**Critique 5: Data Science and Prediction Jagadish Tirumalasetty**

**CS590BD #16143014**

**Summary:**

In the journal paper the authors have emphasized mostly on explaining the actual background behind the word Data Science and how does it has to be interpreted and the history of it. There by the discussion running into effective database management systems, evolution of technologies helping efficient and cheaper ways to analyze data and pull out trends/reports.

As authors have tried giving examples on how the trends in data analysis lead to predictive systems that could allow businesses to take effective and competitive decisions at the right time. The authors have there after emphasized on machine learning with explaining various traits a machine learning system developer or designer should have and also various technological requirements that an analyst should have been mentioned as well.

Further points that found a mention in the paper are various reasons why a machine learning system might fail/not perform as expected, Knowledge discovery and importance of the ability of analyst to be able to see what exactly should be asked to the system.

**Critique:**

* The authors have repeatedly mentioned that the Big Data domain needs people who have the ability of formulating the problems so as to start what to look for, but isn’t the ability to observe patterns also important too. As, after being able to observe the data properly can anyone be able to even think of kind of questions which would hold good for the data.
* No solid discussion on the whole topic of machine learning. The topic has just been touched upon and characteristics required by and individual to take up such tasks is defined but no emphasis on the pros of machine learning discussed in depth.
* Many theories like Popper and Occam’s have been mentioned in the paper on the basis of which some topics have been explained but at least some intro to these would have helped.
* The paper altogether was short and precise in explaining the various qualities one should have to develop machine learning.
* The paper being purely theoretical i.e. unlike many papers that we saw in this semester which had experiments and solid electrical equipment’s/sensors to discuss about, in such comparison the authors could have used many more examples rather than just one example on the medical transactions.

**Questions:**

* Are the knowledge discovery features and predictive features inter-related?
* In the paper it has been mentioned that Big Data is largely used on Passive Data. In this context, is it an overhead to have Big Data system for non-passive data?
* Shouldn’t the ability to select the right data and to read very important for an analyst in order to know what to ask the system.