CSCI 567 Mini-Project

1 Introduction

In this mini-project, you will have the chance to explore interesting machine learning problems by participating in a kaggle competition. Kaggle is a platform for predictive modeling and analytics competitions on which companies and researchers post their data. Machine learning and data mining researchers from all over the world compete to produce the best models.

In this project, we will tackle the **Click-Through Rate Prediction** problem, which is one of the most exciting and important applications in machine learning. Given an advertisement, you need to predict whether or not it will be clicked by users. For more details, please refer to official Kaggle competition website: `https://www.kaggle.com/c/avazu-ctr-prediction`.

2 Grading criteria

Your grade for this project is based on two components, 1) your team's ranking on the leaderboard (60%) and the project report including the code (40%). Note that the leaderboard ranks all teams, but a better ranking will lead to a better grade.

3 What to turn in & deadlines

- Form a team of up to five students by **Nov 5, 11:59pm**. Each team should fill out the form in this google doc. The name of your team on the leadboard should begin with `MLCLASS_`. e.g. MLCLASS_awesome, etc.
- We will grade you based on your ranking on **Nov 30, at 11:59pm**. No later changes will be taken into account.
- Each team needs to submit one project report, which should be written in NIPS format. (6 pages maximum, including the reference; this page limit is strict). The deadline for submitting the hard copy of your paper is **Dec 5, 5pm**. Each team should only submit ONE report to the CS front desk. The NIPS LATEX format can be found here: `http://nips.cc/Conferences/2014/PaperInformation/StyleFiles`
- **Code**. Submitting code via blackboard is mandatory. The deadline for submittng your code is **Dec 5, 11:59pm**. ONLY one of the team members should submit the code.
- Members of the same team will receive the same scores.

4 Collaboration Guidelines

In line with the rules of the competition, you are ONLY allowed to collaborate within your own team. Your code will be analyzed to reproduce the results and compare similarity to other groups' code. All members of the violation group will be given an "F". So please DO obey the collaboration rules.