

# CS 165 Project Abstract

Sha Sha UID: 2104167

Yu Wu UID: 2103504

## 1 Topic

Domain adaption is often used under practical scenarios where source domain and target domain have different but related data distributions. In this project, we are interested in covariate shift and label shift, where assumes the conditional probability does not change while the marginal probability changes.

Active learning is often used in the scenario where manual labelling is expensive. It could help us select the data which includes more information. In this project, active learning will be applied to obtain more information of target domain and help improve the model trained from source domain.

## 2 Previous Approach

According to the previous work under label shift, the main approach is to estimate the importance weights  $q(y)/p(y)$  and further regularize the influence of these estimated weights, which shows good classification performance but with assumptions on distributions made to simply the problem. Another approach is to formulate the label shift problem as a mixture of the class conditional covariate distributions with unknown mixture weights, which studies Bayesian approaches but with heavy computation burden.

## 2 Discussion

As in practical applications, the information of target data is achievable but with expensive cost. Active learning may help choose “useful” small samples of target data to label and learn with good performance based on the additional information, which gives more information on  $Q(Y)$ . Therefore, combining  $P(Y), P(X|Y), Q(X|Y)$  and  $Q(Y)$ , we could map source domain to target domain more accurately, which promises to improve the performance of classifier.