Bryan Bonnet   1/28/14
I pledge my honor that I have abided by the Stevens Honor System.

Title: *Predicting Congressional Voting through Data Analysis Techniques*

Objective: To create a system for training a predictive analysis model to predict the voting of United States congressional members on legislation based on a multitude of inputs.

Technical Description: The systems will use the Computer Engineering discipline of data analysis and machine learning in order to model and determine the most likely vote for each congressmen on a particular subject or bill. The input vector may contain data from, but not limited to the following sources:

- Free-text of congressional acts and bills
- Campaign contributions
- Pertinent news articles
- Pertinent speeches or text from audio clips

Once these inputs are collected and put into a computer-readable format using natural language processing, they will be trained against known voting history. The output of this training sequence will be a computer model for each United States congressman. Future votes will be predicted by running the new inputs into the model and creating a prediction output.

Impact:

This product has a two-fold impact on all those affected by United States politics. The first impact will provide insight into the non-obvious factors which influence politician's decisions when voting on congressional bills. This is exposed by comparing the computer model weighting matrices used by the machine learning algorithms to most accurately predict congressional voting. The second impact is inherit power to predict the likeliness of a congressional member to vote for a particular act or issue. This can give constituents and lobbyists time to impact the member's decision before it is public.

Potential Commercialization and Sales:

The product has the ability to be licensed as intellectual property to several potential stakeholders. News organization may want to have an app or widget on their websites that can accurately predict upcoming congressional votes. Lobbyist may pay for this product as a service to analyze individual congressional members and determine where to focus their efforts. Finally, individual politicians or political parties may use this as a tool to track and predict their peer's votes.

Stakeholders & Potential Requirements:

Assuming a situation where the product would be commissioned to a news-gathering organization and licensed for use on their web site:

User - News consumer, web page user

- The service shall be accessible via web on both desktop and mobile platforms.

- The output of the system shall be human-readable and easily understood by users not familiar with data analysis.

Client - News Organization

- The system shall be easily integrated into their current website via web services.

- The system shall update in real-time without human intervention.

- The system shall require minimal maintenance and up-keep.

Designer – Development team

- The project shall be completed in two semesters.

- The system shall be implemented by adapting existing technologies and proven techniques of data analysis.

- Some level of stakeholder feedback shall be available for incremental development.

Practicality Assessment:

This system can be implemented by adapting existing data analysis tools and techniques. This reduces the risk and makes the project practical to be successfully built and implemented. The project can be developed in incremental pieces based on stakeholder requirements and feedback. Using this agile development approach, makes a two-semester time frame reasonable for significant requirements to be implemented. Meeting several of the performance metrics will indicate satisfactory ability to accurately predict congressional voting. In this case, the system can easily be sold to organizations and improved upon once obtained.

Skills Required:

- An understanding and interest in the United States legislative system and news industry will be needed generate high-level requirements for potential licensing opportunities.

- A team with strong programming skills will be required to implement the system. Useful languages include, but not limited to, Python, R and Matlab.

- Knowledge and experience manipulating and analyzing both structured and unstructured (natural language) information will be essential to the development team.

- Interest or experience in web design and interactive data visualization will be needed for front-end design and development.

SWOT Analysis:

Strength – This combination of computer science and political science can prove to be valuable product for everyone interested in tracking and predicting the outcomes of major policymaking. The project has the ability to be a game-changer in the political news and lobbying industries and therefore extremely sought after by these stakeholders. These stakeholders may seek to uncover insight into the way policy voting works in ways that have previously not been explored.

Weakness – This project can receive push back from potential stakeholders due to the sensitive nature of the information that it will provide. Especially if the system is limited in its level of accuracy, it may be rejected by the target market.

Opportunities – This project has a dynamic market opportunity. The primary target for the project is to build a tool for a news agency but, several other markets are available. Adjusting the final output of the product can attract non-obvious customers such as private businesses and political scientists. This can be accomplished by creating a information asset in the potential outcome of key legislation that can affect their business.

Threats – The ability to adapt existing technologies, such as machine learning, is the main risk to the project. Undergoing the project assumes that data analysis techniques can be used to accurately predict the outcome of congressional voting. In earlier stages of the project, when less inputs are integrated, the system will make critical assumptions regarding the consistency in voting among congressional members. These assumptions can threaten the time line of the project and will require the ability for the team to adapt to changing requirements.