Bayesian Case Model (BCM): Generative Approach for Case-based Reasoning and Prototype Classification

Been Kim, Cynthia Rudin and Julie Shah
{beenkim, rudin, julie_a_shah}@csail.mit.edu

**Motivation**

Develop interactive machine learning framework to bring together data, machine learning algorithms and human experts’ knowledge.

Communication from machine to human: Provide intuitive explanations

First step: communication from human to machine: incorporate feedback

**Why Case-based Reasoning?**

1. **Case-based Reasoning (CBR) is intuitive.**
   - Example of CBR (simplified)
     a. My child is sick. He is having fever, nausea and coughs.
     b. Find the illness that has the most similar symptoms
     c. Apply the solution for flu: rest, ibuprofen, liquid.

2. **Exemplar-based reasoning is the way humans think – machine can better support peoples’ decision-making by representing data in the same way.**
   - Human’s tactical decision is based on exemplar-based reasoning.[1]
   - Skilled fire fighters use recognition-primed decision making — a situation is matched to typical cases. [2,3]

**Limitations of CBR**

- Does not leverage data
- Requires previous solutions
- Does not scale to complex problems


**Bayesian Case Model (BCM)**

Leverage the power of examples (prototypes) and hot features (subspaces) to explain machine learning results.

**Bayesian generative models** + **Case-based reasoning** = **Bayesian Case Model**

**Dataset**

- **a. Recipes**
  - Data from computer cooking contest: liris/cnrs.fr/ccc/ccc2014

- **b. USPS handwritten digits**
  - http://www.cs.nyu.edu/~roweis/data.html

**Prototypes and Subspaces from BCM**

- **1. Learned prototypes and subspaces**
  - Evaluated using features learned by BCM or LDA + SVM
  - Objective measure of human understanding
  - Participant’s task is to assign the ingredients of a specific dish (a new data point) to a cluster.

**Results**

- **1. BCM maintains accuracy while achieving interpretability**
  - Evaluated using features learned by BCM or LDA + SVM
  - Statistical significantly better performance with BCM (85.9% vs. 71.3%)

- **2. BCM improves humans understanding**
  - Objective measure of human understanding
  - Participant’s task is to assign the ingredients of a specific dish (a new data point) to a cluster.
  - Statistically significantly better performance with BCM (85.9% vs. 71.3%)

**An example of BCM**

- **Prototypes**
  - Raw data

- **Subspaces**
  - Cluster A
    - Taco
  - Cluster B
    - Basic crepe
  - Cluster C
    - Chocolate berry tart

**Clusters explained using**

- BCM: ingredients of the prototype recipe for each cluster other recipes name nor subspaces.
- LDA: representative ingredients of each cluster.