Important Questions

- Why not approach classification through regression Link
- Linear vs Logistic regression ?
- CRF Vs HMMs?
- What is maximum likelihood estimates?
- What is Expectation-Maximization algorithm (EM Algorithm)? Can it be used for regressiona or classification ? Does it always converge?
- How SVM work? What are kernels in SVM
- Decision Tree globally or locally optimal? How to avoid overfitting in decision treesgood blog
- MultiCollinearity ? Why bad ? How to detect? How to handle? Link
- L1 vs L2 regularization. What are differences? Which is better ? Can they be termed as feature selection or reduction techniques? Nice Blog
- Optimizations functions in Deep learning ?
- When to use which feature scaling ?
- Similarity measures in Machine learning Nice blog
- Need of Negative sampling and hierarchial softmax in skip gram based word2vec Nice Video1 Nice Video2
- Why Weight initialization ways are important in DNN? Vanishing Gradients and Exploding gradients ?
- Why high dimensionality is a curse?
- what is advantage of higher grams and what is their disadvantage? probability of finding higher grams in test set is less.
- Are features also sampled in Random Forest ? what makes them better than Decision trees?
- How to reduce overfitting without regularization(by processing data)
- startegies to handle imbalanced data
- What happens when no. of features are greter than no. of samples?
- Is random weight assignment better than assigning same weights to the units in the hidden layer?
- Integer encoding vs one hot encoding vs dumy encoding. When to use which?Blog
- Why we use exponentiation in softmax activation funcion instead can we not just divide logit by sum of logits?check why
- Why softmax is computationally complex?

- Which is easy Regression or Classification? Why?