**Why do we do Feature Scaling?**

**Height (cm) Weight (Kg) BMI**

180 78

170 84

Features have magnitude and units

Weight – 78 – magnitude

Weight – Kg – unit

If we plot these points without doing feature scaling, in K Nearest Neighbour, points will be plotted very far – use of Eucledian distance in KNN

These values have huge differences- Eucledian distance is calculated based on these values – thus result will be too far apart. Need to perform feature scaling to standardize these values.

**Compulsory algorithms for feature scaling**

1. Linear Regression

Coefficient in LR is found out wrt Gradient Descent – Coeff should converge in global minima

If we do not scale down these values, the random initial point of the coefficient will be very far from the global minima –thus taking a long time to converge towards GM.

If we scale down these values, the convergence will happen quickly.

1. K Means – Unsupervised technique – Eucledian distance
2. K Nearest Neighbour – Eucledian distance

Algorithms which do not require feature scaling

1. Decision tree
2. Random Forest Classifier – Ensemble technique 🡪 use of Decision trees here
3. Xgboost