Human Pose estimation 101:

## Basics:

* Pose estimation is defined as the problem of localization of human joints.
* Pose estimation is the search for a specific pose in space of all articulated (jointed, connected) poses
* Number of key points varies with the datasets - LSP has 14, MPII has 16, 16 are used in Human3.6m
* Pose estimation is classified into 2d and 3d pose estimation:
  + 2D  pose estimation estimates a 2d pose (x,y) coordinates for each joint in pixel space from a RGB image
  + 3D pose estimation estimation a 3D pose (x,y,z) coordinates in metric space from a RGB image

## Loss function:

* Most commonly used loss function - Mean Squared Error, MSE(Least Squares Loss)
* This is a regression problem. The model will try to regress to the correct coordinates i.e., move to the truth coordinates in small increments. The mode is trained to output continuous coordinates using a Mean Squared Error Loss function

## Evaluation metrics:

### Percentage of correct keypoints

* + - The detected joint is considered  correct if the distance between the  predicted and the true joint is within a certain threshold (threshold varies)
    - PCK@0.5 is when the threshold=50% of the head bone link
    - pck@0.2 == distance between predicted and true joint <0.2\*torso diameter
    - Sometimes 150 mm is taken as the threshold
    - Head, shoulder Elbow, Wrist, Hip, Knee Ankle → keypoints
    - PCK is used as 2D and 3D pose estimation
    - Higher the better

### Average Precision (AP) and Average Recall (AR)

* + - Average precision is used to measure the accuracy of keypoints detection according to precision and recall