Introduction:

- Discriminative dictionaries are effective for classification. However, classification performance depends on:
  - Dictionary size (pre-defined),
  - Relationship between the dictionary atoms and the class labels (pre-defined), and
  - Suitable classifier for the sparse codes (often learned separately).

- A Bayesian approach is proposed to learn a discriminative dictionary jointly with a classifier.

- The learning process:
  - Automatically determines the required dictionary size,
  - Adaptsively learns the relationships between the dictionary atoms and the class labels, and
  - Strongly couples the classifier with the dictionary.

Core intuition:

- In Beta-Bernoulli Process (BP) [1], dictionary atoms relate to the training data under a set of Bernoulli distributions.
- Different sets of Bernoulli distribution are used for separate classes to induce discrimination in the dictionary.
- The same Bernoulli distributions are forced to be used in joint learning of the classifier, to encode class-wise dictionary atom popularity in the classifier.
- The sparse codes of test samples use the popular atoms of the correct class more frequently, thereby easily classified by the classifier.
- The non-parametric nature of the BP is exploited in determining the desired dictionary and classifier sizes.

References:


Acknowledgement:

This research was supported by ARC Discovery grant DP160101458 and it received funding from ARENA as part of ARENA’s Research and Development Programme.