**Projects of CMPT 741, fall 2020. Weight 40% (Version 1, Sept 9 2020).**

**Due: Dec 8, 2020**

**Section 1: Introduction**

**Topics:** This is a group project of group size two. Each group will select one area from the list in Section 2 to work on. Within the chosen area, the student will identify a concrete problem and topic. To start with, some references are listed but students are encouraged to find and read more relevant papers in order to complete the project. **The objective of the project is to write a research report in one of the following forms:**

*Survey*: analyze/compare existing works with your own insights. You should start with checking whether there is already a survey, if so, why you want to write another one, what is the difference between your survey and the existing ones. The focus is your *own analysis and new insights*. Simply reiterating the known analysis in existing works is insufficient.

*New problem formulation*: motivate a new problem statement. You will need to motivate your problem statement, such as why this problem can not be solved by existing solution, even with some transformation, what are the challenges in solving this problem. You need to present your solution, and a concrete solution is preferred to an abstract one. Pseudo code description is acceptable and empirical evidences are highly encouraged. Examples are very helpful.

*Improve existing solutions*: provide a better solution to an existing problem. Unlike the previous one, here you focus on Identifying the limitation of existing solutions to an existing problem, why your solution is better, and provide the improved solution, which can be a more efficient algorithm, a more accurate answer, and so on. Pseudo code description is acceptable and empirical evidences are highly encouraged to support your “better solution” claims. Examples are very helpful.

**Section 2: General Research Areas**

Each group needs to choose one area from the following list. Reference given serves as a starting point. You need to find and read more works through the starting pointer.

**Area 1, Recommender systems against adversarial attack**

References:

Adversarial Attacks on an Oblivious Recommender

**Area 2, Advanced recommender systems dealing with item diversity, substitute items and complementary items, etc.**

References:

Complementary Recommendations: A Brief Survey

Identifying Complements and Substitutes of Products: A Neural Network Framework Based on Product Embedding

Diversity in recommender systems – A survey

**Area 3, Sequential recommendation and recommendation with periodicity**

References:

Personalized Top-N Sequential Recommendation via Convolutional Sequence Embedding.

Learning Graph-based Embedding for Time-Aware Product Recommendation

**Area 4, Explaining black-box model (e.g., DNN)**

References:

Explaining Explanations: An Overview of Interpretability of Machine Learning

Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI

Axiomatic Attribution for Deep Networks

A survey of methods for explaining blackbox models

Visualizing the impact of feature attribution baselines

**Area 5, Federated learning**

References:

TOWARDS FEDERATED LEARNING AT SCALE: SYSTEM DESIGN

A Survey on Federated Learning Systems: Vision, Hype and Reality for Data Privacy and Protection

Federated Learning: Challenges, Methods, and Future Directions

**Area 6, Crowdsourcing against adversarial attack**

References:

Truth Discovery against Strategic Sybil Attack in Crowdsourcing.

**Area 7, Differential Privacy**

References:

The Algorithmic Foundations of Differential Privacy

A Survey on Differentially Private Machine Learning

**Section 3: Evaluation Criteria**

The report should be written with typical structures of a formal research paper. The page limit is 9 with 12 font size, single spacing, including references and all figures. Evaluation is based on :

**Choice of problem and motivation (20%):** is the problem or topic well motivated? Is the objective clearly stated?

**Presentation and clarity (20%):** structural and organizational clarity, well use of examples to illustrate difficult but important points, clarity of main points, English.

**Novelty (20%):** originality of analysis, problem statements, and algorithms. Lack of your own thoughts and insights, just simply reiterating known materials from published works, is considered as low novelty. Highly recommend to state novelty explicitly, such as which part is new contribution (and why) and which part is from which previous works (i.e., not new). Novelty can also be in the form of clever use of established techniques for solving interesting problems.

**Scientific support and depth (20%):** how well your claims on contributions are supported. Ideally, for each claim made, such as a new problem statement, a new algorithm, an improved algorithm, or a new analysis, support these claims, i.e., why it is a new problem, why it improves existing solutions, etc.

**Leverage of literature (20%):** the coverage of relevant literature, demonstration of knowledge about literature, proper and effective use of literature, attribution to literature.

You may be requested for an interview regarding the report in order to complete the evaluation. The interview will verify the work was done by yourself.

**Section 4: A few Points to Note**

1. Start the project as early as possible because it will require a good deal of literature reading and time for picking the topics.
2. Always acknowledge the sources of information used in your report.
3. Write the report in your own words. Cut-and-paste text from published materials and other groups without acknowledgement is considered as plagiarism and will be penalized seriously.
4. Make use of office hours if you need additional help on your projects.