**MLPR Discussion Report Week 3 07/10**

* What is the desired purpose; is it to be completely decisive or used an indicator in situations they are unsure of?
* Is the data representative? As there are equal amounts of red and blue points, while in reality this may not reflect the samples
* Medical history may also affect this, is it something the model can or should consider?
* Which stage of the illness was the data collected? Is there a factor of time to consider?
  + If it can only be predicted well into the onset of the disease?
* Are the features actually good indicators of cancer, are they related to other conditions and can be found more commonly because of this without it actually being cancer?
* Other features to implement? -> Medical professionals may know what are good and useful features (phrased in a sense they can understand without technical jargon)
* How do they want the model to function, case by case or with bulk inputs?
* Can we add future cases we study that get confirmed to improve our model?
* Do we have a large enough data set to accurately create a model to carry out the function?
* What’s an acceptable level of accuracy? -> Do we need to implement a large cost to wrong answers if they want extreme accuracy
* Knowing what our model actually does, can we verify and do we need a model that we can trust?
* Should we do this? -> Ethics
* How was the data collected? Consent, same manner to ensure validity, anonymising etc.
* Which model is a good fit for the function desired?
* What we carry out can massively depend on the data
* “Texture” might be a vague result and may carry an error into the sampling if different doctors would label a result differently, might need to have several evaluations and find an error to carry into the data?
* Is there already a modern model we could compare with?
* Different kinds of cancers might skew the results?