Final Project Check List

Data Analysis Projects

- Quality and Soundness of the Analysis:
    - Clear statement of objectives of the data analysis project, and discussion of the appropriateness of approaches and data analysis tasks used to accomplish the objectives.
    - Discussion of the structure and characteristics of the data, including the results of any exploratory data analysis and data visualization performed on the data.
    - Data driven justification for choices of algorithms, techniques, and parameters used in various data analysis tasks (e.g., based on evaluation using cross validation and other relevant methodologies).
    - Thorough discussion and analysis of data mining results, including an analysis of how the approaches used worked in accomplishing the project objectives.

- Breadth of Analysis:
    - The use of basic statistical analysis, data exploration, and data visualization techniques to provide deeper understanding of the targeted data.
    - The exploration of multiple data analysis tasks on the targeted data, including both supervised knowledge discovery (predictive modeling) as well as unsupervised knowledge discovery for exploratory data analysis.

- Thoroughness of Data Analysis Activities:
    - The use of a a variety of relevant techniques to determine the best approach to accomplish each data analysis tasks (e.g., comparing multiple classifier models for the same classification task, etc.).
    - Experimentation with different parameters to optimize the results of the chosen techniques.
    - Examination of different data preparation and transformation approaches to improve results for the given analysis tasks.

- Appropriate Use of Tools:
    - Correct and appropriate use of tools to accomplish the various data analysis tasks, including tools used for data preparations and libraries or third-party tools used and part of scripts or programs.
    - Appropriate documentation on code segments or scripts used as part of the data analysis.

- An executive summary summarizing the project goals, methods used, and conclusions.

•The main report providing details on the data set(s) used, design decisions, description of how the KDD process was applied, tools used to perform specific data analysis tasks, a detailed description and analysis of the results of data analysis (including the use of visualization, and conclusions. Appendices to the main report including data samples, fully documented code segments, programs, and scripts used as part of the data analysis, and descriptions and references to any third party tools used in the project.

•A final project demo video
>•A short (5 minute) demo or screen recording describing the project, main tasks, components or methods, summary of results, and a demo run (in the case of applications).
>•Audio commentary with your own voice is required.
>•Accepted video formats: mov (QuickTime, MacOS), mpeg, avi (Windows) -- No Flash, please.
>•You may choose to upload your video to the cloud (e.g., upload your video to YouTube) and submit the link. The link should be provided in the comment section of the Dropbox folder for the project.

Application Development Projects

Application Development projects will be evaluated based on the following criteria.

•Correctness and Completeness:
>•System provides the necessary functionality to meet all of the project objectives (including those specified in the project proposal).
>•The application and all of its components work correctly and as expected.
>•Algorithms and techniques used are efficient and scalable.
>•The systems includes all of the relevant components necessary for the underlying application domain or task.

•Design:
>•The use of sound design principles in the system and its components so that the overall architecture of the application is clear and logical.
>•Appropriate use of established software design methodologies to enhance modularity and avoid redundancy.
>•Effective use of interaction design, including appropriate I/O design and effective user interfaces.
>•Different components and pieces are integrated appropriately (so that from the user perspective, there is seamless interaction with the system) .

•Evaluation:
>•The system has been evaluated thoroughly using appropriate data sets.

- •The evaluation methodologies used are appropriate for the specific application domain and the system objectives.
- •The results of the evaluation are clearly presented and discussed, and if necessary compared to baseline systems or algorithms.

- •Documentation:
  - •Adequate evidence (e.g., sample runs on different test inputs) is provided to demonstrate the system functionality as a whole and the functionality of different components.
  - •The source code is fully documented.
  - •The report provides detailed documentation of the system architecture as well as a description of different components and their interactions.

Deliverables for Application Development projects include the following.

- •A detailed project report, including a description of the system system (including specific techniques and algorithms you used), and the interaction between the components (make references to code segments, modules, methods, functions, as necessary). Your write up should also include description of the evaluation of your system demonstrating its correctness, functionality, accuracy, etc., and the description of the data set(s) used in evaluating your system.

- •A final project demo video
  - •A short (5 minute) demo or screen recording describing the project, main tasks, components or methods, summary of results, and a demo run (in the case of applications).
  - •Audio commentary with your own voice is required.
  - •Accepted video formats: mov (QuickTime, MacOS), mpeg, avi (Windows) -- No Flash, please.
  - •You may choose to upload your video to the cloud (e.g., upload your video to YouTube) and submit the link. The link should be provided in the comment section of the Dropbox folder for the project.

- •Appendices to the main report should include:
  - • Complete (actual) sample runs of your program with descriptions, illustrating how your system works, along with any intermediate input or output used for the sample runs.
  - •Fully documented code for the system, including any programs and scripts used for data preparation and analysis.
  - •Binary files (e.g., executables, DLLs, Class files) or other components necessary to run your program.
  - •Samples or detailed descriptions of the data sets used.

•Readme file containing instructions on how to compile, install, and/or run your program, if appropriate, including relevant references to outside resources used.