SMILe: Shuffled Multiple-Instance Learning

Gary Doran
gary.doran@case.edu

Soumya Ray
sray@case.edu

AAAI 2013
Resampling Approaches in Supervised Learning
Bagging
Bagging
Bagging
Bagging
Bagging
Bagging

Classifier 1

Classifier $k$
Bagging

Combine classifiers using voting
Multiple-Instance (MI) Learning
Multiple-Instance (MI) Learning
Multiple-Instance (MI) Learning

Instances

Bags
Multiple-Instance (MI) Learning

Instances

Bags
Multiple-Instance Bagging
Multiple-Instance Bagging
Multiple-Instance Bagging

Resample Entire Bags

(Zhou and Zhang 2003)
SMILEe: Shuffling to create new bags

1. Start with positive bags
SMILe: Shuffling to create new bags

1. Start with positive bags

2. Shuffle positive bag instances
SMILE: Shuffling to create new bags

1. Start with positive bags

2. Shuffle positive bag instances
SMILE: Shuffling to create new bags

1. Start with positive bags
2. Shuffle positive bag instances
3. Sample new bags and label positive
SMILE: Shuffling to create new bags

1. Start with positive bags
2. Shuffle positive bag instances
3. Sample new bags and label positive
SMILe: Shuffling to create new bags

1. Start with positive bags

2. Shuffle positive bag instances

3. Sample new bags and label positive
SMILE: Shuffling to create new bags

1. Start with positive bags

2. Shuffle positive bag instances

3. Sample new bags and label positive
Example: Content-Based Image Retrieval
Example: Content-Based Image Retrieval
Example: Content-Based Image Retrieval
Properties

• Low noise on shuffled bag labels
• Introduces additional constraints
• Additional advantages:
  • Reduced space
  • Improved interpretability
  • Can be used with any classifier
Low noise on shuffled bag labels
Low noise on shuffled bag labels
Low noise on shuffled bag labels
Low noise on shuffled bag labels
Low noise on shuffled bag labels

\[
\frac{3}{6} \times \frac{2}{5}
\]
Low noise on shuffled bag labels

\[
\frac{3}{6} \times \frac{2}{5} \times \frac{1}{4}
\]
Low noise on shuffled bag labels

\[
\frac{3}{6} \times \frac{2}{5} \times \frac{1}{4} = 5\%
\]
Bags as Constraints

“Positive”
Bags as Constraints

“Positive” OR “Positive”
Shuffling adds constraints
Shuffling adds constraints

Shuffled Bag
Shuffling adds constraints

{ "Positive" } OR { "Positive" }
# Variables and Tradeoffs

<table>
<thead>
<tr>
<th>Variable</th>
<th>Advantage</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>↑ Bag Size</td>
<td>Decrease Noise*</td>
<td>Less Powerful Constraint</td>
</tr>
<tr>
<td>↑ # of Bags</td>
<td>More Constraints</td>
<td>Decreasing Marginal Information</td>
</tr>
</tbody>
</table>

*See paper for formal results*
Experiments

• **Hypothesis:** The addition of shuffled bags should increase performance of MI discriminative classifiers.

• **Datasets:** 28 CBIR (image classification)

• **Methodology:** Add various numbers of shuffled bags to an MI dataset, and train using a set kernel classifier (Gärtner et al. 2002).
Shuffling Improves MI Classification

![Graph showing AUC for fox and tiger with labeled axes and legend](image)

- **AUC**
- **Shuffled Bags / Iterations**
- **SMILe**
- **Bagging**
Shuffling Improves MI Active Learning

stripednotebook

Shuffled Bags

Queries

AUC

0
10
20
30
40
50

0
0.70
0.72
0.74
0.76
0.78
0.80
0.82

0
20

Shuffled Bags

0
20
Shuffling helps when datasets are small

5 Initial Bags with 20 Shuffled Bags

Fraction Improved

Queries
Conclusions and Future Work

• SMILe: a new resampling approach for MIL that improves performance beyond bagging
• Helps especially well for small datasets
• Future Work:
  • Expand experiments (more base classifiers)
  • Explore more problem domains
  • Provide stronger theoretical justification