ISE/OR 560 Project Proposal Guidelines
DUE: October 24, 2019 at 11:59pm

Lenovo is concerned with predicting customer and improving customer satisfaction. They would like to be able to understand those factors that are driving customer satisfaction based on the discussions customers have on the web regarding their products and the feedback they receive from customers via the Net Promoter Score (NPS) survey. They currently collect sentiment scores based on third party analysis of weekly web scrapping to understand the customer's view of a product. However, Lenovo uses customers' responses to the NPS survey as their primary measurement of customer satisfaction (surveys are sent to customer in the first 30 days of their purchase but it takes time to collect them and get enough responses to be relevant). Unfortunately, feedback from the NSP survey results arrives rather late (~ six months after product release) in the product life cycle, so Lenovo would ideally like to be able to predict NPS score based on the customer sentiment. Their hope is that they can use this information to tell them what their customers care about and inform timely response to customer concerns. When customers use laptops, telemetry data are sent to Microsoft. Telemetry is an automated communications process by which measurements and other data are collected at remote or inaccessible points and transmitted to receiving equipment for monitoring. Telemetry data starts the moment a machine is powered on. These data are live PC Diagnostics including the following Performance Metrics: Battery Life, Wireless Connectivity, and Reliability (Crashes). Current telemetry data is a 28-day view of system performance. Lenovo wants to explore how they can leverage telemetry data to take action on issues and understand sentiment and NPS. Lenovo has provided data for several consumer and commercial products over a eighteen month to two-year period for the sentiment scores with the corresponding raw customer comments and the NPS survey responses. In addition, Lenovo has provided battery life, wireless connectivity and reliability data. You have been charged by Lenovo with

(i) Utilizing predictive analytics methods (such as regression) to predict customer satisfaction using customer sentiment data. Exploring alternative metrics of measure for quantifying customer satisfaction in addition to pNPS and sentiment based on your analysis of the data.

(ii) Developing a stochastic model of the evolution of sentiment and satisfaction. You are asked to use the provided data to explain the evolution of customer's perceptions of various Lenovo products following release over time and describe the "natural history model" of this process.

(iii) Identify important factors that influence customer satisfaction and how these factors evolve over time (e.g., what factors are important to the customer immediately after release or when they receive the product compared to after they have had the product for three to six months).

(iv) Finding optimal rules for intervening to influence customer satisfaction based on learnings from sentiment.

Lenovo would like you to use the three customer focused datasets to answer the following questions:

1. How can Lenovo use CID sentiment/stars data to predict pNPS survey scores?
2. How does Telemetry data influence pNPS Survey Scores? How does Telemetry data influence sentiment/stars data?
3. Can you predict the pNPS score from CID sentiment data and the influence of telemetry data? Five products will be eliminated from the survey data and your model will be tested on these products.
4. What should Lenovo focus on to improve pNPS? What should Lenovo focus on to improve Total pNPS for all products? What should Lenovo focus on to improve pNPS for the bottom 3 Customer and bottom 3 Commercial products?
5. What impact do the product stars have on the analysis?
6. Do you recommend any changes to Lenovo's data collection methods?
7. Anything else you discover…

As a first step, you are asked to prepare a three to five page proposal (including Figures and Tables).

- Do research on the project topics (including NPS and the role of customer ratings in this industry as well as how telemetry data has been used) and describe the problem with your own words.
- Explain your methodology in detail: how will you perform the statistical analysis part – what methods specifically will you use, how will you use the data to build the elements of your Markov chain and elements of your MDP model, etc.

Your proposal should be written professionally. Proposals should be in Times New Roman font (12 point) and 1.5 spacing. Figures and tables should be embedded in the document and referenced in the text.

The proposal should:

1. Present an overview of the problem and the structure of the model you will develop to address it. Include metrics of measure that you will use to evaluate your strategy and a plan for how to optimize the objective.
2. Explain NPS and its role in the industry and provide an overview regarding the importance of NPS to Lenovo. Explain telemetry, its current use in the industry and its potential for informing NPS.
3. Identify key variables and parameters.
4. Identify and describe critical uncertainties within the model with regards to parameters and decision variables. Identify those uncertainties that may be dynamic (stochastic).
5. Identify states that can be used to develop a Markov chain including the time period associated with a transition. Define states, actions, rewards/costs and how you will use the data to calculate. (This will be used to solve a Markov decision process).
6. Conduct preliminary data analysis – describe each data source, provide summary of data attributes and factors include figures and tables to support your discussion. Figures and Tables must be labeled (number and caption) and referenced within the text.