Software Requirements Specification

For

Virtual Gym Assistant

Date:14th October 2022

Prepared by

Specialisation	SAP ID	Student Name
AI & ML	500075162	ARRA SAI NITISH REDDY
AI & ML	500075774	ARPIT
AI & ML	500075813	BHAVESH
AI & ML	500075953	ARIHANT JAIN

Department of Informatics
School Of Computer Science
UNIVERSITY OF PETROLEUM & ENERGY STUDIES,
DEHRADUN- 248007. Uttarakhand

Table of Contents

1. INTRODUCTION

1.1. Purpose of the Project

Engaging in physical exercises such as weight training and physiotherapy stretching exercises at home requires proper execution and awareness of the exercises. One of the main problems in engaging exercises at home is that there is no proper guidance and feedback provided to align the exercises to the correct movements due to the absence of a physical trainer.

Physical exercises are considered as an important factor that contributes to physical health. They can be categorised into four types depending on the overall effect on the human body. The categories are flexible exercises such as physiotherapy and stretching exercises, aerobic exercises such as walking and running, anaerobic exercises such as weight training exercises and functional training, and muscular endurance and strength exercise

1.2. Target Beneficiary

Strength training exercises are essential for human health. However, when performed incorrectly, humans are at risk of injuring themselves. As such, they often invest in costly personal trainers and group fitness classes in order to have access to a fitness expert who will guide the individual in performing exercises safely. We aimed to increase access to affordable trainers by creating a tool that uses deep learning in order to classify squats as ’correct’ or as one of six common types of ’Incorrect.’These options are inaccessible to individuals with low income and are particularly difficult to use today, in the midst of a pandemic.

1.3. Project Scope

This has become especially relevant with the happening of Covid-19 pandemic in 2020 which affected everyone. Throughout this period, people are prevented from socializing, discouraged from going outside, businesses are shut, gyms are closed, and takeaways are cheap. It becomes very easy for people to forget to look after themselves. Especially where people are expected to carry on with all usual responsibilities such as professional or University work. At Universities, the continually increasing emphasis on mental health from students has become more significant, applying more pressure and responsibilities than ever before to these institutes. On the scale of a community, this can be a seemingly impossible and increasingly complicated problem to manage. To a computer scientist, it is understood that the solution to any problem is also the sum of the solutions to all its counterparts. Even if somebody could not solve the entire problem, targeting some of the problems may bring us much closer to a complete solution. Motion wishes to try and help some of these problems on the scale of the individual. In England, it is estimated that 1 in 4 people will suffer from mental health problems each year (Mind.org.uk, 2020) 17% of these are in the form of anxiety and depression. People who took part in a recent Harvard study have shown that introducing a form of minor activity daily such as an hour of walking, stretching, or any kind of low-intensity movement can reduce the risk of major depression by 26%, and on top of that significantly help individuals from relapsing (Harvard, 2019). This does, however, require individuals to take some responsibility for themselves.

1.4. References

https://arxiv.org/pdf/2006.10204v1.pdf

[1] Ogata R. et al. Temporal distance matrices for squat classification. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2019, pp. 2533-2542.

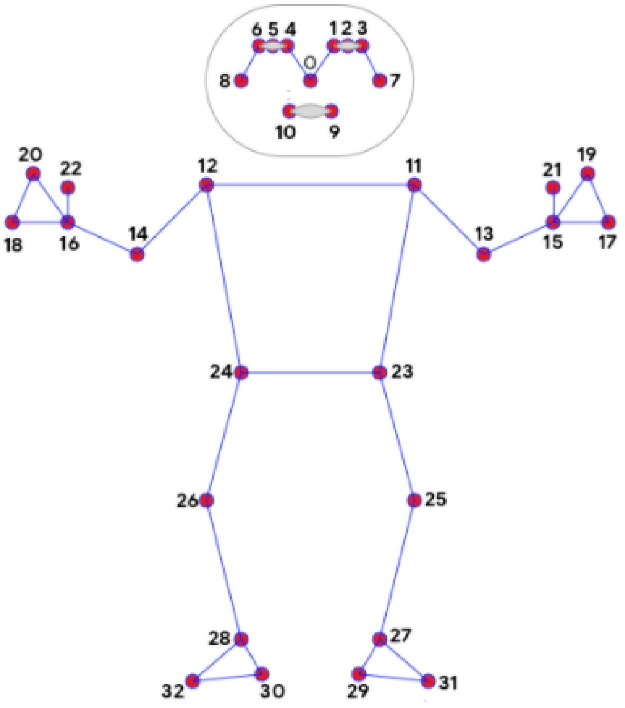
https://google.github.io/mediapipe/

2. PROJECT DESCRIPTION

2.1. Reference Algorithm

- **MediaPipe (Pose detection pipeline):**
  It is an ML solution for high-fidelity body pose tracking, inferring 33 3D landmarks and background segmentation mask on the whole body from RGB video frames utilising the BlazePose research.

  BlazePose is a lightweight convolutional neural network architecture for human pose estimation that is tailored for real-time inference on mobile devices. During inference, the network produces 33 body key points for a single person and runs at over 30 frames per second This makes it particularly suited to real-time use cases like fitness tracking The main features include a novel body pose tracking solution and a lightweight body pose estimation neural network that uses both heatmaps and regression to **keypoint Coordinates**.



0. nose
1. left_eye_inner
2. left_eye
3. left_eye_outer
4. right_eye_inner
5. right_eye
6. right_eye_outer
7. left_ear
8. right_ear
9. mouth_left
10. mouth_right
11. left_shoulder
12. right_shoulder
13. left_elbow
14. right_elbow
15. left_wrist
16. right_wrist
17. left_pinky
18. right_pinky
19. left_index
20. right_index
21. left_thumb
22. right_thumb
23. left_hip
24. right_hip
25. left_knee
26. right_knee
27. left_ankle
28. right_ankle
29. left_heel
30. right_heel
31. left_foot_index
32. right_foot_index

Machine Learning Classification Algorithms

- **SVM**

  A Support Vector Machine, or SVM, is a non-parametric supervised learning model. For nonlinear classification and regression, they utilise the kernel trick to map inputs to high-dimensional feature spaces. SVMs construct a hyper-plane or set of hyper-planes in a high or infinite dimensional space, which can be used for classification, regression or other tasks. Intuitively, a good separation is achieved by the hyper-plane that has the largest distance to the nearest training data points of any class (so-called functional margin), since in general the larger the margin the lower the generalisation error of the classifier. The figure to the right shows the decision function for a linearly separable problem, with three samples on the margin boundaries, called “support vectors".
- **KNN**

  k-Nearest Neighbors is a clustering-based algorithm for classification and regression. It is a type of instance-based learning as it does not attempt to construct a general internal model, but simply stores instances of the training data. Prediction is computed from a simple majority vote of the nearest neighbours of each point: a query point is assigned the data class which has the most representatives within the nearest neighbors of the point.

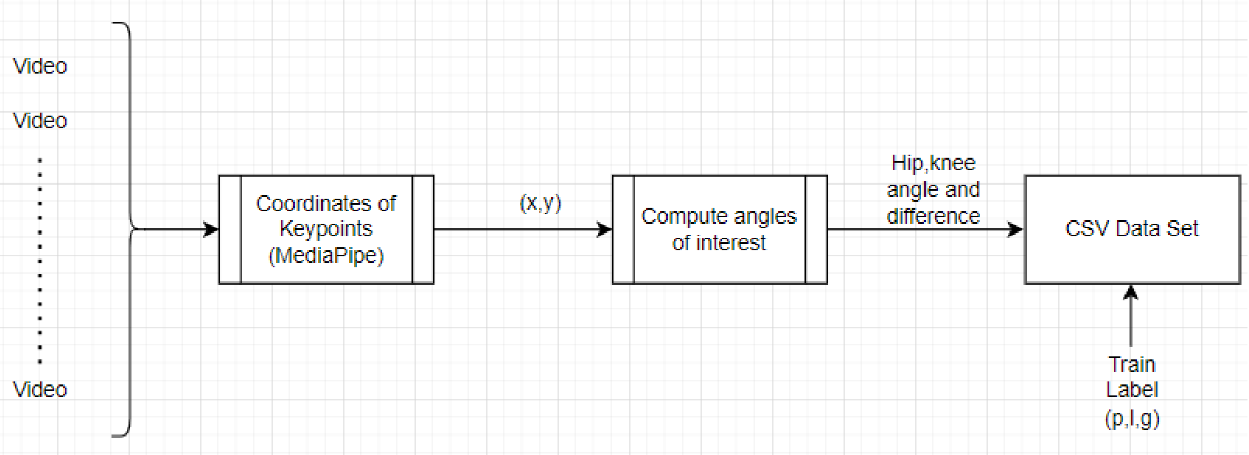
- **Naive Bayes**

  The Bayesian Classification represents a supervised learning method as well as a statistical method for classification. Assumes an underlying probabilistic model and it allows us to capture uncertainty about the model in a principled way by determining probabilities of the outcomes. This Classification is named after Thomas Bayes (1702-1761), who proposed the Bayes Theorem. Bayesian classification provides practical learning algorithms and prior knowledge and observed data can be combined. Bayesian Classification provides a useful perspective for understanding and evaluating many learning algorithms. It calculates explicit probabilities for hypotheses and it is robust to noise in input data. In statistical classification, the Bayes classifier minimizes the probability of misclassification. That was a visual intuition for a simple case of the Bayes classifier, also called: 1)Idiot Bayes 2)Naive Bayes 3)Simple Bayes
- **Neural networks**

  Neural networks, also known as artificial neural networks (ANNs) or simulated neural networks (SNNs), are composed of a node layer, containing an input layer, one or more hidden layers, and an output layer. Each node, or artificial neuron, connects to another and has an associated weight and threshold. If the output of any individual node is above the specified threshold value, that node is activated, sending data to the next layer of the network. Otherwise, no data is passed along to the next layer of the network.

2.2. **Characteristic of Data:**

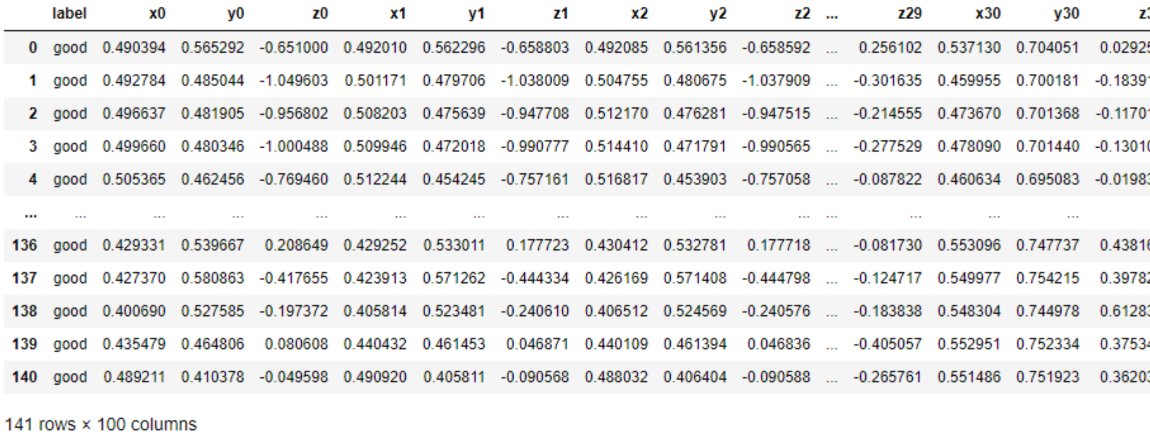
- A video dataset is used for training the model with respected labels below is the flow of data generation


Video
Video
Video
Coordinates of Keypoints (MediaPipe)
(x,y)
Compute angles of interest
Hip,knee angle and difference
CSV Data Set
Train Label (p,l,g)

- Mediapipe is used to get the coordinates of the key points which are used to compute angle of interest

    Hip angle (via keypoints 11, 23, 25)
    The absolute difference between knee (23) and hip (25) key points
    Knee angle (via keypoints 23, 25, 27)

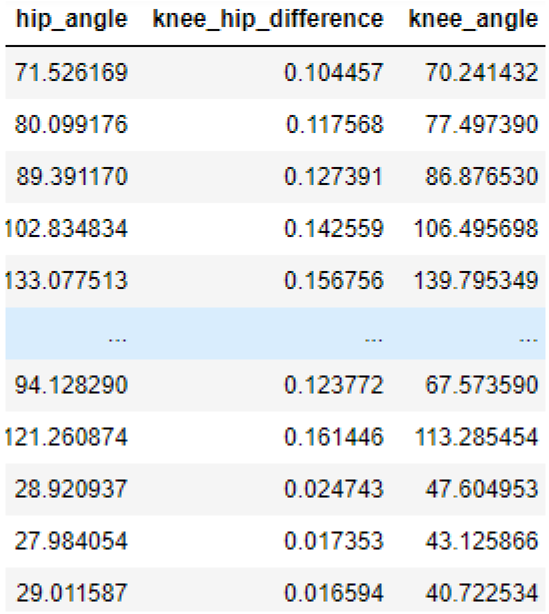
- 33 key points are extracted for the video frame by frame.



	label	x0	y0	z0	x1	y1	z1	x2	y2	z2	...	z29	x30	y30	z3
0	good	0.490394	0.565292	-0.651000	0.492010	0.562296	-0.658803	0.492085	0.561356	-0.658592	...	0.256102	0.537130	0.704051	0.02925
1	good	0.492784	0.485044	-1.049603	0.501171	0.479706	-1.038009	0.504755	0.480675	-1.037909	...	-0.301635	0.459955	0.700181	-0.18391
2	good	0.496637	0.481905	-0.956802	0.508203	0.475639	-0.947708	0.512170	0.476281	-0.947515	...	-0.214555	0.473670	0.701368	-0.11701
3	good	0.499660	0.480346	-1.000488	0.509946	0.472018	-0.990777	0.514410	0.471791	-0.990565	...	-0.277529	0.478090	0.701440	-0.13010
4	good	0.505365	0.462456	-0.769460	0.512244	0.454245	-0.757161	0.516817	0.453903	-0.757058	...	-0.087822	0.460634	0.695083	-0.01983
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
136	good	0.429331	0.539667	0.208649	0.429252	0.533011	0.177723	0.430412	0.532781	0.177718	...	-0.081730	0.553096	0.747737	0.43816
137	good	0.427370	0.580863	-0.417655	0.423913	0.571262	-0.444334	0.426169	0.571408	-0.444798	...	-0.124717	0.549977	0.754215	0.39782
138	good	0.400690	0.527585	-0.197372	0.405814	0.523481	-0.240610	0.406512	0.524569	-0.240576	...	-0.183838	0.548304	0.744978	0.61283
139	good	0.435479	0.464806	0.080608	0.440432	0.461453	0.046871	0.440109	0.461394	0.046836	...	-0.405057	0.552951	0.752334	0.37534
140	good	0.489211	0.410378	-0.049598	0.490920	0.405811	-0.090568	0.488032	0.406404	-0.090588	...	-0.265761	0.551486	0.751923	0.36203

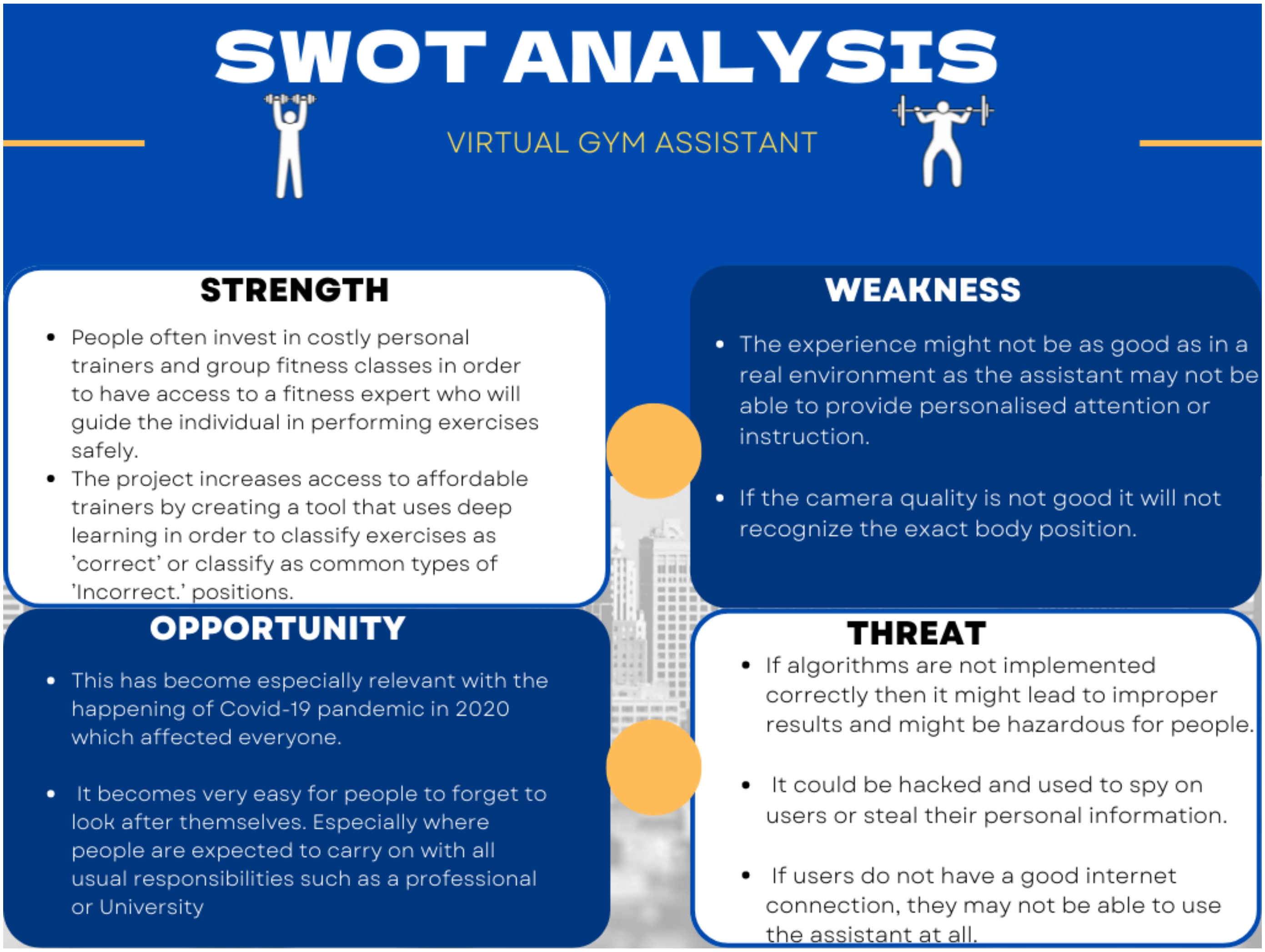
141 rows × 100 columns

- Three angle of interest are computed as above



hip_angle	knee_hip_difference	knee_angle
71.526169	0.104457	70.241432
80.099176	0.117568	77.497390
89.391170	0.127391	86.876530
102.834834	0.142559	106.495698
133.077513	0.156756	139.795349
...	...	...
94.128290	0.123772	67.573590
121.260874	0.161446	113.285454
28.920937	0.024743	47.604953
27.984054	0.017353	43.125866
29.011587	0.016594	40.722534

2.3. SWOT Analysis



SWOT ANALYSIS

VIRTUAL GYM ASSISTANT

STRENGTH

- People often invest in costly personal trainers and group fitness classes in order to have access to a fitness expert who will guide the individual in performing exercises safely.
- The project increases access to affordable trainers by creating a tool that uses deep learning in order to classify exercises as 'correct' or classify as common types of 'Incorrect.' positions.

WEAKNESS

- The experience might not be as good as in a real environment as the assistant may not be able to provide personalised attention or instruction.
- If the camera quality is not good it will not recognize the exact body position.

OPPORTUNITY

- This has become especially relevant with the happening of Covid-19 pandemic in 2020 which affected everyone.
- It becomes very easy for people to forget to look after themselves. Especially where people are expected to carry on with all usual responsibilities such as a professional or University

THREAT

- If algorithms are not implemented correctly then it might lead to improper results and might be hazardous for people.
- It could be hacked and used to spy on users or steal their personal information.
- If users do not have a good internet connection, they may not be able to use the assistant at all.

2.4. Project Features

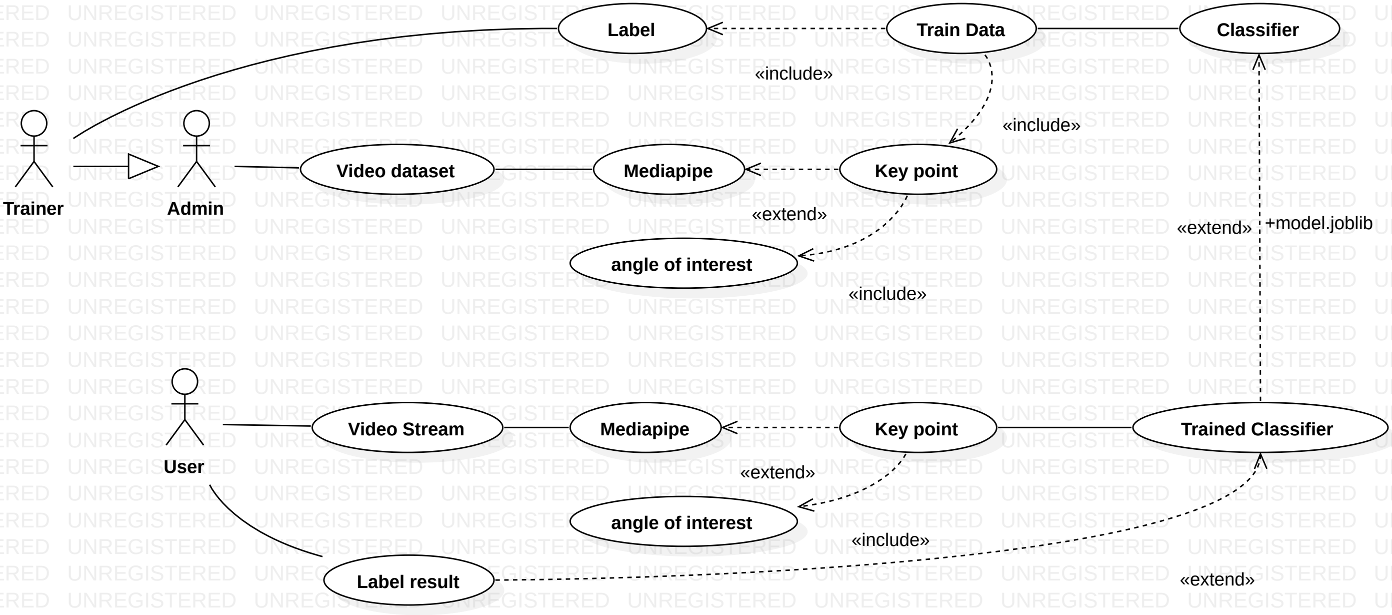
- Live feed analysis of various exercises
- Instant form feedback received to correct exercise execution depending upon common errors in exercise execution
- Supports major exercises like Squats, Overhead Press, Dumbbell Curls etc
- Use body angles and heuristics to check form form correctness

2.5. Design and Implementation Constraints

- Video data has to be manually labelled manually frame-by-frame using utility code in Phase-1 file and converted into .CSV file
- Form depends upon angle analysis between various joints detected by MediaPipe pose detection and form varies from person to person depending on individual joint angles

Model::UseCaseDiagram1

Use Case Diagram



2.7. Assumption and Dependencies

- For form feedback to work, the user's live feedback should be accessible and the video feed should be clear enough to detect user key points - Analysis is dependent upon general form correction and on user body angles

3. SYSTEM REQUIREMENTS

3.1. User Interface: Can be integrated with a web app or a mobile app for easier access

3.2. Software Interface: Software is accessed using Jupyter Notebooks for training and testing, live video feed analysis uses live camera feed from the user to access the training form and suggest instant feedback based on execution of exercise

3.3. Database Interface: A flat file database. CSV is used for storing video-related data

4. NON-FUNCTIONAL REQUIREMENTS

4.1. Performance requirements: Applications should be highly performant for delivering live feedback on exercise form by fastly and accurately examining frames

4.2. Software Quality Attributes

4.3. **Quality Attributes**	**Performance**
adaptability	Medium
availability	Medium
correctness	Medium
flexibility	High
interoperability	Low
maintainability	High
portability	High
reliability	Medium
reusability	High
robustness	High
testability	Medium
usability	High