## ${\bf Section 2: Coding\ wise\ implementation\ (Methodology)}$

Install required libraries and Import dependencies		
<u> </u>		
Dataset in the .xml format. So, I need to extract all the words and their crossponing tags in the CSV		
format. Eg. NN राम		
Remove all the stop words, empty line, handling special characters or punctuation		
Visualize or plot the tags and their crossponding word frequency using bar diagram.  We can see the most repeated words as well as less repeated words		
Ţ.		
Word Embedding		
(Converting text documents into numerical vectors using Word2Vec)		
Shuffle and Splitting Dataset into training and testing set		
Build RNN, LSTM and BiLSTM. And mBERT models with With Hidden layer sigmoid / ReLU activation function, softmax Output activation and 32/64 hidden size		
Compile model with Adam optimizer (with learning rate) and categorical_crossentropy loss		
Fit the model with 50 epochs and 128 hidden size		
Prediction the model and round the prediction		
Create confusion matrix with y_true (test classes) and y_pred (argmax of prediction)		
<b>U</b>		
Evaluate the model predictions		
Now its time to solve underfitting and over fitting problems		
<b>↓</b>		
Experiment with hyper parameters		
Compare models and find the best model		

## **Hyper Parameters**

Hyper-Parameters	Multiclass Classification
Input layer shape( in_feature)	Same as number of features (eg, 4 for age, sex, height, weight)
Hidden layers	Problem specific, minimum = 1, maximum = unlimited
Neurons per hidden layer	Problem specific, generally 10 to 512
Output layer shape (out_features)	1 per class (eg 3 for food, person, dog)
Hidden layer activation	Usually ReLU (rectified linear unit) but can be many others
Output activation	Softmax (torch.softmax in Pytorch)
Loss function	Cross entropy (torch.nn.CrossEntropyLoss in Pytorch)
Optimizer	SGD (stochastic gradient descent) or Adam etc.
Epoch (loop)	50
Batch size (number of neurons)	128
Dropout	0.3

Note: All the above mentioned hyper parameters value and other values are based on previous research paper because that values had given best results in their research.