An investigation of imitation learning algorithms for structured prediction

Imitation learning:
- Learns controllers by observing demonstrations by humans
- Infers the reward function instead of requiring it
- Must account for the dependencies between actions

Structured prediction is similar to imitation learning:
- Complex output spaces are formed by inter-dependent actions
- Labeled data shows us the correct actions to perform a task, but not how to overcome mistakes
- SEARN (Daumé et al., 2009) casted structured prediction as imitation learning successfully.
- DAgger (Ross et al., 2011) achieved better stability

Question: How do they compare on a complex structured prediction task?

SEARN versus DAgger in training:

Input: training data \( S \), optimal policy \( \pi^* \), loss function \( \ell \), learning rate \( \beta \)
Output: Hypothesis \( H_\beta \)
Examples \( E = \emptyset \)

\[
p = (1 - \beta)^{-1}
\]

\[
p = p\pi^* + (1 - p)H_{\pi^*}
\]

for \( i = 1 \) to \( N \) do

- current policy \( \pi = p\pi^* + (1 - p)H_{\pi^*} \)
- SEARN: Examples \( E = \emptyset \)
- SEARN learns from the data of each iteration separately
- Extract features \( \Phi = f(x, \hat{y}, \pi) \)
- foreach possible action \( y'_j \) do
  - SEARN: Predict \( y'_{i+1} = \pi(s | \hat{y}_{i+1}, y'_j) \)
  - DAgger: Predict \( y'_{i+1} = \pi(s | \hat{y}_{i+1}, y'_j) \)
- Estimate cost \( c'_{i, j} = (\|y_{i+1} - y'_{i+1}\|) \)
- Add \( (\Phi, c', E) \) to \( E \)
- Learn a classifier \( h \) from \( E \)
- SEARN: \( H_\beta = \beta\sum_{i=1}^{N} (1 - \beta)^{i-1} h_i \)
- DAgger: \( H_\beta = h \)

Focused costing ameliorates stochastic cost estimation

Experiment results:

F-score

Focused costing improves both SEARN and DAgger when training is stochastic. DAgger though is less sensitive.

DAgger is more stable than SEARN, converging in fewer iterations for a wide range of learning rates (0.1, 0.3, 0.7, 1), thus easier to use.

References


A. Vlachos and M. Craven, 2012. Biomedical event extraction from abstracts and full papers using search-based structured prediction. BICBioinformatics.