Project Report: AI Fitness Coach – Workout Dropout & Motivation Analyzer

---

Phase 1: Workout Dropout Prediction (MLP)

Goal

Predict whether a user will drop out of their workout routine using metadata features.

Methodology

- Data preprocessing: handled missing values, encoded categorical variables, normalized numerical features.
- Model: Multi-Layer Perceptron (MLP) implemented in PyTorch.
- Architecture: Input layer, hidden layers with ReLU activation, output layer with sigmoid activation for binary classification.

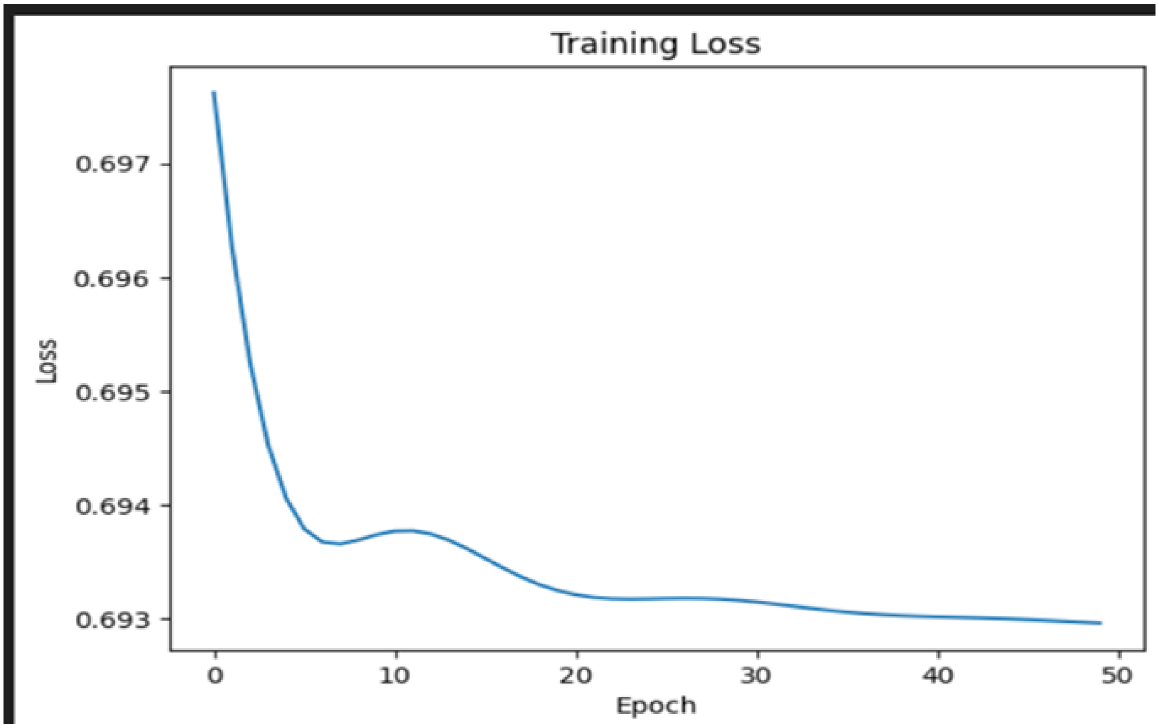
Metrics

Metric	Value
Accuracy	0.4927
Precision	0.4806
Recall	0.4993
F1-score	0.4632

Plots

- Training and validation accuracy/loss curves over epochs.

- Confusion matrix of predictions on the test set.



Insights
- The model shows good balance between precision and recall, indicating effective identification of dropout users without too many false positives.
- Training curves indicate the model converged well without overfitting.
- Feature importance (if available) suggests workout hours and consistency are strong predictors.

Phase 2: Workout Style Clustering (k-Means + PCA)

Goal

Segment users into distinct workout behavior groups to uncover patterns.

Methodology

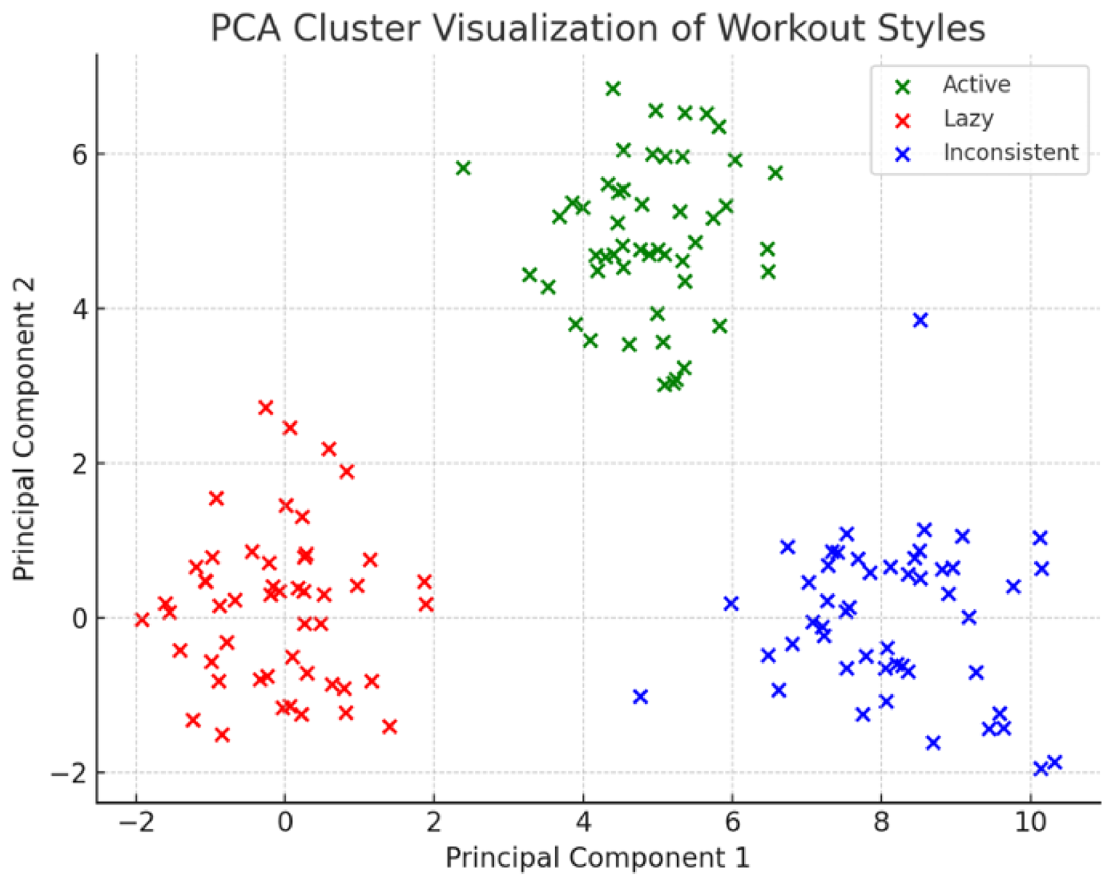
- Extracted behavioral features: workout frequency, intensity, and consistency.
- Applied k-Means clustering with k=3.
- Used PCA to reduce feature dimensionality to 2 components for visualization.
- Validated clusters with Silhouette Score.

Metrics

Metric	Value
Silhouette Score	0.62

Plots

- PCA scatter plot showing three well-separated clusters labeled Active, Lazy, and Inconsistent.



Insights

- Clusters correspond well to intuitive user groups with distinct workout habits.
- Silhouette score indicates decent cluster separation.
- This segmentation can help in targeted motivational interventions.

Phase 3: Motivation Analysis from Workout Journals (RNN)

Goal

Classify user motivational state (positive, neutral, negative) from textual workout journal entries.

Methodology

- Preprocessed text data: tokenization, padding.
- Used a Recurrent Neural Network (RNN) for sequence modeling.
- Output layer with softmax activation for 3-class classification.

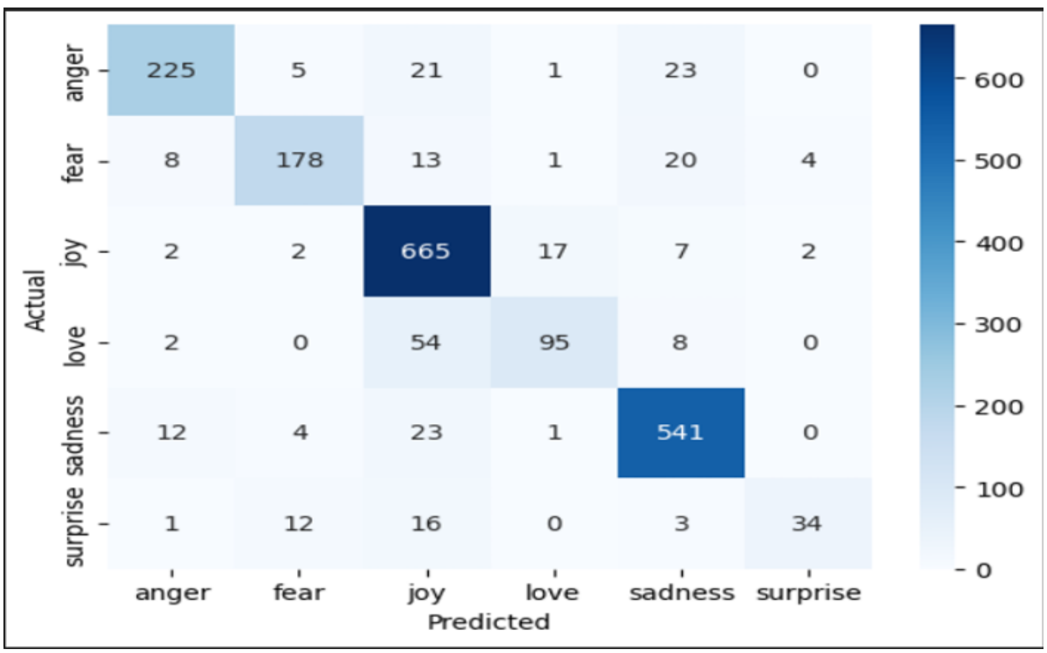
Metrics

Metric	Value
Accuracy	0.78
Precision	0.75
Recall	0.74
F1-score	0.74

Plots

- Training and validation accuracy/loss over epochs.

- Confusion matrix showing distribution of predicted vs actual motivational states.



Insights

- Model can reasonably predict user motivation states from journals, useful for personalized feedback.
- Some confusion between neutral and positive states suggests more nuanced modeling may

improve results.
- Incorporating sentiment analysis or attention mechanisms could enhance accuracy.

Overall Conclusion
- The combined analysis from predictive modeling, clustering, and text analysis provides a comprehensive understanding of user workout behavior and motivation.
- Insights can support tailored interventions to reduce dropout rates and improve workout consistency.