Week of March 29- April11 progress

Lessons learned from Chapter 5: Mathematics

1. Always try to use the standard mathematical notations to avoid ambiguity.
2. Always verify that the proofs given are correct.
3. All the mathematics should be presented in italics, to distinguish it from the other text.
4. Always number all the theorems, propositions, definitions, lemmas, diagrams etc. for ease of reference.
5. Ranges should be written as i=1…..10
6. Using Greek alphabets could be useful.
7. It is easier to understand percentage than probability.
8. All abbreviations should be explained atleast once, like for units.
9. There should be space between units and numbers.
10. Subscripts and superscripts should be used carefully.
11. Do not reuse same notation to explain different concepts.

Lessons learned from Chapter 11: Experimentation, part 1

1. Tests should not be designed to be biased to prove the hypothesis.
2. Always test the hypothesis on the points where they might fail.
3. Experiments should also be conducted on the data set which was not used in the training phase.
4. Results should be understood and explained properly.
5. Anomalies of the experiments / results should be explained and not hidden.
6. Wherever possible use standard resources.
7. The readers should be motivated by carefully explaining the problems that were encountered.
8. The observations made should be verifiable and reproducible by other researchers, else the results are valueless.

Lessons Learned from Chapter 11: Experimentation, part 2:

1. Statistics in an experimental research is a good source of knowledge. Thus statistical approach should be followed.
2. One needs to understand the population of a sample.
3. Those results which are not sensible should not be included in the paper.
4. It is unethical not to report the failed test results.
5. Hypothesis should have the same inputs as that of the test.
6. For experiments, use standard data wherever possible.
7. Conclusions should not be drawn from a small set of samples.
8. Measures of correlation can be used to determine if two variables depend on each other.
9. There are tools available to perform complex mathematical analysis.
10. Always attempt to prove that your hypothesis is incorrect. Once you are unable to prove its incorrect you have reached your goal.

Latest pre-proposal progress

Quite close to finalizing the project topic. Read the following research papers:

1. Stephen Dill , Nadav Eiron , David Gibson , Daniel Gruhl , R. Guha , Anant Jhingran , Tapas Kanungo , Sridhar Rajagopalan , Andrew Tomkins , John A. Tomlin , Jason Y. Zien, SemTag and seeker: bootstrapping the semantic web via automated semantic annotation, Proceedings of the 12th international conference on World Wide Web, May 20-24, 2003, Budapest, Hungary [doi>10.1145/775152.775178]

2. A. Maedche and S. Staab, Ontology learning for the Semantic Web, IEEE Intelligent Systems: Special Issue on the Semantic Web 16 (2001) 72–79.

3. G. Bisson, C. Nedellec and L. Canamero, Designing clustering methods for ontology building – The Mo'K workbench, presented at ECAI Ontology Learning Workshop, Seattle, WA, 2000.

4. Zhou, L. (2007), Ontology Learning: State of the Art and Open Issues, Information Technology and Management, 8(3), 241-252.

5. Noy, N.F.; Sintek, M.; Decker, S.; Crubezy, M.; Fergerson, R.W.; Musen, M.A.; , "Creating Semantic Web contents with Protege-2000," *Intelligent Systems, IEEE* , vol.16, no.2, pp. 60- 71, Mar-Apr 2001 doi: 10.1109/5254.920601
URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=920601&isnumber=19905

6. Decker, S.; Melnik, S.; van Harmelen, F.; Fensel, D.; Klein, M.; Broekstra, J.; Erdmann, M.; Horrocks, I.; , "The Semantic Web: the roles of XML and RDF," *Internet Computing, IEEE* , vol.4, no.5, pp.63-73, Sep/Oct 2000 doi: 10.1109/4236.877487
URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=877487&isnumber=18994