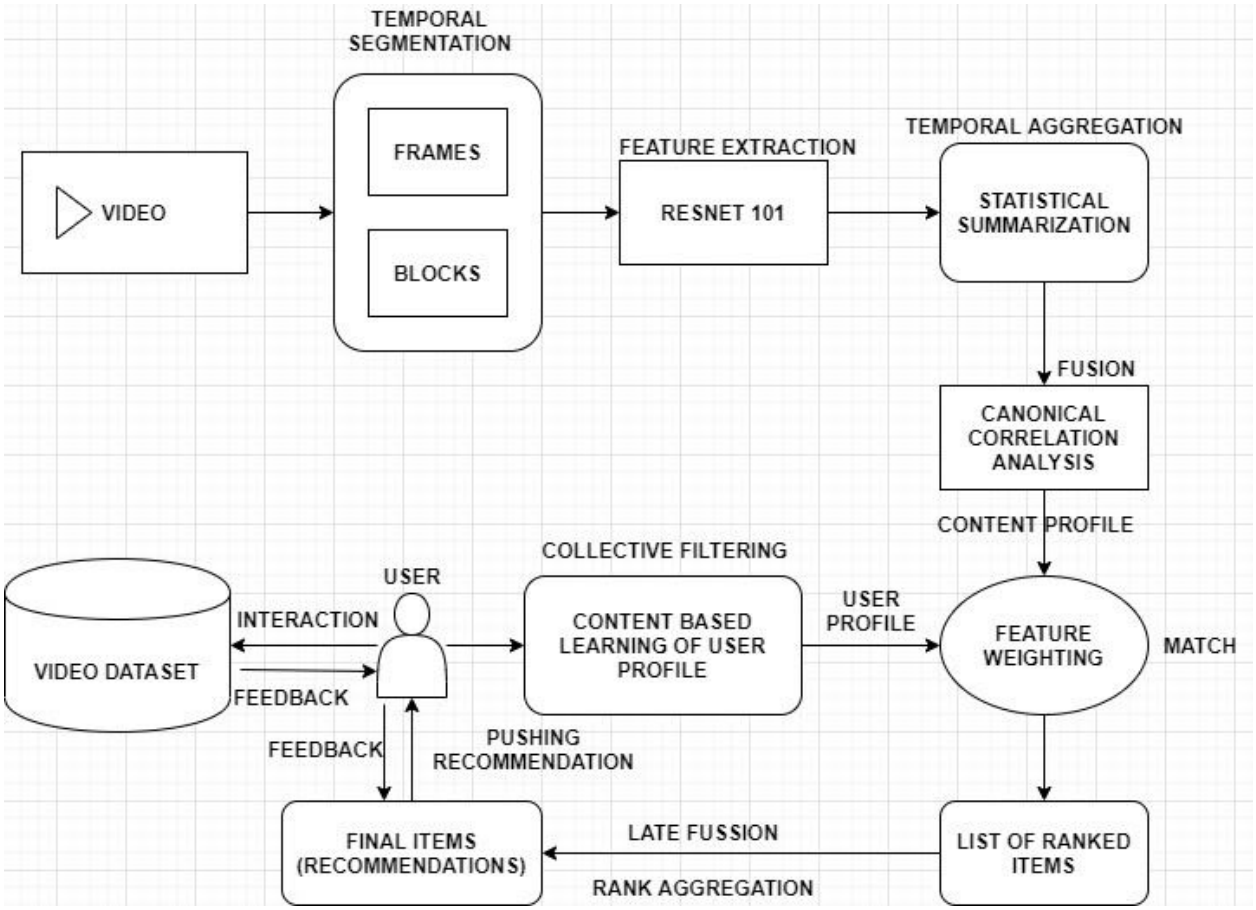

****Since we don't have sufficient labels to provide feed, we take a different approach.****

TEMPORAL SEGMENTATION:

This is the way of collecting human motions from the video through frame by frame format which are processed as blocks.

FEATURE EXTRACTION:

A deep neural network is used to extract features from the video frames and send them for further processing. We use RESNET101 in this system.

TEMPORAL AGGREGATION:

This step involves creating a video-level descriptor by aggregating the features temporally. The following approach is used.
statistical summarization: it is the simplest approach using the operators mean, standard deviation, median, maximum, or combinations thereof, e.g.,means plus covariance matrix to build an item-level descriptor. Upto this a profile about the content is created.

MATCHING:

The info gained from user profile and content profile is matched.

LIST OF RANKED ITEMS:

The content items and the user profile are ranked accordingly and matched with each other for recommendation.

FINAL ITEMS:

The recommendations are now pushed to user and he provides with either +ve, -ve or NULL feedback.

VIDEO DATASET:

The interactions and feedbacks are saved in a dataset and used for the next time when the users uses the platforms.