# PROBLEM STATEMENT

TO CLASSIFY KYPHOSIS DISEASE USING DECISION TREES AND RANDOM FOREST CLASSIFIER

* Kyphosis is an abnormally excessive convex curvature of the spine. The kyphosis data frame has 81 rows and 4 columns. representing data on children who have had corrective spinal surgery. Dataset contains 3 inputs and 1 output

INPUTS:

* Age: in months
* Number: the number of vertebrae involved
* Start: the number of the first (topmost) vertebra operated on.

OUTPUTS:

* Kyphosis: a factor with levels absent present indicating if a kyphosis (a type of deformation) was present after the operation.
* Link to the dataset: <https://www.kaggle.com/abbasit/kyphosis-dataset>
* Source: John M. Chambers and Trevor J. Hastie eds. (1992) Statistical Models in S, Wadsworth and Brooks/Cole, Pacific Grove.

# STEP #0: LIBRARIES IMPORT

# STEP #1: IMPORT DATASET

# STEP #2: VISUALIZE DATASET

# STEP #3: CREATE TESTING AND TRAINING DATASET/DATA CLEANING

# STEP#4: TRAINING THE MODEL

# STEP#5: EVALUATING THE MODEL

# STEP#6: IMPROVING THE MODEL