Cost-sensitive Dynamic Feature Selection

He He[1], Hal Daumé III[1] and Jason Eisner[2]

[1]University of Maryland, College Park

[2]Johns Hopkins University

June 30, 2012

Dynamic Feature Selection

Feature selection in real life is a *sequential decision-making* process.

Dynamic Feature Selection

Feature selection in real life is a *sequential decision-making* process.

Observation	cost	Decision		
coughing	free	cold	flu	H1N1
sore throat	free	cold	flu	H1N1
headache	free	cold	**flu**	H1N1

Dynamic Feature Selection

Feature selection in real life is a *sequential decision-making* process.

Observation	cost	Decision		
coughing	free	cold	flu	H1N1
sore throat	free	cold	flu	H1N1
headache	free	cold	**flu**	H1N1
temperature (101°)	\$1	cold	**flu**	H1N1
nasal swab test (pos.)	\$10	cold	flu	**H1N1**
viral culture test (pos.)	\$50	cold	flu	**H1N1**
molecular test (pos.)	\$100	cold	flu	**H1N1**
blood test (pos.)	\$100	cold	flu	**H1N1**

Dynamic Feature Selection

Feature selection in real life is a *sequential decision-making* process.

Observation	cost	Decision		
coughing	free	cold	flu	H1N1
sore throat	free	cold	flu	H1N1
headache	free	cold	**flu**	H1N1
temperature (101°)	$1	cold	**flu**	H1N1
nasal swab test (pos.)	$10	cold	flu	**H1N1**
viral culture test (pos.)	$50	cold	flu	**H1N1**
molecular test (pos.)	$100	cold	flu	**H1N1**
blood test (pos.)	$100	cold	flu	**H1N1**

Dynamic Feature Selection

Feature selection in real life is a *sequential decision-making* process.

Observation	cost	Decision		
coughing	free	cold	flu	H1N1
sore throat	free	cold	flu	H1N1
headache	free	cold	**flu**	H1N1
temperature (101°)	\$1	cold	**flu**	H1N1
nasal swab test (neg.)	\$10	cold	**flu**	H1N1

Dynamic Feature Selection

Feature selection in real life is a *sequential decision-making* process.

Observation	cost	Decision		
coughing	free	cold	flu	H1N1
sore throat	free	cold	flu	H1N1
headache	free	cold	**flu**	H1N1
temperature (101°)	\$1	cold	**flu**	H1N1
nasal swab test (neg.)	\$10	cold	**flu**	H1N1
viral culture test (pos.)	\$50	cold	flu	H1N1
molecular test (pos.)	\$100	cold	flu	**H1N1**
blood test (pos.)	\$100	cold	flu	**H1N1**

Cost-sensitive Dynamic Feature Selection

Feature Cost

- Computation time
- Data acquisition expense

Dynamic Selection

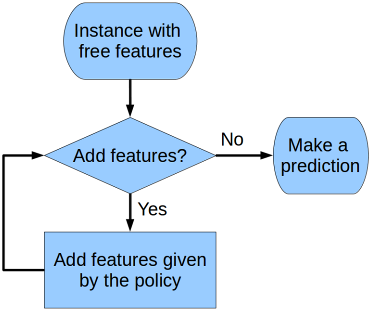
- Based on previous selected features and their values
- Compute features on-the-run

Given a pretrained classifier and feature cost,

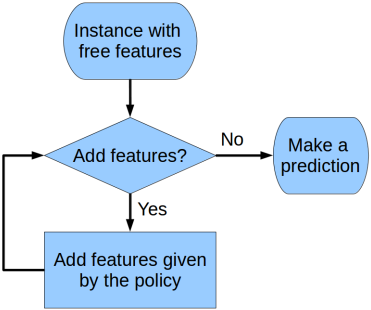
Goal

- Sequentially select features *for each instance at test time*
- Achieve a *user-specified* accuracy-cost trade-off

Dynamic Feature Selection as an MDP



Dynamic Feature Selection as an MDP



At time step t, for one example,

State s_t

Selected features and their values

Action $a_t \in A_t$

Acquire some features or stop

Policy π

Map from state to action: $\pi(s_t) = a_t$

Reward r

$r(s_t, a_t) = \text{margin}(s_t, a_t) - \lambda \cdot \text{cost}(s_t, a_t)$

margin: score of the true class - highest score of other classes

λ: trade-off parameter

Imitation Learning

Oracle

- Demonstrate optimal actions $\pi^*(s) = a_t^*$

Agent

- Learn a policy to mimic the oracle's behavior
- $\pi(s_t) = a_t$

Imitation via Supervised Classification

- Training examples $\{(\phi(s_{\pi^*}), \pi^*(s))\}$
- Feature: $\phi(s)$ label: $\pi^*(s)$ classifier: $\hat{\pi}$
- Minimize a *surrogate loss* $\ell(s, \pi)$ w.r.t. to π^*, e.g. hinge loss in SVM.

Forward Selection Oracle

- Select the feature that yields the maximum immediate reward

Forward Selection Oracle

- Select the feature that yields the maximum immediate reward

$r(s_t, a_t) = \text{margin}(s_t, a_t) - \text{cost}(s_t, a_t)$
$\lambda = 1$, cost scaled to $[0, 1]$, H1N1=positive

order	feat.	marg.	cost	reward
1	coughing, sore throat, headache	-0.20	0.00	-0.10

Forward Selection Oracle

- Select the feature that yields the maximum immediate reward

$r(s_t, a_t) = \text{margin}(s_t, a_t) - \text{cost}(s_t, a_t)$
$\lambda = 1$, cost scaled to $[0, 1]$, H1N1=positive

order	feat.	marg.	cost	reward
1	coughing, sore throat, headache	-0.20	0.00	-0.10
	temperature			
	nasal swab test			
	viral culture test			
	molecular test			
	blood test			

Forward Selection Oracle

- Select the feature that yields the maximum immediate reward

$r(s_t, a_t) = \text{margin}(s_t, a_t) - \text{cost}(s_t, a_t)$
$\lambda = 1$, cost scaled to $[0, 1]$, H1N1=positive

order	feat.	marg.	cost	reward
1	coughing, sore throat, headache	-0.20	0.00	-0.10
	temperature (101°)	-0.10	0.01	-0.11
	nasal swab test			
	viral culture test			
	molecular test			
	blood test			

Forward Selection Oracle

- Select the feature that yields the maximum immediate reward

$r(s_t, a_t) = \text{margin}(s_t, a_t) - \text{cost}(s_t, a_t)$
$\lambda = 1$, cost scaled to $[0, 1]$, H1N1=positive

order	feat.	marg.	cost	reward
1	coughing, sore throat, headache	-0.20	0.00	-0.10
	temperature			
	nasal swab test (pos.)	0.50	0.04	0.46
	viral culture test			
	molecular test			
	blood test			

Forward Selection Oracle

- Select the feature that yields the maximum immediate reward

$r(s_t, a_t) = \text{margin}(s_t, a_t) - \text{cost}(s_t, a_t)$
$\lambda = 1$, cost scaled to $[0, 1]$, H1N1=positive

order	feat.	marg.	cost	reward
1	coughing, sore throat, headache	-0.20	0.00	-0.10
	temperature			
	nasal swab test			
	viral culture test (pos.)	0.60	0.19	0.41
	molecular test			
	blood test			

Forward Selection Oracle

- Select the feature that yields the maximum immediate reward

$r(s_t, a_t) = \text{margin}(s_t, a_t) - \text{cost}(s_t, a_t)$
$\lambda = 1$, cost scaled to $[0, 1]$, H1N1=positive

order	feat.	marg.	cost	reward
1	coughing, sore throat, headache	-0.20	0.00	-0.10
	temperature			
	nasal swab test			
	viral culture test			
	molecular test (pos.)	0.70	0.38	0.32
	blood test			

Forward Selection Oracle

- Select the feature that yields the maximum immediate reward

$r(s_t, a_t) = \text{margin}(s_t, a_t) - \text{cost}(s_t, a_t)$
$\lambda = 1$, cost scaled to $[0, 1]$, H1N1=positive

order	feat.	marg.	cost	reward
1	coughing, sore throat, headache	-0.20	0.00	-0.10
	temperature			
	nasal swab test			
	viral culture test			
	molecular test			
	blood test (pos.)	0.65	0.38	0.27

Forward Selection Oracle

- Select the feature that yields the maximum immediate reward

$r(s_t, a_t) = \text{margin}(s_t, a_t) - \text{cost}(s_t, a_t)$
$\lambda = 1$, cost scaled to $[0, 1]$, H1N1=positive

order	feat.	marg.	cost	reward
1	coughing, sore throat, headache	-0.20	0.00	-0.10
	temperature			
2	nasal swab test (pos.)	0.50	0.04	0.46
	viral culture test			
	molecular test			
	blood test			

Forward Selection Oracle

- Select the feature that yields the maximum immediate reward

$r(s_t, a_t) = \text{margin}(s_t, a_t) - \text{cost}(s_t, a_t)$
$\lambda = 1$, cost scaled to $[0, 1]$, H1N1=positive

order	feat.	marg.	cost	reward
1	coughing, sore throat, headache	-0.20	0.00	-0.10
3	temperature (101°)	0.55	0.05	0.50
2	nasal swab test (pos.)	0.50	0.04	0.46
	viral culture test			
	molecular test			
	blood test			

Forward Selection Oracle

- Select the feature that yields the maximum immediate reward

$r(s_t, a_t) = \text{margin}(s_t, a_t) - \text{cost}(s_t, a_t)$
$\lambda = 1$, cost scaled to $[0, 1]$, H1N1=positive

order	feat.	marg.	cost	reward
1	coughing, sore throat, headache	-0.20	0.00	-0.10
3	temperature (101°)	0.55	0.05	0.50
2	nasal swab test (pos.)	0.50	0.04	0.46
4	viral culture test (pos.)	0.80	0.24	0.56
	molecular test			
	blood test			

Forward Selection Oracle

- Select the feature that yields the maximum immediate reward

$r(s_t, a_t) = \text{margin}(s_t, a_t) - \text{cost}(s_t, a_t)$
$\lambda = 1$, cost scaled to $[0, 1]$, H1N1=positive

order	feat.	marg.	cost	reward
1	coughing, sore throat, headache	-0.20	0.00	-0.10
3	temperature (101°)	0.55	0.05	0.50
2	nasal swab test (pos.)	0.50	0.04	0.46
4	viral culture test (pos.)	0.80	0.24	0.56
	molecular test			
5	blood test (pos.)	0.90	0.62	0.28

Forward Selection Oracle

- Select the feature that yields the maximum immediate reward

$r(s_t, a_t) = \text{margin}(s_t, a_t) - \text{cost}(s_t, a_t)$
$\lambda = 1$, cost scaled to $[0, 1]$, H1N1=positive

order	feat.	marg.	cost	reward
1	coughing, sore throat, headache	-0.20	0.00	-0.10
3	temperature (101°)	0.55	0.05	0.50
2	nasal swab test (pos.)	0.50	0.04	0.46
4	viral culture test (pos.)	0.80	0.24	0.56
6	molecular test (pos.)	0.95	1.00	-0.05
5	blood test (pos.)	0.90	0.62	0.28

Forward Selection Oracle

- Select the feature that yields the maximum immediate reward
- Stop in the global maximum-reward state
  free $\rightarrow$ nasal swab test $\rightarrow$ temperature $\rightarrow$ viral culture test $\rightarrow$ stop

$r(s_t, a_t) = \text{margin}(s_t, a_t) - \text{cost}(s_t, a_t)$
$\lambda = 1$, cost scaled to $[0, 1]$, H1N1=positive

order	feat.	marg.	cost	reward
1	coughing, sore throat, headache	-0.20	0.00	-0.10
3	temperature (101°)	0.55	0.05	0.50
2	nasal swab test (pos.)	0.50	0.04	0.46
4	viral culture test (pos.)	0.80	0.24	0.56
6	molecular test (pos.)	0.95	1.00	-0.05
5	blood test (pos.)	0.90	0.62	0.28

Forward Selection Oracle

- Select the feature that yields the maximum immediate reward
- Stop in the global maximum-reward state
  free $\rightarrow$ nasal swab test $\rightarrow$ temperature $\rightarrow$ viral culture test $\rightarrow$ stop
- Have access to the ground truth, *only available during training*

$r(s_t, a_t) = \text{margin}(s_t, a_t) - \text{cost}(s_t, a_t)$
$\lambda = 1$, cost scaled to $[0, 1]$, H1N1=positive

order	feat.	marg.	cost	reward
1	coughing, sore throat, headache	-0.20	0.00	-0.10
3	temperature (101°)	0.55	0.05	0.50
2	nasal swab test (pos.)	0.50	0.04	0.46
4	viral culture test (pos.)	0.80	0.24	0.56
6	molecular test (pos.)	0.95	1.00	-0.05
5	blood test (pos.)	0.90	0.62	0.28

Policy Features

At step 2, compute $\phi(s_2)$

feat.	cold	flu	H1N1	cost
free	0.20	0.50	0.30	0.00
swab	0.04	0.23	0.73	0.04

Policy Features

- selected features
  e.g. free = coughing, sore
  throat, headache;
  nasal swab test = pos.

At step 2, compute $\phi(s_2)$

feat.	cold	flu	H1N1	cost
free	0.20	0.50	0.30	0.00
swab	0.04	0.23	0.73	0.04

Policy Features

- selected features
  e.g. free = coughing, sore
  throat, headache;
  nasal swab test = pos.
- confidence score and its change
  e.g. $0.04, 0.23, 0.73$;
  $-0.16, -0.27, 0.43$

At step 2, compute $\phi(s_2)$

feat.	cold	flu	H1N1	cost
free	0.20	0.50	0.30	0.00
swab	0.04	0.23	0.73	0.04

Policy Features

- selected features
  e.g. free = coughing, sore
  throat, headache;
  nasal swab test = pos.
- confidence score and its change
  e.g. $0.04, 0.23, 0.73$;
  $-0.16, -0.27, 0.43$
- Does the guess change?
  e.g. Yes

At step 2, compute $\phi(s_2)$

feat.	cold	flu	H1N1	cost
free	0.20	0.50	0.30	0.00
swab	0.04	0.23	0.73	0.04

Policy Features

- selected features
  e.g. free = coughing, sore
  throat, headache;
  nasal swab test = pos.
- confidence score and its change
  e.g. $0.04, 0.23, 0.73$;
  $-0.16, -0.27, 0.43$
- Does the guess change?
  e.g. Yes
- cost and its change
  e.g. 0.04; 0.04

At step 2, compute $\phi(s_2)$

feat.	cold	flu	H1N1	cost
free	0.20	0.50	0.30	0.00
swab	0.04	0.23	0.73	0.04

Policy Features

- selected features
  e.g. free = coughing, sore throat, headache;
  nasal swab test = pos.
- confidence score and its change
  e.g. $0.04, 0.23, 0.73;$
  $-0.16, -0.27, 0.43$
- Does the guess change?
  e.g. Yes
- cost and its change
  e.g. 0.04; 0.04
- cost divided by confidence score
  e.g. 1.00, 5.75, 18.25

At step 2, compute $\phi(s_2)$

feat.	cold	flu	H1N1	cost
free	0.20	0.50	0.30	0.00
swab	0.04	0.23	0.73	0.04

Policy Features

- selected features
  e.g. free = coughing, sore throat, headache;
  nasal swab test = pos.
- confidence score and its change
  e.g. $0.04, 0.23, 0.73$;
  $-0.16, -0.27, 0.43$
- Does the guess change?
  e.g. Yes
- cost and its change
  e.g. 0.04; 0.04
- cost divided by confidence score
  e.g. 1.00, 5.75, 18.25
- current guess
  e.g. H1N1

At step 2, compute $\phi(s_2)$

feat.	cold	flu	H1N1	cost
free	0.20	0.50	0.30	0.00
swab	0.04	0.23	0.73	0.04

Imitation Learning via Classification

s_π: states visited by π T: task horizon
$J(\pi)$: task loss (negative reward) of π

Theorem

Ross & Bagnell (2010) Let $\mathbb{E}_{s_{\pi^*}}[\ell(s,\pi)] = \epsilon$, *then* $J(\pi) \leq J(\pi^*) + T^2\epsilon$.

Imitation Learning via Classification

s_π: states visited by π T: task horizon
$J(\pi)$: task loss (negative reward) of π

Theorem

Ross & Bagnell (2010) Let $\mathbb{E}_{s_{\pi^*}}[\ell(s,\pi)] = \epsilon$, *then* $J(\pi) \leq J(\pi^*) + T^2\epsilon$.

Why do we have quadratically increasing loss? ☹

Imitation Learning via Classification

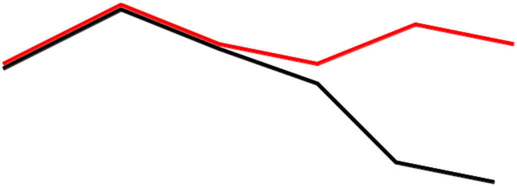
s_π: states visited by π $\quad$ T: task horizon
$J(\pi)$: task loss (negative reward) of π

Theorem

Ross & Bagnell (2010) Let $\mathbb{E}_{s_{\pi^*}}[\ell(s,\pi)] = \epsilon$, *then* $J(\pi) \leq J(\pi^*) + T^2\epsilon$.

Why do we have quadratically increasing loss? ☹

- Trains only under states the oracle visited
- Ignores the difference between the oracle's and the agent's state distribution



Dataset Aggregation (DAgger) (Ross et al. (2011))

Let $\pi_1 = \pi^*$. In iteration i,

execute policy π_i and collect dataset $\mathcal{D}_i = \{(\phi(s_{\pi_i}), \pi^*(s))\}$;

learn π_{i+1} from the aggregated dataset $\mathcal{D}_1 \bigcup \mathcal{D}_2 \bigcup \cdots \bigcup \mathcal{D}_i$.

Return the best policy evaluated on validation set.

Dataset Aggregation (DAgger) (Ross et al. (2011))

Let $\pi_1 = \pi^*$. In iteration i,

execute policy π_i and collect dataset $\mathcal{D}_i = \{(\phi(s_{\pi_i}), \pi^*(s))\}$;

learn π_{i+1} from the aggregated dataset $\mathcal{D}_1 \bigcup \mathcal{D}_2 \bigcup \cdots \bigcup \mathcal{D}_i$.

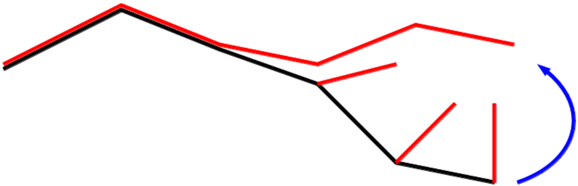
Return the best policy evaluated on validation set.

$Q_t^{\pi'}(s,\pi)$: t-step cost of executing π initially then running π'

$\epsilon_N = \min_{\pi \in \Pi} \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}_{s_{\pi_i}}[\ell(s,\pi)]$

Theorem

Ross et al. (2011) For DAgger, if $Q_{T-t+1}^{\pi^*}(s,\pi) - Q_{T-t+1}^{\pi^*}(s,\pi^*) \le u$ *and* N *is* $\tilde{O}(uT)$, *there is a policy* $\pi \in \pi_{1:N}$ *s.t.* $J(\pi) \le J(\pi^*) + uT\epsilon_N + O(1)$.



Coaching

The oracle's policy can be too good to learn!

- Far from the agent's learning space
- Policy features are insufficient

Coaching

The oracle's policy can be too good to learn!

- Far from the agent's learning space *use kernels*
- Policy features are insufficient *more descriptive features*

Coaching

The oracle's policy can be too good to learn!

- Far from the agent's learning space *use kernels* Overhead!
- Policy features are insufficient *more descriptive features*

Coaching

The oracle's policy can be too good to learn!

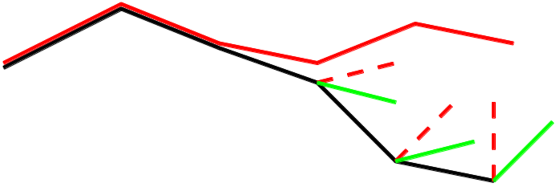
- Far from the agent's learning space *use kernels* Overhead!
- Policy features are insufficient *more descriptive features*

"Hope" action (McAllester et al. (2010); Chiang et al. (2009))

Combines the current policy's preference and the reward:

$$\tilde{a}_t^* = \arg\max_{a \in A_t} \eta \cdot \text{score}_{\pi_i}(a) + r(s_t, a)$$

$$\text{instead of} \quad a_t^* = \arg\max_{a \in A_t} r(s_t, a)$$



Baselines ($|w|$/cost, Forward): add feature statically from a ranked list
Trade-off: $\lambda = (0, 0.1, 0.25, 0.5, 1, 1.5, 2)$
Coaching: initialize η to 1 and decrease by e^{-1} in each iteration

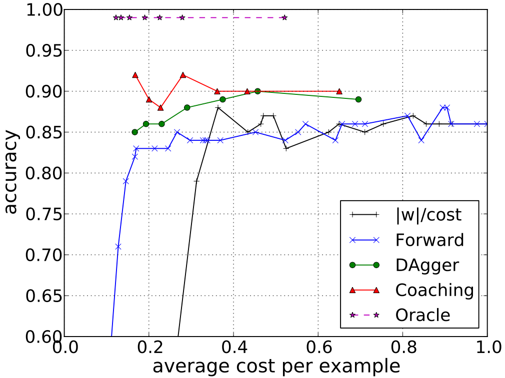


Figure: Radar (binary).

Experimental Results

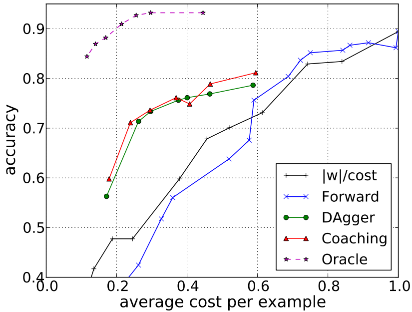


Figure: Digit (10 classes).

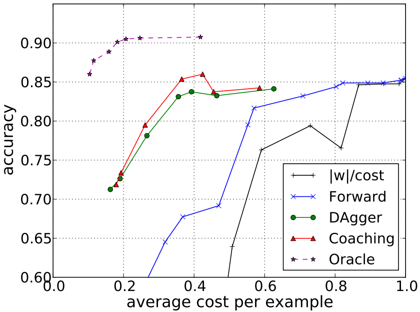


Figure: Segmentation (7 classes).

Conclusion and Future Work

Conclusion

- Feature selection as an MDP
- Imitation learning techniques
- Iterative policy training
- Coaching as a "local update" method

Future Work

- Include feature dependency using feature templates
- Learn feature weights jointly with the policy
- Apply to ensemble learning (select model dynamically)
- Structured prediction problem where
  - policy features might require inference under features selected so far
  - feature cost may need to be inferred at runtime

Bibliography

Chiang, David, Knight, Kevin, and Wang, Wei. 11,001 new features for statistical machine translation. In *NAACL-HLT*, 2009.

McAllester, D., Hazan, T., and Keshet, J. Direct loss minimization for structured prediction. In *NIPS*, 2010.

Ross, Stéphane and Bagnell, J. Andrew. Efficient reductions for imitation learning. In *AISTATS*, 2010.

Ross, Stéphane., Gordon, Geoffrey J., and Bagnell, J. Andrew. A reduction of imitation learning and structured prediction to no-regret online learning. In *AISTATS*, 2011.