

Cameron Palk

Dejing Dou

CIS 453: Data Mining

March 4, 2016

Seattle Crime Analytics and Prediction

My proposal is to investigate and mine data from Seattle police reports taken by officers when responding to incidents around the city. This information is publicly available, supplied by the city of Seattle, and is updated within 6 to 12 hours after a report has been filed. I will utilize this data to make predictions on the type of criminal offense based on its time and location. I will also have a good amount of data generalization to help aid in the understanding of this immense dataset of police reports. Utilizing visualization technics and various analytics, I will make the information more understandable and concise to gain information pertaining to police reports in Seattle. Using data mining tools I will build a prediction for crimes where, based on time and location, I will predict the category of crime.

I will be looking at crimes under the following categories, with a classification of **Other** for any category not present in this list:

A. Narcotics

- A. Found
- B. Sold
- C. Possession
- D. Smuggle
- E. Manufacture

- F. Fraud
- G. Other
- B. Threats
- C. Theft
- D. Robbery
 - A. Street
 - B. Residence
 - C. Business
 - D. and
 - E. Other
- E. Vehicle
- F. Disturbance
- G. Trespass
- H. Burglary
- I. DUI
- J. Homicide
 - A. Premeditated
 - B. Negligence
 - C. Justified
- K.

Content: A brief summary of the proposal

Length: one-third to one-half page, never more than one page

Emphasis: highlighting of the proposed technical and management approach

Table of Contents

Statement of Problem: the “Why?”

Summary of the request by the sponsor (the original problem statement)

Background:

Brief description of company and their business Relevance or importance of problem Background information to educate the reader Previous related work by others—literature review with credible sources Patent search, if applicable Detailed problem description, as you now understand it

Objectives: the “What?”

In the Objectives section, you translate the customer’s quantitative and qualitative needs into clear, objective design specifications. Define the scope of work and clearly state the project objectives, including the following:

- A. Design specifications in specific, quantitative terms. For example, “The plate must be rotated three times at a speed of between 1 and 3 rev/s” or “Control the temperature of a 1 liter non-insulated standard glass beaker of water to $37.5 \pm 0.5^{\circ}\text{C}$ for three hours without temperature deviation.”
- B. Critical design issues, constraints, limitations.

Technical Approach: the “How?”

- A. Although you may not know all the details of the problem solution, you should know a first design on how you will attack the problem, and you should have some design concepts.
- B. The purpose of this section is to present the process by which you will arrive at the final answer.
- C. This section answers the following questions:
 - A. What are the steps in the design process? (Describe and use the nine-step model from Chapter 1 of Hyman’s text)
 - B. What are the benefits and advantages of employing a structured approach to design?
 - C. How will you generate solution concepts?
 - D. How will you analyze the performance of your solution?
 - E. How will you decide on the best alternative?
- C. Specific recommendations for this section include the following:
 - A. First, describe your overall design process in general terms. A one-page synopsis of Chapter 1 in the Hyman text would be appropriate here.
 - B. Provide at least three possible solution alternatives and document your methodology to choose the best alternative. Include illustrations such as Figure 1. Try to be as inclusive and creative as possible with your ideas. Strive to achieve at least one non-conventional or “out of the box” alternative.
 - C. List and describe all the analytical, or computational tools you will employ to analyze your design, such as ProEngineer®, SolidWorks®, MathCAD, and MATLAB.

- D. List and detail all the experimental procedures you will use to test your design concepts.
- E. Evaluate your alternatives based on how well they satisfy the design specifications. Explain the selection criteria by which you will evaluate design alternatives in specific, quantitative terms, such as cost, weight, reliability, ease of use, and ease of manufacture. A matrix table can clearly illustrate this information.
- F. If possible at this time, rank your solution concepts and list the pros and cons of each. At minimum, state what further information or additional work is needed in order to arrive at a final solution alternative.
- G. If any solution is totally unfeasible (or may have been Figure 1. SolidWorks® model of a tried before), state the reason for its elimination. Manual filter wheel with C-Mount adapters.

Team Qualifications: the “Who?”

- A. In a paragraph for each person, establish the team qualifications for the project.
Highlight any specific job or course experiences that are relevant to the project.
- B. Include a one-page resume of each team member in the Appendix. Do not include your hobbies.

- Project Proposals due: Friday March 4 11:59pm
- Project Presentations: Friday March 14 8:00am - 10:00am
- Final Reports due: Friday March 18 11:59pm

Projects are intended to give students the opportunity to explore ideas or directions in Data Mining that we cover in class (e.g., data generalization, data warehousing, association mining, classification and prediction, clustering and outlier analysis), to discover interesting pattern and knowledge from available application data sets (e.g., web data, biomedical data, health data, network data, social network, e-commerce data, and retail data).

Project work will include an 5 minute presentation (with slides) on March 18, Friday 8:00am -- 10:00am (final exam time) and an 8 page research paper for every student, submitted as the final report.

Project proposals are to be 5 pages in length and will include a definition of the direction, the particular problem to be considered in that direction, and a list of several references that will support the project. The bottom of course web page has some useful links for finding a project and existing data mining softwares.

Students should plan on either experimenting existing data mining softwares with collected data or extending existing softwares for specific domains. The projects will be expected to be individual work.

Your final paper and presentation should clarify the following issues:

- what is the problem and data sets addressed by the project
- what are the data mining algorithms or tools used,

- what are the knowledge or pattern discovered from the project,
- what are your evaluations and ideas for extensions and improvements.