**Week3\_Day1**

IT companies are considered SHOPS in India. Thus, they do not come under Companies laws and regulations of India.

The IT sector had self regulations in the 90's. Google was one of the first companies to make its Ethics council & 2 years later the entire council was fired. Fundamentally, companies are incapable of self governance. The aim of any company is to increase its shares on markets. 2 ways to make money: INNOVATE or EXPLOIT. Exploiting is not their choice but it’s their job.

Microsoft also has an ethics council and despite the council they were found to sell for military use.

*Techno solutionism & reification*

**ELON MUSK IS ALWAYS WRONG**

Societal problems are not solved by rushing them with technological advancements.

You cannot replace the structural issue with an accelerant, because it doesn’t change fundamentals or origin of the power.

Thus, Techno Solutionism.

Ex. child malnutrition. Some people, came up with an app that from a photo gives the result if the child is malnourished or not. They made a lot of money, but it does not solve the problem.

Making category and forcing it onto the reality is reification. Entire field on ML is based on reification.

*AI as a scapegoat or the machine never “understand” or “does” policy*

‘Whose fault is it?’

Using humans to do the job, we have to pay humans. We don't want that. Thus ML.

The decision to use the code is human or a corporate decision.

AI does not create problems. People with Power do.

Though tools do have value of their own, which can be dangerous…but the decision to use that tech is always a human decision. Therefore humans are supposed to be held accountable.

**INTRODUCTION TO MACHINE LEARNING.**

For supervised ML, you need training data (get from human beings).   
Two ways magic box works: Sophisticated or Rough. Rough box distinguishes between features. In old times, images were not fed but features were.

Features are important information that the magic box takes from the data.

Basically, training data in N-D space we have 100% accuracy, then in validation they figure out the features themselves.

Having too many features leads to the curse of dimensionality, that the more dimensions there are and the computation becomes more heavy and costly.

5 years back raw data cannot be used as features.

**JARGON**

Cross validation: validation set has to be much bigger than training set.

ex . 80000 sets of images and break them into 10 parts and then take 9 parts to train and then validate on 1 part do this for many times with different 1 part and then take the aggregate of all is called Cross Validation.

Active Learning: Retraining whenever we get new data.  
Ex. if you eat a sandwich everyday you won't learn anything new there for only add data to your magic box that gives knowledge to the magic box. Use an old magic box, wherever the old magic box fails, use these data points to give to the new magic box.

Ensemble learning: take a bunch of models and you join them.

Oracle: Algorithm that automatically creates new images for your magic box to test. Overfitting. ML becomes confident. Data points in the training set are so similar to each other. The model becomes narrow. Such a model will be called brittle. Easy to break with an anomalous data point.

Explainability: decision trees are explainable, Neural networks are not.