

# Algorithm Outline

Omar Elsabrout

May 15, 2018

This document is a scratch book to formalize algorithms and preparing arguments for the internship report. It is meant to keep and update the insights and findings during the research process.

## 1 Preliminaries

$X \subset R^d$  denotes the instance space.

$Y = \{-1, +1\}$  denotes the label space.

$D_S(x, y)$  is the source distribution which is labeled.

$X_S$  is the set of instances in  $D_S$ .

$D_T(x)$  is the target distribution which is not labeled.

$X_T$  is the set of instances in  $D_T$ .

$h_D^*$  represents the hypothesis proposed by the DALC algorithm with the least possible validation risk.

$h_s^*(x) = \{-1, +1\}$  denotes the domain separator hypothesis to classify the points as  $X_S$  or  $X_T$  which is a linear classifier with +1 for source and -1 for target distribution.

$Y_T$  target labels.

## 2 Algorithm

**Result:**  $Y_T$

initialization;

Compute  $h_D^*$  on  $D_S(x, y)$  and  $D_T(x)$ ;

Compute  $h_s^*$  on  $X_S$  and  $X_T$ ;

**while**  $size(Y_T) \neq size(X_T)$  **do**

    instructions;

**if**  $\exists x_T$  where  $h_s^*(x_T) = +1$  **then**

$y_T = h_D^*(x_T)$

$y_T \in Y_T$

**else**

$y_T$  is the closest point to  $h_s^*(x)$

$y_T = h_D^*(x_T)$

**end**

    Update  $h_s^*(x) = \min H_S$  with the lowest classification errors;

    Update  $h_D^*$  on  $D_S(x, y)$  and the new  $D_T(x)$ ;

**end**

**Algorithm 1:** Enforcing active learning on the DALC Algorithm