**CSC529: Advanced Data Mining**

**Paper 1: Ensembles of Classifiers**

**Total: 10 points**

**Due: Sunday, February 18th, 2018, by midnight**

The purpose of this paper review is to get you familiar with ensembles of classifiers. Read the following tutorial paper (posted under the course documents for week 6) and respond the questions below:

Lior Rokach, “Ensemble-based classifiers”, Artificial Intelligence Review, Volume 33 Issue 1-2, 2010.

1. Describe the main components of an ensemble learning framework.
2. Huge databases pose several challenges including computing complexity and storage problems for most machine learning algorithms. Briefly describe three approaches that are variations of AdaBoost which were proposed to handle these challenges.
3. What are the main methods for combining the base-classifiers’ outputs? Briefly describe and compare five of such combination methods.
4. Explain the concept of “diversity” for ensembles of classifiers and briefly describe two approaches for creating diversity on ensembles of classifiers.