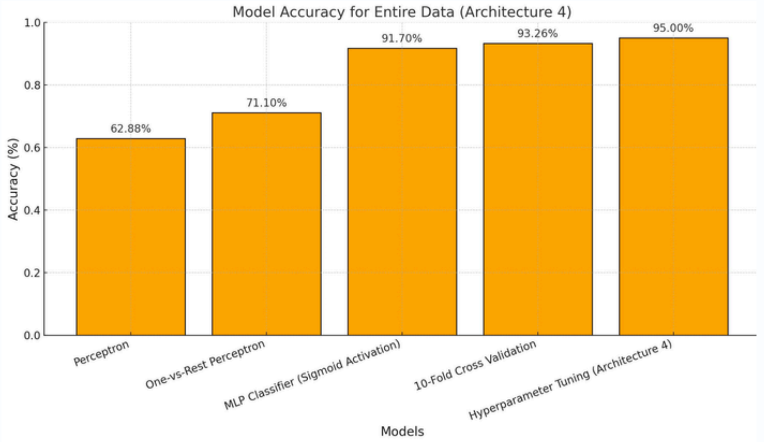






- Hyperparameter-tuned MLP architecture is the best choice, achieving 93.87% accuracy with selected features and 95.00% with the full datasets.
- MLP classifier with sigmoid activation (without hyperparameter tuning) provides a strong balance of simplicity and performance.





R5: Neural Networks Image

Improve emotion classification using CNNs to capture spatial relationships in image data.

Our first CNN Model had a basic requirements but after we optimised it we increased the accuracy from 37% to 65% accuracy, this what we did

Optimized CNN:

- Added more convolutional layers (4 layers with filters like 64, 128, 512) to improve feature extraction.
- Introduced pooling layers to reduce computational complexity while retaining key spatial features.
- Included dropout layers to mitigate overfitting during training.

Training Process:

- Data Augmentation: Generated new training samples using rotation, flipping, and zooming to address class imbalance.
- Batch Normalization: Stabilized learning and improved convergence speed.
- Optimization: Used the Adam optimizer for efficient training.

Aspect	Baseline Models	CNN
Accuracy	~24%–30%	~65%
Feature Extraction	Handcrafted (pixel values)	Automated (learns spatial patterns)
Scalability	Limited to small datasets	Scalable to larger datasets
Computational Cost	Low	High
Generalization	Poor	Better with unseen data

CONCLUSION

Dataset

The results show our data has both linear and nonlinear patterns, with MLP excelling (93.87% accuracy)at capturing complexity. Logistic Regression (81% accuracy) suggests some linear separability, while Naïve Bayes' high specificity (94.13%) points to potential class imbalance. Decision Tree performance highlights varying feature importance. Overall, good preprocessing ensured consistent results, but there's room for feature and model refinement.

Images

The optimized CNN demonstrated superior performance in classifying emotions from images, achieving higher accuracy and better generalization compared to baseline models. By leveraging deeper architectures, advanced techniques like dropout, and effective data augmentation, the model successfully captured complex patterns such as textures and edges. While challenges like class imbalance and computational complexity remain, the results highlight the potential of CNNs for robust image classification tasks. Future improvements, such as transfer learning or larger datasets, could further enhance performance.