Stacking With Auxiliary Features (SWAF)
Nazneen Fatema Rajani and Ray Mooney

Dept. of Computer Science at the University of Texas at Austin

INTRODUCTION

• Stacking (Wolpert, 1992) is a well known ensembling algorithm
• However, it does not adequately discriminate between base systems and input instances
• Stacking With Auxiliary Features (SWAF) integrates information from multiple sources
• Auxiliary Features enable the stacker to leverage relevant information to improve prediction
• We use two types of auxiliary features:
  • Instance features - enable the stacker to discriminate across instances
  • Provenance features - enable the stacker to discriminate across base systems

Figure 1: Our stacking approach to combining system outputs using confidence scores and two types of auxiliary features for improving prediction

TASK OVERVIEW

• We demonstrate SWAF on three very different machine learning problems
• Two of them are in NLP and third is a well known computer vision problem

Slot Filling (SF)

org: Microsoft
1. city_of_headquarters: 2. website: 3. subsidiaries: 4. employees: 5. shareholders: ...

Figure 2: SF involves building a Knowledge Base (KB) from scratch using pre-defined slots. Systems provide confidence score and provenance

Entity Discovery and Linking (EDL)

Hillary Clinton is a US Secretary of State, US Senator, and First Lady of the United States. From 2009 to 2013, she was the US Secretary of State, serving under President Barack Obama.

Figure 3: EDL involves detecting entity mentions in a corpus and linking them to an English KB (FreeBase). Systems provide confidence score and provenance

ImageNet Object Detection

Figure 4: Detect all instances of object categories (total 200) in images and localize using bounding boxes

AUXILIARY FEATURES

Instance Features

• Enables the stacker to discriminate between input instance types
• The intuition is that some systems are better at certain inputs that other systems
• Information about the input type would allow the classifier to make a better prediction
• Slot Filling – slot type (per: age, org: headquarters)
• Entity Detection and Linking – entity type (PER, ORG, GPE)
• Object Detection – object category (200) and VGGNet’s 5x7 features

Provenance Features

• Enables the stacker to discriminate between component systems
• The intuition is that output is reliable if systems agree on the source or provenance
• Information about provenance or source of the output would allow the classifier to make a better prediction
• Slot Filling :-
  1. Document provenance - For a given query and slot, for each system, i, there is a feature \( n_i \) :
     • \( N \) systems provide a fill for the slot.
     • Of these, \( n \) give same provenance document.
   2. Offset provenance - Degree of overlap between systems’ provenance strings. Uses Jaccard similarity coefficient.

RESULTS

Slot Filling

Entity Discovery and Linking

ImageNet Object Detection

Figure 12: Provenance features - the document provenance score. Uses Jaccard similarity coefficient.

Figure 13: Entity Detection and Linking (EDL) - Same as Slot Filling using the entity mention as provenance

Figure 14: Object Detection - Bounding box overlap measured using Jaccard similarity coefficient

Figure 15: ImageNet Object Detection - Detect all instances of object categories (total 200) in images and localize using bounding boxes