

Interacting with VW in active learning

Nikos Karampatziakis
Cloud and Information Sciences Lab
Microsoft

NIPS 2013

Active Learning (in VW)

- ▶ Situation: unlabeled (+little labeled) data

Active Learning (in VW)

- ▶ Situation: unlabeled (+little labeled) data
- ▶ Start with whatever data is available

Active Learning (in VW)

- ▶ Situation: unlabeled (+little labeled) data
- ▶ Start with whatever data is available
- ▶ Learner **interacts** with teacher to learn more

Active Learning (in VW)

- ▶ Situation: unlabeled (+little labeled) data
- ▶ Start with whatever data is available
- ▶ Learner **interacts** with teacher to learn more
- ▶ In VW interaction consists of

Active Learning (in VW)

- ▶ Situation: unlabeled (+little labeled) data
- ▶ Start with whatever data is available
- ▶ Learner **interacts** with teacher to learn more
- ▶ In VW interaction consists of
 - ▶ Teacher gives unlabeled example to VW

Active Learning (in VW)

- ▶ Situation: unlabeled (+little labeled) data
- ▶ Start with whatever data is available
- ▶ Learner **interacts** with teacher to learn more
- ▶ In VW interaction consists of
 - ▶ Teacher gives unlabeled example to VW
 - ▶ VW decides whether and how much it needs that label

Active Learning (in VW)

- ▶ Situation: unlabeled (+little labeled) data
- ▶ Start with whatever data is available
- ▶ Learner **interacts** with teacher to learn more
- ▶ In VW interaction consists of
 - ▶ Teacher gives unlabeled example to VW
 - ▶ VW decides whether and how much it needs that label
 - ▶ Teacher can provide the label (not obliged to)

Active Learning (in VW)

- ▶ Situation: unlabeled (+little labeled) data
- ▶ Start with whatever data is available
- ▶ Learner **interacts** with teacher to learn more
- ▶ In VW interaction consists of
 - ▶ Teacher gives unlabeled example to VW
 - ▶ VW decides whether and how much it needs that label
 - ▶ Teacher can provide the label (not obliged to)
- ▶ The result is an importance weighted dataset

Active Learning (in VW)

- ▶ Situation: unlabeled (+little labeled) data
- ▶ Start with whatever data is available
- ▶ Learner **interacts** with teacher to learn more
- ▶ In VW interaction consists of
 - ▶ Teacher gives unlabeled example to VW
 - ▶ VW decides whether and how much it needs that label
 - ▶ Teacher can provide the label (not obliged to)
- ▶ The result is an importance weighted dataset
- ▶ **No selection bias** so can be used any way you like

active_interactor.py

- ▶ A simple demonstration of how to interact with VW

active_interactor.py

- ▶ A simple demonstration of how to interact with VW
- ▶ Assuming

active_interactor.py

- ▶ A simple demonstration of how to interact with VW
- ▶ Assuming
 - ▶ binary classification

active_interactor.py

- ▶ A simple demonstration of how to interact with VW
- ▶ Assuming
 - ▶ binary classification
 - ▶ examples are in human readable form (text)

active_interactor.py

- ▶ A simple demonstration of how to interact with VW
- ▶ Assuming
 - ▶ binary classification
 - ▶ examples are in human readable form (text)
- ▶ Connects to the host:port VW is listening on

active_interactor.py

- ▶ A simple demonstration of how to interact with VW
- ▶ Assuming
 - ▶ binary classification
 - ▶ examples are in human readable form (text)
- ▶ Connects to the host:port VW is listening on
- ▶ Sends any initially available labeled data

active_interactor.py

- ▶ A simple demonstration of how to interact with VW
- ▶ Assuming
 - ▶ binary classification
 - ▶ examples are in human readable form (text)
- ▶ Connects to the host:port VW is listening on
- ▶ Sends any initially available labeled data
- ▶ Sends unlabeled examples one by one

active_interactor.py

- ▶ A simple demonstration of how to interact with VW
- ▶ Assuming
 - ▶ binary classification
 - ▶ examples are in human readable form (text)
- ▶ Connects to the host:port VW is listening on
- ▶ Sends any initially available labeled data
- ▶ Sends unlabeled examples one by one
 - ▶ Waits for VW's response

active_interactor.py

- ▶ A simple demonstration of how to interact with VW
- ▶ Assuming
 - ▶ binary classification
 - ▶ examples are in human readable form (text)
- ▶ Connects to the host:port VW is listening on
- ▶ Sends any initially available labeled data
- ▶ Sends unlabeled examples one by one
 - ▶ Waits for VW's response
 - ▶ If VW does not want the label, sends the next

active_interactor.py

- ▶ A simple demonstration of how to interact with VW
- ▶ Assuming
 - ▶ binary classification
 - ▶ examples are in human readable form (text)
- ▶ Connects to the host:port VW is listening on
- ▶ Sends any initially available labeled data
- ▶ Sends unlabeled examples one by one
 - ▶ Waits for VW's response
 - ▶ If VW does not want the label, sends the next
 - ▶ Otherwise, VW's response includes an importance

active_interactor.py

- ▶ A simple demonstration of how to interact with VW
- ▶ Assuming
 - ▶ binary classification
 - ▶ examples are in human readable form (text)
- ▶ Connects to the host:port VW is listening on
- ▶ Sends any initially available labeled data
- ▶ Sends unlabeled examples one by one
 - ▶ Waits for VW's response
 - ▶ If VW does not want the label, sends the next
 - ▶ Otherwise, VW's response includes an importance
 - ▶ Asks the user for the label

active_interactor.py

- ▶ A simple demonstration of how to interact with VW
- ▶ Assuming
 - ▶ binary classification
 - ▶ examples are in human readable form (text)
- ▶ Connects to the host:port VW is listening on
- ▶ Sends any initially available labeled data
- ▶ Sends unlabeled examples one by one
 - ▶ Waits for VW's response
 - ▶ If VW does not want the label, sends the next
 - ▶ Otherwise, VW's response includes an importance
 - ▶ Asks the user for the label
 - ▶ If user skips, sends the next

active_interactor.py

- ▶ A simple demonstration of how to interact with VW
- ▶ Assuming
 - ▶ binary classification
 - ▶ examples are in human readable form (text)
- ▶ Connects to the host:port VW is listening on
- ▶ Sends any initially available labeled data
- ▶ Sends unlabeled examples one by one
 - ▶ Waits for VW's response
 - ▶ If VW does not want the label, sends the next
 - ▶ Otherwise, VW's response includes an importance
 - ▶ Asks the user for the label
 - ▶ If user skips, sends the next
 - ▶ Otherwise, we have a new labeled weighted example

active_interactor.py

- ▶ A simple demonstration of how to interact with VW
- ▶ Assuming
 - ▶ binary classification
 - ▶ examples are in human readable form (text)
- ▶ Connects to the host:port VW is listening on
- ▶ Sends any initially available labeled data
- ▶ Sends unlabeled examples one by one
 - ▶ Waits for VW's response
 - ▶ If VW does not want the label, sends the next
 - ▶ Otherwise, VW's response includes an importance
 - ▶ Asks the user for the label
 - ▶ If user skips, sends the next
 - ▶ Otherwise, we have a new labeled weighted example
 - ▶ Sends it to VW (causes update).

active_interactor.py

- ▶ A simple demonstration of how to interact with VW
- ▶ Assuming
 - ▶ binary classification
 - ▶ examples are in human readable form (text)
- ▶ Connects to the host:port VW is listening on
- ▶ Sends any initially available labeled data
- ▶ Sends unlabeled examples one by one
 - ▶ Waits for VW's response
 - ▶ If VW does not want the label, sends the next
 - ▶ Otherwise, VW's response includes an importance
 - ▶ Asks the user for the label
 - ▶ If user skips, sends the next
 - ▶ Otherwise, we have a new labeled weighted example
 - ▶ Sends it to VW (causes update).
 - ▶ Saves it to a file, so **can quit anytime**.