Applied Optimization for Wireless, Machine Learning, Big Data

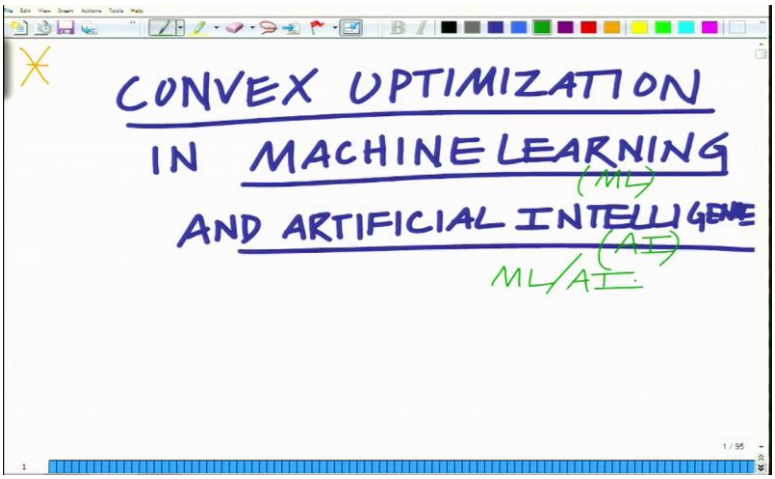
Prof. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture - 60
Practical Application of Machine Learning and Artificial Intelligence: Linear Classification, Overview and Motivation

Keywords: *Machine Learning, Artificial Intelligence, Classifier, Spectrum sensing, Linear classification*

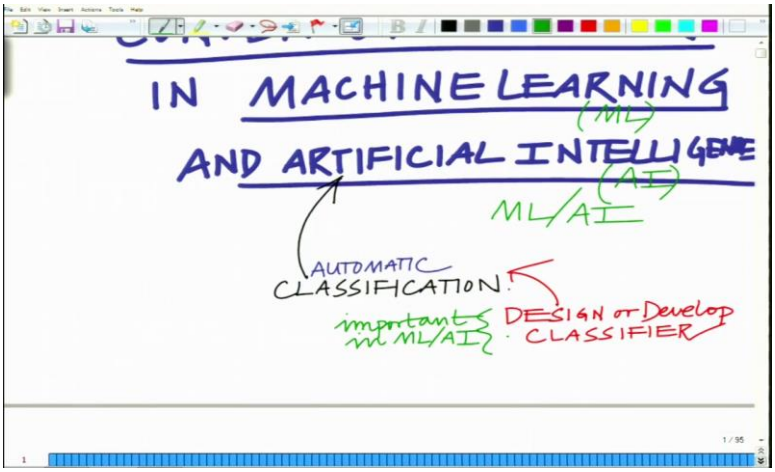
Hello, welcome to another module in this massive open online course. So we are looking at various aspects of optimization. Let us look at another important application of convex optimization that is in the latest evolving field of machine learning and artificial intelligence.

(Refer Slide Time: 00:39)



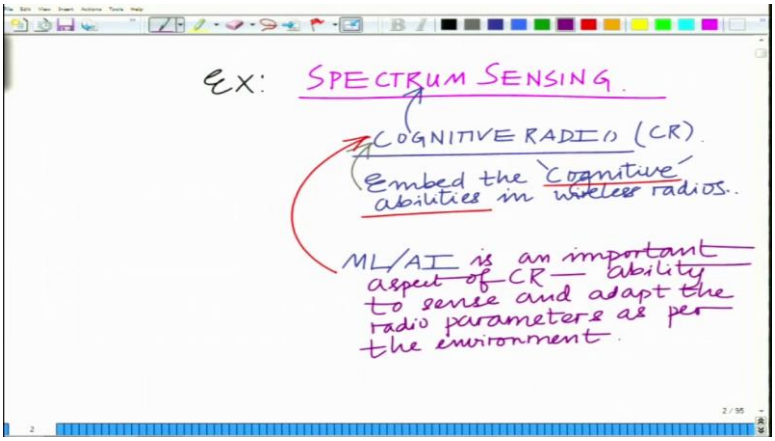
So machine learning is abbreviated as ML and artificial intelligence as AI. And of course this is a huge field, if you look at machine learning or artificial intelligence, there are large number of problems with several interesting applications.

(Refer Slide Time: 02:17)



So we are going to focus on the problem of classification and these problems can lay the foundation for the development of very complicated and sophisticated machine learning algorithm. We have to design or develop as a classifier. And this is an important problem in machine learning or artificial intelligence that is to automatically classify a set of objects belonging to either two different sets or multiple sets. So let us say you have a video or an image, from that video you are trying to classify the objects into various categories. So let us look at a typical example in modern wireless communication and that is of spectrum sensing.

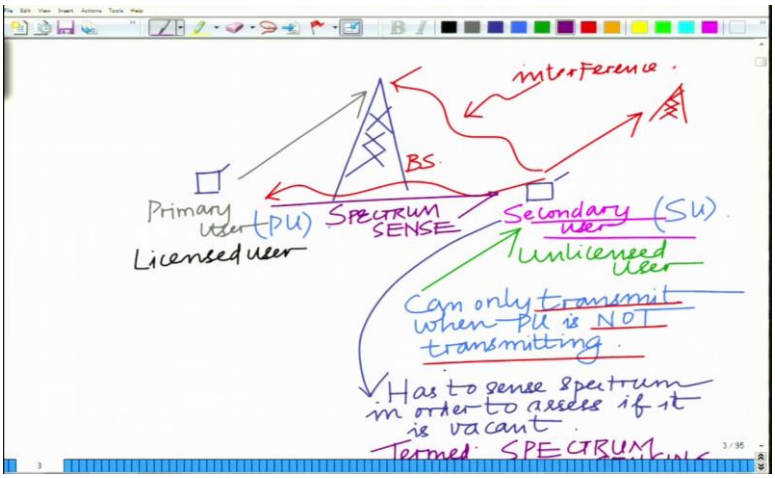
(Refer Slide Time: 04:22)



So the spectrum sensing problem arises in one of the latest wireless technologies which is known as cognitive radio and what happens in a cognitive radio technology is that the wireless device is embedded with intelligence and therefore ML or AI is an important aspect. So cognition is an important aspect of the human brain and the idea is to embed

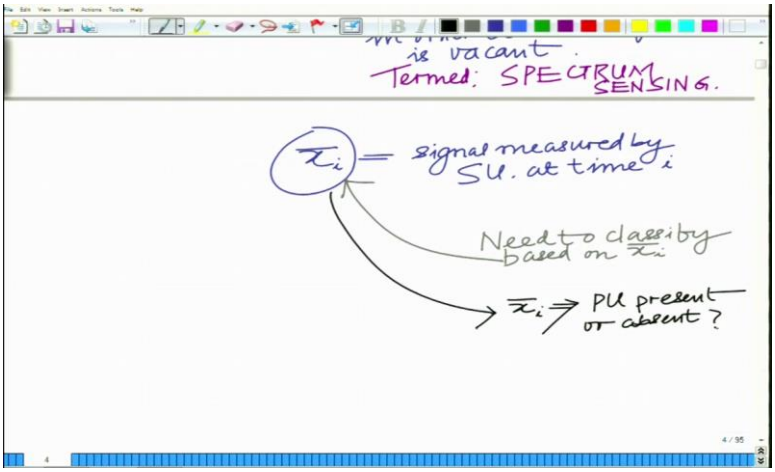
cognition or embed this kind of intelligence in wireless devices or wireless radios. So the ability for the radio to sense the environment and adapt itself is a very important aspect of cognitive radio. And therefore, machine learning or artificial intelligence which basically is concerned with extracting learning rules based on sets of data and in using them later with a very high degree of probability is an important aspect of cognitive radio.

(Refer Slide Time: 08:24)



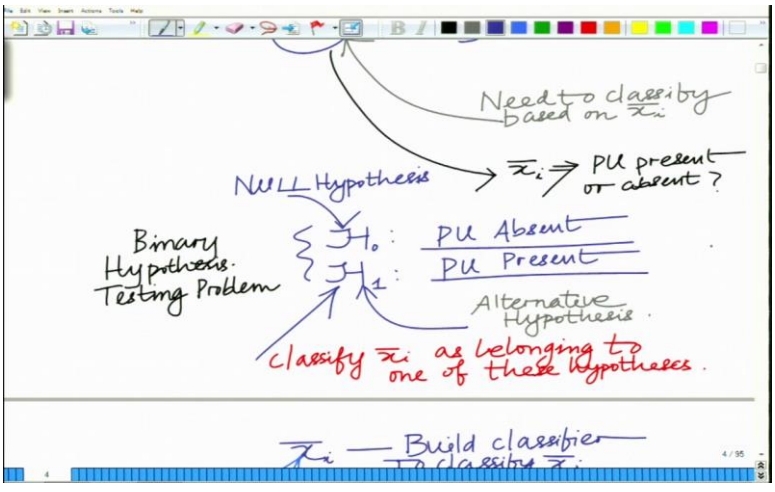
For instance, a simple problem in cognitive radio can be the following and that is when you have a primary user or the user who is licensed to transmit in a certain spectrum and in addition you have a secondary user. Now, this is the unlicensed user and this can be communicating with the different base station but causes interference at the primary base station. So this is as shown in slide. So naturally the secondary user can transmit and it also causes interference to the primary user(PU), therefore the secondary user(SU) can only transmit, when the primary user is not transmitting or the PU is absent. And therefore the SU has to sense the environment and this process is termed the spectrum sensing.

(Refer Slide Time: 11:54)



And this is an important aspect of cognitive radio, because the cognitive radio has to sense the wireless environment and adapt the radio process. And naturally this spectrum sensing process at the secondary user is going to be based on some measurements that are done by the secondary user of the environment, based on the signal let us say $\bar{x}$ that is sensed by the secondary user. So let us say $\bar{x}_i$ is the signal measured by secondary user at time i. Now, we need to classify based on this $\bar{x}_i$ in the sense $\bar{x}_i$ implies PU present or or absent.

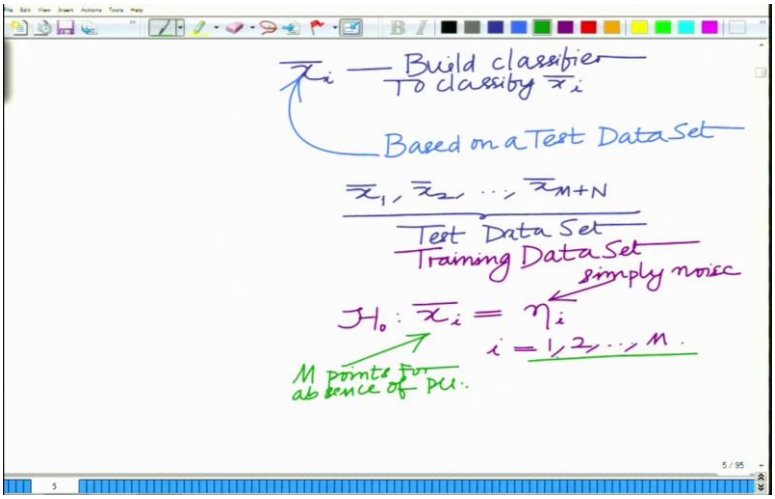
(Refer Slide Time: 14:17)



Now let us consider hypothesis 0 that is primary user is absent and this is termed as the Null hypothesis and primary user is present termed as the alternative hypothesis or hypothesis 1. So we have to assess which of these hypotheses is true and this is known as

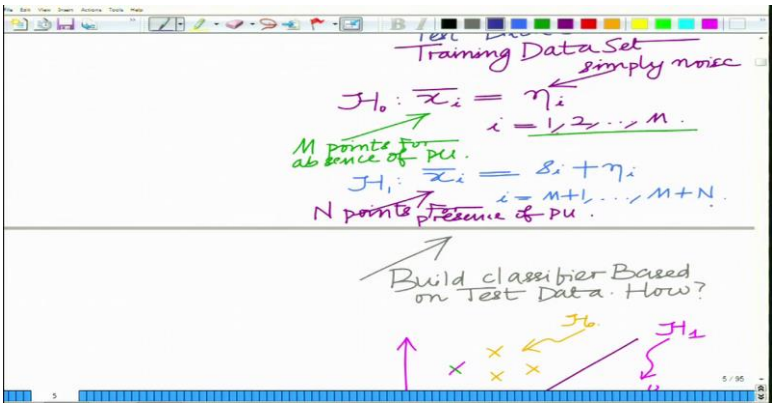
a binary hypothesis testing problem. So the idea is to classify $\overline{x_i}$ as belonging to one of these hypothesis. Now we have to build a classifier and we will build that classifier initially on the base of some data that is available with us.

(Refer Slide Time: 17:00)



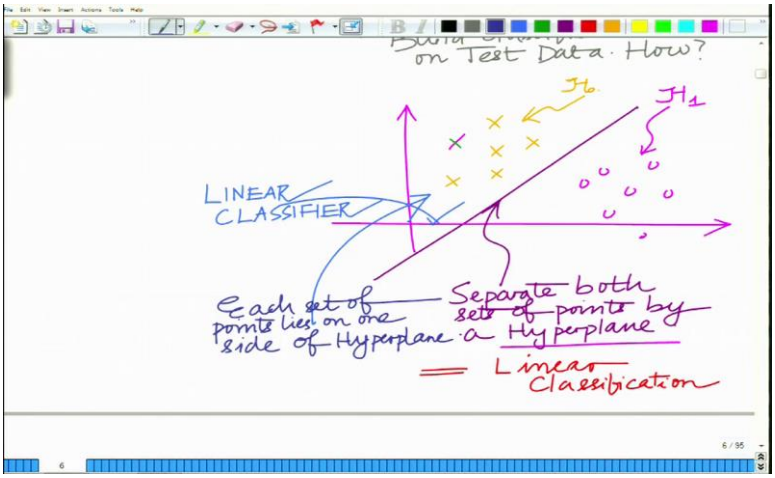
So this building of classifier is based on a test data set. Let us consider that this test data or training data set be $\overline{x_1}, \overline{x_2}, \ldots, \overline{x_{M+N}}$ and the M points corresponds to the absence of primary user that is it simply implies noise as shown in slide.

(Refer Slide Time: 19:07)



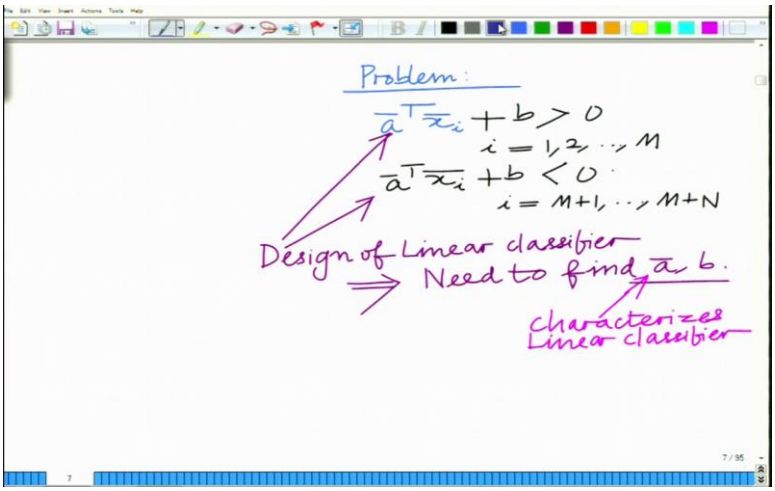
So we have $H_0: \overline{x_i} = \eta_i$ $i = 1, 2, \ldots, M$ and $H_1: \overline{x_i} = s_i + \eta_i$ $i = M+1, \ldots, M+N$. So in the test data set M points corresponds to absence of the primary user, the null hypothesis and N points correspond to the presence of the primary user or the alternative hypothesis and we have to build our classifier based on this test data set.

(Refer Slide Time: 20:55)



So let us plot this, let us say we have a 2 dimensional plot. So if all of them occur in a single cluster, then it would be difficult to classify or separate them. So you expect to see some logical separation between these points. So to separate both the set of points we can use a hyper plane and this is basically termed as linear separation or linear classification. This hyper plane which is separating them is known as a linear classifier or a linear discriminant. One set lies on one side of the hyper plane and the other set lies on the other side of the hyper plane. So the problem can be represented as follows.

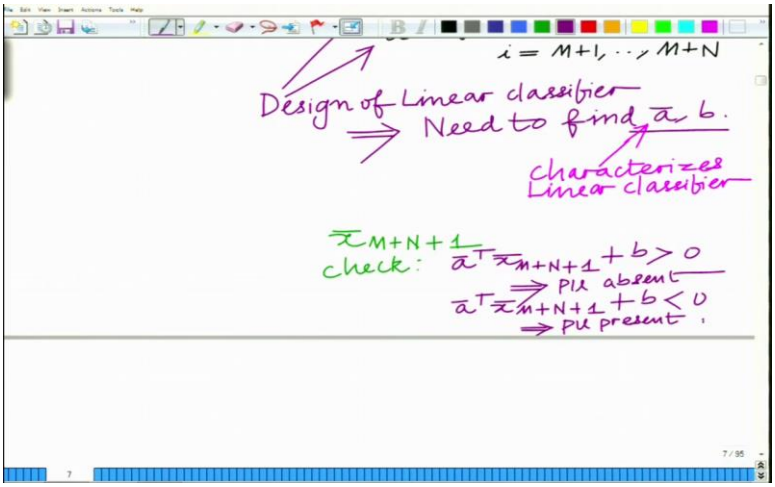
(Refer Slide Time: 24:22)



So we have $\bar{a}^T \bar{x}_i + b > 0$, $i = 1, 2, \ldots, M$ $\bar{a}^T \bar{x}_i + b < 0$ $i = M + 1, \ldots, M + N$. And finding such a classifier is nothing but finding the hyper plane which is characterized by these parameters $\bar{a}$ and b. Therefore this implies that we have to estimate $\bar{a}$ and b. So basically it is very interesting that we

have boiled down this machine learning or artificial intelligence problem of classification in this cognitive radio system into the design of a hyper plane which achieves this separation such that all the points belonging to the presence of the primary user lie on one side of the hyper plane and all the points belonging to the absence of the primary user lie on the other side of the hyper plane. So if you find these parameters $\bar{a}$ and b corresponding to such a hyper plane, then we would build this classifier and this classifier can then be used to basically classify the further things. So let us say you make a measurement $\bar{x}_i$ at time $M + N + 1$. Now we have $\bar{a}^T \bar{x}_{M+N+1} + b > 0$ implies the primary user is absent and $\bar{a}^T \bar{x}_{M+N+1} + b < 0$ implies primary user is present.

(Refer Slide Time: 29:08)



So thereby you build a classifier and then you can subsequently use the classifier to classify the measure signal and to eventually sense the spectrum that decides, if the primary user is present or if the primary user is absent. So we will stop here and we will see how convex optimization helps in building the classifier in the subsequent modules. Thank you very much.